

Univerzita Karlova v Praze

Filozofická fakulta

Katedra sociologie

Sociologie

Petra A n ý ž o v á

**Mezinárodní srovnatelnost postojových škál
sociálního kapitálu a politického odcizení
v Evropském sociálním výzkumu**

**Cross-Country Comparability of Social
Capital and Political Disaffection Attitude
Scales in the European Social Survey**

Disertační práce

vedoucí práce - PhDr. Jiří Vinopal, Ph.D.

2014

Poděkování:

Ráda bych poděkovala svému školiteli PhDr. Jiřímu Vinopalovi, Ph.D. za jeho čas, ochotu, podnětné poznámky a jeho důvěru, která mi poskytla dostatek volného prostoru pro vypracování této práce. Dále jsem mu také nesmírně vděčná za inspiraci a podporu při směřování do metodologie sociálních výzkumů. Velice si cením materiální podpory ze strany Vysoké školy finanční a správní a děkuji za užitečné rady při psaní této práce mým kolegům. V neposlední řadě bych ale ráda vyjádřila úctu a díky své rodině za její neutuchající psychickou oporu, energii a ohledy a za poskytnutí tak komfortního prostředí k napsání této disertace.

Prohlášení:

„Prohlašuji, že jsem tuto disertační práci napsala samostatně s využitím pouze uvedených a řádně citovaných pramenů a literatury a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.“

V Praze dne 14. února 2014

Petra Anýžová

Abstrakt

Práce se zabývá metodologickým zázemím komparativních mezinárodních výzkumů, jejichž význam v současné globální době nabývá na důležitosti. Sociální vědci analyzují atraktivní mezinárodní data čím dál častěji, bohužel často bez vědomí toho, jaké úrovně srovnatelnosti tato data vlastně dosahují a jakým způsobem je s nimi tedy možné v analýze dat statisticky zacházet a dále je interpretovat v komparativní perspektivě. Tato práce představuje ekvivalenci dat jako jeden z nejdůležitějších aspektů kvality mezinárodního výzkumu a zaměřuje se speciálně na problematiku postojových škál, jejichž srovnatelnost je v komparativních výzkumech ohrožena nejvíce. Předmětem rozboru je zde konkrétně mezinárodní srovnatelnost dvou v současnosti velmi často využívaných robustních postojových škál, a to škály sociálního kapitálu a politického odcizení, tak jak jsou měřeny v Evropském sociálním výzkumu. Oba dva tyto sociálně vědní koncepty se v sociologii i politologii řadí mezi velmi významné, řadu let teoreticky pojednáváné a empiricky zkoumané sociální jevy. Nežřídká mají i reálný dopad na sociální a politickou praxi, když uvážíme, že podávají vysvětlení k povaze sociální soudržnosti a k příčinám nízké volební účasti v současných moderních společnostech. Avšak jen málokteré empirické studie doplňují své komparativní závěry statisticky ověřenou a potvrzenou úrovní srovnatelnosti těchto postojových škál, což je z metodologického hlediska nedostačující. Tato studie na základě strukturního modelování dokládá, že i v evropských zemích, které jsou si značně kulturně, historicky i sociálně blízké, lze sice komparativně porovnávat vztahy mezi těmito postojovými škálami a jinými vysvětlujícími proměnnými, nikoli však již žádné jiné jednodušší statistické ukazatele typu aritmetický průměr škály či položky, atp. Tudíž že i komparace takto robustních postojových škál má určité limity, které je nutné brát ať už v primární či sekundární analýze dat v potaz.

Klíčová slova: mezinárodní výzkum, ekvivalence, srovnávání, postojová škála, měřicí položka, strukturní modelování, Evropský sociální výzkum, sociální kapitál, politické odcizení

Abstract

The doctoral thesis deals with the methodological perspectives of comparative cross-national research. Currently, in the globalization period, the importance of these perspectives has been increasing. Unfortunately, social scientists analyse attractive international data more and more often without realising the levels of data comparability. Therefore they are not fully aware the best approach to statistical analysis of these data and their convenient comparative interpretation. This thesis introduces the issue of data equivalence as one of the most important quality aspects of cross-national research and focuses especially on attitude scales owing to the fact that their comparability is endangered the most in comparative research. In particular, the subject of the analysis is the international comparability of two of the most frequent robust attitude scales: namely social capital and political disaffection scales as they are measured in the European Social Survey first data set. In sociology and political science, these two social science concepts are very significant and they have been studied both theoretically and empirically as social phenomena for very long time. Quite often, they have factual social and political impact as well when we realise that they can explain social cohesion character and low voter turnout causes in modern societies. However, only a few empirical studies and their comparative outcomes are complemented appropriately with statistically verified and proven comparability levels of these attitude scales which is far from being sufficient from the methodology point of view. On the basis of structural modelling, this thesis documents that even in the European countries which are considerably culturally, historically, and socially close, there is a possibility to carry out comparative juxtaposition of the relations between these attitude scales and other clarifying variables from the data set, however it is not possible to compare any other statistical indicators like arithmetic mean of scale or items etc. Consequently, the comparison of so robust attitude scales has its certain limits which are necessary to be taken into consideration both in the primary and secondary data analyses.

Key words:

cross-national survey, equivalence, comparison, attitude scale, measurement indicator, structural modelling, The European Social Survey, social capital, political disaffection

OBSAH

SEZNAM POUŽITÝCH ZKRATEK A SYMBOLŮ	11
ÚVOD.....	15
1 KOMPARATIVNÍ POVAHA SOCIÁLNÍCH VÝZKUMŮ	17
2 KONCEPT EKVIVALENCE V KOMPARATIVNÍCH VÝZKUMECH	19
2.1 Kategorizace ekvivalence.....	21
2.1.1 <i>Kategorizace ekvivalence na teoretické úrovni</i>	<i>21</i>
2.1.2 <i>Kategorizace ekvivalence na úrovni měření</i>	<i>22</i>
2.2 Základní předpoklady ekvivalence položek.....	24
2.2.1 <i>Výběr etických a emických položek.....</i>	<i>25</i>
2.2.2 <i>Technika překladu dotazníku</i>	<i>26</i>
2.2.3 <i>Metody využívající kognitivních přístupů</i>	<i>28</i>
2.3 Techniky testování ekvivalence položek	29
2.3.1 <i>Statistická analýza prvního a druhého stupně</i>	<i>30</i>
2.3.2 <i>Mnohorozměrné statistické analýzy.....</i>	<i>31</i>
2.3.2.1 <i>Strukturní modelování.....</i>	<i>31</i>
2.3.2.2 <i>Víceúrovňové modelování</i>	<i>33</i>
2.3.2.3 <i>Teorie odpovědi na položku.....</i>	<i>34</i>
2.3.2.4 <i>Analýza latentních tříd</i>	<i>35</i>
3 POSTOJOVÉ ŠKÁLY SOCIÁLNÍHO KAPITÁLU A POLITICKÉHO ODCIZENÍ	37
3.1 Koncept sociálního kapitálu a jeho dimenze.....	38
3.2 Koncept politického odcizení a jeho dimenze	42
4 ZÁKLADNÍ ANALÝZY SROVNATELNOSTI VYBRANÝCH POSTOJOVÝCH ŠKÁL.....	45
4.1 Korelační analýza.....	45
4.2 Analýza reliability	49
4.3 Explorativní faktorová analýza	53
5 TESTOVÁNÍ SROVNATELNOSTI ŠKÁL POMOCÍ STRUKTURNÍHO MODELOVÁNÍ	59
5.1 Analýza ekvivalence měření ve faktorovém modelu měření.....	61
5.2 Analýza strukturní ekvivalence ve strukturním modelu	66
6 ANALÝZA ZÁKLADNÍHO STRUKTURNÍHO MODELU POSTOJOVÝCH ŠKÁL	69
6.1 Základní předpoklady testování strukturního modelu.....	71
6.2 Systematické chyby měření položek z postojových škál	80
7 SEPARÁTNÍ ANALÝZA SROVNATELNOSTI POSTOJOVÝCH ŠKÁL	84
8 SIMULTÁNNÍ ANALÝZA SROVNATELNOSTI POSTOJOVÝCH ŠKÁL.....	86

9	SIMULTÁNNÍ ANALÝZA KULTURNĚ HOMOGENNÍCH SKUPIN ZEMÍ	98
9.1	Simultánní analýza zemí střední a jihovýchodní Evropy.....	98
9.1.1	Česká republika, Maďarsko, Polsko a Slovinsko	98
9.1.2	Německo, Rakousko a Švýcarsko.....	106
9.2	Simultánní analýza zemí jižní a západní Evropy	112
9.3	Simultánní analýza zemí severní Evropy	120
9.4	Simultánní analýza zemí západní Evropy	127
9.4.1	Spojené království, Irsko a Nizozemsko	127
9.4.2	Benelux – Nizozemsko, Belgie a Lucembursko.....	133
9.4.3	Francie, Belgie a Lucembursko	139
9.5	Shrnutí výsledků simultánní analýzy dat	146
10	SIMULTÁNNÍ ČASOVÁ ANALÝZA SROVNATELNOSTI ŠKÁL DŮVĚRY	153
	ZÁVĚR.....	167
	LITERATURA.....	170
	PŘÍLOHY	181
	Příloha 1: znění vybraných otázek.....	181
	Příloha 2: diagram strukturního modelu škál sociálního kapitálu a politického odcizení.....	185
	Příloha 3: syntax pro výpočet polychoriální korelační matice a test základního modelu metodou odhadu DWLS v LISRELU	186
	Příloha 4: grafické srovnání latentních průměrů sociálního kapitálu a politického odcizení.	187
	Příloha 5: separátní analýzy základního modelu škál důvěry v šesti vlnách šetření ESS	188
	SUMMARY	194

SEZNAM POUŽITÝCH ZKRATEK A SYMBOLŮ

ESS	Evropský sociální výzkum / European Social Survey
ISA	Mezinárodní sociologická asociace / International Sociological Association
ASA	Americká sociologická společnost / American Sociological Association
EB	Eurobarometr / Eurobarometer
GSS	Obecný sociální výzkum / General Social Survey
LSMS	Studie životní úrovně / Living Standards Measurement Study
WVS	Světový výzkum hodnot / World Value Survey
EVS	Evropský výzkum hodnot / European Social Survey
ISSP	Mezinárodní sociální výzkumný program / International Social Survey Programme
PISA	Program pro mezinárodní hodnocení žáků / Programme for International Student Assessment
PIAAC	Mezinárodní výzkum dospělých / Programme for International Assessment of Adult Competencies
OECD	Organizace pro hospodářskou spolupráci a rozvoj / Organisation for Economic Co-operation and Development
EQLS	Evropský výzkum kvality života / European Quality of Life Survey
ASQ	Model překladu stejných otázek / „ask-the same-question“
ADQ	Model překladu různých otázek / „ask-different-question“
TRAPD	Metoda překladu - Translation (překlad), Review (revize), Adjudication (posouzení), Pretesting (předběžné testování), Documentation (dokumentace)
MTMM	Multitrait-multimethod design výzkumu / Multitrait-multimethod model
CFA	Konfirmační faktorová analýza / Confirmatory Factor Analysis
MGCFA	Několika výběrová konfirmační faktorová analýza / Multigroup Confirmatory Factor Analysis
EFA	Explorativní faktorová analýza / Exploratory Factor Analysis
PCA	Analýza hlavních komponent / Principal Component Analysis
KMO	Kaiser-Meyer-Olkinova míra / Kaiser-Meyer-Olkin Measure
CATPCA	Kategorická analýza hlavních komponent / Categorical Principal Component Analysis
SEM	Strukturní modelování / Structural Equation Modelling
MSEM	Víceúrovňové strukturní modelování / Multilevel Structural Equation Modelling
ANCOV	Analýza modelů kovariančních struktur / Analysis of Covariance Structures
MACS	Analýza průměrů a kovariančních struktur / Mean and Covariance Structure Analysis
BIC	Bayesovské informační kritérium / Bayes Information Criterion
CFI	Komparativní index shody / Comparative fit index
GFI	Index dobré shody / Goodness-of-Fit Index
AGFI	Adjustovaný index dobré shody / Adjusted Goodness-of-Fit Index
RMSEA	Index odmocniny z průměrného čtverce chyby odhadu / Root Mean Square Error of Approximation
TLI	Tucker-Lewisův index / Tucker-Lewis Index
χ^2/df	Relativní chí-kvadrát / (Minimum Discrepancy/Degrees of freedom)
M.I.	Modifikační index / Modification Index
Par Change	Změna velikosti parametru / Parameter Change
S.E.	Standardní (střední) chyba průměru / Standard Error of the Mean
FIML	Metoda maximální věrohodnosti s úplnou informací / Full Information Maximum Likelihood
ML	Metoda maximální věrohodnosti / Maximum Likelihood
EM	Metoda očekávané maximalizace / Expected Maximalization
DWLS	Metoda diagonálně vážených nejmenších čtverců / Diagonally Weighted Least Squares
C.I.	Interval spolehlivosti / Confidence Interval
IRT	Teorie odpovědi na položku / Item Response Theory
CTT	Klasická teorie testů / Classical Test Theory
DIF	Odlíšné fungování položek (bias) / Differential Item Functioning
ICC	Charakteristická křivka položky / Item Characteristic Curve

GRM	Model stupňovaných odpovědí / Graded Response Model
LCM	Analýza latentních tříd / Latent Class Modeling
CAPI	Osobní dotazování za pomoci počítače / Computer Assisted Personal Interviewing
CAWI	Dotazování na webovém formuláři / Computer Assisted Web Interviewing
SQP	Webová aplikace pro testování kvality otázek / Survey Quality Prediction
SRC	Výzkumné centrum michiganské školy / Survey Research Centre of the University of Michigan
ITRUST	Individuální sociální kapitál / Individual Social Capital
SOCNET	Kolektivní sociální kapitál / Collective Social Capital
INSTRUST	Institucionální důvěra / Institutional Trust
INDIS	Individuální politické odcizení / Individual Disaffection
INSTDIS	Institucionální politické odcizení / Institutional Disaffection
INTEREST	otázka 1: „Jak moc se zajímáte o politiku?“ / How interested would you say you are in politics?
DIFFICULT	otázka 2: „Jak často se politika zdá být tak složitá, že vlastně nemůžete pochopit, co se děje?“ / How often does politics seem so complicated that you can't really understand what is going on?
POLACTIV	otázka 3: „Myslíte, že byste mohl(a) hrát aktivní roli ve skupině, která se zabývá politickými otázkami?“ / Do you think that you could take an active role in a group involved with political issues?
OPINION	otázka 4: „Jak je pro Vás těžké či lehké udělat si názor na politické otázky?“ / How difficult or easy do you find it to make your mind up about political issues?
VOTE	otázka 5: „Řekl(a) byste, že politici se zajímají jen o získávání hlasů voličů spíše než o názory lidí?“ / Would you say that politicians are just interested in getting people's votes rather than in people's opinions?
CARE	otázka 6: „Myslíte si, že se politici obecně starají o to, co si lidé jako Vy myslí?“ / Do you think that politicians in general care what people like you think?
POLIT	otázka 7: „Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: politici“ / Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: politicians
PARLAM	otázka 8: „Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: Český parlament“ / Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: parliament
LEGAL	otázka 9: „Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: právní systém“ / Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: the legal system
TRUSTED	otázka 10: „Obecně vzato, řekl(a) byste, že se většině lidí dá důvěřovat, nebo že člověk nemůže být při jednání s lidmi nikdy dost opatrný?“ / Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?
FAIR	otázka 11: „Myslíte si, že by se Vás většina lidí snažila podvést, pokud by měli tu možnost, nebo by se snažili být poctiví?“ / Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?
HELPFUL	otázka 12: „Řekl(a) byste, že se lidé většinou snaží pomáhat druhým, nebo že se většinou starají o sebe?“ / Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves?
FRIEND	otázka 13: „Jak důležité jsou ve Vašem životě následující věci?: přátelé“ / How important are friends in your life?
VOLUNORG	otázka 14: „Jak důležité jsou ve Vašem životě následující věci?: dobrovolné organizace“ / How important are voluntary organizations in your life?
MEET	otázka 15: „Jak často se ve svém volném čase společensky scházíte s přáteli, příbuznými či kolegy?“ / How often do you meet socially with friends, relatives or work colleagues?

SOCACTIV	otázka 16: „Když se srovnáte s jinými lidmi Vašeho věku, jak často se ve svém volném čase účastníte společenských aktivit?“ / Compared to other people of your age, how often would you say you take part in social activities?
AT	Rakousko / Austria
BE	Belgie / Belgium
CZ	Česká republika / Czech Republic
DE	Německo / Germany
DK	Dánsko / Denmark
ES	Španělsko / Spain
FI	Finsko / Finland
FR	Francie / France
UK	Spojené království / United Kingdom
GR	Řecko / Greece
HU	Maďarsko / Hungary
CH	Švýcarsko / Switzerland
IE	Irsko / Ireland
IL	Izrael / Israel
IT	Itálie / Italy
LU	Lucembursko / Luxembourg
NL	Nizozemí / Netherland
NO	Norsko / Norway
PO	Polsko / Poland
PT	Portugalsko / Portugal
SE	Švédsko / Sweden
SI	Slovinsko / Slovenia

Symboly

$I(x)$	indikátor / položka
$c(\tau)$	intercept / konstanta
$f(\Lambda)$	směrnice / faktorová zátěž / regresní koeficient / váha
$L(\xi)$	latentní proměnná
$u(\delta)$	chyba měření / reziduum
Cov	kovariance faktorů
Var(V)	variance faktoru
$M(L)$	latentní průměr
$M(I)$	průměr položky
χ^2	chí-kvadrát test

ÚVOD

Kvalitně provedený kvantitativní sociální výzkum realizovaný na národní úrovni má být reliabilní a validní. Z hlediska reprezentativity šetření by měly být pravděpodobnosti výběru jednotek rovnocenné a/nebo přinejmenším známé a nenulové, stejně tak musí být různé subskupiny společnosti rovnocenně zastoupené. Dalším aspektem kvality takového výzkumu je široce srovnatelný význam otázek a škál odpovědí pro všechny respondenty. Výzkumníci si musí být jisti, že variance mezi daty pochází z rozdílů mezi odpověďmi respondentů na otázky v dotazníku, spíše než z rozdílů v interpretaci samotných otázek. Pouze když dosáhnou národní výzkumy úspěchů v těchto bodech, tak se jejich výsledky pravděpodobněji blíží sociální realitě. Žádný národ není ale homogenní, pokud jde o slovní zásobu obyvatel, způsob vyjadřování, úroveň vzdělání atp., proto je položka v dotazníku vždy vystavena riziku určitého zkreslení chybou měření, kterou jak inferenční statistika, tak i výzkumník při analýze a interpretaci dat bere v potaz. I když je tato heterogenita populace často prezentována jako problém realizace národních výzkumů, je zdrojem bariér především úspěšně, rozumějme kvalitně provedeného mezinárodního výzkumu [Jowell 1998]. Velké rozdíly v jazycích, kultuře i sociálních strukturách jednotlivých zemí, které činí mezinárodní výzkumy tak analyticky hodnotné, jsou totiž zároveň zásadní překážkou pro srovnatelnost teoretických konceptů a jejich měřicích položek mezi těmito zkoumanými společnostmi [Smith 2003: 73]. Srovnatelnost se tak vedle reliability a validity stává jedním z nejdůležitějších aspektů kvality mezinárodních šetření, kterého je ale velmi náročné dosáhnout.

Globální provázanost světa a multikulturní snahy o vzájemné porozumění ale činí v současnosti z mezinárodních komparativních výzkumů zaměřených na dlouhodobé mapování postojů a změn v lidském chování nezbytnost sociálního zkoumání. Z tohoto důvodu nelze od mezinárodní komparace ustoupit pod tíhou metodologické kritiky, ale spíše přistoupit k různým variantám řešení dané problematiky. Ty zmíněný problém srovnatelnosti cizojazyčných otázek sice zcela neeliminují, ale přispívají k vyšší transparentnosti při vykazování zjištěných kulturních rozdílů v mezinárodních šetřeních. Navíc srovnatelná a přesná měření nabývají v těchto výzkumech jak vědecké, tak i politické důležitosti, pokud mají výsledná zjištění očekávatelný vliv na mezinárodní i národní politickou praxi. V tomto kontextu nemusí být výsledky měření založené na mylných analýzách pouze nesolidní a nespolehlivé, ale i potenciálně devastující v aplikaci do sociálního světa [Herkenrath 2002: 517].

Tématem této metodologicky zaměřené práce je oblast aplikace postojových škál do mezinárodních dotazníků a to obzvláště těch robustních, propracovaných, mnohokrát testovaných a opakovaně využívaných. Mezinárodní dotazníky jsou složené převážně z takovýchto škál, kterým mezinárodní výzkumné týmy věnovaly již mnoho pozornosti z hlediska ověřování validity, reliability a jejich konzistence. Metodologickým rámcem, ve kterém se hlavní část této práce pohybuje, je proto oblast srovnatelnosti, resp. ekvivalence takovýchto postojových škál, na jejíž ověřování se při meritorní analýze dat často zapomíná. Přitom je ale právě srovnatelnost postojových škál v komparativních výzkumech ohrožena nejvíce. Výklad zmíněného problému zde nebude předložen ani tak v dikci psychometrických přístupů – tj. co vše je nezbytné vykonat v průběhu přípravy výzkumného nástroje na počátku výzkumu, aby byly škály srovnatelné – ale spíše v dikci statistických přístupů k analýze sebraných dat.

Předmětem rozboru je zde mezinárodní srovnatelnost dvou v současnosti velmi často využívaných postojových škál a to škály sociálního kapitálu a politického odcizení tak, jak jsou měřeny v Evropském sociálním výzkumu. Oba dva tyto sociálně vědní koncepty se v sociologii i politologii řadí mezi velmi významné, řadu let teoreticky pojednávané a

empiricky zkoumané sociální jevy. Nezřídka mají i reálný dopad na sociální a politickou praxi, když uvážíme, že podávají vysvětlení k povaze sociální soudržnosti a k příčinám nízké volební účasti v současných moderních společnostech. Empirické studie, které pravidelně tyto koncepty analyzují a to i v komparativním rámci, ale někdy zanedbávají obecné limity srovnatelnosti postojových škál a svá zjištění nepodkládají žádnými metodologickými studii, které by bylo ověřeno, zda lze vůbec k takovéto široké mezinárodní komparaci dat přistoupit [např. Eurobarometer 2005]. To je ale v případě mezinárodních výzkumů zcela kontraproduktivní, neboť hrozí, že zjištěné rozdíly mezi zeměmi nemají reálný podklad v datech a jsou jen uměle vykonstruované.

Cíle, které tato práce sleduje, lze rozdělit do několika oblastí. První oblastí je ohodnocení mezinárodní srovnatelnosti důležitých sociálně vědních konceptů tak, jak jsou měřeny v jednom z nejpracovnějších komparativních projektů dnešní doby. Apel je zde kladen na to, že alespoň jedním z předložených způsobů je nezbytné ověřovat srovnatelnost každé postojové škály ve všech mezinárodních šetřeních ať už v primární či sekundární analýze dat. Druhou více teoretickou oblastí je představení konceptu ekvivalence v kontextu komparativních výzkumů české sociologie. V českých sociálních vědách zatím nebyla problematika ekvivalence nijak kontinuálně rozebírána a prezentována v odborných publikacích, spíše se jen obsahově skrývá za témata kvality sociálně vědních výběrových šetření a chyb měření v komparativních výzkumech [např. Kostecký, Čermák 2003; Krejčí 2008]. Je to o to překvapivější, že se Česká republika po roce 1989 do řady kontinuálních mezinárodních projektů aktivně zapojila, jak také názorně ilustruje Krejčí [2006].

Struktura práce odpovídá těmto dvěma cílům. Obecně lze říci, že první část se zaměřuje na představení problematiky ekvivalence v komparativních šetřeních a uvedení vhodných metod k jejímu ověřování, což v logice výkladu poslouží jako praktický úvod ke druhé části práce, kde je hlavním výzkumným cílem ohodnotit mezinárodní srovnatelnost postojových škál sociálního kapitálu a politického odcizení v datech Evropského sociálního výzkumu a vyvodit z toho univerzálnější doporučení pro meritorní komparativní práci s postojovými škálami. Struktura kapitol pak vypadá následovně: v první je nastíněn rozvoj komparativních výzkumů ve 20. století, ve druhé kapitole je představen koncept ekvivalence v mezinárodních výzkumech, zavedeno základní pojmosloví, navržena stručná kategorizace základních druhů ekvivalence a načrtnuto portfolio statistických technik, které jsou používány k ověřování jednotlivých úrovní dosažené ekvivalence postojových škál v datech. Třetí kapitola potom ukazuje teoretické zázemí konceptů sociálního kapitálu a politického odcizení a způsob jejich měření v Evropském sociálním výzkumu. V tomto kontextu je také poukázáno na povahu systematických chyb měření, které dokážou významně narušit srovnatelnost jednotlivých měřicích položek.

Konečně čtvrtá kapitola už podrobně popisuje základní techniky ověřující srovnatelnost vybraných postojových škál, porovnává je mezi sebou a ukazuje první výsledky hodnocení mezinárodní ekvivalence škál. Pátá kapitola představuje pokročilejší techniku strukturního modelování a detailně vykládá způsob a postup testování různých druhů ekvivalence na úrovni měření. V šesté kapitole je vystaven a předběžně ověřen základní model postojových škál, který bude dále testován v analýzách srovnatelnosti. Sedmá, osmá a devátá kapitola představují separátní a simultánní analýzy srovnatelnosti vybraných postojových škál a doplňují je bohatým empirickým materiálem, na základě něhož dochází k zevrubnému hodnocení mezinárodní srovnatelnosti jednotlivých položek z postojových škál. Desátá kapitola vysvětluje povahu a podává výstupy simultánní časové analýzy srovnatelnosti položek z postojových škál mezilidské a institucionální důvěry a účelově tak doplňuje závěry předchozích analýz. V závěru jsou shrnuty nejdůležitější poznatky ze všech aplikovaných analýz srovnatelnosti a vyvozeny univerzální doporučení pro komparativní práci s postojovými škálami.

1 KOMPARATIVNÍ POVAHA SOCIÁLNÍCH VÝZKUMŮ

„Research is always an adventure... But the adventure is even more hazardous when entering a field as new as cross-national research.“ [Christiansen, Hyman, Rommetveit 1951]

Komparace je pro vědecké bádání a myšlení zcela nepostradatelná a to obzvláště v sociálních vědách, kde se porovnávají postoje, jednání a hodnoty různých skupin jedinců a kde žádný sociální fenomén není studován izolovaně [Øyen (Ed.) 1990: 4; Arts, Halman 1999: 1]. Komparativní perspektiva zkoumání sociálních jevů je neoddelitelnou součástí sociologického bádání již desítky let, ale až přibližně v 50. letech 20. století se mezinárodní komparativní sociologie jako taková začala systematicky rozvíjet a etablovat mezi sociologickými disciplínami [podrobněji Marsh 1967; Payne 1973]. Není náhoda, že v roce 1949 byla založena také Mezinárodní sociologická asociace (ISA), která dodnes funguje pod záštitou UNESCO, neboť vedení ISA chtělo totiž už v této době výslovně upozornit na mimořádný význam, který má oblast komparativních výzkumů pro celý obor sociologie [Krejčí 2006: 150].¹

V těchto letech se také odehrál průlom v oblasti komparativního sociálního výzkumu, neboť Almond a Verba představili ve své knize *The Civic Culture* [1963] mezinárodní šetření, které dnes již lze označit za první pokus o systematický sběr komparativních dat a mezinárodní analýzu na poli komparativní politiky a občanské kultury. Jako první sociální vědci navrhli Almond a Verba promyšlený výzkum politického chování pro pět značně odlišných zemí a názorně demonstrovali, že možnosti komparativních analýz nespočívají pouze v popisu odlišností a podobností mezi zeměmi (resp. politickými systémy), ale rozšiřují se až k empirickému testování obecných hypotéz a ke generování nových teorií. Dokázali tak podtrhnout hlavní rozdíly mezi požadavky kladenými na realizaci monokulturních a mezinárodních komparativních výzkumů. Po drásající kritice uměle slepovaných datových souborů z národních výzkumů [Harkness, Mohler, Van de Vijver 2003: 4], což bylo u (ne)komparativních výzkumů do 60. let 20. století běžným nešvarem, přišel se studií Almonda a Verby velký boom a strategicky promyšlené mezinárodní výzkumy se staly na poli empirických šetření jednou z nejvyšších výzkumných ambic sociálních vědců.

Analytické problémy, se kterými se potýkali Almond a Verba, byly zdrojem několikaleté kritiky, která nakonec přibližně v 70. – 80. letech vyústila v komplexní odvětví zabývající se metodologií komparativních mezinárodních šetření. Tehdy se Melvin L. Kohn stal prezidentem Americké sociologické společnosti (ASA) a na každoroční kongres v Chicagu vybral jako hlavní téma diskuze „mezinárodní výzkumy v sociologii“. Kohn se ještě v 80. letech 20. století cítil v této metodologické oblasti jako misionář a jeho primárním cílem bylo ve vedoucí pozici prezidenta ASA vytvořit mezinárodní sociologii. Viděl tento směr bádání jako výzvu pro americkou sociologii, kterou chtěl vymanit ze své provinciální perspektivy a vsadit Spojené státy do kontextu rozvoje ostatních zemí. V roce 1989 pak publikoval v díle „Cross-national Research in Sociology“ několik příspěvků, jež na zmíněné konferenci zazněly, ve snaze o osvětu ohledně možných způsobů využití komparativních výzkumů. Jeho cílem bylo především upozornit na to, že komparativní šetření sice mají na jednu stranu neodiskutovatelný potenciál, co se týče srovnávání různých sociálních jevů mezi zeměmi, kulturami a jinými sociálními skupinami anebo co se týče sledování proměn sociálních fenoménů v průběhu času, ale na druhou stranu jsou tyto výzkumy obzvláště náchylné k chybám měření. Ve své knize Kohn mírně ironicky podotkl, že to byl právě

¹ Kohn [1989: 20] definuje komparativní výzkumy jako „studie, které systematicky využívají srovnatelná data z dvou či více národů.“ Toto pojetí postupem času převládlo a komparativní výzkum se již standardně definuje jako studium sociálních jevů ve dvou či více společnostech.

komparativní výzkum, který vyprodukoval nejvíce metodologických pouček bez adekvátní odezvy [Kohn (Ed.) 1989: 78].

Významy jednotlivých sociálních fenoménů se totiž liší pro různé lidi, v různých dobách a především v různých kulturách v závislosti na mnoha kontextuálních proměnných jako je sociální struktura společnosti, její kultura, převažující hodnotová orientace populace, právní a politický systém, ekonomický stav země, atp. V komparativních výzkumech se objevují i jiné problémy spojené s obecnou tendencí jedinců manifestovat takové chování, které je sociálně přijatelné a dobře hodnocené (tj. sociální desirabilita), anebo s různými způsoby zpracování otázek v myslích jedinců (tj. kognitivními procesy), které mohou být navíc také kulturně specifické a proměnlivé.

Logicky se tak jako jedna z prvních zásadních otázek v oblasti metodologie komparativních výzkumů časem objevilo to, zda jsou sledované sociální fenomény v různých společnostech vůbec srovnatelné či nikoli; zda lze vůbec sociální jevy měřit ve výzkumu stejně, rovnocenně a ekvivalentně ve všech participujících zemích. Relevantnost této metodologické otázky se potvrzuje obzvláště tváří v tvář častým pochybnostem, zda mezinárodní data opravdu reflektují rozdíly v odpovědích respondentů a nikoliv pouze rozdíly v jejich interpretaci otázek a zda tak nejsou zjištěné odlišnosti mezi kulturami jen něčím uměle vykonstruovaným při analýze dat [podrobněji Jowell 1998]. Všechny tyto obavy pak ještě více zdůrazňují, nakolik je důležité věnovat pozornost metodologickému pozadí mezinárodních výzkumů a jejich přípravě.

Rámeček 1 Rozvoj komparativních výzkumů

Rozvoj mezinárodních výzkumů souvisí především s rozšířením informačních technologií v 60. letech a systematickým vývojem komparativních metod. Prvním příkladem již poměrně kvalitního výzkumu z hlediska provedení je Eurobarometr (EB), který monitoruje sociální a politické postoje a hodnotovou orientaci dvakrát ročně v zemích Evropské unie od roku 1973. Stejnou logiku sleduje výzkum General Social Survey (GSS), který se každoročně realizuje už od roku 1972 a sleduje strukturu a vývoj americké společnosti včetně mapování trendových dat postojů a sociálních změn v Americe. V roce 1980 byl Světovou bankou iniciován vznik výzkumného projektu Living Standards Measurement Study (LSMS) z důvodů lepšího porozumění sociálním a ekonomickým vlivům v rozvojových zemích a také kvůli vylepšení kvality sbíraných dat z jednotlivých domácností. V roce 1981 započala World Value Survey (WVS) a European Value Study (EVS) jako dlouhodobá komparativní šetření hodnot v sociální, kulturní a politické oblasti.

Následoval v roce 1985 každoročně prováděný nejrozsáhlejší mezinárodní výzkum International Social Survey Programme (ISSP), který se zaměřuje na výzkum postojových a behaviorálních otázek. Na jeho počátku stály čtyři státy, jejichž četnost se časem rozrostla až do počtu téměř 40 participujících zemí v současnosti. V roce 1985 vznikla také skupina Cross-National Research Group s cílem poskytnout fórum pro diskuzi a výměnu názorů a zkušeností mezi výzkumníky z různých sociálně vědních disciplín zapojených do mezinárodních komparativních projektů. V přelomové dekádě 20. století vznikly mezinárodní projekty, které mají podobu buď kontinuálních šetření (OECD, ESS atp.) nebo jednorázových výzkumů (EQLS atp.). Komparativní výzkumy se dnes realizují jak v komerční sféře, tak i v akademické oblasti. I Česká republika se zejména po roce 1989 zapojila do řady těchto mezinárodních projektů [stručný přehled poskytuje Krejčí 2006: 150].

2 KONCEPT EKVIVALENCE V KOMPARATIVNÍCH VÝZKUMECH²

Ve většině odborných publikací věnujících se metodologii mezinárodních výzkumů je možné se dočíst, že základem a také nutnou podmínkou kvalitního mezinárodního šetření je kromě detekce chyb měření, mapování rozsahu chybějících odpovědí, návratnosti, testování validity a reliability též tzv. ekvivalence měření. Na rozdíl od validity (přesnosti) nebo reliability (spolehlivosti) měření se ekvivalencí rozumí stálost či neměnnost měření, respektive jeho srovnatelnost.³ Horn a McArdle [1992: 117] definují neměnnost měření jako situaci, kdy operacionalizace konstruktů vyústí v měření zcela totožných znaků a to i za různých okolností studia daného fenoménu. Různými okolnostmi se má na mysli různý čas měření, měření různých populací anebo použití různých metod sběru dat. Již v definici ekvivalence se tak odráží tři různé úrovně srovnatelnosti a to 1) srovnatelnost teoretického konceptu, 2) srovnatelnost položek, které daný koncept měří, a 3) srovnatelnost jejich škál měření (tj. kategorií odpovědí).

Srovnatelnosti měření je většinou dosaženo poté, kdy respondenti vnímají, chápou a interpretují daný koncept, měřicí položku a současně také škálu měření úplně stejným nebo alespoň velmi podobným způsobem i přesto, že pocházejí z různých kulturních prostředí a obsahy položek jsou přeloženy do různých jazyků. Ukazuje se, že právě toto je velmi podstatný problém, který se projevuje obzvláště při formulaci postojových otázek, které by měly zachycovat hodnoty, názory a stanoviska respondentů k různým společenským tématům. Protože na rozdíl od tzv. objektivních proměnných jako je dotaz na vzdělání či příjem respondenta, je percepce těchto položek mnohem více závislá na kontextu dotazování – tj. na kulturních, společenských, politických, právních a ekonomických poměrech dané země v daném čase, na rodném jazyku, tradicích, sociální struktuře, atp. – a tudíž jsou tyto položky více náchylné k systematické chybě měření.

Otázkou pak zůstává, zda v situaci, kdy se dotazujeme například na problematiku násilí na ženách, antisemitismu, individualismu a kolektivismu, žebříček hodnot, sociální kapitál, důvěru v politické instituce, autority či nepotismus, hovoříme o tom samém jako respondent. Eldad Davidov nebo Janet Harkness ukazují za desítek let své výzkumné praxe řadu konkrétních příkladů, z nichž lze usoudit, že tomu tak vždy není. Například pojem „být tradiční“ má v různých zemích jiný význam, například v Bavorsku to znamená být religiózní („být katolík“), jinde může mít spíše charakter životního stylu („mít rodinu, dům, děti“). Jiným příkladem může být různé vnímání a hodnocení frekvence návštěv na základě územní rozlohy země. V územně rozsáhlých zemích (např. Německu nebo Spojených státech) dochází v porovnání s územně malými státy k jinému vnímání otázky „Navštěvujete často své děti?“, neboť zde většinou není možné tak často navštěvovat své příbuzné kvůli velké vzdálenosti. Odpověď „ano, často“ tak může pro Čecha v tomto kontextu znamenat jednou týdně a pro Američana čtyřikrát za rok [Davidov 2012].

Výjimečně choulostivé jsou v dnešní době často diskutované pojmy spojené s přistěhovalectvím nebo etnickými skupinami. Například interpretace slova „imigrant“ je obzvláště specifická podle kulturní situace v dané zemi. Nejčastěji se uvádějí rozdíly mezi evropskými zeměmi, kde imigrant je vnímán jako přistěhovalec a znamená pro většinu obyvatel nové sociální problémy, a Izraelem, kde imigrant je Žid se všemi právy a povinnostmi vůči své zemi) [Davidov, Schmidt 2007].⁴ Pojem „etnická skupina“ má pro většinu zemí ve střední a východní Evropě podtext romské etnické skupiny, která je zde

² Obsah následující kapitoly byl v neúplné verzi publikován také v [Anýžová 2013].

³ Ekvivalence je také často označována jako invariance měření (measurement invariance).

⁴ Tato skutečnost byla zjištěna na základě výzkumu ESS, na kterém Izrael již od prvních vln šetření opakovaně participuje.

také odlišně posuzována [více příkladů viz www.csdiworkshop.org]. Pomíjím v tuto chvíli až nepřekonatelné rozdíly mezi kulturně zcela odlišnými zeměmi (Afrika versus Evropa, Amerika), kde výzkumy některých sociálních jevů jako je například studium médií (četnost sledování TV, poslouchání rozhlasu, čtení novin a získávání informací z internetu) zcela pozbývá smysl.

V případě, že je dotazník opravdu kvalitně zpracován a respondenti v různých zemích dokonce konceptu a položené otázce podobně rozumí, ještě může dojít k velmi rozdílnému použití měřicí škály, která slouží k ohodnocení postojového výroku a poskytuje analytikovi číselný skór (tj. odpověď) respondenta. Respondenti z některých zemí mají například silnější tendenci k volbě krajních hodnot škály, a zdá se, že tato tendence může být kulturně determinovaná [Smith, Fischer 2008]. Tímto způsobem se problémy se srovnatelností postojových škál postupně nabalují na sebe a vzájemně násobí.

V současné době začínají převážně sociální vědci upozorňovat na to, že je nezbytné se věnovat z hlediska kulturní srovnatelnosti vedle postojových (tj. subjektivních) otázek také objektivním položkám (resp. sociodemografickým charakteristikám) [Mohler, Johnson 2010; Scheuch 1993]. Koncept ekvivalence byl dlouhou dobu rozvíjen hlavně v komparativní psychologii, kognitivních přístupech, teorii překladu a modelu celkové výzkumné chyby. A v této oblasti zkoumání takto po dlouhou dobu upozadovala psychologie potřeby sociologie, protože techniky škálování, které měří postoj či chování jedince souborem mnoha otázek, jednoduché objektivní položky nevyužívají, a proto jim nebyla věnována dostatečná pozornost [Mohler, Johnson 2010: 26].

Scheuch toto varování podkládá kritikou plošné standardizace sociodemografických charakteristik mezi zeměmi založenou na příkladu věkových kategorií, které jsou v každé zemi rozdělené většinou do stejných intervalů. Proměnná věk podle Scheuche [1993: 115-116] neodkazuje pouze k fyzickému stáří člověka, ale značí také rozdílné sociální role a rozdílné vzorce chování. Pokud je ale ve výzkumu fyzický věk indikátorem takovýchto rozdílných sociálních rolí, stejný věk může znázorňovat různé sociální role a životní etapy pro lidi z různých zemí, a není tudíž srovnatelným indikátorem. Pojem „starý“ jako sociální definice je obvykle vztažen k různému fyzickému věku napříč různými kulturami. Dosažená hranice 40 let může mít jiné sociální důsledky a označovat jinou pozici ve společnosti pro ženu ze Sicílie a ženu z Kalifornie. A nejedná se jen o proměnnou věku, stejný problém postihuje například i indikátor příjmu člověka, počtu členů v domácnosti, počtu let strávených ve škole, atp., o které by se měli sociologové začít zajímat.

Tematika srovnatelnosti sociodemografických údajů je součástí mnohem komplexnější kritiky systematického přejímání metod a pojmosloví z psychologických přístupů v oblasti zkoumání ekvivalence ze strany jiných sociálních věd. Mohler a Johnson [2010: 26] upozorňují, že psychologický model ekvivalence je příliš striktní a je nutné ho pro potřeby dalších sociálních věd upravit. Navržené změny se týkají jednak konceptuální oblasti, kde by bylo vhodné nahradit pojmy ekvivalence, identita a invariance za více relativní pojmy jako je srovnatelnost a podobnost, neboť například nalezení zcela identického indikátoru, který bude mít úplně stejný význam v různých zemích, je v sociálních vědách téměř nedosažitelný cíl. A jednak se týkají i metodické oblasti, kde se například poměrně striktní pravidla posuzování ekvivalence v konfirmační faktorové analýze - tento nástroj pro posuzování ekvivalence poprvé představil švédský psychometrik Jöreskog (1971) - začínají postupně relativizovat díky novému softwarovému nástroji Jrule [van der Veld, Saris 2009], který na místo se statistikami vhodnosti modelu pracuje detailněji s indexy modifikace pro parametry jednotlivých položek jako alternativní verzi hodnocení modelu v sociálních vědách. Důkazem převahy psychologických publikací v této oblasti metodologie je přehledová studie Vandenberg a Lance [2000] a aktualizovaná studie Schmitta a Kuljanina [2008].

2.1 Kategorizace ekvivalence

Konceptu ekvivalence se tedy výzkumníci z oblastí psychologie, sociologie, vzdělávání, statistiky a výzkumu organizací věnují jak z teoretického hlediska, tak i z empirického již přes 40 let. Za tu dobu byla navržena řada kategorizací různých druhů ekvivalence a jejich definic. Názorným příkladem je zmatečný přehled Johnsona [1998: 6-7], který nashromáždil přes 50 různých typů ekvivalence figurujících v odborné literatuře, ze kterých nakonec klasifikoval ekvivalenci do dvou souhrnných kategorií; interpretativní a procedurální ekvivalence. Ze všech sumarizací a kategorizací, které je možné v odborné literatuře nalézt, lze ale nakonec rozdělit a definovat ekvivalenci jednoduše na dvou úrovních, které se pohybují na kontinuu od teorie ke konkrétnímu měření.

2.1.1 Kategorizace ekvivalence na teoretické úrovni

Na teoretické úrovni dochází nejčastěji k důslednějšímu vymezení tzv. konceptuální, operacionalizační a položkové ekvivalence [Hui, Triandis 1985]. Ty označují nutnost dosáhnout ve výzkumu srovnatelnosti 1) konceptu, 2) konstruktů a 3) měřicí položky. Konceptuální (conceptual) ekvivalence označuje situaci, kdy dochází k sociálnímu, emocionálnímu i kognitivnímu přenosu významu jednotlivých konceptů. Pokud respondenti připisují konceptům stejný význam, značí to, že jejich typické chování se v dané věci příliš kulturně neliší (např. vnímání a projevoování úcty k rodičům, chápání procesu seberealizace, moderní výchova dětí, atp.) [Straus 1969; Lonner 1985; Singh 1995; van de Vijver, Leung 1997]. To znamená, že zkoumaný sociální jev (resp. jeho reprezentace v podobě konceptu) je kulturně srovnatelný, může být smysluplně diskutován mezi všemi respondenty a je mu podobně rozuměno.

Obrázek 1 Proces operacionalizace sociálního jevu; zaměření ekvivalence na teoretické úrovni na koncept, konstrukt a položku



Operacionalizační (operationalization) ekvivalence značí, že konstrukt je měřen v různých skupinách stejně interpretovanými indikátory a že je opět stejně smysluplný pro všechny respondenty. Položková (item) ekvivalence s tím úzce souvisí a říká, že nástroj měření (tj. dotazník), který měří daný konstrukt, obsahuje měřicí položky opět se stejným významem pro jednotlivé respondenty z různých kultur [Hui, Triandis 1985]. Tyto typy ekvivalence jsou v podstatě nutnou podmínkou pro jakékoli statistické srovnávání mezinárodních dat a nepřipouští se tudíž žádná varianta částečné ekvivalence tohoto druhu [Donahue 2006: 2].

Nezbytnou součástí předchozích typů ekvivalencí je tzv. překladová (translation) ekvivalence nebo jinak lingvistická (linguistic) ekvivalence, která řeší konkrétní formulaci otázek, přenos významu jednotlivých pojmů v procesu překladu, totožnou interpretaci otázek napříč světovými jazyky a také čtivost a přirozenost položek v jednotlivých jazykových verzích dotazníku [Lonner 1985; Scheuch 1993; Blair, Piccinino 2005; Ægisdóttir et al. 2008]. Překladová ekvivalence většinou zaručuje, že položka i její hodnotící škála budou mít i po překladu stejný význam. Tento typ ekvivalence bývá zajišťován zejména kvalitním zpětným překladem a psychometrickými technikami před sběrem dat, v některých případech

bývá při kontrole překladu v pilotáži dotazníku testována podobnost faktorové struktury dat anebo aplikováno statistické třídění dat prvního a druhého stupně.⁵

Rámeček 2 Koncept, konstrukt a indikátor

V tomto kontextu je důležité si vymezit pojmy indikátor, konstrukt a koncept, neboť spolu úzce souvisí. V obecném metodologickém schématu tyto pojmy hovoří o vztahu mezi teorií a empirií [Harkness, Mohler, Van de Vijver 2003: 11; Harkness et al. 2010: 41]. Teoretickým konceptem se většinou myslí sociální fenomén, který je ve výzkumu zkoumán a který sociální badatele zajímá (např. sociální kapitál); ten ale nikdy není přímo měřitelný.

Latentní konstrukt by měl tento teoretický koncept co nejlépe reprezentovat. Konstrukt reprezentuje základní teoretický jev, který je předmětem měření, ale musí být z důvodu zjednodušení měření převeden do měřitelných empirických indikátorů, protože sám o sobě přímo měřitelný není, je příliš složitý (např. obecná mezilidská důvěra, která vyjadřuje sociální kapitál jedince).

Manifestní indikátory, již lze přímo měřit. Jsou v dotazníku většinou jednoduše vyjádřeny položkou (tj. otázkou), přičemž se doporučuje využít k měření latentního konstruktů větší množství položek, resp. baterii položek (např. otázka na míru důvěry/opatrnosti člověka; otázka na poctivost druhých lidí, otázka na míru ochoty pomáhat ostatním). Ty by měly pokrýt pokud možno všechny dimenze sociálního jevu a dohromady měřit latentní konstrukt.

Rozdíl mezi indikátorem a položkou prozaicky vysvětlují Mohler a Johnson [2010: 23] a to tak, že indikátor nemusí být na jedné straně nutně otázkou/položkou, může se jednat také o jednoduchý číselný údaj například z oficiálních statistik (např. roční spotřeba alkoholu v dané zemi). Na druhou stranu mohou otázky formulované v různých jazycích významově odkazovat ke stejnému indikátoru. V mezinárodních šetřeních někdy není možné použít ve všech kulturních skupinách stejný indikátor pro měření daného latentního konstruktů, neboť by ho neměřil dostatečně adekvátně, a je nutné jich využít více s různě formulovanými položkami, aby došlo ke kompletnímu zachycení sociálního jevu [Harkness et al. 2010: 42].

2.1.2 Kategorizace ekvivalence na úrovni měření

Na úrovni měření jde zejména o to, ověřit si míru podobnosti vybraných položek a především jejich hodnotících škál mezi skupinami, která se zákonitě odrazí v možnostech statistického srovnání dat mezi kulturami. Nejedná se tedy již o zkoumání srovnatelnosti konceptů a konstruktů na úrovni teorie, ale o srovnatelnost škál měření položek a kategorií odpovědí. Škály měření položek jsou srovnatelné, pokud je respondenti stejně interpretují, každá slovní i numerická kategorie odpovědi má pro respondenty stejný význam a jejich rozsah využívají všichni zcela shodným způsobem. Podobnosti škál měření se ale dosahuje poměrně obtížným způsobem, protože jsou velmi často zatíženy systematickou chybou měření. Saris [2003: 275] v otázce získávání a zaznamenávání odpovědí skepticky podotýká, že proces, kterým respondenti vybírají odpovědi mezi nabízenými kategoriemi, se mezi kulturami vždy liší, a že odpovědi v podstatě nejsou dobrým indikátorem respondentova úsudku. Při hodnocení srovnatelnosti škál měření se využívá třech úrovní, kterých mohou škály dosáhnout. Podle těchto úrovní se tak hierarchicky rozlišuje skalární ekvivalence položek s nejmenší systematickou chybou měření, metrická ekvivalence a konfigurační ekvivalence položek s největší systematickou chybou měření.

⁵ Druhú ekvivalence, které označují, že v překladu byly přeneseny adekvátní významy pojmů a že otázky budou respondenti shodně interpretovat, je mnoho, avšak jejich definice nejsou vždy jednoznačné a mohou se i značně překrývat [viz Johnson 1998]. Což může být jedním z důvodů, proč nejsou tyto druhy ekvivalencí natolik systematicky v odborné literatuře pojednávány; většinou se shrnují pod jeden obecný pojem - konceptuální ekvivalence.

Obrázek 2 Proces operacionalizace sociálního jevu; zaměření ekvivalence na úrovni měření na škálu měření dané položky



Nejvyšší míru podobnosti mají tedy tzv. skalárně (scalar) ekvivalentní položky. V tomto případě má škála měření položky v dotazníku ve všech zemích stejný rozsah a stejnou jednotku měření (např. Lickertova škála na měření postojů) a navíc mají jednotlivé skóry na škále pro respondenty stejný význam a jsou totožně interpretovány; škála není systematicky zkreslena žádnými subjektivními vlivy. Pokud jsou splněny všechny podmínky skalární ekvivalence, lze pak v analýze dat přímo srovnávat průměrné skóry jedinců v dané položce uvnitř i mezi kulturními skupinami (např. T-test, analýza rozptylu) [van de Vijver, Leung 1997; Steenkamp, Baumgartner 1998; Vandenberg, Lance 2000; Ægisdóttir et al. 2008].

Typy měření, které vykazují metrickou (metric) ekvivalenci, jinak také ekvivalenci měřicí jednotky [van de Vijver, Leung 1997], se vyznačují stejnou jednotkou na škále měření a stejným rozsahem škály, ale subjektivně jsou rozsah a jednotky škály vnímány respondenty z různých kulturních skupin odlišně. Tato situace nastává například při obecném hodnocení souhlasu s předloženým výrokiem na stupnici od jedné do pěti. Čechům může tato škála připomínat klasické známkování známé ze školních lavic, kde jedna znamená nejlepší hodnocení a pět nejhorší, pro Rusy by ale tato škála byla převrácená, protože jejich školní klasifikace má opačnou logiku.⁶ Jiným příkladem je hodnocení riskantnosti užívání alkoholu před jízdou automobilem na škále od jedné do pěti - skór dva může znamenat pro Němce mnohem nebezpečnější chování řidiče než pro Čecha. Průměrné skóry jedinců z takovéto škály mohou být v analýze dat přímo porovnávány pouze uvnitř jedné kulturní skupiny (např. T-test), ale ne mezi kulturami [Ægisdóttir et al. 2008]. Mezi kulturami ale mohou být porovnávány vztahy testovaných manifestních či latentních proměnných s jinými proměnnými z datového souboru, u kterých byla též ověřena metrická ekvivalence (např. korelační a regresní analýza, rozdíly bodů mezi skóry jedinců, rozdíly bodů různých skupin či kategorií atp.).

Nejmenší míru podobnosti vykazují měřicí položky, které jsou tzv. konfiguračně (configural) ekvivalentní, jinak též konstruktově (construct) ekvivalentní. Pokud jsou data konfiguračně ekvivalentní, předpokládá se, že ačkoli v jednotlivých dotaznících není použita stejná položka (např. nestejně znějící otázka) a ani měřicí škála o stejném rozsahu a stejné jednotce (záleží na kulturně vhodném použití), všechny jejich varianty odkazují ve všech kulturních skupinách ke stejnému sociálnímu jevu. V případě konfigurační ekvivalence dat není možné v analýze dat porovnávat mezi skupinami ani vztahy latentních či manifestních proměnných s jinými proměnnými v datových souborech, ani průměrné skóry jedinců. Skóry z takovéto škály mohou být opět srovnávány pouze uvnitř jedné skupiny.

Tato terminologie různých úrovní ekvivalence měření vychází z navržené série testů modelů měření v konfirmační faktorové analýze od Jöreskoga [1971], který takto doporučil ověřovat ekvivalenci indikátorů (resp. škál měření) na úrovni měření, a zavedli ji do praxe sociálních výzkumů Horn a McArdle [1992].⁷ Ti jako první označili testovaný model podobné faktorové

⁶ Tradiční ruské známkování má tuto podobu (5 = nejlepší známka/výborně, 1 = nejhorší známka/nedostatečně). Některé modernější systémy známkování v Rusku dokonce používají stupnici od 10 do 1 (1 = nejhorší známka).

⁷ Jedná se tedy o testování vztahů mezi měřicí položkou a základním latentním konstruktem v konfirmační faktorové analýze.

struktury jako konfigurační ekvivalenci, resp. slabou ekvivalenci, a druhý testovaný model shody faktorových zátěží jako metrickou ekvivalenci, resp. silnou ekvivalenci. Meredith [1993] tento koncept ekvivalence na úrovni měření ještě rozšířil a zavedl označení silná ekvivalence pro model shody faktorových zátěží a konstant položek a striktní ekvivalence pro model shody chyb měření testovaných položek. Steenkamp a Baumgartner [1998] se pokusili tuto terminologii blíže objasnit a také poprvé označili model shody konstant položek za skalární ekvivalenci. Rozruch, který se okolo skalární ekvivalence a otázky validity srovnávání průměrů položek rozpoutal, byl způsoben především ohromujícím množstvím vágních srovnávacích tabulek, které se začaly v 90. letech 20. století objevovat v různých sekundárních analýzách zpracovávající nepřehledné množství dat z mezinárodních výzkumů. Tyto kategorizační tabulky jen řadily velké množství zemí za sebou podle účelově vybraných proměnných a směřovaly tím k ploché interpretaci bez potřebného kontextu. Navíc pokud tyto sekundární analýzy pracovaly s nesrovnatelnými daty, byla tato komparace značně nesolidní a mohla být zneužita k manipulaci výstupů výzkumu [více D. Lievesley 2001: 378]. Proto i Madden [2004: 10] později navrhl využívat spíše různých poměrů proměnných a jiných složených statistických ukazatelů jakožto adekvátních komparativních statistik.

Steenkamp a Baumgartner se také jako jedni z prvních po významné studii Byrne, Shavelsona a Muthéna [1989] zabývali blíže konceptem částečné ekvivalence na úrovni měření, která přináší možnost vzájemně srovnávat koncepty a měřicí položky v různých skupinách, i když škály měření položek mezi nimi vykazují na základě určitých parametrů mírné odlišnosti [Donahue 2006: 13]. Van de Vijver a Leung [1997] zakomponovali koncept ekvivalence do teorie chyb měření a Vandenberg a Lance [2000] sepsali přehledovou studii o výzkumech, kde byla ekvivalence položek na úrovni měření ověřována, jakým způsobem a s jakými výsledky. Tuto studii v současnosti zaktualizovali Schmitt a Kuljanin [2008].

2.2 Základní předpoklady ekvivalence položek

Koncept ekvivalence je často spojován (ne-li zaměňován) s problematikou dosahování celkové kvality v mezinárodních výzkumech. K budování kvalitního šetření ale existují tři poměrně odlišné přístupy – statistický, technicko-administrativní a psychometrický – které dokážou až ve vzájemné kombinaci vygenerovat kvalitně provedený výzkum. Statistický přístup se ve výzkumech projevuje v podobě řady technických standardů, jež by měly všechny země účastníci se mezinárodního projektu dodržovat. Vycházejí z něj především principy výběrového designu a výpočty chyb měření. Druhým podobným pojetím je technicko-administrativní přístup, který se věnuje stanovování pravidel ohledně aplikace technik sběru dat a tvorby standardní dokumentace výzkumu. Třetím významným zdrojem v oblasti dosahování kvality komparativního výzkumu jsou psychometrické přístupy, které se soustředí především na konstrukci kvalitního nástroje měření. Řešení problematiky ekvivalence je v kompetenci až tohoto třetího psychometrického přístupu ke kvalitě dat.⁸

Aby se tedy mohlo v datovém souboru pracovat se srovnatelnými indikátory a komparovat tak mezinárodní data, musí být podstatný díl úsilí věnován právě fázi formulování a překladu otázek před sběrem dat v terénu. Jedná se konkrétně o navržení kvalitního výzkumného nástroje, jeho pilotáž a případnou modifikaci ve všech participujících zemích, což je přirozeně časově i finančně značně náročné [Mohler, Johnson 2010: 23]. Důraz je kladen především na přenesení co nejpodobnějších významů jednotlivých pojmů položek v procesu překladu a

⁸ Zde platí, že není možné dosáhnout ekvivalentního měření bez jistoty, že výběrový design a sběr dat neproběhl ve všech zemích srovnatelným způsobem – to vše by totiž mohlo být dalším potenciálním zdrojem systematického zkreslení měření a příčinou nesrovnatelnosti škál položek. Na druhou stranu, naplnění těchto kvalitativních parametrů organizace sběru dat ještě samo o sobě nezaručuje, že koncepty, položky a jejich škály měření budou kulturně srovnatelné.

na zajištění jejich shodné interpretace respondenty. Nejčastěji se v tomto kontextu hovoří o dosahování ekvivalence na konceptuální úrovni, ale bez kvalitní přípravy dotazníku nemůže být dosaženo ani ekvivalence položek na úrovni měření. Při řešení problematiky ekvivalence před sběrem dat se nejčastěji hovoří o třech technikách, které by měly vyprodukovat při ideální souhře kvalitní a srovnatelné otázky. Těmi jsou a) dodržení rovnováhy při výběru etických a emických položek, b) relevantní překlad dotazníku a c) použití kognitivních metod ověřování kvality otázek.

2.2.1 *Výběr etických a emických položek*

Ještě před překladem dotazníku musí být rozhodnuto při jeho sestavování, zda budou zkoumány emické či etické koncepty a zda budou k jejich měření použity emické či etické položky. V mezinárodních výzkumech je téměř výhradně nezbytné, aby se alespoň v centrálních částech dotazníků společných pro všechny participující země vyskytovaly pouze etické koncepty, které popisují názory, vzorce myšlení a chování společné všem lidem v jakémkoli kulturním prostředí (např. mezilidské vztahy, sociální distance) [více Davidson et al. 1976; Harris 1976; Ægisdóttir et al. 2008]. Na druhou stranu emické koncepty se týkají způsobu myšlení, vzorců chování a témat, která jsou specifická pro dané konkrétní kulturní prostředí (např. postoje k imigrantům, nepotismus, vymezení blízké sociální skupiny), a je vhodné je zkoumat ve specifické části dotazníku, kterou si vytvořila daná země jen pro své výzkumné potřeby [viz praxe mezinárodního výzkumu ISSP; www.issp.org] anebo v monokulturních výzkumech.⁹

Komplikovanost komparativních výzkumů tkví v tom, že etické koncepty mohou být měřeny etickými otázkami (jinak kulturně univerzálními položkami) i emickými otázkami (jinak kulturně specifickými otázkami) a touto volbou, která se vždy odvíjí od kulturně specifické situace a případných lingvistických potíží při překladu, je také ovlivněna úroveň ekvivalence položky na úrovni měření. Konceptuální ekvivalence lze v datech dosáhnout i tak, že etický koncept, který chápou všechny kultury podobně, může být měřen emickými položkami, které v každém jazyce různě znějí a nejsou doslovně přeloženy. Tyto položky většinou nedosahují nejvyšší míry srovnatelnosti mezi kulturami, ale díky jejich použití jsou alespoň srovnatelné teoretické etické koncepty; nižší míra ekvivalence položek se tak projeví jen při konkrétním výběru statistické analýzy dat (např. upřednostněním regresní analýzy před analýzou průměrů). Naopak skalární ekvivalence lze většinou dosáhnout pouze měřením etického konceptu etickými otázkami, které jsou více méně doslovně přeložené. K nejhorší situaci dochází, pokud je emický koncept (kulturně zcela specificky vnímaný – např. nepotismus) měřen etickými otázkami – doslovně přeloženými v každé participující zemi. V tomto případě bylo zanedbáno kulturně specifické myšlení a vytváření postojů k dané problematice a doslova přeložené stejné otázky generují systematicky zkreslené odpovědi, které v případě nesprávné interpretace v rámci analýzy dat produkují uměle vytvořené nerelevantní kulturní odlišnosti. Z tohoto důvodu je emické koncepty vhodné měřit téměř výhradně emickými položkami i za cenu jejich nesrovnatelnosti.

To, jak velkou roli hraje proces překladu dotazníku při dosahování co nejvyšší míry ekvivalence dat, naznačili Van de Vijver a Leung [1997: 265], když navrhli rozdílné přístupy k překladu emických a etických otázek. Tři varianty překladu, které navrhuje (tj. application – adaptation – assembly) se liší především v míře modifikace původní otázky ze zdrojového

⁹ Pojmy vycházejí z lingvistické teorie o specifických znacích jazyka používaných v určitém kulturním prostředí. Fonemika (phonemics) se zabývá hláskami, které existují pouze v jediném konkrétním jazyce (tj. emickými) a fonetika (phonetics) hláskami existujícími univerzálně ve všech jazycích (tj. etickými) [Harris 1976]. Ve stejném duchu rozlišovali typy otázek i Przeworski a Teune [1970]: „core items/culture general“ (pro etické otázky) a „cultural specific items“ (pro emické otázky).

dotazníku. Doslovný překlad otázky je podle nich vhodný v případě šetření velmi podobných zemí se stejným kulturním zázemím a jazykem (tedy pro etické položky měřící etický koncept). Jakékoli další varianty jazykové a obsahové modifikace položek či dokonce tvoření nových, kulturně vhodných (emických) otázek jsou adekvátní ve výzkumu značně rozdílných zemích, v případě lingvistických problémů v překladu a tam, kde se předem očekává větší systematické zkreslení odpovědí respondentů.¹⁰

Jak etický, tak i emický přístup k tvorbě dotazníku má několik výhod i nevýhod. Hlavní výhodou etického přístupu je snížení nákladů na náročnost překladu a v případě, že je daná otázka totožně interpretována ve všech kulturních skupinách a odkazuje ke stejnému sociálnímu jevu, také dosažení značně vysoké míry ekvivalence měřících položek a tím pádem otevření možností statisticky neomezené komparace dat. Emický přístup naproti tomu dokáže zachytit kulturně odlišné aspekty zkoumaného jevu a je významným přínosem z hlediska kulturního porozumění, především pokud zkoumáme velmi rozdílné země; tehdy etický přístup selhává. Bohužel emické otázky většinou nedosahují příliš vysokých úrovní ekvivalence na úrovni měření, což značně omezuje statistické srovnávání dat.

Kvůli množství nevýhod obou přístupů se v současnosti navrhuje kombinovaný „emicko-etický“ přístup k mezikulturnímu použití dotazníkových metod. Tento přístup preferuje vybrat k souboru etických otázek společných všem přeloženým dotazníkům v každé zemi kulturně specifické otázky, protože ve správném poměru dokreslí kulturní obrázek o daném jevu (viz výzkumná praxe ISSP). Ve stejné míře se v současnosti doporučuje využívat potenciálu baterií otázek, spíše než jednoduchých otázek, sestávajících i z několika položek, protože jen tímto způsobem je možné statisticky zachytit chyby měření. Brislin [1976: 220] navíc zdůrazňuje aktivní spolupráci expertů v multikulturních týmech, kteří díky hluboké znalosti vlastní země reprezentují svou kulturu při přípravě šetření a snadněji tak nalézají etické a emické položky a možné zdroje jejich systematického zkreslení.

2.2.2 *Technika překladu dotazníku*

Jako jeden z prvních metodologických problémů se při sestavování co nejkvalitnějšího designu mezinárodního šetření objevilo téma relevantního, precizního a ekvivalentního překladu otázek do cizích jazyků a je to i dnes v odborné literatuře snad nejčastěji zmiňovaná překážka komparativních výzkumů [Harkness 2003: 38, Scheuch 1993: 107]. Proto byla také první pozornost věnována především lingvistické a funkční ekvivalenci v naději, že pečlivý překlad dotazníku povede k odstranění všech systematických zdrojů zkreslení odpovědí a tím pádem k plné ekvivalenci dat [van de Vijver 1998: 51]. Implicitně se tedy téma vývoje srovnatelných otázek a dotazníků promítá již po dlouhou dobu především do problematiky adekvátního překladu mezinárodních dotazníků [Johnson 1998; Harkness 2003; Smith 2003] a překrývá se s oblastí využití psychometrických přístupů při zvyšování kvality mezinárodních šetření.

Jowell [1998] uvádí, že každá druhá země musí dotazník přeložit do vlastních funkčně ekvivalentních slov a frází a tento proces s sebou přináší nemalé problémy. V počátcích budování kvalitního designu mezinárodního výzkumu se stal výzkumníkům trnem v oku doslovný překlad otázek bez přenesení jejich sociálního a kognitivního významu, který se liší

¹⁰ Tento přístup se shoduje se dvěma nejznámějšími modely formulace otázek. První model ASQ („ask-the-same-question“/„ptej se stejnou otázkou“) operuje s doslovně přeloženými položkami, tedy jejich přesnou replikací v každé kulturní skupině. Model ADQ („ask-different-questions“/ „ptej se odlišnou otázkou“) naopak upřednostňuje výběr specificky kulturních indikátorů podle dané situace za podmínky, že srovnatelnost dat bude uchována alespoň na akceptovatelné úrovni. Tyto položky jsou pak tzv. funkčně ekvivalentní [Scheuch 1993], neboť jejich vztah k latentnímu konstruktů je stejně silný, ale lépe zachycují všechny dimenze teoretického konceptu [Harkness et al. 2010: 46-49].

pro různé sociální skupiny, což podstatně komplikuje vzájemné srovnávání otázek. Scheuch [1993: 107 - 114] uvádí řadu konkrétních příkladů z vlastních výzkumů, kdy je doslovný překlad pojmů zcela nepřipustný. Například Bogardova škála sociální distance, která umožňuje měřit distanci jedince vůči sociální (většinou etnické či národnostní) skupině, musela být v německé verzi dotazníku modifikována. Jeden ze stupňů Bogardovy škály je položka „přijetí příslušníka určité skupiny za souseda“ [“have as a neighbor in the same street”], přičemž, aby mohl být daný význam sociální distance vyjádřen ekvivalentně i v německém jazyku, bylo zapotřebí změnit formulaci dané položky na „přijetí příslušníka určité skupiny za známého (se kterým se sice zdravím, ale moc dobře ho neznám)“ [“have as a greeting acquaintance”]; pojem souseď měl totiž pro Němce jiný sociální a emocionální význam.¹¹

Nejnámější překladatelský přístup používaný v mezinárodních výzkumech je tzv. technika zpětného překladu (back-translation model), která velmi pečlivě kontroluje překladovou ekvivalenci [Harkness, Schoua-Glusberg 1998: 97; Harkness 2003: 41].¹² Základní logika zpětného překladu spočívá v přeložení originálního (zdrojového) dotazníku do rodného jazyka obyvatelů dané země. Poté požaduje zpětný překlad přeložené verze dotazníku do původního jazyka, porovnat zpětný překlad s originálem a případně opět pokračovat v procesu překládání z této druhotně vytvořené verze dotazníku. V procesu překládání se má takto postupovat až k dosažení nejlepšího, nejbližšího možného výsledku překladu. Navíc každý tento krok překladu má ideálně za povinnost provést jiný překladatel, resp. tým spolupracovníků, což umožňuje snížit možná zkreslení. O osobách překladatelů se předpokládá, že jsou opravdu znalé kultury země, do jejíhož jazyka se dotazník překládá, podchytí všechny významy zkoumaného jevu a vyhnou se tak doslovnému překladu; někdy se dokonce zvažuje účast bilingvních překladatelů [Brislin 1976: 221].

Dnes již existují různé variace a vylepšení tohoto modelu zpětného překladu, např. metoda TRAPD - akronym pro Translation (překlad), Review (revize), Adjudication (posouzení), Pretesting (předběžné testování), Documentation (dokumentace) - která je využívána kupříkladu v projektu Evropský sociální výzkum. Tato metoda spočívá v provedení dvou nezávislých překladů, jejich revizi a finálním posouzení. Proces překladu postupuje podle specifického protokolu, je pečlivě zaznamenáván každý krok a jednotlivé rozhodovací úseky pro budoucí citace, odkazy a diskuze překladatelů a výzkumníků. Finančně zajištěn bývá standardně tým tří zkušených překladatelů pro překlad, kontrolu a úpravu překladu [Harkness 2003; European Social Survey 2010].

Zda se podařilo překladatelům přenést význam jednotlivých otázek do všech jazykových verzí dotazníků, se většinou zjišťuje pilotáží. Způsobů, jak ohodnotit pravděpodobnou srovnatelnost přeložených otázek v pilotáži, je několik. Lze předložit originální a přeloženou verzi dotazníku bilingvním osobám nebo malé skupině respondentů a porovnávat průměrné skóre položek v obou verzích dotazníku statistickými metodami (např. T-test). Pokud jsou výsledky významně rozdílné, dochází pak k přeformulování otázek nebo jejich vyloučení. Stejně tak může být posuzována souběžná, obsahová a konstruktová validita měřicího nástroje a mezipoložková reliabilita, nebo mohou být jednotlivé otázky v obou verzích překladu hodnoceny na základě podobného významu nebo své jednoznačnosti či kulturně vhodného použití za pomoci kognitivních rozhovorů [Mullen 1995; Ægisdóttir et al. 2008; více o moderních přístupech k překladu Harkness, Villar, Edwards 2010].

¹¹ Jak dále Scheuch uvádí, daná položka sice neměla v Německu zcela totožnou pozici na stupnici sociální distance, ale měla v analýze dat stejnou ordinální pozici na škále (tzn. stejné umístění v pořadí sedmi položek Bogardovy stupnice), a proto byla celá škála pro potřeby analýzy dat tzv. konstruktově ekvivalentní a tudíž částečně srovnatelná.

¹² Termínu „překladová ekvivalence“ se překladatelé velmi brání a spíše preferují termín kvalita překladu, vhodnost překladu či adekvátnost překladu [Harkness, Schoua-Glusberg 1998: 94].

2.2.3 *Metody využívající kognitivních přístupů*

V průběhu pilotáže nemusí docházet pouze k testování kvality překladu otázek do cizích jazyků, lze zkoumat daleko širší problematiku – kvalitu celého výzkumného nástroje. Lze se tak dostat až za hranice sestavování tzv. dobrých otázek („good“ question), pro které se vymezují obecně stylistická, syntaktická a gramatická doporučení, do oblasti kognitivních přístupů. Kognitivní přístupy se zabývají psychologickými stránkami získávání informací od respondentů ve standardizovaných dotazníkových šetření, v centru jejich pozornosti stojí problematika interpretace otázky na individuální a interpersonální úrovni [Vinopal 2008: 11]. Nejen že díky kognitivním přístupům je možné mapovat, jakým způsobem respondent reaguje na výzkumné otázky a jak chápe a zpracovává položky v dotazníku, také lze identifikovat reálné či potenciální zdroje zkreslení odpovědí na úrovni interakce a komunikace respondenta s tazatelem; v mezinárodních šetřeních dokonce i na úrovni kulturních a jazykových kontextů [Braun, Harkness 2005; Schwarz et al. 2010].¹³

V mezinárodních výzkumech je tak tematiku kognitivních přístupů obzvláště třeba chápat v širším kontextu výzkumné chyby, neboť, jak již bylo řečeno výše, chyby měření jsou významným zdrojem narušování různých úrovní ekvivalence a podstatně tak ztěžují vzájemné srovnávání položek. Vinopal [2008, 2009] uvádí, že kognitivní přístupy se věnují chybám pozorování, které vznikají na úrovni výzkumného nástroje, tazatele, respondenta a modu sběru dat.¹⁴ A jejich praktickým účelem je mimo jiné tuto chybu redukovat a pokusit se zvýšit kvalitu otázek i celých dotazníků [Vinopal 2009: 399].

Vinopal [2008: 49] podotýká, že každá kognitivní technika testování dotazníků může sloužit různým cílům, přináší jiný typ poznatků a má různý potenciál ve smyslu možného vylepšení cílových otázek. Techniky, které lze v současnosti použít k ověření kvality dotazníku, jsou kognitivní rozhovor, debriefing, kódování chování účastníků rozhovoru, posouzení kognitivními experty, měření reakčního času, atd. [více Vinopal 2008: 51-57]. Pro účely mezinárodního srovnávání je nejčastěji využíváno kognitivních rozhovorů (cognitive interview) a metody debriefingu.

Kognitivní rozhovor odkrývá výzkumníkům to, co je ukryté pod běžně sledovaným povrchem odpovědí na věcný dotaz [Vinopal 2008: 51], umožňuje poznat, proč respondent odpověděl na otázku daným způsobem. Otázky typu - „Jakým způsobem jste dospěl k Vaší odpovědi?“, „Jak byste otázku vyjádřil vlastními slovy?“, „Co podle Vás tato otázka/výrok znamená, jaký má podle Vás význam?“ [Pan et al. 2010: 92] – jsou základním stavebním kamenem kognitivních rozhovorů, které mají většinou podobu hloubkových rozhovorů a využívají techniky „přemýšlení nahlas (thinkaloud)“ či „zpětného dotazování (retrospective probing)“ [více Vinopal 2008: 51-53]. V mezinárodních výzkumech se tyto rozhovory většinou realizují s bilingvními tazateli za účasti kombinace bilingvních a monolingvních respondentů [více Goerman, Caspar 2010], neboť se nejčastěji testuje překladová ekvivalence. Navíc se předpokládá, že sladění tazatele s respondentem obzvláště v sociodemografických charakteristikách, jazyku a etniku hraje při kognitivních rozhovorech ještě větší roli z hlediska systematického zkreslování výsledků než při standardizovaném dotazníkovém šetření v terénu, proto se na tuto souhru klade o to větší důraz [Goerman, Caspar 2010: 79].

Metoda debriefingu pracuje již většinou přímo s tazateli, kteří mají za úkol zhodnotit průběh rozhovoru a odhalit slabá místa v dotazníku [více Vinopal 2008: 54]. Tazatelé sdělují

¹³ Zvláštní zřetel je brán na překladovou ekvivalenci škál odpovědí, efekt kontextu nabídky odpovědí a preferenci jedné kategorie odpovědí před jinými.

¹⁴ Chyby pozorování mohou mít jak systematický, tak i nesystematický charakter, přičemž v problematice ekvivalence je zdůrazňován především vliv systematických chyb měření na srovnatelnost jednotlivých otázek [více van de Vijver, Leung 1997].

výzkumníkům své zkušenosti z rozhovorů, komentují interpretaci jednotlivých slov v otázce, hodnotí, zda je otázka kulturně vhodně zvolená, uvádějí své dojmy z respondentů a z jejich reakcí na otázky. V mezinárodních výzkumech jsou to většinou jazykoví experti anebo rodilí mluvčí, kteří zaznamenávají kulturní normy mluvení a preferovaný styl komunikace respondentů v různých zemích [Pan et al. 2010: 99].

2.3 Techniky testování ekvivalence položek

Ukazuje se, že první přípravná fáze výzkumu – vývoj srovnatelných otázek, jejich překlad a tvorba dotazníků - většinou přináší centrálním či národním koordinačním týmům daných projektů, které se problematice ekvivalence již po dlouhou dobu vytrvale a systematicky věnují.¹⁵ Mnoho mezinárodních projektů v současnosti již také rutinně spolupracuje se společností cApStAn [www.capstan.be/content/home.html], zaměřující se na kontrolu kvality překladu v případě multilingvních projektů a zajištění lingvistické a sémantické ekvivalence obsahu textu dotazníku. Avšak navzdory tomuto velkému úsilí vytvořit a předložit respondentům metodologicky kvalitní dotazník s plně ekvivalentními otázkami a srovnatelnou nabídkou odpovědí se tyto snahy nemusí setkat se stoprocentní úspěšností, obzvláště v případě velkého počtu participujících zemí.

Systematické chyby měření mohou zkreslit škálu měření do různé míry a rozsah statistické srovnatelnosti dat tak nelze předem odhadnout. Z mnoha příkladů z praxe také plyne, že i mezinárodní výzkumy se kontinuálně rozvíjejí a kvalita jejich výsledků roste společně s množstvím opakovaných šetření, neboť se jejich metodologie systematicky vylepšuje na základě chybných kroků z minulé vlny šetření. Pokud tedy výzkumník přistoupí k sekundární analýze prvních sérií sběru dat či se pokusí získat časový vývoj sociálního jevu, je možné, že se začne potýkat s problémem, jak komparovat ne zcela ekvivalentní indikátory.¹⁶ Ukazuje se, že i tam, kde se to neočekává, nelze občas připustit určitý druh statistické komparace mezinárodních dat nebo je nezbytné vyřadit určitou zemi z komparativní analýzy, aby nedocházelo ke zkreslování kulturních rozdílů mezi zeměmi. Nelze tedy srovnatelnost položek jen očekávat na základě kvalitně odvedené práce překladatelů při tvorbě dotazníků¹⁷, je většinou nezbytné dosaženou úroveň ekvivalence indikátorů v průběhu analýzy dat statisticky otestovat; respektive zjistit, zda se jedná o skutečné kulturní rozdíly v náhledu na daný teoretický koncept, anebo o rozdíly mezi specifickými interpretacemi položených otázek a jejich škál měření.

Způsobům testování ekvivalence indikátorů byla v zahraniční odborné literatuře věnována již velká pozornost, avšak až o mnoho let později než problematice překladu dotazníků. Hlavní příčina by se dala přisoudit především nedostatečnému rozvoji výpočetní techniky v 60. letech a bleskovému zdokonalování softwarových možností v současnosti, neboť většina technik testování dosažené ekvivalence indikátorů na úrovni měření vyžaduje vícerozměrnou statistickou analýzu. Časová a finanční náročnost takového testování a nejistota většiny sociálních výzkumníků ohledně možností jeho realizace posunula analytické a metodologické práce na tomto tématu až do posledních přibližně 30 – 40 let [Johnson 1998: 2]. Zájem o to,

¹⁵ Příkladem může být struktura organizace realizace mezinárodního sociálního výzkumu ESS (European Social Survey) s centrálním koordinačním týmem ve Velké Británii [více www.europeansocialsurvey.org]. Stejně tak je tomu u projektů ISSP (International Social Survey Programme), PISA (Programme for International Student Assessment), PIAAC (Programme for International Assessment of Adult Competencies), ESV (European Social Value), atd.

¹⁶ Na tuto problematiku upozorňoval van Deth [2003: 296], podle něhož je nutné myslet na všechny problémy spojené s analýzou sekundárních dat. Ty lze nalézt ve třech metodologických oblastech a to odlišné operacionalizaci položek, nevhodném měření klíčových jevů a neadekvátní dokumentaci.

¹⁷ Jak podotýká Davidov [2012]: „Dobrý překlad je nutnou, avšak ne dostačující podmínkou k získání srovnatelných indikátorů. Pro některé idiomy či fráze není možné nalézt v některých jazycích adekvátní výrazy.“

jak nejlépe mezinárodní šetření kontrolovat, se ale stále zvyšuje, z tohoto důvodu se také testování ekvivalence dat postupně stává v současnosti nedílnou součástí mnoha mezikulturních studií a to nejen metodologické povahy.

Ekvivalence se v současnosti testuje nejčastěji u postojových otázek, protože u nich je hrozba zkreslení chybou měření nejvyšší. Navíc je zapotřebí klást tyto otázky v dotazníku v tzv. baterii otázek, to znamená v souboru několika dalších otázek, které všechny směřují k zachycení jednoho jevu, jedné dimenze (např. mezilidské důvěry).¹⁸ Tato dimenze se v jiné terminologii nazývá latentní proměnná, kterou nelze měřit přímo, ale jen zprostředkovaně prostřednictvím nejlépe několika měřících indikátorů. Tento základní předpoklad pro testování ekvivalence položek je a) teoreticky logický – postojové (i jiné) otázky se v mezinárodních dotaznících nejčastěji používají pro ověřování teorie a přirozeně zachycují latentní proměnné (neboli teoretické koncepty), b) metodologicky pochopitelný – větší množství položek měřící jednu latentní proměnnou zpravidla vyústí v přesnější měření (ve smyslu redukce chyb měření) a c) empiricky nezbytný – neboť v současnosti lze za prvé kontrolovat chybu měření pouze u baterie otázek, nikoli u jednoduchých položek, a za druhé je zvyklostí testovat ekvivalenci těchto položek na základě jejich vztahu k latentní proměnné.¹⁹

Obecně lze konstatovat, že v případě testování srovnatelnosti položek hrají primární roli statistické techniky. Zde uvedený přehled je řadí od těch nejzákladnějších až po ty nejsložitější analytické přístupy. Byl sestaven na základě výčtu technik Brauna a Johnsona [2010: 377] a Johnsona [1998: 25-29] s podporou poznatků z řady dalších empirických studií, kde byly dané techniky aplikovány v praxi při analýze dat nebo vzájemně porovnávány. Kritériem pro výběr vhodné techniky může být kromě obvyklé volby, zda do analýzy vybrat kategoriální či kardinální proměnnou, také její analytická a potažmo časová náročnost, počet zemí a počet proměnných, které lze do analýzy zahrnout, anebo míra podrobností, které je zapotřebí zjistit o nesrovnatelných položkách v dané zemi a příčinách této situace [Braun, Johnson 2010: 376].

2.3.1 Statistická analýza prvního a druhého stupně

Jednoduchá analýza položek není při testování jejich srovnatelnosti mezi zeměmi příliš běžnou volbou, ale velmi lehce se na data aplikuje a umožňuje bezprostřední a rychlé pochopení toho, zda obecně pracujeme se srovnatelnými či nesrovnatelnými daty. Toto je obzvláště vhodné pro výzkumníky, které zajímá především zkoumaný problém a ne metodologická analýza [Braun, Johnson 2010: 375], ale přirozeně mají tyto techniky svá slabá místa. Konkrétně se jedná například o analýzu procentuálního rozložení (distribuce) jednotlivých kategorií odpovědí a jeho variability napříč zeměmi. Důležitou informací je také údaj o respondentech, kteří na danou otázku neodpověděli (tzn. chybějící případy), protože právě tento údaj může poukazovat na problém s interpretací či překladem dané otázky [Braun, Johnson 2010: 379]. Tyto varianty testů jsou přirozeně ošidné při větším množství zkoumaných zemí, protože výstupy jsou velmi nepřehledné [Braun, Johnson 2010: 378].

Dále se mohou porovnávat průměry ordinálních a kardinálních proměnných mezi dvěma zeměmi (T-test) anebo mezi více zeměmi (analýza rozptylu). Braun a Johnson [2010: 380] ale důrazně upozorňují na to, že tento test lze realizovat pouze, pokud jsme si jisti tím, že dané

¹⁸ Není nutné klást otázky v dotazníku doslova za sebou či přímo v nějaké baterii otázek – podstatné je moci vytvořit tuto sadu otázek při analýze dat.

¹⁹ Velice speciálním případem je testování ekvivalence a chyby měření jednoduché položky měřenou třemi různými metodami sběru dat (MTMM design), kde tyto tři odlišně měřené položky mohou v modelu „měřit“ jednu latentní proměnnou. V této podobě je pak možné otestovat jejich kulturní srovnatelnost a efekt metody sběru dat prostřednictvím konfirmační faktorové analýzy [podrobněji Revilla, Saris 2011].

položky jsou skalárně ekvivalentní. Proto navrhuji porovnávat průměry různých proměnných mezi zeměmi a na základě toho sledovat pouze vytvořené pořadí zemí. Pokud mají dané proměnné měřit stejný jev (tzn. latentní proměnnou), měly by také podle průměrů vytvořit stejné pořadí zemí. Jestliže tomu tak není, položky pravděpodobně neměří stejný jev a mohou tak výrazně zkreslit výsledky kulturního srovnávání, protože ty se budou lišit v závislosti na vybraném indikátoru [Braun, Johnson 2010: 380].²⁰ Byrne a Watkins [2003] použili analýzu rozptylu i jako alternativní techniku k identifikaci položkové systematické chyby měření. Třetí variantou připadající v úvahu je srovnávání korelace měřicích položek reprezentujících latentní proměnnou s identifikačními proměnnými, s nimiž by měly vykazovat podobné vztahy napříč zeměmi, pokud mají mít ve všech kulturách stejný význam [více Johnson 1998: 25; Braun, Johnson 2010: 380 - 381].

2.3.2 *Mnohorozměrné statistické analýzy*

Složitější statistické techniky vyžadují od výzkumníka daleko více trpělivosti a matematických dovedností. Většinou na zjištění toho, že se v datovém souboru vyskytují nesrovnatelné otázky, připadá mnohem větší díl analytické práce než při aplikaci jednoduchých technik, což ještě v součtu s nepříliš uživatelsky příjemnými statistickými programy činí z těchto metod značně nepopulární analytický nástroj. Navíc nedodržování poměrně striktních pravidel při testování dat mnohorozměrnými technikami může vést až k takovému zkreslení výsledků, že se na problém s neekvivalentními daty jednoduše nepříjde [Braun, Johnson 2010: 376]. Jejich hlavní výhoda ale tkví především v podrobnosti zjištění, která tyto techniky o (ne)srovnatelných položkách generují, a počet zemí a počet proměnných, které můžeme do analýzy zahrnout.

Konkrétních analýz, které lze pro účely testování ekvivalence položek použít, je mnoho: explorační faktorová analýza, testování reliability škál, mnohorozměrná korespondenční analýza, mnohorozměrné škálování, strukturní modelování, víceúrovňové modelování, analýza latentních tříd a teorie odpovědi na položku. Není zde z praktických důvodů ani možné a ani efektivní a potřebné detailně technicky popisovat a vysvětlovat všechny zmíněné druhy analýz, z tohoto důvodu zde budou dále podrobněji rozebírány pouze ty analýzy, které jsou nejčastěji a s nejlepšími výsledky aplikovány na mezinárodních datech (viz poslední čtyři ve výčtu). Pro informační účely jsou ale u všech metod v Tabulce 1 souhrnně uvedeny alespoň odborné studie, kde se lze o povaze jejich metodologického využití při testování ekvivalence položek dozvědět více.

2.3.2.1 *Strukturní modelování*

Strukturní model je formální (grafickou a matematickou) reprezentací určité teorie (komplexní hypotézy) o kauzálních vztazích mezi proměnnými, přičemž se často očekává, že sledované jevy a kauzální vztahy mezi nimi reflektují hlubší souvislosti z reality [Matějů 1989: 402]. Výhodou strukturního modelování je především to, že umožňuje pracovat nejen s manifestními (přímo měřenými) proměnnými (v tom se příliš neliší od regresní analýzy), ale též s latentními proměnnými, které jsou definovány přímo v modelu (podobně jako ve faktorové analýze), proto lze na strukturní modelování pohlížet jako na kombinaci faktorové a regresní analýzy s tím, že je obvykle posuzováno spíše jako konfirmační než explorativní technika [Matějů 1989; Hox, Bechger 1998: 1].²¹

²⁰ Podobným způsobem lze analyzovat další míry centrální tendence (modus, medián), střední chybu průměru a i jiné míry variability (resp. míry rozptýlenosti), interval spolehlivosti, atp.

²¹ Do souboru technik strukturního modelování lze zahrnout také několik tradičních mnohorozměrných analýz jako speciální případy modelování, např. korelační analýzu, regresní analýzu, kanonickou korelační analýzu, diskriminační analýzu a faktorovou analýzu [Hox, Bechger 1998: 1].

Modely jsou typicky zobrazovány grafickým diagramem vycházející z pěšinkové analýzy (path analysis), kde jsou měřicí položky umístěny do obdélníků, latentní proměnné do oválů a chyby měření do kroužků. Základem je odlišení modelu měření (measurement model), který přesně stanovuje, která měřicí položka náleží k dané latentní proměnné, a strukturního modelu (structural model), který popisuje vzájemné vztahy mezi těmito latentními proměnnými a je již oproštěn od variance chyb měření daných proměnných, což je jeho velkou výhodou [více Matějů 1989: 408]. Tento statistický model je pak obvykle matematicky reprezentován řadou strukturních (lineárních regresních) rovnic, které představují dané kauzální vazby. V průběhu analýzy se odhadují nejen parametry modelu (zpravidla regresní koeficienty, chyby odhadů, vyčerpané variance, atp.), ale posuzuje se i jeho celková vhodnost, obvykle na základě chí-kvadrát testů, které jsou ale příliš náchylné na velikost výběrového souboru.²² Z tohoto důvodu se dnes prosazují i jiné statistiky vhodnosti modelu (např. BIC, CFI, RMSEA), které zohledňují za prvé ještě akceptovatelnou jednoduchost modelu, chyby měření a také míru, do jaké testovaný model reprodukuje výchozí kovarianční matici [více Meade, Johnson, Braddy 2006].

Ve strukturních modelech je ale také navíc možné otestovat hypotézu o ekvivalenci položek na úrovni měření prostřednictvím několika výběrové konfirmační faktorové analýzy (MGCFAs), což je v současnosti nejvyužívanější postup ověřování ekvivalence dat.²³ V prvním kroku analýz je klasicky ověřena konstruktová, resp. konfigurační ekvivalence dat. Konkrétně se testuje podobnost modelů měření, které by měly mít stejnou faktorovou strukturu v jednotlivých skupinách/výběrech, tzn. statisticky významně velké faktorové zátěže měřicích položek a stejný počet faktorů ve všech zemích [Meredith 1993; Steenkamp, Baumgartner 1998, Vandenberg, Lance 2000]. Je to základní typ ekvivalence mezi daty, který ale ještě neumožňuje přímé statistické porovnávání zemí mezi sebou v daných položkách; pouze z něj vyplývá, že zkoumaný teoretický jev je podobně chápán ve všech kulturách a je také podobně měřen v každé zemi.

Druhým krokem je analýza metrické ekvivalence, kde se testuje shoda velikostí faktorových zátěží (směrnic) položek v modelech měření ve všech zemích [Vandenberg, Lance 2000: 37]. Pokud je metrická ekvivalence dat potvrzena, škály měření daných položek jsou srovnatelné mezi zeměmi z hlediska jednotky a rozsahu, ale ještě stále není jisté, zda je škála měření používána respondenty stejným způsobem. Proto přichází v úvahu pouze porovnávání vztahů latentních proměnných a měřicích položek s ostatními proměnnými z datového souboru, u kterých byla prokázána metrická ekvivalence, napříč zeměmi (např. korelační a regresní analýza). Nejvyšší úroveň, tzv. skalární ekvivalence dat, se ověřuje testem shody velikosti konstant měřicích položek v modelu měření.²⁴ Skalární ekvivalence dat je nutnou podmínkou pro komparaci takových statistických ukazatelů, jako je aritmetický průměr měřicí položky či průměr latentních proměnných (resp. faktorů). Posledním typem ekvivalence je tzv. striktní ekvivalence, která vypovídá o shodě variancí chyb měření položek v daných skupinách, a na základě ní lze ověřit podobnost reliability položek.²⁵

²² To je značný problém při zařazení velkého množství zemí do analýzy, protože i když model měří data velmi dobře, statistiky vhodnosti založené na chí-kvadrát testech mají tendenci tento model zamítnout [Hox, de Leeuw, Brinkhuis 2010: 397- 98].

²³ Strukturní analýzy, které testují, zda jsou parametry daného modelu pro danou populaci skutečně signifikantně odlišné od parametrů získaných na jinak definované populaci, se nazývají několika výběrové analýzy (multisample analysis) [Matějů 1989: 406].

²⁴ V tuto chvíli je zřejmá kombinace regresní a faktorové analýzy, neboť dochází k testování parametrů v regresních rovnicích (tj. směrnic a konstant), které reprezentují vztahy mezi faktory a měřicími položkami [více Hox, Bechger 1998: 1].

²⁵ Je možné dále v analýzách pokračovat a testovat srovnatelnost strukturních částí modelů, jako je ekvivalence kovariancí a variancí faktorů, ale tyto testy vypovídají více o konstruktní validitě měřicího nástroje spíše než o ekvivalenci položek na úrovni měření.

Velmi zvláštním případem je ve strukturním modelování model MTMM (Multitrait-Multimethod model), který má v případě testování ekvivalence položek speciální úlohu a to odhalit možný zdroj systematického zkreslení položek v metodě dotazování. Lze tak vytvořit za prvé ve strukturním modelování model měření sestávající z jedné latentní proměnné, která je měřena nejméně třemi měřicími položkami pocházejících ze stejného sběru dat (např. CAPI), a porovnat jej na základě konfirmační faktorové analýzy s modelem měření založeném na položkách z jiného metodologického sběru dat (např. CAWI). Za druhé lze za podmínky měření položek různými metodami sběru dat ověřit ekvivalenci a chyby měření jednoduchých položek, které nejsou součástí baterie položek; latentní proměnnou potom měří tři stejné položky sebrané v šetření různými metodami sběru dat. Tento přístup je poměrně novátorský a konečně umožňuje ověřovat srovnatelnost jednoduchých položek [více Revilla, Saris 2011].

Rámeček 3 Multitrait-Multimethod design

Poprvé byl MTMM design představen Campbellem a Fiskem v roce 1959, kteří zjistili, že otázky v dotaznících interagují s konkrétní metodou sběru dat a to má pak za následek rozdílnou interpretaci otázek či narušení vztahů mezi proměnnými [Saris 2003: 268]. Vznikají tedy náhodné a systematické chyby ve výzkumu, které od sebe nelze odlišit, pokud se nepřistoupí k více experimentálnímu designu výzkumu. Základem tohoto nového MTMM designu jsou opakovaná měření několika proměnných různými metodami současně (např. dotazování pomocí papírových dotazníků, telefonu či počítače), většinou v kombinaci minimálně tří otázek v dotazníku a tří využitých metod sběru dat [Řehák 1998b]. Postupně se provádí metaanalýza dosavadních poznatků z MTMM experimentů (ty k současnému datu otestovaly více jak 3 000 otázek), na jejímž základě se v současnosti buduje webová aplikace Survey Quality Prediction [SQP 2.0; více www.sqp.nl]. Ta umožňuje výzkumníkům vytvářet otázky a předpovídat následně jejich kvalitu z hlediska validity, reliability a efektu metody z různých parametrů daných otázek [více Saris, Gallhofer 2007].

Nevýhoda strukturních modelů tkví v tom, že do modelu nejsou zahrnuty různé úrovně měření tak, jak je tomu v případě víceúrovňového modelování, čímž výzkumník přichází o možnost vysvětlovat rozdíly mezi průměry latentních proměnných nějakou kontextuální proměnnou (např. zemí, kulturou) a nacházet tak možné příčiny systematického zkreslení položek [Hox, de Leeuw, Brinkhuis 2010: 400]. Z tohoto důvodu se začínají v současnosti prosazovat pro testování ekvivalence položek víceúrovňové strukturní modely (multilevel structural equation modelling, MSEM), kde se testuje ekvivalence modelů měření a strukturních modelů na každé hierarchické úrovni dat zvlášť (tzn. položka, respondent, země).

2.3.2.2 Víceúrovňové modelování

Víceúrovňové (hierarchické) modelování (Multilevel Modelling) je metoda speciálně vyvinuta pro statistickou analýzu dat, které mají hierarchickou či shlukovou strukturu [Hox, de Leeuw, Brinkhuis 2010: 401].²⁶ Principiálně vychází z konceptu kontextuální analýzy, resp. ze skutečnosti, že mnohé proměnné nemá smysl měřit na individuální úrovni (micro level), ale až na úrovni větších celků (macro level), a že působení faktorů na těchto dvou úrovních může být různé [Soukup 2006a: 988]. Soukup [2006a: 989-994] navíc zdůrazňuje, že až v případě, že zohledníme hierarchickou strukturu v datech a budeme využívat v analýze hodnot měřených na mikro- i makroúrovni zároveň, lze správně zobecňovat závěry analýzy na celý základní soubor (např. jedince, třídy a školy).

²⁶ Většinou se jedná o data, která vycházejí z vícestupňového náhodného výběru, a jejich analýza tak musí plně respektovat strukturu (design) výzkumu [Soukup 2006a: 993].

Víceúrovňové modely lze chápat ze statistického hlediska jako zobecnění klasické lineární regresní analýzy, kdy se ale nezůstane jen při modelování jedné (první) úrovně (např. jedinec/žák), ale modelujeme každou relevantní úroveň zvlášť (jedinec, třída, škola, země, atp.), protože nemůžeme jednotlivá pozorování v rámci jednotlivých skupin považovat za vzájemně nezávislá [Soukup 2006a: 990]. Jsou tedy založeny na víceúrovňové regresní analýze a reprezentuje je hierarchický systém regresních rovnic [Hox, de Leeuw, Brinkhuis 2010: 401]. Jejich velkou výhodou je možnost zařadit do analýzy velké množství zemí.

Pokud je v případě komparativního výzkumu cílem testovat srovnatelnost položek prostřednictvím analýzy víceúrovňového modelování, postupuje se při modelování vztahů tak, že se odpověď, bez ohledu na jakou položku, považuje za závislou proměnnou. V tom případě pak položky představují první úroveň, respondenti druhou úroveň a zkoumané země třetí úroveň modelu. Tímto způsobem lze zjistit, která z těchto tří úrovní způsobuje varianci odpovědí, resp. položkovou systematickou chybu [Braun, Johnson 2010: 387]. Nevýhoda této metody tkví ve špatné identifikaci odlehlých pozorování, například, ve které zemi došlo k neekvivalentnímu překladu otázky. Proto je pro účely testování ekvivalence vhodnější zkombinovat strukturní a víceúrovňové modelování [podrobněji Braun, Johnson 2010: 386-390].

2.3.2.3 Teorie odpovědi na položku

Teorie odpovědi na položku (Item Response Theory) je moderním metodologickým přístupem, který je již dlouhodobě využíván v psychometrii [Stout 2001, 2002].²⁷ Přispívá ke zvyšování kvality sociálních výzkumů, protože umožňuje uspokojivě zodpovídat otázky ohledně funkce každé položky v dotazníku, její rozlišovací schopnosti a informačního přínosu pro test (tj. dotazník) [Rudá 2012: 7]. V oblasti metodologie mezinárodních výzkumů ji lze efektivně využít při sledování toho, zda byl zachován v průběhu překladu dotazníku význam daných položek, při posuzování relevance výběru kategorií a jejich optimálního počtu do škály odpovědí a především při analýze položkové systematické chyby [Van de Vijver 2003b: 216; Meade, Lautenschlager 2004].

Matematické modely vniklé na základě IRT znázorňují vztah mezi měřenou latentní proměnnou a odpovědí na konkrétní položku, což IRT značně odlišuje od klasické teorie testů (CTT), které hodnotí validitu a reliabilitu na základě jednoduchých statistických ukazatelů součtů a průměrů.²⁸ IRT modely jsou založeny na pravděpodobnosti (tj. šancích), s jakou určitý respondent odpoví na konkrétní položku v souladu s jeho úrovní měřené latentní proměnné; tím se podobají logistické regresi [Stout 2002; Rudá 2012]. V modelech jsou opět odhadovány určité parametry stejně, jako je tomu u konfirmační faktorové analýzy (resp. strukturního modelování). Hlavní rozdíl mezi nimi ale spočívá v tom, že IRT pracuje většinou s každou položkou odděleně, ale CFA odhaduje parametry všech měřících položek (tj. celé baterie položek) v dané latentní proměnné najednou.

IRT sleduje ekvivalenci položek především z hlediska tzv. odlišného fungování položek (Differential item functioning, DIF), což je v terminologii IRT vyjádření pro systematickou chybu měření (tj. bias). Systematické zkreslení může být v kontextu teorie odpovědi na položku identifikováno dvěma různými způsoby. Jedním z nich je porovnání parametrů

²⁷ V české odborné literatuře je IRT také známá jako teorie položkové analýzy.

²⁸ Klasická teorie testů (CTT), teorie odpovědi na otázku (IRT) a G teorie (generalizability theory) jsou dnes považovány za tři překrývající se přístupy k testování reliability měření. Klasická teorie testů pracuje s typickou představou o zkreslení skutečného skóru chybou měření (s jediným zdrojem) a používá k analýze reliability již dlouho zavedené přístupy testu a retestu, analýzy mezipoložkové reliability, vnitřní konzistence škály, atp. G teorie se jako mnohem modernější přístup snaží identifikovat různé zdroje zkreslení položky (tj. v terminologii G teorie aspekt/facet) a zvažuje i jejich vzájemnou interakci [Mushqash, O'Connor 2006; více o G teorii Brennan 2001].

pomoci klasického chí kvadrátu²⁹, druhým je porovnáním plochy oblasti mezi charakteristickými křivkami položek (ICC) v různých kulturních skupinách. V praxi je také důležitá vizuální analýza, porovnání tvarů funkce pro každou skupinu odděleně [Jelínek, Květoň, Vobořil 2011: 89]. Tyto křivky reprezentují pravděpodobnost označení dané kategorie odpovědi v rámci sledované položky, a pokud si křivky vzájemně odpovídají ve vybraných populacích, tzn. pravděpodobnost výběru odpovědi na danou otázku je totožná ve všech skupinách pro danou latentní proměnnou, je potvrzeno, že jednotlivé položky měří pro každou ze skupin stejný konstrukt a že jsou tudíž srovnatelné [Bhalla, Lin 1987: 282; Johnson 1998: 26; Woehr, Meriac 2010: 423].

Tato metoda pracuje se statistickým modelem s jedním základním předpokladem o jednodimenzionalitě latentní proměnné reprezentované několika měřicími položkami [Jelínek, Květoň, Vobořil 2011]. Právě striktní požadavek na jednodimenzionalitu latentní proměnné je hned vedle dalšího požadavku na velký počet položek zařazených do analýzy, které by měly být měřeny ideálně dichotomickou škálou odpovědí, podnětem řady kritik na účet IRT.³⁰ Navíc je teorie odpovědi na položky krajně nepraktická v případě velkého množství zemí zařazených do analýzy [Braun, Johnson 2010: 390].³¹ Na druhou stranu, několik studií již srovnávalo výsledky testování ekvivalence prostřednictvím CFA i IRT a docházejí k tomu, že a) tyto dvě metody mají jiné předpoklady (linearita/nelinearita modelů, jednodimenzionalita/vícemimenzionalita faktorů), b) testy mohou identifikovat i velmi rozdílné problematické položky, c) v některých případech může IRT generovat podrobnější a užitečnější informace o povaze systematického zkreslení položek, d) žádná z těchto metod není bez nevýhod, a e) nejlepších výsledků lze dosáhnout při aplikaci jich obou [Meade, Launtenschlager 2004; Raju, Laffitte, Byrne 2002].

2.3.2.4 Analýza latentních tříd

Mnohorozměrná analýza latentních tříd (Latent Class Modeling), která vychází z analýzy latentních struktur, významné metodologické inovace P. Lazarsfelda [více Jeřábek, Soukup 2008], obecně hledá latentní třídy podobných respondentů ve sledované populaci a odhaduje vhodný počet těchto tříd. Respondenti jsou rozřazováni do latentních tříd zpravidla na základě kategoriálních měřících položek - v tom je její obrovská výhoda oproti jiným technikám, i když v současnosti se možnosti této analýzy rozrůstají i o volbu ze spojitých proměnných [Hox, de Leeuw, Brinkhuis 2010: 403-404]. Analýza latentních tříd dokáže zručně kombinovat některé výhody strukturního modelování a víceúrovňového strukturního modelování tím, že modeluje rozdíly mezi zeměmi jako rozdíly mezi latentními třídami respondentů, jejichž počet je ale menší než počet zemí [Hox, de Leeuw, Brinkhuis 2010: 396].

V případě, že tuto vícerozměrnou metodu používáme při testování srovnatelnosti položek, postup ověřování ekvivalence se značně shoduje s přístupem konfirmační faktorové analýzy v modelu měření. Předpokládáme, že latentní model má více méně podobu konfirmačního faktorového modelu, ve kterém zjišťujeme rozdíly mezi latentními třídami [Hox, de Leeuw, Brinkhuis 2010: 404]. Dále už tato metoda funguje z metodologického (ne statistického)

²⁹ Tzv. odlišné fungování položky se prokáže, pokud má položka statisticky významně odlišný odhadnutý parametr a , který vyjadřuje schopnost diskriminace položky a někdy je také označován jako směrnice charakteristické křivky položky, nebo odlišný parametr b , který značí obtížnost položky a někdy je také připodobňován ke konstantě položky v CFA modelu, i když jeho funkce a interpretace je zcela odlišná.

³⁰ Zahraněční literatura na téma různých statistických variací tohoto modelu IRT je ale v současnosti už poměrně bohatá. Nejznámější IRT model pro ordinální proměnné je model stupňovaných odpovědí (GRM), viz Stout [2001, 2002], May [2006], atp.

³¹ Kritika někdy zaznívá i ze strany psychometrie, například podle Borga [1998: 147] nemůže žádný statistický model zajistit správný překlad otázek, protože položky s podobnými statistickými charakteristikami ještě nemusí zachycovat zkoumaný sociální jev se všemi specifickými kulturními kontexty.

hlediska na obdobném principu jako strukturní modelování. Dochází k postupnému testování shody několika modelů a jejich parametrů – homogenního, částečně homogenního a heterogenního - lišících se mírou shody latentních struktur, silou vztahu mezi latentní třídou a položkami a vlivem kontextuální proměnné země (příslušnost k dané skupině) na latentní proměnnou a potažmo měřicí položky.

Tabulka 1 Vybrané mnohorozměrné statistické techniky pro testování ekvivalence postojových škál po sběru dat

Metoda	Vhodný typ proměnné	Empirické studie – aplikace metody na datech	Typ ekvivalence
Explorační faktorová analýza	Kardinální data Ordinální data* Nominální data (CATPCA)	Welkenhuysen-Gybels, van de Vijver 2001	Konstruktová
Testování reliability škál	Kardinální data Ordinální data*	Devins et al. 1997	Překladová
Mnohorozměrná korespondenční analýza	Kategoriální data (grafická metoda)	Blasius, Thiessen 2006	Konstruktová
Mnohorozměrné škálování	Ordinální data (grafická metoda)	Braun, Scott 1998; Fontaine 2003	Konstruktová
Víceúrovňové modelování	Kardinální data Ordinální data*	van de Vijver, van Hemert, Poortinga (eds.) 2008	Konceptuální
Strukturní modelování	Kardinální data Ordinální data* Nominální data (za podmínky vypočítání matice jiných než Pearsonovských korelací, např. polychorických, PRELIS)		Konfigurální Metrická Skalární
a) strukturní modelování		Mullen 1995; Billiet 2003; Byrne, Watkins 2003	
b) víceúrovňové strukturní modelování		van Hemert et al. 2002; van de Vijver, Poortinga 2002; Davidov et al. 2012	
c) konfirmační faktorová analýza		Hsueh et al. 2005; Lievens et al. 2007; Davidov 2010	
d) MTMM design		Saris 2003; Saris, Gallhofer 2007; Revilla, Saris 2011	Konceptuální
Analýza latentních tříd	Kategorická i kardinální data	Kankaras, Moors 2009	Konfigurální Metrická Skalární
Teorie odpovědi na položku	Dichotomická data Ordinální proměnné	Wang, Russel 2005; May 2006; Woehr, Meriac 2010	Překladová Skalární

Zdroj: vlastní zpracování podle Brauna a Johnsona [2010: 377]

* Ordinální proměnné jsou přípustné, pouze pokud je splněna základní podmínka, že vzdálenosti (intervaly) mezi jednotlivými kategoriemi (číslly) je možné považovat za stejné, rovnocenné. Takových proměnných není v sociologii mnoho, ale občas se vyskytují; v ideálním případě jde o proměnné intervalové.

3 POSTOJOVÉ ŠKÁLY SOCIÁLNÍHO KAPITÁLU A POLITICKÉHO ODCIZENÍ

Jak již bylo v úvodu práce řečeno, postojové škály v mezinárodních dotaznících, které se v procesu analýzy dat dále vyhodnocují, je nutné ověřovat z hlediska kvality a srovnatelnosti. Vedle hodnocení jejich validity a reliability je tak nezbytné posoudit i úroveň jejich ekvivalence napříč všemi zahrnutými zeměmi, aby bylo možno dále rozhodnout, v jakých zemích a jakým statistickým způsobem lze jejich výsledky komparovat. Tato metodologická práce ověřování postojových škál by neměla být jednorázová záležitost ve fázi testování obsahu a formy dotazníku před zahájením první vlny mezinárodního šetření. Měla by probíhat kontinuálně a držet krok se všemi realizovanými vlnami daného výzkumu. Proto i významné, robustní postojové škály, které se roky formovaly k dnešní podobě a pravidelně se ověřovala jejich reliability a validita, by se měly testovat z hlediska ekvivalence ve všech realizovaných vlnách šetření, neboť výsledky ohledně dosažené úrovně jejich srovnatelnosti se mohou v čase měnit a je pak nutné zvolit jiný statistický přístup k jejich analýze.

Cílem této práce je v tomto duchu tedy ověřit na jednom z nejprestižnějších komparativních projektů – Evropský sociální výzkum (dále jen ESS), zda budou prokazovat tak robustní postojové škály jako je škála měřící sociální kapitál a politické odcizení, potom co vykázaly v prvotních pilotážích vysokou míru reliability a validity [Saris, Gallhofer (n.d.)], také vysokou míru mezinárodní srovnatelnosti. Jako zdroj mezinárodních dat byl cíleně vybrán projekt ESS, protože patří mezi metodologicky velmi ceněné kontinuální mezinárodní výzkumy. Je zde z hlediska kvality zajištěný jak design výběru vzorku, tak i výpočet chyb měření (návrtnost, chybějící údaje, designový efekt, atp.), adekvátní výběr technik sběru dat, podrobná dokumentace a precizní překlad dotazníku, a jeho přípravě se věnuje rozsáhlý mezinárodní tým vědců [Dorer 2012; European Social Survey 2012]. Všechny tyto statisticko-technické a administrativní aspekty výzkumu mohou ovlivňovat úroveň dosažené ekvivalence dat, proto vysoká míra jejich kvality značí, že pro příčiny potenciálních obtíží se srovnatelností škál se lze obrátit do psychometrické oblasti k nástroji měření – dotazníku – a jeho formulaci, překladu a interpretaci respondenty. Navíc pokud by byla identifikována problematická srovnatelnost postojových škál v takovémto šetření, o to větší důraz by měl být potom kladen na ověřování srovnatelnosti měřících položek v méně metodologicky zpracovaných výzkumech.

Postojové škály měřící sociální kapitál a politické odcizení byly vybrány primárně kvůli neoddiskutovatelné důležitosti těchto dvou sociálně vědních konceptů pro současné demokratické společnosti. Již 20 let se hovoří o vzájemném propojení sociálního kapitálu se stabilitou občanské společnosti [Putnam 1995, 2000] a ideálu demokracie, kdy sociální důvěra vede k dobrovolné občanské participaci a naopak [Paxton 2002]. Není tedy důvod nepředpokládat poměrně silný vzájemný vztah mezi nízkým sociálním kapitálem a pocitem politického odcizení a zároveň vztah mezi vysokým sociálním kapitálem a pocitem politické efektivity jedince, a i mezi jejich jednotlivými subdimenzemi. Navíc jsou tyto dva koncepty tradičně kvůli (re)formulaci sociálních a politických teorií analyzovány i ve vztahu k jiným důležitým sociálním jevům (např. k sociálním normám a hodnotám, politické angažovanosti, politické orientaci, životní spokojenosti, občanské morálce a tendenci ke korupčnímu jednání, atp.) a tyto výsledky jsou pravidelně komparovány napříč demokratickými společnostmi [např. viz Meulemann (Ed.) 2008; Ferragina (Ed.) 2012]. Je tedy nezbytné znát míru jejich mezinárodní srovnatelnosti.

ESS tyto dva koncepty do svého výzkumu zařazuje už od první vlny šetření a obzvláště sociálnímu kapitálu se věnuje kontinuálně až dodnes. I řada dalších mezinárodních šetření se měřením těchto sociálních jevů systematicky zabývá (např. ISSP, WVS, EVS, EB), což také nepřímo potvrzuje jejich důležitost a aktuálnost v současné sociologii a dalších sociálních vědách. Navíc položky, které tyto koncepty tradičně měří, jsou z hlediska ekvivalence

po dlouhou dobu kritizovány jako značně problematické a subjektivní. Vše nasvědčuje tomu, že tyto dva koncepty jsou skutečně příhodnými adepty pro ověření toho, zda jsou kulturně srovnatelné anebo nikoliv.

Rámeček 4 Představení projektu Evropský sociální výzkum (ESS)

Už od roku 2001 ESS každé dva roky mapuje dlouhodobé postoje, hodnoty a změny v chování v kontextu evropského sociálního, politického, kulturního a morálního klimatu. Sebraná data mají za úkol odhalit neobvyklé kontrasty napříč 30 zeměmi, většiny zemí Evropské unie, řady dalších evropských zemí a některých mimoevropských zemí (Turecko, Izrael, Rusko). ESS se věnuje výzkumu politických témat a problémů, postojům k demokracii, politické participaci a volebnímu chování, sociální a politické orientaci, využívání masmédií a komunikaci, sociálně politické identitě, prosperitě, spokojenosti se životem, hodnotovým žebříčkům, náboženství a jiným tématům. V šetřeních jsou v jednotlivých zemích sbírány sociodemografické proměnné (věk, pohlaví, rasa/etnikum, občanství atp.) a sociostrukturní charakteristiky (sociální pozice, vzdělání, pracovní život, sociální exkluze atp.) dotázaných a jejich domácností. Kromě toho ESS poskytuje prostředky, díky kterým mohou země posuzovat a hodnotit samy sebe, také částečně podle toho, jak se jejich obyvatelé cítí nebo jak se vyslovují ke své aktuální existenční situaci. Klíčovým cílem ESS není výzkum konkrétních témat, ale postupné budování databáze časově a mezinárodně srovnatelných indikátorů, snadno a rychle dostupných všem výzkumníkům pro různé výzkumné účely [Krejčí 2008: 92].

I když tento evropský projekt sdílí jistou podobnost s jinými velkými dlouhodobými mezinárodními komparativními programy (viz. EVS, ISSP, EB atp.), v žádném případě je nenahrazuje [Plecitá-Vlachová, Kalvas 2005]. Tento exaktní komparativní výzkum je využíván k informování a doplňování akademických a politických debat. Poskytuje badatelům, zákonodárcům evropských vlád, agenturám a ostatním zainteresovaným skupinám možnost opětovně měřit a interpretovat změny v lidských hodnotách a postojích. Je to první sociálně vědní projekt, který získal prestižní cenu „Descartes Prize“ (2005) za excelentní vědecký výzkum prováděný ve spolupráci vědců. Oceněn byl jejich postup na poli komparativních šetření v oblasti systematického využívání přísných metodologických a akademických standardů. Akceptováním striktních přístupů ohledně pravděpodobnostního vzorku, testování otázek, zaznamenávání kontextových událostí, precizního překladu a vylepšení míry návratnosti se ESS stal směrodatným a oficiálním zdrojem informací, týkajících se změn sociálních hodnot v transformující se Evropě.

3.1 Koncept sociálního kapitálu a jeho dimenze

Fenomén sociálního kapitálu se v sociálních vědách studuje a rozpracovává od konce osmdesátých let, přičemž skutečný rozmach zaznamenává z hlediska publikovaných teoretických pojednání i empirických výzkumů přibližně v posledních deseti letech [Šafr, Sedláčková 2006: 12].³² A byla to právě četnost empirických studií a různorodost operacionalizací tohoto konceptu, které měly pravděpodobně za následek spletitost odlišných definicí a teoretických přístupů ke studiu sociálního kapitálu. Na začátku zkoumání stála dvě nezávislá pojetí sociálního kapitálu. Jedno poměrně negativní v práci ekonoma Glenna Loury [1977], který vnímá sociální kapitál jako sadu zdrojů s diskriminačním potenciálem determinující individuální šance na úspěch v životě. A druhé pro sociologii mnohem podstatnější individualistické pojetí sociálního kapitálu Pierra Bourdieu [1986] jako množiny

³² V české sociologii je to především dílo Iva Možného [1991], empirické studie Petra Matějů [1993, 2002] a Matějů a Vitáskové [2006], a přehledové studie Šafra a Sedláčkové [2006] a Stachové, Bernarda a Čermáka [2009].

potenciálních zdrojů vyplývajících ze sociálních vztahů a známostí, které může jedinec využít v sociálním světě jako svůj vlastní nemateriální kapitál. Kruh se uzavírá empirickým rozpracováním sociálního kapitálu v oblasti výchovy a vzdělávání Jamesem Colemanem [1988], jehož snahy započaly dlouhý proces vývoje modelu měření, který by zachytil všechny dimenze sociálního kapitálu validním, opakovatelným a komparativním způsobem.

Opravdový průlom ve studiu sociálního kapitálu se udál v 90. letech, kdy politolog Robert Putnam vydal své populární sociálně-politické pamflety o italských politických institucích [Putnam et al. 1993] a americké společnosti pozbývající sociálního kapitálu [1995, 2000], které měly za následek nebývalý vzrůst zájmu o tuto tematiku obzvláště mezi politiky a ekonomy. Putnam propojuje neobvyklým způsobem sociální kapitál (jako je důvěra, normy a sítě) se sociálně soudržnou, občansky demokratickou společností, ke které svým dílem sociální kapitál přispívá a udržuje ji, pokud jsou lidé občansky angažovaní. Vytváří tak teorii kolektivního kapitálu, který nemá přínos jen pro jedince, ale i pro celou občanskou komunitu. Takto vznikla dvě odlišná paradigmatu teoretického i empirického zkoumání a konceptualizace již z definice stejného jevu – individuálního a kolektivního sociálního kapitálu [Šafr, Sedláčková 2006: 12].

Individuální sociální kapitál je jakýmsi soukromým statkem, kterým disponuje jedinec na základě sociálních známostí a v průběhu života ho pečlivě shromažďuje. V případě potřeby je možné kdykoli tento prostředek zmobilizovat a využívat ve svůj prospěch většinou za podmínky reciprocity, tedy vzájemnosti [Matějů, Vitásková 2006: 495]. Toto jednání se většinou vnímá jako instrumentální, jedná se tedy o naplňování vlastních osobních cílů jedince.³³ V případě kolektivního sociálního kapitálu se jedná spíše o kapitál celé společnosti, který se využívá pro společenské blaho. Opět se jedná o sítě vztahů, které ale umožňují lidem vzájemně spolupracovat a občansky jednat, což zvyšuje produktivitu jejich práce a stmeluje občanskou společnost dohromady. Základem takového kapitálu je kromě sociálních vazeb také vysoká míra mezilidské důvěry a princip reciprocity, které také zároveň zvyšují důvěru v politické instituce a demokratické uspořádání [Matějů, Vitásková 2006: 495].³⁴

Matějů a Vitásková [2006: 495-496] i Šafr a Sedláčková [2006: 23] upozorňují na to, že ekonomové a politologové mají ve zvyku akcentovat spíše Putnamův komunitaristický přístup kvůli jeho vazbě na růst produktivity práce společnosti, resp. ekonomický růst společnosti (viz operacionalizace tohoto konceptu ve výzkumech Světové banky, Mezinárodního měnového fondu, OECD, atp.) a sociologové zase individualistický přístup k výběru proměnných osobní prestiže, neformální sítě konexí a sociálního statusu. Ani jeden z těchto přístupů se nezdá být v takovémto separátním stavu dostačující, obzvláště ne v případě operacionalizace sociálního kapitálu při přípravě komparativních empirických šetření, protože nedokážou postihnout všechny jeho dimenze.

Na jednu stranu se tedy zdá, že je nezbytné snažit se ve výzkumu postihnout obě podoby sociálního kapitálu, na druhou stranu už Matějů a Vitásková [2006] upozorňovali na to, že efektivnost a smysluplnost jejich výzkumného využití se může značně lišit podle typu zkoumané společnosti, např. západní společnosti versus post-komunistické společnosti. Stachová, Bernard a Čermák [2009: 13-14] definují jiné podrobnější subdimenze sociálního kapitálu - formální a neformální sociální sítě, důvěru (v ostatní lidi i v instituce) a sdílené normy a hodnoty, což více méně koresponduje i s operacionálním pojetím konceptu sociálního kapitálu u Halmana a Luijkxe [2006: 76]. V některých jiných konceptualizacích sociálního kapitálu jsou uváděny další jeho dimenze jako je například občanská angažovanost

³³ Bourdieu tento koncept propojuje s teorií reprodukce sociálních tříd, kde je sociální kapitál využíván především privilegovanými jedinci či skupinami k budování statusové pozice za využití vlastní sítě konexí. Takto podle něj dochází k reprodukci sociálních nerovností [Šafr, Sedláčková 2006: 20].

³⁴ Blíže o typologii a dalších formách sociálního kapitálu Šafr a Sedláčková [2006].

či smysl pro komunitu (resp. ochotu nezištně pomáhat ostatním) nebo v poměrně moderním pojetí sociálního kapitálu dimenze četnosti virtuálních sítí respondentů (sledování TV, rádia, internetu a čtení novin) [Parts 2008: 32-35].

Konceptualizace sociálního kapitálu je tedy očividně značně komplikovaná a zdá se, že koordinátoři různých mezinárodních šetření do dotazníků zahrnují zpravidla jen některé dimenze. Na základě dostupných proměnných z mezinárodního šetření ESS, které bylo zvoleno jako hlavní zdroj dat pro tuto studii, je možné v modelu měření sociálního kapitálu odlišit dimenzi individuálního sociálního kapitálu měřenou indikátory generalizované (obecné lidské) důvěry, dimenzi institucionální důvěry v politické instituce a dimenzi kolektivního sociálního kapitálu měřenou neformálními sociálními sítěmi. Takže i v této studii budou jako při klasických analýzách účinků sociálního kapitálu rozlišovány dvě jeho úrovně a to kolektivní a individuální.³⁵ Subdimenze hodnot musela být zanedbána kvůli příliš rozsáhlému měření v dotazníku ESS, stejně tak subdimenze norem, neboť ta není ve výzkumu ESS měřena vůbec. Strukturní dimenze R. Putnama, která má zachycovat formální sítě a je měřena indikátory členství v dobrovolných organizacích a mírou participace, musela být také vyřazena, neboť charakter těchto položek není pro testování ekvivalence žádoucí. Tyto položky jsou naneštěstí behaviorální, ukazují na konkrétní chování či jednání člověka – respondent uvádí pouze, zda je či není členem daného sdružení (tj. dichotomická proměnná s odpověďmi ano/ne) – proto není vhodné testovat jejich srovnatelnost napříč kulturami, to je potřebné pouze u postojových dotazů.

Koncept sociálního kapitálu je tedy v analyzovaném modelu měření rozdělen na dvě, resp. tři latentní proměnné a to individuální sociální kapitál (ITRUST), kolektivní sociální kapitál (SOCNET) a důvěru v instituce (INSTRUST). V roce 2002 bylo ve výzkumu ESS dotazováno 9 položek měřících sociální kapitál, z nichž 3 by měly reprezentovat dimenzi důvěry v instituce (INSTRUST), 3 individuální sociální kapitál (ITRUST) a 3 kolektivní sociální kapitál (SOCNET). Ve všech případech měl respondent vyjádřit míru souhlasu s danými výroky nejčastěji pomocí pětistupňových a jedenáctistupňových škál.

Latentní proměnná individuálního sociálního kapitálu (ITRUST) odpovídá jinak často citovanému konceptu generalizované důvěry, která je měřena sadou tří standardizovaných otázek zaměřující se na důvěru respondenta v lidi obecně (viz níže položka TRUSTED, FAIR, HELPFUL v Tabulce 2).³⁶ Stachová, Bernard a Čermák [2009: 18] upozorňují, že u této škály důvěry není jisté, zda respondenti v různých zemích do této skupiny řadí stejný okruh lidí, ani to, jestli samotný pojem „důvěra“ chápou stejným způsobem. Navíc různé další zdroje uvádějí, že právě u takovýchto postojových otázek měřících například míru důvěry hrozí riziko narušení jejich srovnatelnosti napříč zeměmi kvůli obtížnému překladu a specificky kulturní interpretaci položek respondenty [Stachová, Bernard a Čermák 2009: 16-17]. Což nás přivádí ke zřejmým kulturním problémům s měřením sociálního kapitálu. Ekvivalenci této škály v první a druhé vlně ESS testoval již Reeskens a Hooghe [2008] s tím výsledkem, že tato škála není uspokojivě srovnatelná v několika evropských zemích a nejhorší výsledky vykazuje položka HELPFUL (tzn. pomoc druhým), proto by se s ní mělo zacházet při komparaci velmi opatrně.³⁷

Měření konceptu kolektivního sociálního kapitálu (SOCNET) je ve výzkumu ESS poněkud problematické – odpovídá již dříve zmíněným dimenzím rozsahu formálních a neformálních

³⁵ Určitou míru nezávislosti těchto dvou podob sociálního kapitálu na českých datech ISSP (2001) už prokázaly Matějů a Vitásková [2006].

³⁶ Otázky pocházejí z původně pěti-položkové Rosenbergovy škály pro měření důvěry [Stachová, Bernard a Čermák 2009: 26].

³⁷ Van der Veld a Saris [2009] došli na základě vylepšeného nástroje Jrule k lepším výsledkům o částečné srovnatelnosti škály generalizované důvěry ve všech zemích ve výzkumu ESS.

sítí respondenta. Zapojení respondentů do formálních sdružení je v první vlně ESS zjišťováno stupňovitě podle přiznaného členství, participace, donátorství a vykonávané dobrovolné práce v různých občanských sdruženích, popř. výskytu přátel na těchto místech. Naneštěstí jsou tyto položky behaviorální a ne postojové a z výše uvedených důvodů není ani vhodné, ani potřebné testovat jejich srovnatelnost. Z tohoto důvodu byla dimenze formálních sítí nahrazena shrnující otázkou VOLUNORG, která zjišťuje, do jaké míry jsou dobrovolné organizace pro respondenta důležité. Zbýlé tři proměnné (FRIEND, MEET, SOCACTIV) se zaměřují na zapojení respondentů do neformálních sociálních sítí a dotazují se na důležitost přátel v respondentově životě, společenské aktivity a množství času stráveného s kolegy a příbuznými.³⁸

Tabulka 2 Znění vybraných položek v modelu měření

Dimenze	Kód	Položka
Individuální odcizení /INDIS	INTEREST	„Jak moc se zajímáte o politiku?“
	DIFFICULT	„Jak často se politika zdá být tak složitá, že vlastně nemůžete pochopit, co se děje?“
	POLACTIV	„Myslíte, že byste mohl(a) hrát aktivní roli ve skupině, která se zabývá politickými otázkami?“
	OPINION	„Jak je pro Vás těžké či lehké udělat si názor na politické otázky?“
Institucionální odcizení /INSTDIS	VOTE	„Řekl(a) byste, že politici se zajímají jen o získávání hlasů voličů spíše než o názory lidí?“
	CARE	„Myslíte si, že se politici obecně starají o to, co si lidé jako Vy myslí?“
	POLIT	„Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: politici“
Důvěra v instituce /INSTRUST	POLIT	„Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: politici“
	PARLAM	„Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: Český parlament“
	LEGAL	„Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: právní systém“
Individuální sociální kapitál /ITRUST	TRUSTED	„Obecně vzato, řekl(a) byste, že se většině lidí dá důvěřovat, nebo že člověk nemůže být při jednání s lidmi nikdy dost opatrný?“
	FAIR	„Myslíte si, že by se Vás většina lidí snažila podvést, pokud by měli tu možnost, nebo by se snažili být poctiví?“
	HELPFUL	„Řekl(a) byste, že se lidé většinou snaží pomáhat druhým, nebo že se většinou starají o sebe?“
Kolektivní sociální kapitál /SOCNET	FRIEND	„Jak důležité jsou ve Vašem životě následující věci?: přátelé“
	VOLUNORG	„Jak důležité jsou ve Vašem životě následující věci?: dobrovolné organizace“
	MEET	„Jak často se ve svém volném čase společensky scházíte s přáteli, příbuznými či kolegy?“
	SOCACTIV	„Když se srovnáte s jinými lidmi Vašeho věku, jak často se ve svém volném čase účastníte společenských aktivit?“

Poznámka: podrobněji viz Tabulka 3, Příloha 1

Součástí sociálního kapitálu je také manifestovaná důvěra ve veřejné instituce. V první vlně ESS se mohou respondenti vyjadřovat k šesti různým institucím, z nichž klíčové jsou pro náš výzkum, jak dokazuje Eva Parts faktorovou analýzou [2008: 36], první čtyři – parlament,

³⁸ Latentní proměnná individuálního sociálního kapitálu by mohla být měřena podle Parts [2008] i položkou HLPPPL (aktivní pomoc lidem mimo rodinu, práci a dobrovolné organizace). Dále byla zvažována do škály měření sociálních sítí položka IMPFML (důležitost rodiny pro respondenta) a INMDISC (diskutování osobních záležitostí). Všechny tyto položky se ale ukázaly být v modelu měření nevýznamné z hlediska velikosti faktorové zátěže a po jejich vyloučení se výrazně zlepšili hodnoty statistik vhodnosti modelu jako je CFI, AGFI a RMSEA.

politici, právní systém a policie. Výzkum ESS do dotazníku nezahrnuje žádné otázky zjišťující závaznost norem občanského soužití, proto nemohly být tyto položky do modelu pro testování ekvivalence zahrnuty; baterii těchto položek tradičně obsahují pouze výzkumy EVS a WVS, jak ilustrují Stachová, Bernard a Čermák [2009: 28].

3.2 Koncept politického odcizení a jeho dimenze

Po tom, co Traxlerová a Rabušic [2008: 8] upozornili na absenci standardně využívaného nástroje, který by zachycoval subjektivní pocit možnosti vlivu jedince na společensko-politické události, představil Linek [2011] české sociologii známý koncept politického odcizení, resp. politické efektivity a možnosti jeho měření. Pocit politické efektivity (jeho opakem je pocit politického odcizení či bezmoci)³⁹ je termín běžně užívaný v politických a sociálních teoriích od 50. let a označuje subjektivní pocit jedince, že může svým (politickým) jednáním jako občan participovat na politických rozhodnutích, aktivitách a politických změnách a ovlivňovat tak celý politický proces a že má tudíž smysl hlasovat ve volbách a splnit tak svou občanskou povinnost [Campbell et al. 1954: 187]. Naopak pocit politického odcizení je pocitem odloučení jedince od politických institucí, politických hodnot a zástupců voličů, což vyúsťuje ve vnímání sebe sama jako politického outsidera [Citrin et al. 1975: 2-3].

Tento koncept byl vyvinut, aby vysvětlil kolísavou míru účasti voličů v amerických volbách [Balch 1974: 1] a mnoho politologů a sociologů se ve svých analýzách snaží opakovaně ověřit, zda je pocit politické efektivity opravdu fundamentálně důležitý pro chod demokratické společnosti a politickou participaci [viz Almond, Verba 1963; Madsen 1978; Abramson 1983; Craig, Niemi, Silver 1990; Karp, Banducci 2008].⁴⁰ Tato důležitost politické efektivity pro politický život většinou vyplývá až z plastického popisu důsledků politického odcizení, které má tendenci vyústit ve společnosti po delší době v celkovou apatii a politickou pasivitu a vést ke ztrátě loajality ke státnímu aparátu a občanské důvěry v politický (demokratický) systém, který tak postupně pozbývá legitimitu a ohrožují ho politické ideologie [Halperin, Sulitzeanu-Kenan 2010: 6].

Empirické studie už v 50. a 60. letech prokázaly, že koncept politické efektivity tvoří dvě částečně nezávislé komponenty a tím je vnitřní politická efektivity a vnější politická efektivity [Balch 1974].⁴¹ Dimenze vnitřní politické efektivity odkazuje k pocitu vlastní schopnosti porozumět politice, udělat si vlastní názor a efektivně se tak podílet na politickém životě. Vnější politická efektivity pak odráží lidské vnímání občanské odpovědnosti politiků a státního aparátu a jejich reflektování potřeb a požadavků občanů ve stávajícím režimu. Linek [2010, 2011] v současnosti předkládá teoretické i empirické důkazy k tomu, že obě tyto dimenze jsou součástí širšího konceptu politického odcizení, který patří do souboru negativních postojů k politickému režimu a jeho představitelům.⁴² Tento koncept opět rozděluje na dvě nezaměnitelné dimenze individuálního odcizení a institucionálního odcizení.

³⁹ V překladu z anglického „sense of political efficacy“ nebo „sense of political disaffection“ – o problému nepřesného překladu pojmu sense či feeling nebo pojmu odcizení blíže Traxlerová a Rabušic [2008]. O rozdílech mezi pojmy „disaffection“ a „alienation“ blíže Citrin et al. [1975].

⁴⁰ Almond a Verba [1963: 257] tento pocit ztotožňují s vírou ve vlastní občanské kompetence a schopnosti a se sebedůvěrou občana, která je klíčová k formování politických postojů a potažmo k aktivní účasti na politickém dění.

⁴¹ Nezaměnitelnost těchto dvou konceptů je prokazována rozdílným vztahem k dalším konceptům jako je politická aktivita či důvěra k politickým institucím nebo zájem o politiku [více Balch 1974; Hayes, Bean 1993]. Dokonce spolu tyto dvě dimenze nemusejí souviset vůbec [Linek 2011: 13]

⁴² Linek [2010, 2011] koncept politického odcizení na základě teorie i empirie striktně odděluje od dalších politických postojů jako je politická legitimita (resp. legitimita režimu) či politická nespokojenost [viz také Montero, Gunther, Torcal 1997].

Jeho přidanou hodnotou je propojení indikátorů pocitu vnitřní politické efektivity s faktorem politického zapojení (resp. zájmem o politiku), které měří koncept individuálního odcizení, a indikátorů vnější politické efektivity s důvěrou v politické instituce, které měří koncept institucionálního odcizení [2011: 54-60]. Navíc tuto konceptualizaci neuzavírá a ponechává ji otevřenou možnostem dalšího doplnění.

Koncept politické efektivity je stejně tak často citován v teoretických statích jako v empiricko-metodologických studiích, kde se testuje a ověřuje především validita a reliabilita položek vybraných k měření tohoto konceptu. Po léta nejpoužívanější škála měření sestávala z položek vyvinutých michiganskou školou volebního chování SRC (Survey Research Centre). Uplatňovala se v řadě výzkumů i mimo Spojené státy a dotazovala se respondentů na jejich percepci vlastních schopností ovlivňovat chod politického systému, kterou měli vyjádřit mírou souhlasu či nesouhlasu se čtyřmi až pěti výroky [Seligson 1980: 631]. Poté co Balch [1974] empiricky ověřil dvě zřetelně odlišné dimenze politické efektivity a tudíž narušil představu o jednodimenzionalitě škály měření SRC, snesly se na tradiční přístup k měření politické efektivity zdrcující kritiky týkající se dvojznačnosti položek, jejich sociálně desirabilní povahy, jen dočasné stability z hlediska reliability a jejich náchylnosti k systematickým chybám měření (jako je např. response style, tendence k souhlasu bez ohledu na obsah položky, atp.) [blíže Seligson 1980: 631-632]. To ale nezabránilo výzkumníkům zařazovat některé z nich i v současnosti do mezinárodních dotazníků (viz ISSP, ESS, atp.), což může být jedna z příčin, proč se škála politické efektivity (resp. politického odcizení) ukazuje jako velmi problematická při kulturní komparaci.⁴³

Rozšířený koncept politického odcizení od Linka [2011] se jeví jako velmi perspektivní obzvláště z hlediska větších možností při výběru měřicích položek do analytické studie. V roce 2002 bylo ve výzkumu ESS dotazováno 7 položek měřících politické odcizení, z nichž 4 by měly reprezentovat dimenzi individuálního odcizení (INDIS), 3 institucionální odcizení (INSTDIS). Latentní proměnná individuálního politického odcizení je v první vlně ESS měřena standardizovanými třemi otázkami (DIFFICULT, POLACTIV, OPINION) vystihující odcizení jedince způsobeného reflexí vlastní politické (ne)kompetence a schopnosti orientace v politice. Linek [2011: 16] zvažuje možnost měření i dalšími indikátory, např. zájmem o politiku, který je také v ESS pravidelně zjišťován (INTEREST), takže ho lze do škály měření zakomponovat (viz Tabulka 2).

Dimenze institucionálního politického odcizení je v první vlně ESS obtížněji identifikovatelná. ESS měří koncept politické efektivity tradičním způsobem podle vzoru otázek v amerických národních studiích volebního chování [Thomassen (n.d.)] a rozlišuje vnitřní a vnější politickou efektivitu. Indikátory měřící institucionální odcizení jsou položky vnější politické efektivity vyjadřující postoje respondenta vůči fungování politického systému, resp. je to odcizení orientováno na politické instituce/politiky (položky VOTE, CARE). Linek [2011: 17] proto k pocitu vnější politické efektivity navíc přidává možnost zahrnout i indikátor důvěry v politické instituce. V této souvislosti se zdá být ideální vzájemně propojit tyto dvě dimenze, proto se třetí položkou institucionálního odcizení stala důvěra v politiku (POLIT).⁴⁴

⁴³ Z původních čtyř výroků (1) „Hlas lidí jako já nemá v tom, co dělá vláda, žádnou váhu.“; 2) „Nemyslím si, že se vláda zajímá o to, co si myslí lidé jako já.“; 3) „Někdy se politika zdá být tak komplikovaná, že lidé jako já nemohou opravdu rozumět tomu, co se děje.“; 4) „Hlasování je jediná možnost, jak mohou lidé jako já ovlivnit to, co dělá vláda.“ se ve výzkumu ISSP stále používají první tři položky a ve výzkumu ESS druhá a třetí položka (v podobě otázky, ne výroku).

⁴⁴ Další vysvětlující proměnné jako občanská angažovanost, politická participace, levo-pravá politická orientace respondenta, pocity životního štěstí, občanská morálka atp. nebyly do analýzy zahrnuty, protože by již příliš komplikovaly model měření, který má být co nejdůležitější v rámci (výzkumných) možností a ve vztahu k teorii.

Dimenze důvěry v instituce (INSTRUST) se prostřednictvím této položky propojuje s dimenzí politického odcizení, což zcela reflektuje teoretický předpoklad, že sociální kapitál a pocit politického odcizení spolu budou vzájemně souviset a ovlivňovat se navzájem, v tomto případě skrze dimenzi institucionální důvěry. Kvůli teoreticky podloženému vztahu sociálního kapitálu a politického odcizení jsou v navrženém strukturním modelu zvažovány také vzájemné vztahy mezi jednotlivými latentními proměnnými (názorněji viz Diagram 8, Příloha 2). Cíl propojit tyto dva koncepty v modelů měření a otestovat jejich vzájemný vztah a jejich mezinárodní srovnatelnost současně rozhodl i o výběru první vlny šetření výzkumu ESS, neboť pouze tato vlna a tento mezinárodní výzkum měří vybrané položky společně (viz Tabulka 4).⁴⁵

Tabulka 4 Dotazování položek měřících sociální kapitál a politické odcizení v jednotlivých vlnách výzkumu ESS

koncept	položka	První vlna	Druhá vlna	Třetí vlna	Čtvrtá vlna	Pátá vlna	Šestá vlna
		ESS (2002)	ESS (2004)	ESS (2006)	ESS (2008)	ESS (2010)	ESS (2012)
Individuální politické odcizení	INTEREST	x	x	x	x	x	x
	DIFFICULT	x	x	x	x		
	POLACTIV	x					
	OPINION	x	x	x	x		
Institucionální politické odcizení	VOTE	x					
	CARE	x					
	POLIT	x	x	x	x	x	x
Důvěra v instituce	POLIT	x	x	x	x	x	x
	PARLAM	x	x	x	x	x	x
	LEGAL	x	x	x	x	x	x
Individuální sociální kapitál	TRUSTED	x	x	x	x	x	x
	FAIR	x	x	x	x	x	x
	HELPFUL	x	x	x	x	x	x
Kolektivní sociální kapitál	FRIEND	x					
	VOLUNORG	x					
	MEET	x	x	x	x	x	x
	SOCACTIV	x	x	x	x	x	x

Poznámka: Křížek znamená měření dané položky v konkrétní vlně výzkumu.

⁴⁵ Tohoto šetření se v roce 2002 účastnilo celkem 22 zemí, dohromady 42 359 respondentů [více www.europeansocialsurvey.org]. Konkrétně: Rakousko, Belgie, Česká republika, Německo, Dánsko, Španělsko, Finsko, Francie, Velká Británie, Řecko, Maďarsko, Švýcarsko, Irsko, Izrael, Itálie, Lucembursko, Nizozemí, Norsko, Polsko, Portugalsko, Švédsko a Slovinsko. Data dostupná z: www.europeansocialsurvey.org/data/.

4 ZÁKLADNÍ ANALÝZY SROVNATELNOSTI VYBRANÝCH POSTOJOVÝCH ŠKÁL

V předchozí části studie byla nastíněna dimenzionalita konceptů sociálního kapitálu a politického odcizení z teoretického hlediska. Dále byly ukázány možnosti jejich měření v rámci první vlny mezinárodního šetření ESS a konkrétní forma operacionalizace jednotlivých postojových škál. Následující část studie bude mít za cíl 1) ověřit vnitřní konzistenci zkoumaných postojových škál a 2) provést základní analýzy srovnatelnosti zvolených měřicích položek a z nich vyvodit první soudy o míře jejich ekvivalence na úrovni měření. K tomuto účelu budou využity tři různé vícerozměrné statistické techniky: korelační analýza, analýza reliability a explorativní faktorová analýza.

4.1 Korelační analýza

Vnitřní konzistenci postojových škál je vhodné prozkoumat nejdříve pomocí korelační analýzy. U položek měřicích individuální sociální kapitál se hodnoty korelačního koeficientu pohybují v rozmezí od 0,44 do 0,51 s průměrnou hodnotou korelačního koeficientu 0,47 (viz Tabulka 5). V případě kolektivního sociálního kapitálu spolu korelují položky o trochu méně a nacházejí se v intervalu mezi 0,1 až 0,3 s tím, že průměrný korelační koeficient dosahuje velikosti 0,2. V této dimenzi dosahuje nejnižších hodnot položka VOLUNORG a ve většině případů je mezipoložkový korelační koeficient menší než 0,15. Pokud by bylo cílem vytvořit co nejlepší škálu měření kolektivního sociálního kapitálu, bylo by oprávněné zvažovat vyřazení této položky z měřicí škály. Primárním cílem této studie je ale otestovat srovnatelnost této měřicí položky napříč 22 zeměmi v první vlně šetření ESS a z tohoto důvodu jsou zjištěné hodnoty postačující. Mezipoložkový korelační koeficient položek měřicích důvěru respondentů ve veřejné instituce osciluje mezi 0,5 až 0,66 s průměrnou hodnotou korelačního koeficientu 0,56, což je velice dobrý výsledek.

Položky měřicí individuální politické odcizení dosahují hodnot korelačního koeficientu od 0,29 do 0,44 s tím, že průměrný korelační koeficient je 0,39 (viz Tabulka 6). U položek měřicích institucionální politické odcizení dosahuje mezipoložkový korelační koeficient hodnot od 0,49 do 0,58 s průměrným korelačním koeficientem o velikosti 0,52. Jednou z položek měřicí institucionální politické odcizení je položka POLIT, která je zároveň součástí měřicí škály důvěry v instituce. Z pohledu na korelační koeficienty je zřejmé, že výběr této konkrétní položky (ve srovnání s ostatními položkami měřicích dimenzi důvěry v instituce) byl oprávněný, neboť se zbývajícími dvěma položkami měřicích institucionální odcizení koreluje nejvíce ($r = 0,5$). Všechny korelační koeficienty položek měřicích danou latentní proměnnou mají vždy stejné znaménko a nedosahují příliš velkých hodnot, což koresponduje s obecným doporučením odborné literatury. Tyto výsledky ukazují, že teoretický předpoklad ohledně dimenzionality konceptu sociálního kapitálu a politického odcizení byl opodstatněný a že byly dané dimenze zvoleny adekvátně.

K základní analýze mezinárodní srovnatelnosti jednotlivých položek lze využít údajů korigované korelace položky s celkovým součtem položek vycházejících z korelační analýzy provedené v každé participující zemi první vlny šetření ESS zvlášť. Jednotlivé položky jsou korelovány s příslušnými postojovými škálami; výsledky jsou zaznamenány jako velikost ukazatele korelačního koeficientu Pearsonova r . Při výpočtu korelace skóru položky se skórem škály je navíc daná položka vyloučena z výpočtu skóru škály - proto se označuje jako korigovaná korelace - aby se předešlo problému nadhodnocení odhadu [Urbánek, Šimeček 2001]. Na základě těchto údajů je možné zjistit, zda lze položky tvořící jednotlivé dimenze sloučit do jedné škály, zda je tato škála konzistentní a do jaké míry korelační koeficienty dané položky mezi zeměmi varíují. Vysoká variabilita hodnot těchto koeficientů by měla ukazovat na problematické položky z hlediska mezinárodního srovnávání.

Tabulka 5 Matice interkorelací položek měřících sociální kapitál v roce 2002 (Pearsonovo r)

	Důvěra v instituce			Individuální sociální kapitál			Kolektivní sociální kapitál			
	POLIT	PARLAM	LEGAL	TRUSTED	FAIR	HELPFUL	FRIEND	VOLUNORG	MEET	SOACTIV
POLIT	1	0,66	0,50	-0,30	-0,27	-0,26	-0,06	-0,12	-0,06	-0,10
PARLAM		1	0,51	-0,28	-0,24	-0,23	-0,06	-0,11	-0,07	-0,09
LEGAL			1	-0,25	-0,22	-0,22	-0,07	-0,10	-0,06	-0,08
TRUSTED				1	0,51	0,44	0,14	0,08	0,12	0,12
FAIR					1	0,46	0,17	0,04	0,11	0,11
HELPFUL						1	0,14	0,05	0,10	0,08
FRIEND							1	0,16	0,26	0,21
VOLUNORG								1	0,10	0,13
MEET									1	0,30
SOACTIV										1

Zdroj: ESS 2002 (vážená data). Celkový soubor 22 zemí – analýza všech zemí dohromady. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Tabulka 6 Matice interkorelací položek měřících politické odcizení v roce 2002 (Pearsonovo r)

	Individuální politické odcizení				Institucionální politické odcizení		Důvěra v instituce		
	INTEREST	DIFFICULT	POLACTIV	OPINION	VOTE	CARE	POLIT	PARLAM	LEGAL
INTEREST	1	0,38	0,44	0,43	0,23	0,28	0,23	0,18	0,12
DIFFICULT		1	0,29	0,42	0,15	0,19	0,14	0,14	0,10
POLACTIV			1	0,37	0,19	0,22	0,15	0,13	0,11
OPINION				1	0,12	0,14	0,08	0,08	0,06
VOTE					1	0,58	0,50	0,39	0,28
CARE						1	0,49	0,41	0,27
POLIT							1	0,66	0,50
PARLAM								1	0,51
LEGAL									1

Zdroj: ESS 2002 (vážená data). Celkový soubor 22 zemí – analýza všech zemí dohromady. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

V případě individuálního sociálního kapitálu dosahuje korelační koeficient mezi jednotlivými položkami a jim příslušnou škálou dostatečně vysokých hodnot 0,4 až 0,6 ve všech zemích, což značně převyšuje minimální požadovanou hladinu korelace o velikosti 0,3 (viz Tabulka 7). Pouze v případě položky HELPFUL v Lucembursku a Norsku se hodnoty této korelace začínají blížit hodnotě 0,3, ale stále to nenaznačuje žádný problém v konzistentnosti dané škály v těchto zemích. Největší variabilitu hodnot korelací mezi zeměmi vykazují položky FAIR a HELPFUL, u nichž by se na základě těchto výsledků daly očekávat potíže ve srovnatelnosti v šetřených zemích.

U položek měřících kolektivní sociální kapitál značně oscilují hodnoty korigované korelace jednotlivých položek a navíc kolísají kolem limitní hodnoty 0,3, často jí ani nedosahují. Ani v jedné zemi nedosahují všechny položky škály hodnoty nad 0,3 zároveň, což značí, že tato škála měření není nikde dostatečně homogenní a konzistentní. Na druhou stranu vykazují všechny položky velmi podobnou strukturu hodnot více méně ve všech zemích, což naznačuje, že mezinárodně srovnatelná tato škála měření pravděpodobně alespoň někde je. Již teď lze předpokládat výskyt potíží se srovnatelností Lucemburska s ostatními zeměmi, protože předložené hodnoty korigovaných korelací položek jsou v porovnání s ostatními zeměmi nejnižší. Největší problém je možné vidět v položce VOLUNORG, průměrně se její korelace se škálou pohybuje okolo hodnoty 0,19, proto je možné také očekávat, že právě tato položka není ve všech zemích ekvivalentní. Největší variabilitu mezi zeměmi vykazují položky MEET a SOCACTIV – dosahuje vyšší hodnoty než 0,3 bodu rozdílu - proto je nutné celé této škále věnovat v průběhu další analýzy ekvivalence větší pozornost.

Tabulka 7 Analýza korigované korelace položky s celkovým skóre škály – sociální kapitál

země	Individuální sociální kapitál			Kolektivní sociální kapitál			
	TRUSTED	FAIR	HELPFUL	FRIEND	VOLUNORG	MEET	SOCACTIV
Rakousko	0,61	0,69	0,60	0,37	0,20	0,41	0,37
Belgie	0,52	0,52	0,44	0,28	0,33	0,25	0,39
Česká republ.	0,64	0,61	0,52	0,28	0,22	0,37	0,37
Německo	0,47	0,52	0,46	0,30	0,23	0,36	0,39
Dánsko	0,57	0,58	0,44	0,24	0,20	0,23	0,36
Španělsko	0,62	0,60	0,51	0,35	0,30	0,33	0,26
Finsko	0,55	0,57	0,49	0,31	0,25	0,28	0,38
Francie	0,46	0,44	0,41	0,33	0,19	0,28	0,29
Spoj. králov.	0,55	0,61	0,52	0,33	0,18	0,32	0,35
Řecko	0,60	0,62	0,57	0,24	0,23	0,38	0,38
Maďarsko	0,60	0,59	0,58	0,43	0,22	0,49	0,49
Švýcarsko	0,52	0,55	0,43	0,32	0,19	0,30	0,34
Irsko	0,55	0,61	0,51	0,22	0,11	0,26	0,36
Izrael	0,57	0,62	0,44	0,30	0,12	0,25	0,30
Itálie	0,55	0,61	0,54	0,39	0,22	0,26	0,24
Lucembursko	0,48	0,49	0,36	0,25	0,08	0,18	0,18
Nizozemí	0,57	0,59	0,47	0,27	0,24	0,30	0,35
Norsko	0,49	0,52	0,37	0,27	0,18	0,24	0,37
Polsko	0,48	0,49	0,42	0,28	0,21	0,37	0,36
Portugalsko	0,61	0,57	0,47	0,31	0,24	0,32	0,17
Švédsko	0,52	0,55	0,44	0,27	0,23	0,25	0,33
Slovinsko	0,55	0,59	0,52	0,27	0,28	0,40	0,41
Všechny země	0,56	0,58	0,52	0,31	0,19	0,34	0,32
Variabilita⁴⁶	0,18	0,25	0,24	0,21	0,25	0,31	0,32

Zdroj: ESS 2002 (vážená data), 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

⁴⁶ Variabilita je vypočítána jako rozdíl nejvyšší a nejnižší hodnoty korelace položky se škálou v dané zemi vždy pro konkrétní proměnnou.

Z hlediska škály důvěry v instituce korelují jednotlivé položky v dostatečné míře s celkovým součtem položek měřících daný faktor, hodnoty se pohybují od 0,5 do 0,7 (viz Tabulka 8). Škála se tak jeví jako konzistentní a hodnoty korelací jsou si v jednotlivých zemích velice podobné. Variabilita položek POLIT a PARLAM patří mezi jedny z nejnižších, vyšší variabilitu vykazuje položka LEGAL. Bylo by tedy nakonec možné z této korelační analýzy odvodit, že se srovnatelností těchto položek napříč zeměmi by problém být neměl.

Tabulka 8 Analýza korigované korelace položky s celkovým skóre škály – institucionální důvěra

země	Důvěra ve veřejné instituce		
	POLIT	PARLAM	LEGAL
Rakousko	0,63	0,71	0,56
Belgie	0,71	0,71	0,62
Česká republika	0,67	0,73	0,61
Německo	0,65	0,67	0,51
Dánsko	0,68	0,71	0,53
Španělsko	0,65	0,69	0,65
Finsko	0,70	0,75	0,58
Francie	0,70	0,66	0,59
Spojené království	0,73	0,70	0,55
Řecko	0,65	0,72	0,60
Maďarsko	0,68	0,71	0,69
Švýcarsko	0,65	0,70	0,60
Irsko	0,66	0,69	0,54
Izrael	0,61	0,68	0,48
Itálie	0,58	0,63	0,43
Lucembursko	0,65	0,75	0,63
Nizozemí	0,73	0,67	0,61
Norsko	0,65	0,68	0,52
Polsko	0,64	0,64	0,57
Portugalsko	0,61	0,62	0,57
Švédsko	0,71	0,71	0,58
Slovinsko	0,64	0,71	0,61
Všechny země	0,67	0,68	0,56
Variabilita	0,15	0,13	0,26

Zdroj: ESS 2002 (vážená data), 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Co se týče homogenity škál měření politického odcizení, obě se jeví ve všech zemích jako značně konzistentní. Položky individuálního odcizení korelují s položkami celé škály v rozmezí od 0,4 do 0,6 a taktéž položky institucionálního odcizení (viz Tabulka 9). Jako možný zdroj problémů se srovnatelností se jeví Izrael, který v případě individuálního odcizení vykazuje nejmenší velikosti korelací položek pohybujících se okolo limitní hodnoty 0,3. Největší variabilitu hodnot korigované korelace mezi zeměmi vykazují položky POLACTIV a OPINION v případě individuálního politického odcizení a položka POLIT v případě institucionálního odcizení. Opět by ale bylo možno na základě předložených výsledků usoudit na bezproblémovou mezinárodní komparaci vybraných položek.

Tabulka 9 Analýza korigované korelace položky s celkovým skóre škály – politické odcizení

země	Individuální politické odcizení				Institucionální odcizení		
	INTEREST	DIFFICULT	POLACTIV	OPINION	VOTE	CARE	POLIT
Rakousko	0,55	0,49	0,43	0,52	0,58	0,59	0,46
Belgie	0,51	0,43	0,49	0,46	0,53	0,59	0,55
Česká republika	0,52	0,50	0,46	0,57	0,60	0,62	0,54
Německo	0,55	0,46	0,39	0,55	0,64	0,64	0,60
Dánsko	0,51	0,54	0,47	0,54	0,67	0,64	0,55
Španělsko	0,56	0,52	0,48	0,57	0,63	0,60	0,46
Finsko	0,50	0,43	0,39	0,50	0,61	0,62	0,59
Francie	0,54	0,49	0,45	0,48	0,58	0,54	0,54
Spoj. království	0,50	0,47	0,46	0,41	0,60	0,64	0,60
Řecko	0,51	0,45	0,63	0,66	0,56	0,57	0,45
Maďarsko	0,54	0,45	0,33	0,56	0,62	0,55	0,51
Švýcarsko	0,58	0,51	0,48	0,53	0,54	0,59	0,53
Irsko	0,53	0,44	0,47	0,45	0,62	0,63	0,53
Izrael	0,38	0,35	0,36	0,43	0,55	0,49	0,44
Itálie	0,57	0,42	0,52	0,59	0,62	0,60	0,47
Lucembursko	0,53	0,45	0,43	0,47	0,55	0,54	0,46
Nizozemí	0,53	0,51	0,48	0,54	0,60	0,64	0,56
Norsko	0,50	0,44	0,50	0,35	0,60	0,60	0,56
Polsko	0,51	0,50	0,45	0,56	0,57	0,55	0,47
Portugalsko	0,58	0,58	0,58	0,61	0,57	0,59	0,41
Švédsko	0,52	0,45	0,44	0,46	0,61	0,63	0,61
Slovinsko	0,43	0,49	0,46	0,51	0,49	0,42	0,38
Všechny země	0,56	0,48	0,48	0,55	0,62	0,62	0,55
Variabilita	0,20	0,23	0,30	0,31	0,18	0,22	0,23

Zdroj: ESS 2002 (vážená data), 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

4.2 Analýza reliability

Vnitřní konzistenci měřicích škál a mezinárodní srovnatelnost položek lze ohodnotit také analýzou reliability. Pod pojmem reliability bychom měli rozumět podle Kreidla [2004] „přesnost, konzistentnost měření, tj. schopnost dosáhnout stejného výsledku měření v případě, že se stav pozorovaného předmětu nezměnil“. Pro účely této studie je klíčová tzv. mezipoložková reliability, která měří konzistentnost v odpovědích na baterii otázek, resp. vnitřní konzistentnost škály, a nejčastěji ji lze měřit Cronbachovým koeficientem alfa [o dalších typech reliability a metodách jejich odhadu viz Řehák 1998a; Kreidl 2004; Schubert 2010]. Lze takto ověřit i to, zda položky opravdu měří různé aspekty jednoho a téhož konstruktů a zda mohou vytvořit souhrnnou kvalitní škálu měření [Řehák 1998a; Soukup 2006b].

Schubert podotýká, že je-li baterie otázek složena z více položek, které mají měřit jednu latentní proměnnou, tak by jednotlivci měli na tyto položky odpovídat podobně a měla by mezi nimi existovat korelace [2010: 114]. Cronbachovo alfa má několik interpretací, ale nejčastější představou je, že se jedná právě o průměrnou korelaci, resp. kovarianci položek v rámci jedné škály v závislosti na tom, zda jsou položky standardizovány či nikoliv. Koeficient alfa nabývá hodnot 0 až 1. Krajiní hodnota 0 znamená, že jednotlivé položky nejsou korelované, naopak krajiní hodnota 1 znamená, že položky jsou absolutně korelované. Koeficient alfa však jedničky nikdy nedosáhne, jen se k ní limitně blíží [Soukup 2006b]. Záporné hodnoty tohoto koeficientu jsou posuzovány jako problém v logice kódování otázek (tj. je nutná transformace) anebo v tom, že položky měří více dimenzí najednou.

Pro Cronbachovo alfa platí, že jeho hodnota je silně závislá na počtu položek [Soukup 2006b; Schubert 2010]. V literatuře je doporučována velikost Cronbachova alfa blíží se hodnotě 0,7

(určuje spodní hranici spolehlivosti testu), avšak této hodnotě je nutné přiřadit význam právě v souvislosti s počtem položek škály měření.⁴⁷ Často je také důležité vědět, jaký vliv má každá položka na celkovou reliabilitu škály. K tomu se využívá údaj velikosti Cronbachova alfa za podmínky, že danou položku ze škály vyloučíme; dochází k opakovanému výpočtu ukazatele bez jednotlivých položek. Pokud po vyjmutí některé položky hodnota Cronbachova alfa narostla anebo se jen minimálně změnila, je možné zvážit její vyjmutí ze škály.

Pro potřeby této studie jsou v následujících tabulkách uvedeny hodnoty Cronbachova alfa po vyjmutí položky z dané škály pro danou položku a Cronbachova alfa pro celou škálu vždy podle jednotlivých dimenzí sociálního kapitálu a politického odcizení a podle zemí. Položky, resp. škály by měly mít velmi podobné hodnoty Cronbachova koeficientu alfa napříč zeměmi, pokud je splněn základní předpoklad o jejich mezinárodní srovnatelnosti.

Při pohledu na hodnoty koeficientu alfa po vyloučení položek ze škály měřící individuální sociální kapitál lze vidět, že v řadě zemí by po vyloučení položky HELPFUL nedošlo k příliš velké změně reliability škály a v případě čtyř zemí (Dánsko, Izrael, Lucembursko a Norsko) by došlo dokonce ke zvýšení reliability škály (viz Tabulka 10). Tato položka vykazovala problémy již v korelační analýze, analýza reliability jen tyto výsledky potvrzuje. Největší variabilitu Cronbachova koeficientu alfa ale vykazuje položka TRUSTED. Celkově je Cronbachovo alfa této škály ale nadmíru dobré pokud zvážíme, že dosahuje velikosti v různých zemích od 0,63 do 0,79 u škály složené ze tří proměnných (viz Tab. 11).

Tabulka 10 Analýza mezipoložkové reliability škály měřící sociální kapitál – Cronbachův koeficient alfa po vyřazení položky ze škály

země	Individuální sociální kapitál			Kolektivní sociální kapitál			
	TRUSTED	FAIR	HELPFUL	FRIEND	VOLUNORG	MEET	SOCATIV
Rakousko	0,74	0,66	0,75	0,44	0,59	0,42	0,45
Belgie	0,55	0,56	0,65	0,48	0,44	0,50	0,37
Česká republ.	0,63	0,65	0,75	0,48	0,52	0,39	0,38
Německo	0,60	0,52	0,60	0,47	0,53	0,42	0,40
Dánsko	0,56	0,56	0,73	0,39	0,44	0,40	0,26
Španělsko	0,61	0,64	0,74	0,39	0,46	0,41	0,48
Finsko	0,61	0,58	0,69	0,44	0,49	0,46	0,35
Francie	0,50	0,52	0,56	0,34	0,48	0,39	0,39
Spoj. králov.	0,66	0,59	0,69	0,39	0,52	0,40	0,36
Řecko	0,68	0,66	0,71	0,50	0,51	0,37	0,37
Maďarsko	0,67	0,68	0,69	0,54	0,65	0,48	0,49
Švýcarsko	0,57	0,53	0,68	0,39	0,51	0,41	0,35
Irsko	0,65	0,58	0,70	0,36	0,48	0,31	0,21
Izrael	0,59	0,53	0,75	0,30	0,48	0,35	0,29
Itálie	0,68	0,60	0,69	0,29	0,45	0,42	0,45
Lucembursko	0,48	0,46	0,64	0,18	0,37	0,26	0,26
Nizozemí	0,59	0,58	0,72	0,44	0,47	0,42	0,36
Norsko	0,50	0,46	0,68	0,38	0,47	0,41	0,28
Polsko	0,54	0,52	0,62	0,47	0,51	0,37	0,39
Portugalsko	0,56	0,62	0,73	0,33	0,40	0,32	0,46
Švédsko	0,57	0,53	0,67	0,39	0,44	0,42	0,33
Slovinsko	0,65	0,60	0,68	0,54	0,52	0,42	0,42
Všechny země	0,64	0,61	0,68	0,41	0,51	0,38	0,40
Variabilita	0,26	0,22	0,19	0,36	0,28	0,24	0,28

Zdroj: ESS 2002 (vážená data, Z skórov standardizace dat). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

⁴⁷ Soukup [2006b] uvádí, že zatímco hodnota 0,7 u škály z pěti položek je dobrý výsledek, u škály z 30 položek nikoliv.

Analýza reliability škály měřící kolektivní sociální kapitál opět kopíruje výsledky předešlé korelační analýzy. Jako problematická se opakovaně jeví položka VOLUNORG, jejíž vyloučení ze škály by přineslo buď zvýšení reliability škály (viz Rakousko, Irsko, Švýcarsko, atp. v Tabulce 10) anebo jen velmi malou či žádnou změnu ve velikosti indexu. Největší variabilitu Cronbachova koeficientu alfa mezi zeměmi ale vykazuje položka FRIEND. Velikost Cronbachova alfa celé škály není zcela uspokojivá, nejnižší hodnota 0,33 v Lucembursku je více méně nevyhovující pro další práci s touto měřicí škálou (viz Tabulka 11). Nejvyšších hodnot 0,62 dosahuje tento koeficient v Maďarsku, tato velikost se při zvážení přítomnosti pouhých čtyř proměnných zdá být již postačující. Variabilita velikosti Cronbachova alfa pro celou tuto škálu napříč zeměmi je podstatně větší než u individuálního sociálního kapitálu (0,29 bodu oproti 0,16), proto lze očekávat, že položky měřící kolektivní sociální kapitál budou činit při komparaci větší problémy než položky měřící individuální sociální kapitál.

Tabulka 11 Analýza mezipoložkové reliability škál - Cronbachův koeficient alfa pro dané škály

Cronbachův koeficient alfa pro škálu	Individuální sociální kapitál	Kolektivní sociální kapitál	Důvěra ve veřejné instituce	Individuální politické odcizení	Institucionální politické odcizení
země					
Rakousko	0,79	0,55	0,79	0,71	0,72
Belgie	0,68	0,52	0,82	0,69	0,73
Česká republika	0,76	0,52	0,82	0,73	0,76
Německo	0,67	0,53	0,77	0,70	0,79
Dánsko	0,71	0,45	0,79	0,73	0,77
Španělsko	0,75	0,51	0,81	0,74	0,74
Finsko	0,72	0,51	0,82	0,67	0,77
Francie	0,63	0,47	0,80	0,70	0,73
Spojené království	0,74	0,49	0,81	0,68	0,78
Řecko	0,76	0,51	0,81	0,76	0,70
Maďarsko	0,76	0,62	0,83	0,68	0,73
Švýcarsko	0,69	0,49	0,80	0,73	0,73
Irsko	0,73	0,42	0,79	0,69	0,76
Izrael	0,72	0,43	0,76	0,60	0,68
Itálie	0,74	0,48	0,72	0,73	0,74
Lucembursko	0,63	0,33	0,82	0,69	0,70
Nizozemí	0,72	0,50	0,82	0,73	0,77
Norsko	0,65	0,46	0,78	0,66	0,76
Polsko	0,65	0,51	0,78	0,72	0,71
Portugalsko	0,73	0,45	0,77	0,78	0,70
Švédsko	0,69	0,47	0,81	0,68	0,78
Slovinsko	0,73	0,55	0,80	0,69	0,62
Všechny země	0,73	0,50	0,79	0,73	0,77
Variabilita	0,16	0,29	0,11	0,18	0,17

Zdroj: ESS 2002 (vážená data, Z skór standardizace dat). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Pokud se podíváme na výsledky analýzy škály měřící důvěru ve veřejné instituce, zjistíme, že velikost Cronbachova alfa položek je v jednotlivých zemích nadprůměrná a blíží se doporučené limitní hranici 0,7 (viz Tabulka 12). Jako sporná položka se opět objevuje proměnná LEGAL, při jejímž vyloučení by se reliability škály ve většině zemí ještě zvýšila anebo by zůstala stejná. Variabilita velikostí ukazatele Cronbachova alfa pro celou škálu napříč zeměmi je ale minimální (0,11 bodu v Tabulce 11) a hodnoty alfa jsou značně vysoké

(> 0,7), tudíž se zdá, že tato škála by měla být bezproblémová jak z hlediska konzistence, tak i mezinárodní srovnatelnosti.

Tabulka 12 Analýza mezipoložkové reliability škály měřící důvěru ve veřejné instituce – Cronbachův koeficient alfa po vyřazení položky ze škály

země	Důvěra ve veřejné instituce		
	POLIT	PARLAM	LEGAL
Rakousko	0,72	0,62	0,73
Belgie	0,72	0,73	0,82
Česká republika	0,75	0,68	0,81
Německo	0,65	0,62	0,80
Dánsko	0,67	0,64	0,82
Španělsko	0,76	0,72	0,76
Finsko	0,72	0,68	0,84
Francie	0,67	0,72	0,79
Spojené království	0,66	0,69	0,84
Řecko	0,75	0,66	0,79
Maďarsko	0,78	0,75	0,77
Švýcarsko	0,73	0,67	0,79
Irsko	0,67	0,65	0,80
Izrael	0,65	0,56	0,79
Itálie	0,59	0,53	0,76
Lucembursko	0,78	0,68	0,80
Nizozemí	0,68	0,74	0,81
Norsko	0,66	0,62	0,80
Polsko	0,67	0,67	0,75
Portugalsko	0,67	0,67	0,72
Švédsko	0,70	0,69	0,82
Slovinsko	0,74	0,67	0,78
Všechny země	0,69	0,67	0,80
Variabilita	0,19	0,22	0,12

Zdroj: ESS 2002 (vážená data, Z skór standardizace dat). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Obě dvě škály měřící koncept politického odcizení se jeví z hlediska reliability stabilně. Žádná z položek měřících individuální politické odcizení se ve většině zemí nezdá být vhodným adeptem k vyloučení ze škály; individuální případy zvýšení reliability škály jsou spíše náhodného rázu. Největší variabilitu Cronbachova koeficientu alfa mezi zeměmi vykazují položky OPINION a DIFFICULT (viz Tabulka 13). V případě individuálního politického odcizení ukazuje Cronbachovo alfa škály na její konzistentnost ve všech zemích. Jeho hodnoty se pohybují od 0,6 do 0,78 s tím, že baterie sestává pouze ze tří položek. Variabilita hodnot tohoto indexu mezi zeměmi je pro individuální politické odcizení pod hranicí 0,2 (viz Tabulka 11).

Naopak v případě institucionálního odcizení lze zauvažovat nad položkou POLIT, jejíž vyřazení by zvýšilo reliability této škály v šesti zemích. Lze očekávat, že tato položka nemusí být v těchto zemích ekvivalentní při komparaci s ostatními zeměmi. Vysokou variabilitu koeficientu alfa nacházíme u položky VOTE, proto je vhodné věnovat pozornost i jí z hlediska srovnatelnosti (viz Tabulka 13). Variabilita hodnot Cronbachova alfa pro celou škálu je i v případě institucionálního odcizení pod hranicí 0,2 a navíc jeho velikost ve většině zemí přesahuje hranici 0,7 (viz Tabulka 11).

Tabulka 13 Analýza mezipoložkové reliability škály měřící politické odcizení - Cronbachův koeficient alfa po vyřazení položky ze škály

země	Individuální politické odcizení				Institucionální odcizení		
	INTEREST	DIFFICULT	POLACTIV	OPINION	VOTE	CARE	POLIT
Rakousko	0,61	0,65	0,69	0,63	0,59	0,57	0,73
Belgie	0,60	0,65	0,61	0,63	0,67	0,61	0,65
Česká republika	0,66	0,67	0,69	0,63	0,66	0,64	0,73
Německo	0,60	0,65	0,70	0,60	0,70	0,70	0,74
Dánsko	0,68	0,66	0,69	0,65	0,62	0,67	0,77
Španělsko	0,66	0,68	0,71	0,65	0,57	0,60	0,77
Finsko	0,57	0,63	0,65	0,57	0,69	0,68	0,71
Francie	0,61	0,64	0,66	0,65	0,61	0,66	0,65
Spoj. království	0,58	0,60	0,61	0,64	0,72	0,67	0,71
Řecko	0,73	0,76	0,67	0,65	0,57	0,54	0,72
Maďarsko	0,57	0,63	0,70	0,56	0,58	0,66	0,71
Švýcarsko	0,64	0,68	0,70	0,67	0,65	0,59	0,67
Irsko	0,58	0,65	0,62	0,64	0,65	0,64	0,75
Izrael	0,53	0,55	0,55	0,49	0,52	0,58	0,65
Itálie	0,64	0,74	0,68	0,63	0,58	0,61	0,75
Lucembursko	0,58	0,63	0,65	0,62	0,56	0,58	0,68
Nizozemí	0,66	0,67	0,69	0,65	0,68	0,64	0,73
Norsko	0,56	0,60	0,56	0,66	0,65	0,66	0,70
Polsko	0,65	0,66	0,69	0,62	0,57	0,59	0,69
Portugalsko	0,73	0,73	0,73	0,72	0,55	0,53	0,75
Švédsko	0,59	0,63	0,64	0,62	0,71	0,68	0,71
Slovinsko	0,65	0,62	0,63	0,60	0,42	0,53	0,59
Všechny země	0,64	0,69	0,69	0,65	0,66	0,66	0,74
Variabilita	0,20	0,21	0,18	0,23	0,30	0,17	0,18

Zdroj: ESS 2002 (vážená data, Z skór standardizace dat). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

4.3 Explorativní faktorová analýza

K odlišení jednotlivých dimenzí konceptů sociálního kapitálu a politického odcizení, potvrzení konzistentnosti škál a ohodnocení mezinárodní srovnatelnosti položek lze využít také explorativní faktorovou analýzu (PCA: Principal Component Analysis). Tento typ faktorové analýzy posuzuje korelace většího množství přímo měřitelných (tzn. manifestních) proměnných a na základě toho určuje, zda je možné zastoupit dané proměnné co nejmenším počtem nepřímo měřených latentních proměnných (jinak dimenzí, komponent, faktorů atp.). Mezi takto nalezenými latentními proměnnými mohou být různě silné vztahy vzájemné či jednostranné závislosti, avšak v běžně provedené faktorové analýze jsou faktory zpravidla zkonstruovány tak, aby byly nekorelované a standardizované.

Klíčovým výstupem faktorové analýzy je zde tabulka, ve které nalezneme ve sloupcích vypočtené faktory a faktorové zátěže, které reprezentují vztah dané položky k vytvořenému faktoru; matematicky jsou vyjádřeny Pearsonovým korelačním koeficientem vypočítaným mezi danou položkou a faktorem. Hodnoty faktorových zátěží se pohybují od nuly do jedné (se znaménky plus i minus), kde nula znamená nezávislost položky a faktoru (tj. proměnná daný faktor nijak nevysvětluje) a směrem k jedničce se síla jejich vztahu zvětšuje [viz Hendl 2006: 474-477]. Vytvořené faktory byly dále v analýze transformovány pomocí rotace Varimax, tak aby je bylo možné snadněji interpretovat.⁴⁸

⁴⁸ Varimax je ortogonální rotace, která minimalizuje počet proměnných, které mají vysoké zátěže s každým společným faktorem. Relativně snadná interpretace faktorů tkví v tom, že faktorové zátěže v rotaci Varimax nabývají hodnot blízkých buď jedné, nebo nule. To znamená, že každá položka je silně korelována jen s některými faktory a s ostatními faktory je korelována slabě [více Hendl 2006: 476-477].

Tabulka 14 Dimenze sociálního kapitálu a politického odcizení – výsledky faktorové analýzy PCA s rotací Varimax

	Individuální odcizení	Institucionální odcizení	Důvěra v instituce	Individuální sociální kapitál	Kolektivní sociální kapitál
INTEREST	,665	,148	,097	-,082	-,042
DIFFICULT	,562	,054	,076	-,049	,002
POLACTIV	,555	,119	,070	-,002	-,117
OPINION	,675	-,003	-,006	-,056	-,049
VOTE	,148	,659	,277	-,163	-,076
CARE	,206	,664	,283	-,151	-,078
POLIT	,108	,367	,707	-,189	-,068
PARLAM	,101	,198	,766	-,163	-,079
LEGAL	,068	,094	,590	-,178	-,087
TRUSTED	-,101	-,088	-,183	,640	,130
FAIR	-,070	-,088	-,117	,719	,112
HELPFUL	-,018	-,105	-,142	,597	,101
FRIEND	-,016	-,008	-,006	,150	,435
VOLUNORG	-,011	-,031	-,119	,014	,246
MEET	,003	-,022	,001	,067	,560
SOCACTIV	-,146	-,048	-,025	,041	,504
Vysvětlená variance (%)	10,3	7,1	10,6	9,2	5,7
Vysvětlená variance celkem (%)	42,76				
Míra KMO	0,82				

Zdroj: ESS 2002 (vážená data). Celkový soubor 22 zemí - analýza všech zemí dohromady. Výběr případů do analýzy metodou Listwise; vlastní výpočty.
 Poznámka: Kaiser-Meyer-Olkinova míra (KMO) se používá pro posouzení vhodnosti položek pro faktorovou analýzu (její hodnoty se pohybují mezi 0 – 1).
 Hodnota mezi 0,8 – 0,9 se považuje za vysoké hodnocení vhodnosti provedené faktorové analýzy.

Výsledky faktorové analýzy jsou založené na analýze, v níž byl podle teorie stanoven požadavek na extrakci pěti faktorů (resp. latentních proměnných).⁴⁹ Faktorová analýza ukázala, že je zcela oprávněné hovořit o několika dimenzích sociálního kapitálu i politického odcizení. Položky, které měly měřit dané faktory na základě teoretických předpokladů, je opravdu měří, všechny sytí pouze jeden faktor. V tomto faktoru mají položky relativně vysoké hodnoty faktorových zátěží, zatímco v ostatních faktorech poměrně nízké. Nejproblematictější se opět jeví položka VOLUNORG v dimenzi kolektivního sociálního kapitálu a POLIT v institucionálním odcizení (viz Tabulka 14), neboť mají nejmenší faktorové zátěže v daných faktorech. Z hlediska vysvětlené variance vycházejí faktorové analýzy ve všech zemích velmi podobně. Nejvyšších hodnot dosahuje Rakousko, Česká republika, Řecko, Maďarsko (nad 45 %) a nejnižších hodnot Francie, Izrael, Lucembursko a Polsko (pod 40 %). KMO míra posuzuje vhodnost faktorové analýzy také podobně, koeficient osciluje od 0,72 (Izrael) do 0,83 (Nizozemí a Belgie), jedná se tudíž o středně vysoké až vysoké hodnocení vhodnosti použití daných položek.

V základním testování mezinárodní srovnatelnosti položek bylo nutné porovnat velikosti jejich faktorových zátěží v příslušném faktoru v jednotlivých zemích mezi sebou; položky s největší mírou variability lze pokládat za nejproblematictější. Prezentované výsledky zde dokládají samozřejmě jen část získaných údajů. Bylo by možné porovnávat velikosti faktorových zátěží dané položky mezi zeměmi i v jiných faktorech, ale vygenerované údaje z takovéto analýzy již nelze uspokojivě a přehledně prezentovat a interpretovat – to je jedna z nevýhod použití explorativní faktorové analýzy k ohodnocení srovnatelnosti položek. Ke komplexní analýze srovnatelnosti faktorových struktur je tudíž téměř nezbytné využít konfirmační faktorovou analýzu, která má schopnost vyprodukovat mnohem přehlednější výstupy. Jsou zde z tohoto důvodu dále posuzovány pouze vztahy položek k vybranému, z teorie definovanému faktoru a ostatní faktorové zátěže jsou zanedbány.

Faktorové zátěže položek měřících faktorů individuálního sociálního kapitálu jsou ve všech zemích značně vysoké, jejich velikost se pohybuje v intervalu od 0,4 do 0,8. Z hlediska země dosahují položky nejslabších výsledků ve Francii a z hlediska položky se jako problematičtější jeví HELPFUL s největší variabilitou hodnot faktorových zátěží napříč zkoumanými zeměmi (viz Tabulka 15).

V případě škály kolektivního sociálního kapitálu se jako nejslabší položka jeví VOLUNORG, jejíž hodnoty faktorových zátěží kolísají od 0,08 do 0,48 (viz Tabulka 15). V některých zemích tak tato položka sytí daný faktor dostatečně (např. Belgie, Španělsko, Itálie, atp.) a v některých zemích daný faktor nevysvětluje téměř vůbec (např. Irsko, Izrael, Řecko). Na první pohled se také zdá, že pokud položka VOLUNORG vysvětluje daný faktor uspokojivě, jsou faktorové zátěže jednotlivých položek ve faktoru více méně vyrovnané, pohybují se okolo hodnot 0,4 – 0,5. Pokud daný faktor nevysvětluje vůbec anebo jen velmi slabě, velikost faktorových zátěží položek MEET a SOCACTIV se začne pohybovat okolo 0,6 – 0,7. Zdá se tak, že roli dobrovolných organizací v tomto případě přebírají při vysvětlování kolektivního sociálního kapitálu spíše přátelské a příbuzenské sociální sítě. Nevyváženost příspěvku položky VOLUNORG do faktoru kolektivního sociálního kapitálu by tak mohla způsobovat značné problémy při její mezinárodní komparaci. Navíc je variabilita velikosti faktorových zátěží jednotlivých položek v zemích příliš vysoká (více než 0,4 bodu rozdílu mezi nejvyšší a nejnižší hodnotou), proto lze i z explorativní faktorové analýzy usoudit na jejich problematické srovnávání napříč zeměmi.

⁴⁹ Může být stanoveno i jiné kritérium například na základě požadavku, aby eigenvalue (tzv. vlastní hodnota) dosahovala hodnoty jedna, což je i běžnější postup při aplikaci faktorové analýzy.

Výsledky zatím ukazují, že údaje všech tří technik si vzájemně odpovídají. Škála individuálního sociálního kapitálu je poměrně robustní a největší překážky v mezinárodní komparaci můžeme na základě korelační analýzy a explorativní faktorové analýzy očekávat od položky HELPFUL a na základě analýzy reliability od položky TRUSTED. Škála kolektivního kapitálu se nezdá být obecně příliš konzistentní, a co se týče srovnatelnosti, největší problém činí položka VOLUNORG. Ale ani ostatní položky nejsou z hlediska variability koeficientů mezi zeměmi ideální – analýzy ukazují komplikace též v případě položek FRIEND a SOCACTIV.

Tabulka 15 Výsledky faktorové analýzy PCA s rotací Varimax pro faktory sociálního kapitálu (faktorové zátěže)

země	Individuální sociální kapitál			Kolektivní sociální kapitál			
	TRUSTED	FAIR	HELPFUL	FRIEND	VOLUNORG	MEET	SOCACTIV
Rakousko	0,67	0,81	0,68	0,45	0,21	0,66	0,56
Belgie	0,66	0,67	0,53	0,41	0,48	0,38	0,61
Česká republika	0,76	0,72	0,59	0,38	0,28	0,62	0,59
Německo	0,54	0,74	0,56	0,39	0,27	0,60	0,59
Dánsko	0,72	0,78	0,52	0,36	0,28	0,40	0,63
Španělsko	0,78	0,70	0,56	0,61	0,39	0,48	0,36
Finsko	0,64	0,72	0,57	0,48	0,29	0,47	0,55
Francie	0,60	0,58	0,53	0,50	0,23	0,51	0,45
Spoj. království	0,66	0,76	0,61	0,49	0,22	0,53	0,56
Řecko	0,71	0,76	0,66	0,24	0,19	0,69	0,61
Maďarsko	0,69	0,67	0,67	0,46	0,21	0,74	0,68
Švýcarsko	0,64	0,71	0,49	0,46	0,24	0,54	0,55
Irsko	0,64	0,81	0,59	0,24	0,08	0,59	0,71
Izrael	0,72	0,81	0,45	0,30	0,09	0,53	0,62
Itálie	0,62	0,73	0,64	0,71	0,40	0,31	0,30
Lucembursko	0,61	0,68	0,45	0,52	0,28	0,32	0,38
Nizozemí	0,72	0,74	0,50	0,41	0,34	0,46	0,56
Norsko	0,63	0,76	0,43	0,42	0,29	0,43	0,60
Polsko	0,64	0,66	0,46	0,30	0,22	0,65	0,57
Portugalsko	0,76	0,63	0,55	0,48	0,42	0,54	0,23
Švédsko	0,64	0,71	0,53	0,43	0,37	0,40	0,51
Slovinsko	0,69	0,76	0,58	0,30	0,30	0,65	0,64
Všechny země	0,64	0,72	0,60	0,44	0,25	0,56	0,50
Variabilita	0,24	0,23	0,25	0,47	0,40	0,43	0,48

Zdroj: ESS 2002 (vážená data). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Poznámka: V tabulce jsou z úsporných důvodů prezentovány pouze faktorové zátěže položek sytící nejvíce daný faktor.

Výsledky faktorové analýzy položek vysvětlujících dimenzi důvěry respondentů ve veřejné instituce jsou opět velice uspokojivé. Jejich faktorové zátěže jsou dostatečně vysoké a pohybují se v rozmezí od 0,5 do 0,8 (viz Tabulka 16). V žádné zemi a ani u žádné položky se neobjevuje extrémní výchylka. Největší variabilitu vykazuje položka LEGAL (0,27 bodu rozdílu), která by mohla při mezinárodní komparaci činit potíže. V případě škály důvěry ve veřejné instituce všechny tři typy analýz potvrzují, že škála je velmi kvalitní a konzistentní, avšak nemalé problémy může způsobovat právě tato položka LEGAL. Analýza reliability je jediná, která naopak ukazuje v rámci této škály na potíže s položkou PARLAM.

Tabulka 16 Výsledky faktorové analýzy PCA s rotací Varimax pro faktor důvěry ve veřejné instituce (faktorové zátěže)

země	Důvěra ve veřejné instituce		
	POLIT	PARLAM	LEGAL
Rakousko	0,66	0,85	0,63
Belgie	0,72	0,76	0,61
Česká republika	0,69	0,80	0,67
Německo	0,69	0,80	0,54
Dánsko	0,71	0,78	0,56
Španělsko	0,72	0,76	0,73
Finsko	0,66	0,80	0,56
Francie	0,77	0,68	0,63
Spojené království	0,80	0,76	0,59
Řecko	0,68	0,87	0,66
Maďarsko	0,67	0,75	0,75
Švýcarsko	0,63	0,83	0,62
Irsko	0,70	0,76	0,61
Izrael	0,69	0,86	0,54
Itálie	0,66	0,79	0,48
Lucembursko	0,65	0,84	0,66
Nizozemí	0,77	0,73	0,57
Norsko	0,71	0,71	0,51
Polsko	0,70	0,72	0,63
Portugalsko	0,69	0,71	0,70
Švédsko	0,72	0,77	0,59
Slovinsko	0,69	0,81	0,62
Všechny země	0,71	0,77	0,59
Variabilita	0,17	0,19	0,27

Zdroj: ESS 2002 (vážená data). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Poznámka: V tabulce jsou prezentovány pouze faktorové zátěže položek sytící nejvíce daný faktor.

Co se týče konceptu politického odcizení, v obou faktorech se ukazuje, že položky je vysvětlují dostatečně. Jejich faktorové zátěže jsou až na drobné výjimky poměrně vysoké (viz Tabulka 17). V případě individuálního politického odcizení variabilita faktorových zátěží jednotlivých položek ukazuje na to, že obzvláště při mezinárodní komparaci položek POLACTIV a OPINION napříč zeměmi tak možná budou vznikat problémy. Ve faktoru institucionálního odcizení má nejnižší hodnoty faktorových zátěží proměnná POLIT a největší variabilitu velikostí faktorových zátěží vykazuje položka VOTE. Opět v tomto případě všechny tři typy analýz posuzují škálu politického odcizení jako vnitřně konzistentní, ale z hlediska ekvivalence se v individuálním odcizení ukazují jako problematické položky OPINION a POLACTIV a v institucionálním odcizení POLIT a VOTE.

Závěrem lze říci, že pro výzkumníka je při použití těchto typů analýz důležité pozorovat spíše variabilitu velikostí jednotlivých indexů napříč zeměmi, spíše než hodnotit obecné spodní hranice přípustné velikosti daných koeficientů (jako např. 0,7 u Cronbachova alfa a 0,3 u korelačního koeficientu). Pokud nedochází k extrémním výchylnkám, které jsou jasnou známkou toho, že škála neměří stejný koncept. Z předložených údajů lze usuzovat, že na hodnotu variability 0,2 až 0,3 (a výše) bodu rozdílu mezi velikostmi jednotlivých koeficientů (korelací, faktorových zátěží, atp.) v daných zemích se lze dívat jako na první ukazatel možných komplikací při mezinárodní komparaci položek.

Z pohledu identifikace zemí, kde by bylo obtížné vybrané položky obecně komparovat - nelze prozatím jednoznačně stanovit, které země by byly v komparaci vychýlené, v tomto ohledu

jsou tyto techniky nepraktické. Jako nejproblematičtější se ale prozatím ukazují Lucembursko, Izrael, Irsko a Norsko.

Tabulka 17 Výsledky faktorové analýzy PCA s rotací Varimax pro faktory politického odcizení (faktorové zátěže)

země	Individuální politické odcizení				Institucionální odcizení		
	INTEREST	DIFFICULT	POLACTIV	OPINION	VOTE	CARE	POLIT
Rakousko	0,68	0,61	0,49	0,65	0,70	0,67	0,31
Belgie	0,63	0,53	0,60	0,57	0,55	0,63	0,38
Česká republika	0,58	0,61	0,55	0,73	0,59	0,68	0,38
Německo	0,67	0,56	0,44	0,72	0,66	0,62	0,40
Dánsko	0,60	0,65	0,58	0,65	0,76	0,64	0,35
Španělsko	0,65	0,60	0,57	0,67	0,71	0,68	0,28
Finsko	0,63	0,54	0,48	0,66	0,67	0,63	0,43
Francie	0,65	0,61	0,52	0,61	0,51	0,63	0,31
Spoj. království	0,65	0,57	0,58	0,52	0,55	0,65	0,34
Řecko	0,56	0,48	0,74	0,82	0,65	0,71	0,32
Maďarsko	0,65	0,52	0,37	0,72	0,72	0,59	0,36
Švýcarsko	0,69	0,61	0,56	0,62	0,63	0,59	0,37
Irsko	0,68	0,52	0,61	0,53	0,70	0,68	0,36
Izrael	0,48	0,48	0,49	0,62	0,62	0,62	0,33
Itálie	0,67	0,50	0,61	0,71	0,75	0,65	0,30
Lucembursko	0,60	0,56	0,58	0,53	0,63	0,67	0,37
Nizozemí	0,60	0,60	0,57	0,67	0,65	0,67	0,33
Norsko	0,65	0,54	0,65	0,42	0,61	0,71	0,40
Polsko	0,61	0,57	0,55	0,68	0,65	0,65	0,31
Portugalsko	0,64	0,65	0,68	0,73	0,72	0,72	0,24
Švédsko	0,60	0,57	0,54	0,59	0,70	0,58	0,42
Slovinsko	0,51	0,63	0,54	0,65	0,68	0,53	0,31
Všechny země	0,67	0,56	0,56	0,68	0,66	0,66	0,37
Variabilita	0,21	0,17	0,37	0,4	0,25	0,19	0,19

Zdroj: ESS 2002 (vážená data). 22 zemí – každá analýza byla provedena samostatně. Výběr případů do analýzy metodou Listwise; vlastní výpočty.

Poznámka: V tabulce jsou prezentovány pouze faktorové zátěže položek sytící nejvíce daný faktor.

5 TESTOVÁNÍ SROVNATELNOSTI ŠKÁL POMOCÍ STRUKTURNÍHO MODELOVÁNÍ

Základním předpokladem při komparaci dat mezinárodních šetření je to, že v různých skupinách respondentů jsou teoretické koncepty chápány a interpretovány stejným způsobem, že mají tyto koncepty ve všech skupinách stejnou dimenzionální strukturu (tj. stejný počet a podobu latentních proměnných) a posléze že jsou tyto latentní proměnné také měřeny položkami podobným způsobem [Byrne 2004].⁵⁰

Jeden z přístupů, jak k dané problematice statisticky přistoupit a ověřit daný předpoklad, je otestovat podobnost faktorové struktury daného konceptu mezi jednotlivými skupinami [Byrne 2008] – v případě této studie faktorovou strukturu sociálního kapitálu a politického odcizení mezi 22 zeměmi. K tomu se v současnosti nejčastěji využívá strukturní modelování, jehož základy vycházejí z logiky konfirmační faktorové analýzy.⁵¹ Protože je nezbytné testovat ekvivalenci měřicích položek v různých skupinách populace (tzn. ve věkových, genderových, kulturních, rasových, jazykových či jiných sociálních skupinách) anebo v různých časových okamžicích (např. kontinuální výzkumy ESS) či mezi různými metodami sběru dat (např. MTMM design), aplikují se tzv. několika výběrové analýzy (multisample analysis) [Matějů 1989: 406], v současnosti spíše označované jako několika skupinové analýzy (multigroup analysis).⁵²

Strukturní modelování (resp. konfirmační faktorová analýza) se začalo dynamicky rozvíjet kvůli narůstající potřebě ověřování nepřeborného množství teoretických koncepcí.⁵³ Při jeho aplikaci se data modelují různými modely latentních proměnných [Hendl 2006: 483].⁵⁴ Výzkumník má na rozdíl od explorativní faktorové analýzy předběžnou představu o počtu a vzájemných vztazích latentních proměnných a tyto hypotézy dále ověřuje. Ze statistického hlediska se hypotézy týkají toho, zda původní kovarianční matice může být reprezentována kovarianční maticí, jež je odvozena na základě teoretického předpokladu, že měřené položky vznikly jako specifické lineární kombinace faktorů [Hendl 2006: 480]. Nebo jednodušeji, zda můžeme pozorovaná data vysvětlit navrženým teoretickým modelem. Pomocí statistik vhodnosti a velikosti koeficientů (resp. parametrů) ověřujeme míru podobnosti těchto dvou kovariančních matic a souběžně také (statistickou) kvalitu testovaného modelu.

Strukturní model je tedy formální reprezentací systému kauzálních vazeb mezi manifestními a latentními proměnnými. Klasicky je model zobrazován grafickým diagramem vycházejícím z path analýzy (tj. path diagram) [více Matějů 1989: 400; Hox, Bechger 1998: 2], kde jsou měřicí položky umístěny do obdélníků, latentní proměnné do elips (oválů), chyby měření do kroužků a šipky znázorňují vztahy mezi faktory. Základem je odlišení modelu měření

⁵⁰ To zahrnuje stejné vnímání a výklad obsahu jednotlivých položek ve všech skupinách, jejich adekvátní formát, podobný podíl položky na vysvětlení latentní proměnné, podobné vztahy mezi latentními proměnnými v daných skupinách atp. Pokud chceme tyto teoretické koncepty porovnávat v různých skupinách, musí být tento předpoklad platný.

⁵¹ O přehledu empirických studií využívajících k testování ekvivalence měření strukturní modelování více Vandenburg a Lance [2000] a aktualizovaně Schmitt a Kuljanin [2008].

⁵² Technika využívaná v této studii se proto nazývá „multigroup confirmatory factor analysis (MGCF)” – je to konfirmační faktorová analýza testující hypotézu o shodě teoretického modelu s daty v několika skupinách/výběrech zároveň. Tyto metody vhodné pro testování ekvivalence mezi několika skupinami vychází z významné práce švédského psychometrika Jöreskoga (1971).

⁵³ Z konfirmační analýzy se vyvinuly metody analýzy modelů kovariančních struktur (analysis of covariance structures (ANCOV/COVS), dnes známých pod souhrnným názvem strukturní modelování - jinak modely strukturních rovnic (structural equation models; SEM), strukturní modely (structural models) či strukturní modelování (structural modelling). V současnosti se čím dál více prosazují tzv. analýzy průměrů a kovariančních struktur (mean and covariance structure analysis; MACS). Pro potřeby sociologie tyto statistické techniky představili Bielby a Hauser v 70. letech [viz Bielby, Hauser 1977].

⁵⁴ V tomto ohledu následuje strukturní modelování odkaz psychometrických přístupů a hlavní důraz klade právě na nepřímo pozorované znaky, tj. latentní proměnné.

(measurement model) a strukturního modelu (structural model). Model měření je v podstatě variantou konfirmačního faktorového modelu, který přesně stanovuje, jaká položka měří konkrétní latentní proměnnou. Strukturní model následně stanovuje vzájemné vztahy mezi těmito latentními proměnnými a je již oproštěn od variance chyb měření (resp. reziduální variance měřících položek), což je jeho velkou výhodou (viz Diagram 1) [více Matějů 1989: 408; Hox, de Leeuw, Brinkhuis 2010: 397]. Ve strukturním modelování se vždy usiluje o model, který je co nejúspornější, pokud jde o počet kauzálních vazeb a počet proměnných [tj. princip parsimony, více Matějů 1989; Urbánek 2000].

Za grafickým vyjádřením modelu se ale skrývá řada regresních rovnic, které představují dané kauzální vazby. Každý indikátor je reprezentován touto rovnicí:

$$I = c + f * L + u \quad (1)$$

Kde I znamená přímé měření neboli indikátor/položka, c je konstanta neboli intercept, L zastupuje nepřímé měření neboli latentní proměnnou, f je regresní koeficient/váha (tzv. směrnice) neboli faktorová zátěž v pojmech faktorové analýzy a nakonec u je reziduum neboli chyba měření (viz Diagram 1) [Schmitt, Kuljanin 2008: 211].⁵⁵ Teď již lze za použití vysvětlené regresní rovnice reformulovat základní hypotézu o invariantním (tj. ekvivalentním) měřením: Pokud mají osoby z různých výběrů (skupin) stejné postavení v daném teoretickém konceptu L (resp. latentní proměnné), měly by obdržet stejný skóre (hodnocení, známku) I a tudíž parametry položek (směrnice f a konstanta c) by měly být také shodné [Robert, Lee, Chan 2006: 79; Schmitt, Kuljanin 2008: 211].⁵⁶

Testování ekvivalence měření prostřednictvím strukturního modelu (resp. konfirmační faktorové analýzy) je hierarchický stupňovitý proces. Strategie, která byla pro toto testování navržena Jöreskogem [1971], se vyznačuje vzrůstajícími restrikcemi, která je kladena na testovaný model. Vzrůstající restrikce se myslí rostoucí počet parametrů modelu (tj. faktorové zátěže, konstanty, chyby měření atd.), u kterých se požaduje, aby byly totožné ve všech zkoumaných skupinách. Na parametry základního (prvního) modelu se nekladou žádné restrikce a postupně se přidáváním přísnějších požadavků v logickém pořadí vytváří hierarchicky řazená řada modelů se stejnými proměnnými a vztahy mezi nimi, ale jinými restrikcemi uvalenými na dané parametry. Takto definované modely se pak testují a zkoumá se, zda odpovídají reálným datům z výzkumu [podrobněji Byrne 2010; Cheung, Rensvold 1999, 2002; Steenkamp, Baumgartner 1998; Vandenberg, Lance 2000].

Obecně dochází nejdříve k ověřování ekvivalence položek na úrovni modelu měření a posléze k testování ekvivalence latentních proměnných na úrovni strukturního modelu. Na úrovni modelu měření dochází k testování invariance faktorové struktury, faktorových zátěží, interceptů, chyb měření a někdy také korelace chyb měření. Na úrovni strukturního modelu se testuje shoda variancí a kovariancí latentních proměnných a shoda průměrů latentních proměnných. Obecně se má za to, že shoda na úrovni modelů měření je nezbytnou podmínkou

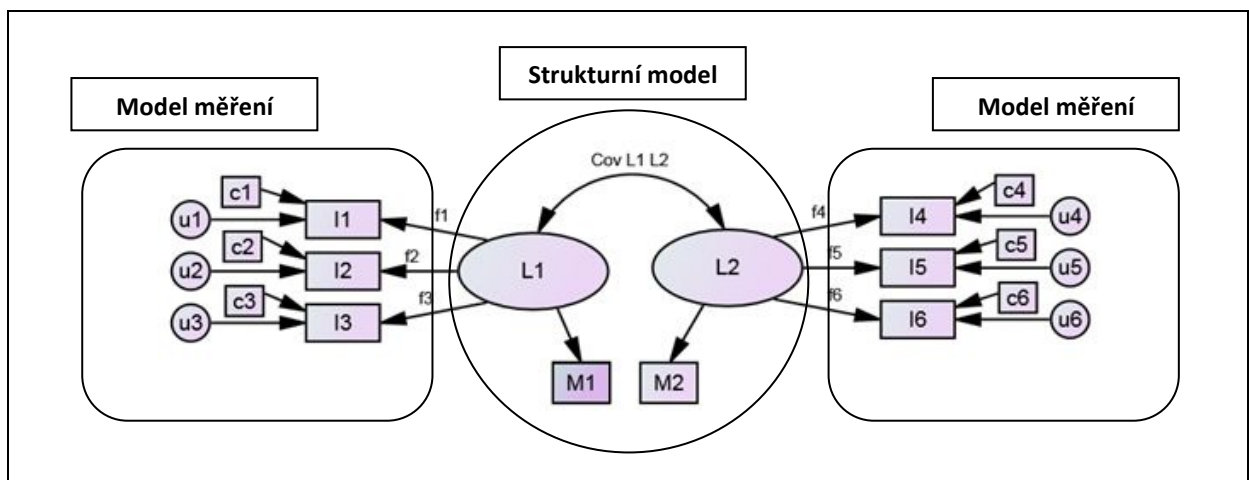
⁵⁵ Regresní rovnice bývá nejčastěji popsána řeckými písmeny ($x^{(g)} = \tau^{(g)} + \Lambda^{(g)} * \xi^{(g)} + \delta^{(g)}$). V tomto případě, protože tato metodologická studie není jen čistě statistického rázu, je upřednostněno použití latinky kvůli lepší přehlednosti. Formulace předloženého vzorce je částečně využita ze Schmitta a Kuljanina [2008].

⁵⁶ Stejným postavením v konceptu se rozumí stejný názor na danou věc – např. stejnou míru důvěry v instituce, stejný sociální kapitál, podobnou míru politického odcizení atp. Invarianci narušuje situace, kdy dva lidé, kteří mají například stejnou míru důvěry ve veřejné instituce (tj. shodnou se na svých názorech), manifestují v dotazníku jiné hodnocení dané důvěry (tj. odpovídají na otázku odlišně). Důvodů, proč se tomu tam mohlo stát, je mnoho (viz systematické chyby měření, kapitola 5.4). Částečně tento výklad koresponduje s výkladem reliability měření, kdy pozorovaná hodnota, resp. indikátor závisí na skutečné hodnotě (tj. hodnotě latentní proměnné) a chybě měření. Hodnoty pozorované v indikátorech/položkách nemusí tedy představovat pouze hodnotu latentní proměnné [Schubert 2010: 106-107].

pro komparaci položek a latentních proměnných mezi skupinami a výzkumník jí musí věnovat pozornost vždy, na druhou stranu shoda parametrů na úrovni strukturního modelu je více méně předmětem volby a výzkumných potřeb a výzkumník by měl znát důvod, proč je pro něj nezbytná [Milfont, Fischer 2010: 113].

Podle této logiky strukturního modelu existují i specifické názvy jednotlivých druhů daných ekvivalencí. V základní rovině se dostáváme k vymezení ekvivalence měření (pro model měření) a strukturní ekvivalence (pro strukturní model), které ale v tomto kontextu souvisí svým pojmenováním čistě s pojmoslovím strukturního modelování.⁵⁷ Další druhy ekvivalence budou dále definovány podle pořadí jednotlivých modelů, které budou v následující analýze pak také hierarchicky testovány. Z každé úrovně testování ekvivalence položek pak jasně vyplývá, jak lze mezinárodní data interpretovat, a především, jak je s nimi možné dále v komparativních analýzách zacházet a srovnávat je.

Diagram 1 Příklad strukturního modelu



Poznámka: Strukturní model se dvěma latentními proměnnými (L1 a L2), každou reprezentovanou třemi měřicími indikátory (I1 až I3 a I4 až I6). V regresní rovnici jsou další parametry: intercepty (c1 až c6), faktorové zátěže (f1 až f6) a rezidua (u1 až u6). Mezi faktory je vzájemná vazba (Cov L1 L2) a každá latentní proměnná má svou varianci a svůj průměr M1, M2.

Poznámka 2: V modelu měření je testována ekvivalence měření, ve strukturním modelu strukturní ekvivalence.

Pro aplikaci strukturního modelování na data bylo vyvinuto několik statistických softwarů – LISREL (Linear Structural Relations, autoři K. Jöreskog a Dag Sörbom), EQS (autor P. M. Bentler), AMOS (Analysis of Model Structures; distribuce IBM software, autor J. L. Arbuckle), Mplus (autor B. Muthén), atd. Pro ilustraci realizace strukturního modelování nebo konfirmační faktorové analýzy v jednotlivých softwarech lze doporučit publikace B. Byrne [1998, 2006, 2010, 2012] a Albrighta a Parka [2009].

5.1 Analýza ekvivalence měření ve faktorovém modelu měření

Ačkoli Jöreskog navrhoval jako první krok testování ekvivalence ověřit shodu kovariančních matic daných proměnných mezi jednotlivými skupinami, neboť tato shoda by měla ukázat na zcela ekvivalentní měření ve všech skupinách a nebylo by zapotřebí dalšího testování, Barbara Byrne tvrdí, že tento test je sporný a často vede k matoucím závěrům [podrobněji

⁵⁷ Je důležité si neplest strukturní ekvivalenci v modelování se strukturní ekvivalencí, o níž se hovořilo v kapitole 2.1.1 pod pojmem konceptuální ekvivalence – ta totiž označuje stejnou interpretaci a shodu ve vnímání sociálního jevu, který zkoumáme, v různých kulturních či sociálních skupinách [Milfont, Fischer 2010: 113].

Byrne 2010: 198-199]. Proto, při následování logiky jejího výkladu začíná testování ekvivalence měření tzv. separátní analýzou, ta funguje jako prerekvizita pro další multivariční testování ekvivalence položek.⁵⁸

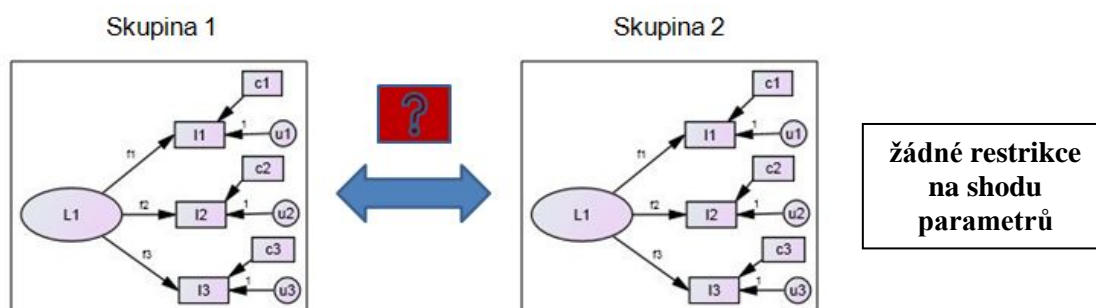
Díky univariční analýze se otestuje tzv. *základní model* v každé skupině zvlášť. V něm se ještě nekladou žádné požadavky (tj. restriktce) na shodu velikosti parametrů jednotlivých položek. Zkoumá se, zda základní model reprezentuje data z každé skupiny vhodným způsobem a podobně (tj. stejný počet faktorů a jejich vzájemné vztahy a stejné položky měřící dané faktory), zda je smysluplný a odpovídá principu úspornosti; to vše se posuzuje na základě statistik vhodnosti modelu. V procesu dalšího testování jsou postupně kladeny požadavky na shodu jednotlivých parametrů, proto je dále nezbytné analyzovat data ve všech skupinách zároveň (tj. aplikovat simultánní přístup).

Konfigurální/konstruktová ekvivalence (Model 1)

Konfigurální model se testuje ve všech zemích najednou a ještě stále v něm nejsou kladeny žádné požadavky na shodu parametrů kromě stejné podoby modelu (viz Diagram 2). Je proto naprosto totožný se základním modelem, jediný rozdíl je v simultánním přístupu k jeho testování. Ověřuje se, zda model adekvátně reprezentuje data ze všech skupin při daném počtu faktorů a zda jsou velikosti faktorových zátěží ve všech skupinách dostatečně veliké [podrobněji Meredith 1993; Vandenberg, Lance 2000: 36]. Jeho význam je převážně statistický, neboť umožňuje analyzovat data ze všech zemí současně a zároveň poskytuje tzv. referenční neboli základní, souhrnnou hodnotu statistik vhodnosti modelu, se kterou se porovnávají výsledky hodnocení dalších více restriktivních modelů [Byrne 2008: 873].

Tímto testem je možné zjistit, že si jsou faktorové struktury dat podobné, ale ne nutně ekvivalentní ve zkoumaných skupinách [Byrne 2008: 873]. Proto ještě není při meritorní analýze dat možné přímé statistické porovnávání skupin mezi sebou v daných proměnných [Hox, de Leeuw, Brinkhuis 2010: 399]. Konfigurální ekvivalence, která je takto otestována a potvrzena, ukazuje, že škály měří stejný teoretický koncept v každé skupině a že tento jev je podobně sémanticky vymezen, tj. vysvětlován stejnými kategoriemi, pojmy a významy.

Diagram 2 Testování konfigurálního modelu při analýze konstruktové ekvivalence dat



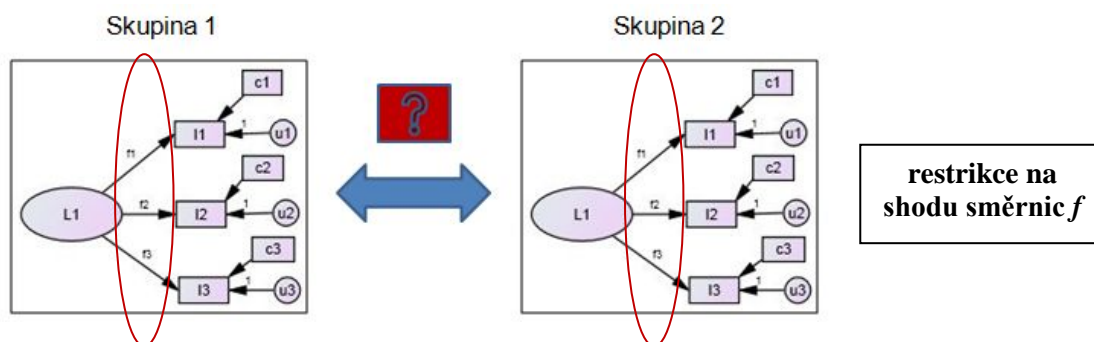
Metrická ekvivalence/ekvivalence měřící jednotky (Model 2)

Při zjišťování metrické ekvivalence se v konfigurálním modelu zadají požadavky na shodu faktorových zátěží každé měřící položky na všech faktorech ve všech zkoumaných skupinách a dále se testuje, zda takto definovaný model reprezentuje data z každé skupiny v šetření. Statisticky se jedná o testování shody směrnic (neboli faktorových zátěží) f v regresních rovnicích ($I = c + f \cdot L + u$) ve všech skupinách, kterých je g (viz Diagram 3, směrnice f).

$$f_1 = f_2 = f_3 = f_4 = \dots = f_g \quad (2)$$

⁵⁸ V zahraniční literatuře separátní analýzu většinou najdeme pod názvem „single analysis“ a simultánní analýzu pod názvem „multi-group analysis“.

Diagram 3 Testování metrického modelu při analýze metrické ekvivalence dat



Metrická ekvivalence nalezená v datech značí, že obsah každé položky, tak jak je v jednotlivých jazycích předložena, je vnímán a interpretován velmi podobným způsobem ve všech skupinách a že respondenti z různých skupin, pokud zastávají podobný názor, na otázky odpovídají stejně [Byrne 2008: 873; Milfont, Fischer 2010: 115]. Měřicí položky a potažmo celá škála pak měří konstrukt všude podobně a tedy i srovnatelně. Důležité je to, že výzkumník může v meritorní analýze dat přistoupit ke statistickému porovnávání vztahů daných latentních proměnných s jinými vysvětlujícími proměnnými, u kterých byla identifikována metrická ekvivalence, mezi zeměmi (tj. provádět například korelační a regresní analýzu). Metrická ekvivalence je nutnou podmínkou pro testování všech dalších typů ekvivalencí [Hox, de Leeuw, Brinkhuis 2010: 399]. Meredith [1993] nazývá tento typ ekvivalence jako „slabou ekvivalenci“ a následující skalární ekvivalenci jako „silnou ekvivalenci“.

Ne všechny faktorové zátěže musí být shodně velké ve všech zemích. Téma *částečné (parciální) ekvivalence měření* poprvé představili Muthén a Christofferson [1981] a její řešení v praxi ukázali Byrne, Shavelson a Muthén [1989]. Když všechny parametry (faktorové zátěže nebo konstanty či variance chyb) splňují danou restrikcí na jejich shodu ve všech skupinách, nazývá se tento stav plná ekvivalence. Pokud nikoli, za určitých podmínek mohou dosáhnout úrovně alespoň částečné ekvivalence. Základní myšlenka je taková, že na některé parametry v modelu může být uvalena restrikcí o shodě mezi skupinami a některé parametry se mohou mezi skupinami lišit. Druhý pohled je takový, že testované parametry mohou být zcela shodné jen v určitých zemích a v jiných shodné být nemusí [Vandenberg, Lance 2000]. Jedná se tedy buď o uvolnění některých položek, nebo některých zemí z testované skupiny. Mnoho autorů se kloní k názoru, že při určitém povolení restrikcí shody některých parametrů nemusí nutně dojít k tomu, že měřicí položku označíme za neekvivalentní a vyřadíme ji. Můžeme ji považovat za částečně ekvivalentní s tím, že v případě rozumné míry narušení její srovnatelnosti lze dále pokračovat v testování více restriktivních modelů a lze mezi některými skupinami položky či faktory do určité míry smysluplně porovnávat.

Pravidla tohoto kompromisu částečné ekvivalence měření (resp. povolení míry zkreslení) nejsou jasně a definitivně stanovena, ale platí, že pro narušenou ekvivalenci položek musí existovat teoretické nebo empirické vysvětlení [Byrne 2010]. Robert, Lee a Chan [2006: 90-91] dodávají, že tyto nesrovnatelné položky mohou dále měřit latentní proměnnou, pokud neohrožují její konstruktovou validitu a obsah. Nebo jinak, je nutné rozumně zvážit vyřazení položek na základě jejich vychýlených parametrů, neboť to může mít nedozírné následky na vnitřní integritu škály měření a potažmo i teoretického konceptu.⁵⁹

⁵⁹ Například v testovaném modelu (viz Diagram 8, Příloha 2) není možné vyřadit ani jednu položku, protože už takto jsou latentní proměnné měřeny velmi úspěšně třemi až čtyřmi položkami. Jejich vyřazení by zcela jistě narušilo konstruktovou validitu modelu, který by navíc nebyl ani tzv. identifikovatelný.

Co se týče praktičtějších návodů, Steenkamp a Baumgartner [1998] uvádějí na příkladech, že v případě částečné metrické ekvivalence měření musí být zcela invariantní mezi skupinami jedna referenční proměnná⁶⁰, jejíž vazba se k latentní proměnné fixuje a faktorová zátěž je typicky pevně stanovena na velikost jedna, a alespoň ještě jedna další libovolná proměnná pro daný faktor. Byrne, Shavelson, Muthén [1989] doporučují trochu nepřesně, že některé parametry se mohou v modelu lišit. Vandenberg a Lance [2000] upřesňují, že neekvivalentní měřicí položky nesmí v daném faktoru převažovat a že pro testování ekvivalence více restriktivních modelů musí být splněny podmínky konfigurální a alespoň částečně metrické ekvivalence. Donahue [2006] a Steinmetz [2013] doplňují podle svých studií, že by měly být více jak dvě třetiny (ne-li téměř většina) položek ekvivalentní a že dopad částečné ekvivalence položky na její srovnatelnost silně závisí na velikosti testovaného souboru dat, velikosti modelu, velikosti faktorových zátěží a na velikosti rozdílu mezi konstantami.

Pokud se přistoupí na jakýkoli z těchto kompromisů a dále se pracuje s částečně ekvivalentním modelem, položky, jejichž parametry nejsou srovnatelné díky odlišné velikosti, jsou tzv. uvolněny z restrikce shody parametrů a velikosti jejich parametrů jsou „volně“ odhadnuty (tj. vypočítány). Tyto uvolněné parametry se v modelu ponechají a i ve více restriktivních modelech jsou jejich velikosti vždy volně odhadnuty. Lze si i při bližším zkoumání ověřit dopad takovýchto nesrovnatelných faktorových zátěží a konstant položek na výsledky dalších analýz. Neekvivalentní faktorové zátěže mají vliv především na odhad kovariancí částečně ekvivalentní latentní proměnné s ostatními latentními proměnnými, je tedy možné si porovnat rozdíl mezi velikostmi kovariancí faktorů v modelu plné metrické ekvivalence a částečné metrické ekvivalence [Robert, Lee, Chan 2006: 90]. V případě neekvivalentních konstant položek lze ověřit jejich dopad komparací rozdílů průměrů latentních proměnných měřených jednou s problematickými položkami, a jednou bez nich [více Chan 2000].

Po tom, co je identifikována metrická ekvivalence v datech, se nemusí dále pokračovat níže navrženým způsobem. Pokud je primárním cílem práce srovnávat velikost průměru latentních proměnných nebo manifestních proměnných, je nezbytné testovat po metrické ekvivalenci skalární ekvivalenci. Pokud je záměrem zjistit úroveň konstruktové validity měřicího nástroje nebo porovnávat vztahy mezi latentními proměnnými, může se přejít rovnou na testování shody variance a kovariance latentních proměnných a skalární ekvivalenci přeskočit [Schmitt, Kuljanin 2008: 211]. Z toho vyplývá, že následující předložené modely již nejsou hierarchicky seřazeny a lze z nich libovolně vybírat. Tuto hierarchickou posloupnost ilustrují Steenkamp a Baumgartner [1998: 83], Vandenberg a Lance [2000: 56] a nově Milfont a Fischer [2010: 116].

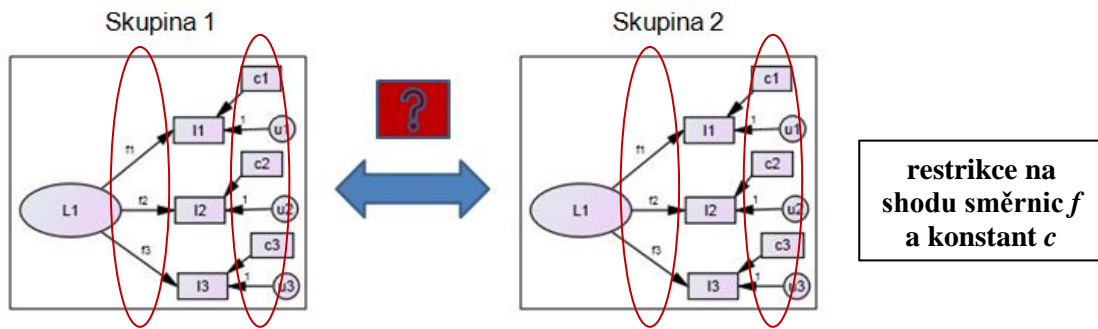
Skalární ekvivalence (Model 3)

Při testování skalární ekvivalence se v modelu ponechávají restrikce shody na všech faktorových zátěží, u kterých nebyl identifikován problém s ekvivalencí (všechny ostatní zůstávají uvolněné), a přidávají se požadavky na shodu všech konstant měřicích položek ve všech skupinách [Vandenburg, Lance 2000: 38]. Statisticky se jedná o testování shody konstant (neboli interceptů) c v regresních rovnicích ($I = c + f \cdot L + u$) ve všech skupinách, kterých je g (viz Diagram 4, konstanta c).

$$c_1 = c_2 = c_3 = c_4 = \dots = c_g \quad (3)$$

⁶⁰ Tato proměnná poskytuje jednotku škály měření pro latentní proměnnou, resp. určuje její varianci.

Diagram 4 Testování skalárního modelu při analýze skalární ekvivalence dat



Skalární ekvivalence je základním předpokladem pro testování rozdílu mezi průměry latentních proměnných (např. průměrná míra důvěry ve veřejné instituce) a také mezi průměry měřicích položek.⁶¹ Je tomu tak proto, že průměr měřicí položky, resp. latentní proměnné se vypočítá následujícím způsobem:

$$M(I) = c + f * M(L) \quad (4)$$

Kde $M(I)$ je průměr měřicí položky, c je intercept položky, f je faktorová zátěž položky a $M(L)$ je průměr latentní proměnné. Pokud je záměrem testovat rozdíly mezi průměry latentních proměnných (jednu neznámou veličinu), je nezbytné ověřit shodu faktorových zátěží a interceptů položky ve všech skupinách [Davidov 2012].

V případě plné skalární ekvivalence jsou skóry daných indikátorů mezi kulturami přímo statisticky porovnatelné [Hox, de Leeuw, Brinkhuis 2010: 400]. Lze tedy srovnávat jednotlivá hodnocení daných položek mezi skupinami a pozorované rozdíly mezi položkami pokládat za reálné odlišnosti daných skupin [Milfont, Fischer 2010: 115]. Tento typ ekvivalence ukazuje, že jedinci, kteří měli stejný skór (hodnocení) v latentní proměnné, také stejně skórovali v dané měřicí proměnné (udávali stejnou odpověď) bez ohledu na to, v jaké skupině se nacházejí [Milfont, Fischer 2010: 115]. Jinak řečeno, že rozdíly mezi průměry měřicích indikátorů napříč zeměmi jsou způsobeny rozdíly mezi průměry latentních proměnných a ne ničím jiným, viz vzorec 4 výše [Steenkamp, Baumgartner 1998: 80]. Z tohoto důvodu je nutná nejvyšší opatrnost při interpretaci průměrů manifestních i latentních proměnných, jestliže je v datech nalezena jen částečná skalární ekvivalence položek, neboť nestejně velké konstanty položek mohou reflektovat jak kulturně rozdílné názory, tak i systematickou chybu měření, která s postojem skupiny vůbec nesouvisí [Robert, Lee, Chan 2006: 92; více Steinmetz 2013].⁶²

Ekvivalence variancí chyb měření (Model 4)

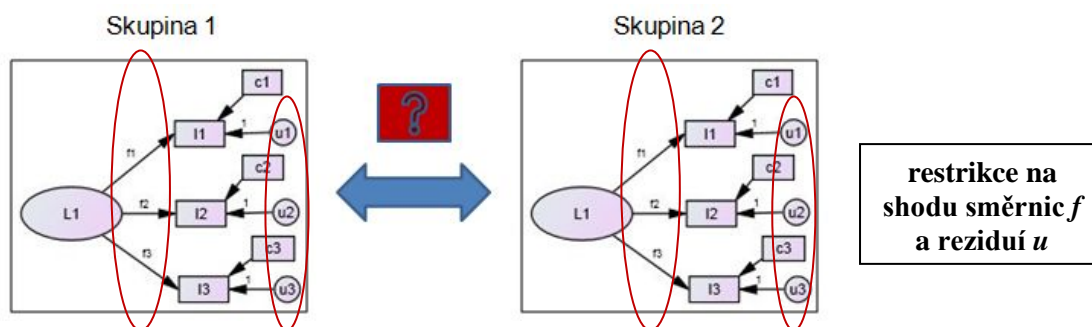
Pokud se testuje shoda velikostí chyb měření všech měřicích položek, k restrikcím ohledně faktorových zátěží (a případně dalším restrikcím) se připojí požadavek na shodu všech variancí chyb měření pro každou položku v každé skupině. Statisticky se jedná o testování shody reziduí (neboli chyb měření) u v regresních rovnicích ($I = c + f * L + u$) ve všech skupinách, kterých je g (viz Diagram 5, reziduum u).

$$u_1 = u_2 = u_3 = u_4 = \dots = u_g \quad (5)$$

⁶¹ U porovnávání průměrů měřicích položek je navíc nutnou podmínkou dosažení plné skalární ekvivalence, v případě komparace průměrů latentních proměnných je postačující částečná skalární ekvivalence.

⁶² Proto se někdy průměr přímo měřené položky považuje za funkci konstanty položky, faktorových zátěží a průměru latentní proměnné [Steinmetz 2013: 12].

Diagram 5 Testování modelu variancí chyb měření při analýze ekvivalence variancí chyb měření



Obecně je test shody chyb měření všech položek v jednotlivých skupinách považován za příliš restriktivní a běžně se při testování ekvivalence nepoužívá. Někdy je tato ekvivalence nazývána jako „striktní či přísná“ [Meredith 1993] a většina výzkumníků se shoduje na tom, že je velice náročné ji v měření dosáhnout. Tento test je vhodný jen tehdy, pokud stojí v centru pozornosti studie ověřování shody (mezi-)položkové reliability, kterou lze na základě velikosti chyby měření vypočítat [Byrne 2008; Schmitt, Kuljanin 2008; Vandenberg, Lance 2000; pro příklad tohoto testování Byrne 1988], anebo v případě, že není zájem dané položky použít v modelu, ale v jiných analýzách, kde nebude latentní proměnná oprostěna od variancí chyb měření [Robert, Lee, Chan 2006: 80].

Je také možné do základního modelu měření zahrnout vzájemnou korelaci chyb měření dvou proměnných.⁶³ I v tomto případě je možné testovat shodu korelací chyb měření daných položek ve všech skupinách. Tento test je poměrně běžný (není však nezbytný) a slouží k ověření shody modelů měření.

5.2 Analýza strukturní ekvivalence ve strukturním modelu

Při testování strukturní ekvivalence se pozornost soustředí výhradně na nepřímo pozorované latentní proměnné, a tudíž se dotýká strukturní části modelu. Zjišťování tohoto typu ekvivalencí dat je zcela libovolné a mělo by být dostatečně zdůvodněné. Nejčastěji se uvádí důležitost testování strukturní ekvivalence v případě zájmu o analýzu konstruktové validity postojových škál, tehdy jsou tyto testy modelů nezbytné [Byrne 2008: 874; Schmitt, Kuljanin 2008: 211]. Takto testované modely nemají hierarchickou strukturu a nemusí být testovány v uvedeném pořadí [Milfont, Fischer 2010: 115]. Lze si vybrat jeden z testů podle vlastních výzkumných záměrů a jeho výsledky porovnat se statistikami vhodnosti metrického modelu (viz Model 2). Nutnou podmínkou je pouze dosažení plné nebo částečné metrické ekvivalence v datech.

Ekvivalence variancí faktorů/latentních proměnných (Model 5)

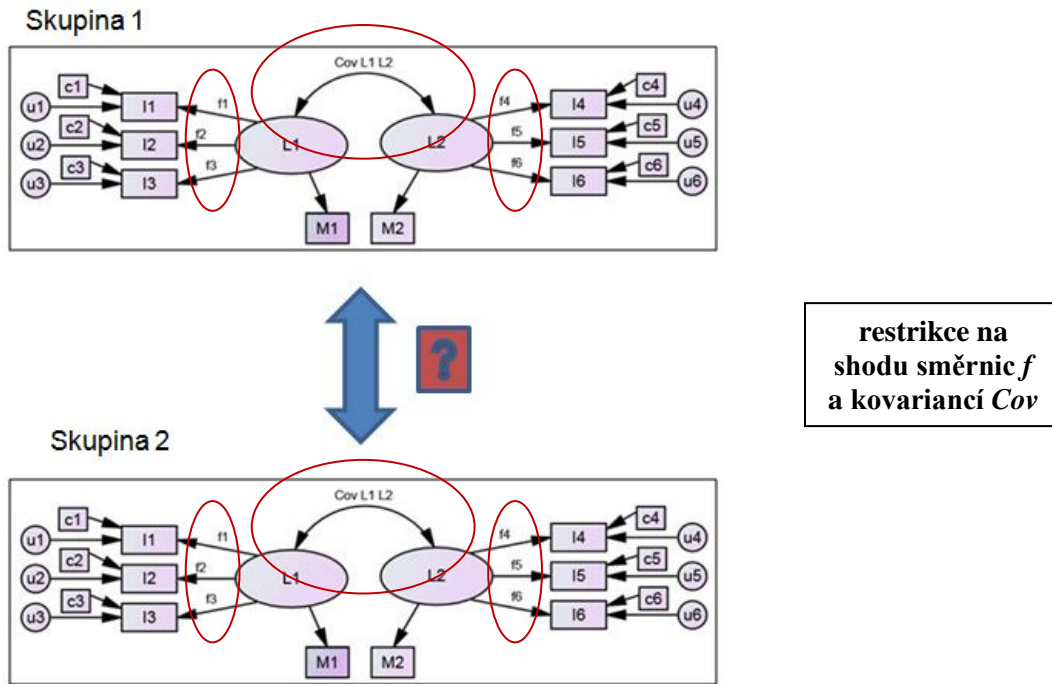
Tento model je testován přidáním požadavků na shodu variancí příslušných faktorů ve všech skupinách, které se připojují k restrikcím shody faktorových zátěží položek. Není to příliš běžný test. Ekvivalence variancí faktorů značí, že rozsah skóre latentních proměnných se mezi skupinami neliší [Milfont, Fischer 2010: 115]. V případě, že je tento typ ekvivalence narušen, ukazuje to na přítomnost metodologické systematické chyby měření (např. problémy s rozsahem škály, volba krajních kategorií odpovědí, atp.) [Robert, Lee, Chan 2006: 68].

⁶³ Ke korelaci chyb měření dochází v případě, kdy se dvě proměnné navzájem silně (obsahově a významově) překrývají, popř. mají střední až vysokou míru korelace. Tato podoba by měla vycházet ze sémantického výkladu položek a lze ji očekávat [Byrne 2004: 281].

Ekvivalence kovariancí faktorů (Model 6)

Tato ekvivalence je ze všech strukturních ekvivalencí analyzována nejčastěji. Je zjišťována přidáním požadavků na shodu všech kovariancí mezi faktory ve všech skupinách (viz Diagram 6, kovariance $Cov L1 L2$). Pokud je ekvivalence kovariancí faktorů potvrzena, vztahy mezi latentními proměnnými jsou srovnatelné ve všech zkoumaných skupinách a škála měří dimenzionální strukturu teoretického konceptu ekvivalentním způsobem, to vše je důkazem konstruktové validity nástroje měření [více Byrne, Watkins 2003].

Diagram 6 Testování modelu kovariancí faktorů při analýze strukturní ekvivalence dat

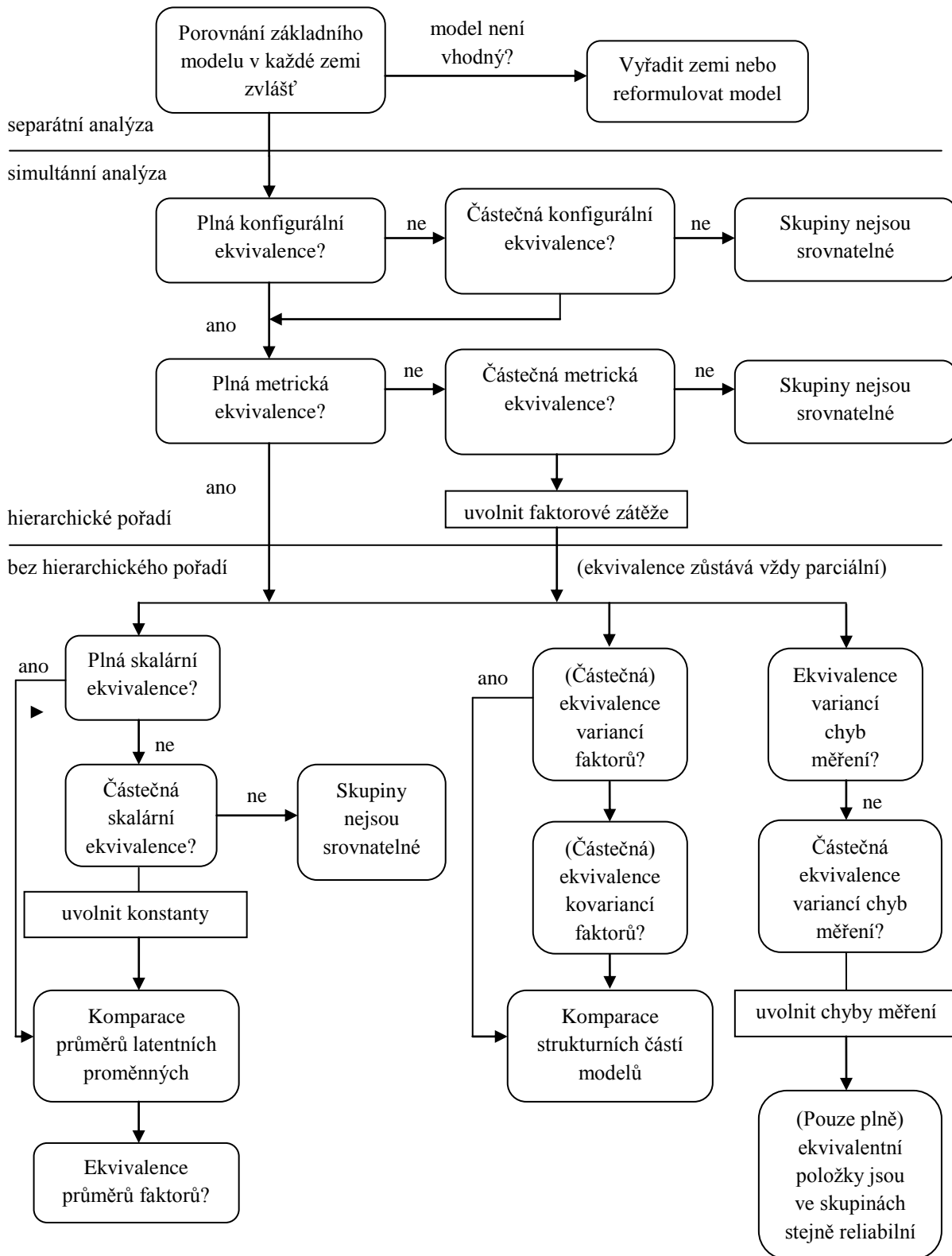


Ekvivalence průměrů faktorů (Model 7)

Tento poslední model lze považovat za zcela dobrovolný. Díky přidání restrikcí na shodu průměrů latentních proměnných je možné otestovat, zda se jednotlivé skupiny liší ve své pozici ve zkoumaném teoretickém konceptu (např. stejné míře mezilidské důvěry) [Milfont, Fischer 2010: 115]. Protože se většinou očekává, že skupiny se liší, přistupuje se rovnou ke konkrétnímu odhadu velikosti průměrů faktorů a ty se pak vzájemně porovnávají v jednotlivých skupinách. Toto však není možné, pokud není splněn požadavek alespoň částečné skalární ekvivalence dat; jinak není možné průměry vzájemně komparovat.

Hierarchickou posloupnost testování jednotlivých modelů, resp. druhů ekvivalence ilustruje Diagram 7, který je navržen na základě kompilace předloh podobných diagramů Steenkampa a Baumgartnera [1998: 83], Vandenberg a Lance [2000: 56] a nové revize Milfonta a Fischera [2010: 116]. Diagram ilustruje logiku testování navrženou B. Byrne od analýzy základního modelu v každé skupině zvlášť (single analysis) po simultánní analýzu restriktivních modelů ve všech skupinách najednou. Také přidává nový vizuální pohled na hierarchii testování. Je nutné dodržet posloupnost testování konfigurálního a metrického modelu a porovnávat je mezi sebou, ale pořadí testování dalších modelů je možné si libovolně vybrat podle vlastních výzkumných záměrů. Poslední upozornění se týká vlivu částečné metrické ekvivalence na všechny další druhy ekvivalencí – ty už budou také pouze parciální [viz Steenkamp, Baumgartner 1998: 83].

Diagram 7 Navržená posloupnost testování modelů při ověřování ekvivalence měření a strukturní ekvivalence



Zdroj: volně podle Steenkampa a Baumgartnera [1998: 83], Vandenberg a Lance [2000: 56] a Milfonta a Fischera [2010: 116]

6 ANALÝZA ZÁKLADNÍHO STRUKTURNÍHO MODELU POSTOJOVÝCH ŠKÁL

Pro ověření ekvivalence položek vybraných postojových škál prostřednictvím konfirmační faktorové analýzy, resp. strukturního modelování, bylo nezbytné sestavit model měření, kde by byly graficky (a v pozadí taktéž matematicky) naznačeny vztahy mezi měřicí položkou a latentní proměnnou, popř. vztahy mezi jednotlivými latentními proměnnými. Dimenzionalita daných teoretických konceptů byla již rozebírána, stejně tak výběr měřicích položek.

Model, který byl na základě teorie sestaven, v předběžné analýze ukázal, že některé zamýšlené položky, i když byl jejich výběr teoreticky opodstatněný, neměří dané faktory dostatečně dobře, a proto musely být tyto položky z modelu vyloučeny. Například se ukázalo, že mezi položkou důvěry v právní systém a důvěry v policii je příliš silná vazba (Pearsonovo $r > 0,6$), která narušuje celkové výsledky modelování, proto byla položka důvěry v policii z modelu vyřazena a zbyly tři položky měřící faktor institucionální důvěry POLIT, LEGAL, PARLAM. Dále musela být mezi položkami MEET a SOCACTIV vytvořena vazba, v modelu naznačena oboustrannou šipkou (viz Diagram 8, šipka mezi err3 a err4), která znamená, že jejich chyby měření (u) jsou vzájemně korelovány. Jak již bylo řečeno výše, tato korelace chyb měření nejčastěji ukazuje na obsahový překryv daných položek. Při pohledu na znění obou položek je očividné, že se obě dotazují na četnost společenského setkávání ve volném čase respondenta. Navíc modifikační indexy (M.I. = 24,703; Par Change = 0,239) předběžného základního modelu ukázaly, že pokud bude tato korelace do modelu zahrnuta, zlepší se celkové hodnocení modelu.⁶⁴

Model se tedy nakonec skládá z pěti faktorů, které jsou ve vzájemných vztazích a které měří celkem 16 vybraných položek (viz Diagram 8), což je i z hlediska principu úspornosti a smysluplnosti modelu zcela dostačující. Model je ještě stále přijatelně jednoduchý, takže i míra ohrožení podceněním velikosti statistik vhodnosti modelu není tak velká. Obecné pravidlo zní, že čím složitější a komplexnější model výzkumník vytvoří, tím složitější a nemožnější bude dokazování ekvivalence jednotlivých položek, neboť statistiky vhodnosti mají tendenci být nižší [Byrne 2008: 873]. Vytvořený model byl testován v programu AMOS (verze 21.0.) [Arbuckle 2012] ve 22 zemích (tzn. skupinách/výběrech) na datech z výzkumu ESS z první vlny šetření (2002).

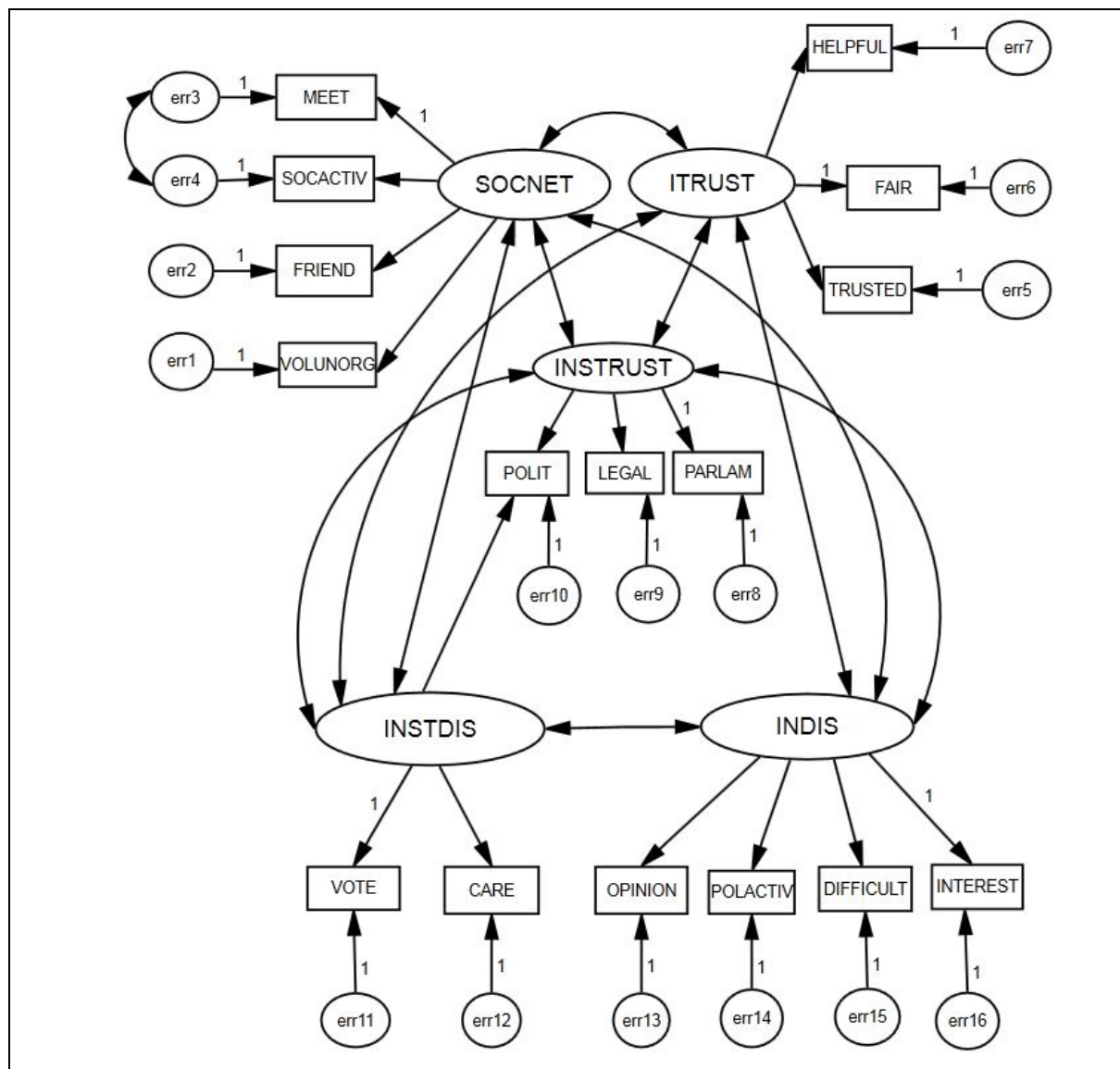
Problematika hodnocení modelů je značně široká a nelze ji zde komplexně zachytit. Její součástí je pečlivé zvážení všeho od teoretického podkladu modelu a stanovených hypotéz až po velikost a signifikanci jednotlivých odhadů parametrů, modifikačních indexů a statistik vhodnosti modelu. Tradičně se u jednotlivých modelů posuzuje to, zda teoretický model odpovídá pozorovaným datům z výzkumu v každé skupině, zda je reprezentuje adekvátně. Hodnotí se tak celková vhodnost základního modelu a všech dalších více restriktivních modelů obvykle na základě chí-kvadrát testů (tradiční přístup), které jsou ale příliš náchylné na velikost výběrového souboru. To je značný problém při zařazení velkého množství zemí do analýzy, protože i když model měří data velmi dobře, statistiky vhodnosti založené na chí-kvadrát testech mají tendenci tento model zamítnat [Hox, de Leeuw, Brinkhuis 2010: 397-398].

Z tohoto důvodu se dnes prosazují i jiné (méně striktní) statistiky vhodnosti modelu (např. χ^2/df , BIC, CFI, RMSEA) v tzv. pragmatickém či moderním přístupu, které zohledňují za prvé (ještě) akceptovatelnou jednoduchost modelu, chyby měření a také míru, do jaké testovaný model reprodukuje výchozí kovarianční matici [podrobněji o aspektech jednotlivých indexů Byrne 2010: 73–84; Hox, Bechger 1998: 8–10; Jöreskog, Sörbom 1993; Meade, Johnson, Braddy 2006; Urbánek 2000: kap. 9]. Právě tyto statistiky byly také

⁶⁴ Pro každý potenciálně zahrnutelný parametr lze zjistit, zda by se model jeho přidáním zlepšil – tuto informaci poskytují tzv. modifikační indexy [více Byrne 2010: 86-91].

zpravidla využívány při hodnocení modelů jednotlivých úrovní ekvivalencí postojových škál v této studii.

Diagram 8 Testovaný strukturální model postojových škál sociálního kapitálu a politického odcizení



Poznámka: Vysvětlení ke zkratkám položek a faktorů nabízí souhrnně Tabulka 3 v Příloze 1, chyby měření jsou v tomto modelu oproti teorii uvedené výše označeny místo písmenem *u* zkratkou *err* (jako error).

Jmenovitě byl vybrán k hodnocení vhodnosti modelů v této studii index RMSEA (Root mean square error of approximation), jehož velikost menší anebo rovna 0,05 ukazuje velmi dobré výsledky shody modelu s daty, od 0,06 do 0,08 středně vyhovující výsledky a nad hodnotu 0,1 nevyhovující výsledky [Byrne 2010: 80]. Pro posouzení byl také využit 90 % interval spolehlivosti pro index RMSEA a index PCLOSE, který musí být větší anebo roven hodnotě 0,5.⁶⁵ Jako další hodnotící index byl vybrán komparativní index shody CFI (Comparative fit index), jehož hodnoty větší anebo rovny 0,95 znamenají, že je model považován za dostatečně

⁶⁵ Hodnota PCLOSE funguje podobně jako *p* value pro udání signifikance odhadu velikosti indexu RMSEA. Pokud je jeho hodnota větší než 0,5, potvrzuje signifikanci odhadu velikosti indexu RMSEA, který je menší než 0,05.

vhodný, a hodnoty mezi 0,92–0,94 značí přiměřeně vhodný model.⁶⁶ Další indexy většinou potvrzující výsledky předešlých indexů jsou index dobré shody GFI (Goodness-of-Fit Index) a adjustovaný index dobré shody AGFI (Adjusted Goodness-of-Fit Index), pro jejichž ideální velikost platí stejná pravidla jako pro index CFI ($\geq 0,95$). Poslední Tucker-Lewisův index TLI (Tucker-Lewis Index) je opět známkou dobrého výsledku, jestliže se jeho hodnoty v případě velkého výběrového souboru rovnají anebo jsou větší než 0,95.

Tyto statistiky vhodnosti modelu jsou zpravidla používány při porovnávání hierarchicky seřazených modelů testujících ekvivalenci, které probíhá tak, že model s většími restrikcemi je srovnáván s modelem „pod ním“, který má méně restrikcí (např. Model 2 s Modelem 1, Model 3 s Modelem 2). Cheung a Rensvold [2002] pro tento případ komparace navrhli tři indexy, z nichž byl v této studii sledován především CFI index. Logika porovnávání modelu tkví v tom, že pokud je změna velikosti tohoto indexu (Δ CFI) větší než 0,01, více restriktivní model musí být zamítnut, protože nedostatečně reprezentuje data, a daná ekvivalence není potvrzena. Také je možné posuzovat hodnotu indexu RMSEA více restriktivního modelu, která musí spadat do intervalu spolehlivosti indexu RMSEA předchozího méně restriktivního modelu. Cieciuch a Davidov [2012: 39] uvádí konkrétní velikost změny (Δ RMSEA), která nesmí u indexu RMSEA přesáhnout hodnotu 0,015.⁶⁷

6.1 Základní předpoklady testování strukturního modelu

Při konstrukci modelu měření či strukturního modelu bylo nezbytné sledovat a dodržet několik základních pravidel. Kromě několika principů již zmíněných (úspornost modelu, podklad v teorii) se jedná například o potřebu identifikovatelnosti (určenosti) modelu [podrobněji Byrne 2010: 33–35; Urbánek 2000: kap. 4].⁶⁸ Existuje řada dalších pravidel, které je třeba dodržovat, aby bylo možné model odhadnout či testovat, ale za nejzákladnější se považuje to, že každá latentní proměnná by měla být měřena alespoň třemi indikátory; tuto podmínku základní model splňuje. Dále je nutné „zafixovat“ vazbu jednoho indikátoru s danou latentní proměnnou typicky na jednotku a pomocí toho určit rozptyl latentní proměnné. Urbánek [2000: 73] vysvětluje, že pokud je hodnota rozptylu latentní proměnné rovna jedné, splní se podmínka, která stanovuje škálu latentní proměnné, neboť ta jinak žádnou přirozenou škálu nemá.

Diskuzí ohledně volby této referenční položky je řada [např. Cheung, Rensvold 2002; Vandenberg 2002]. Nejčastěji se doporučuje provést explorační faktorovou analýzu a tu položku, která má v daném faktoru největší faktorovou zátěž, zvolit za referenční položku. Vandenberg [2002] navrhuje provést explorační faktorovou analýzu dokonce v každé zkoumané skupině a nalézt položky, jejichž hodnoty faktorových zátěží na daném faktoru nejméně variiují mezi zeměmi. Byrne [2010] doporučuje zvolit jako referenční položku i tu s nejvyšší reliabilitou. Více méně je ale volba referenční položka na volbě výzkumníka, neboť

⁶⁶ Ve starších textech najdeme dokonce akceptovatelnost hodnoty $CFI \geq 0,90$.

⁶⁷ V současnosti existují i modernější přístupy k posuzování vhodnosti restriktivních modelů testujících ekvivalenci. William van der Veld vymyslel software Jrule (pro program LISREL, Daniel Oberski pro Mplus), který má ohodnotit pouze smysluplné (tedy vážné) odchylky od invariantního měření. Jrule je vyvinut pro určení modifikačních indexů a standardizovaných reziduí a je novou alternativou k obecnému hodnocení modelu na základě indexů. Tento program bere v úvahu a vypočítává sílu testu odhadu parametrů, na něž je dána restrikce. Jeho nevýhodou je výpočet pouze síly tohoto testu parametrů, ale ne síly statistik vhodnosti modelu. Program byl již využit při studii Sarise, Satorra a van der Velda (2009) nebo van der Velda a Sarise (2009) a je k dispozici jako freeware verze na osobní požádání doktora van der Velda (kontakt: w.vanderveld@bsi.ru.nl) [e-mailová komunikace s Williamem van der Veldem, stručný popis van der Veld, Saris (2009)].

⁶⁸ Model musí být identifikovatelný, tj. minimálně musí platit, že počet odhadovaných parametrů není větší než počet vstupních momentů kovarianční matice [vice Byrne 2010: 33; Matějů 1989].

ještě stále není jasné, jaký dopad má volba dané položky na výsledek provedených analýz [Schmitt, Kuljanin 2008: 216].

V základním modelu je jako referenční položka zvolena na faktoru SOCNET (kolektivní sociální kapitál) proměnná *MEET*, na faktoru ITRUST (individuální sociální kapitál) proměnná *FAIR*, na faktoru INSTRUST (důvěra v instituce) proměnná *PARLAM*, na faktoru INDIS (individuální politické odcizení) proměnná *INTEREST* a na faktoru INSTDIS (institucionální politické odcizení) proměnná *CARE*. Jejich volba vycházela především z analýzy reliability; byly vzaty v úvahu průměrné koeficienty realibility (Cronbachovo alfa), variabilita velikostí koeficientů a také výsledky explorační faktorové analýzy.

Kromě identifikovatelnosti modelu musela být věnována pozornost také dalším dvěma problémům a to chybějícím položkám (missing values) a normalitě dat. Nekompletní data mohou způsobovat vážné systematické odchylky ve statistických závěrech, pokud není jejich výskyt náhodný. Je třeba je neignorovat a kontrolovat a v některých případech se pokusit zjistit příčinu, proč některá data chybí. Chybějící hodnoty jsou ve strukturním modelování překážka a datový soubor, se kterým se při modelování pracuje, musí být ve většině statistických softwarů bez chybějících hodnot (systémových i uživatelsky definovaných).⁶⁹ Byrne [2010: 355-359] hovoří o několika nejčastějších způsobech, jak se s nekompletními daty vypořádat: vyloučení neúplně popsanych objektů (listwise/casewise deletion), vyloučení pouze neúplných párů (pairwise deletion) a imputace dat, avšak žádné z nich nejsou bez nevýhod.

Neexistuje jasně daný milník, od jaké hodnoty se potýkáme už s příliš velkým množstvím chybějících hodnot. V některých publikacích se dočteme, že by množství chybějících hodnot nemělo přesáhnout 10 % dat, resp. velikosti výběru [Kline 1998: 75; obecně o problematice chybějících hodnot Little, Rubin 1987]. Analyzovaný agregovaný datový soubor ESS z roku 2002 obsahuje čtyři proměnné, u nichž se podíl chybějících hodnot pohybuje od dvou do čtyř procent; v případě většiny ostatních proměnných se tento podíl pohybuje do dvou procent (viz Tabulka 18).

Tabulka 18 Podíl chybějících hodnot v jednotlivých testovaných proměnných v agregovaném souboru dat

proměnné	chybějící data (respondenti)	chybějící hodnoty (%)
MEET	105	0,25
SOCACTIV	827	1,95
FRIEND	187	1,44
VOLUNORG	702	1,66
HELPFUL	271	0,64
FAIR	409	0,97
TRUSTED	193	0,46
PARLAM	1582	3,73
LEGAL	1207	2,85
POLIT	914	2,16
VOTE	760	1,79
CARE	879	2,08
OPINION	799	1,89
POLACTIV	439	1,04
DIFFICULT	718	1,70
INTEREST	148	0,35

Zdroj: ESS 2002; vlastní výpočty

⁶⁹ V programu LISREL je možné chybějící hodnoty označit nějakým specifickým číslem jako v programu SPSS, v programu AMOS musí být chybějící hodnoty z datového souboru úplně odstraněny nebo nahrazeny nějakým jiným vhodným statistickým způsobem.

Kumulativně nicméně dochází k nárůstu počtu chybějících hodnot až na 13 % v agregovaném datovém souboru, tedy dohromady ve všech zkoumaných zemích. V jednotlivých zemích pak podíl chybějících hodnot osciluje mezi 1,9 % až 35 %. Do rozmezí deseti procent se nachází osm zemí, od 10 do 20 % se pohybuje 12 zemí a 20 % chybějících hodnot přesahují tři země (viz Tabulka 19).

Tabulka 19 Kontrola chybějících dat (missing values) a přehled počtu respondentů v jednotlivých zemích

země	počet respondentů	počet respondentů po vyloučení chybějících hodnot	chybějící data (respondenti)	chybějící hodnoty (%)
Rakousko	2257	1895	362	16,04
Belgie	1899	1636	263	13,85
Česká republika	1360	1128	232	17,06
Německo	2919	2744	175	6,00
Dánsko	1506	1360	146	9,69
Španělsko	1729	1320	409	23,66
Finsko	2000	1930	70	3,50
Francie	1503	1419	84	5,59
Spojené království	2052	1959	93	4,53
Řecko	2566	2258	308	12,00
Maďarsko	1685	1368	317	18,81
Švýcarsko	2040	1787	253	12,40
Irsko	2046	1780	266	13,00
Izrael	2499	2079	420	16,81
Itálie	1207	1055	152	12,59
Lucembursko	1552	1004	548	35,31
Nizozemí	2364	2191	173	7,32
Norsko	2036	1997	39	1,92
Polsko	2110	1734	376	17,82
Portugalsko	1511	1168	343	22,70
Švédsko	1999	1825	174	8,70
Slovinsko	1519	1307	212	13,96
Všechny země	42359	36944	5415	12,78

Zdroj: ESS 2002; vlastní výpočty

V současnosti se podařilo zdokonalit i několik metodologických přístupů k analýze nekompletních dat založených na modelování, např. použít metody odhadu velikosti parametrů FIML (metoda maximální věrohodnosti s úplnou informací) anebo EM (metoda očekávané maximalizace), obě založené na principu metody maximální věrohodnosti (ML) [Byrne 2010: 358]. Ty jsou dnes považovány za nejúčinnější řešení výskytu chybějících hodnot v datovém souboru [viz Arbuckle 1996]. Metodu odhadu velikosti parametrů FIML lze použít v programu AMOS přímo na data s chybějícími hodnotami a Byrne [2010: 365] ukazuje, že vykazuje velmi dobré výsledky i na datech s 25 % chybějících dat.

Prezentované výsledky kompletnosti dat nebyly nijak ideální, obzvláště v případě Lucemburska, Španělska a Portugalska, proto byla dále provedena předběžná analýza základního modelu, kde 1) byly chybějící hodnoty z datového souboru vyloučeny (za použití metody listwise deleted) a jako metoda odhadu parametrů modelu byla zvolena metoda maximální věrohodnosti (ML) a 2) chybějící hodnoty zůstaly v analyzovaném datovém souboru a jako metoda odhadu parametrů byla zvolena metoda maximální věrohodnosti s úplnou informací (FIML). Výsledky hodnocení vhodnosti modelu těchto dvou metod odhadu byly vzájemně porovnány, viz Tabulka 20a, 20b, 20c.

Tabulka 20a Separátní analýza základního modelu - srovnání výsledků tří metod odhadu parametrů modelu ML, FIML a DWLS

země	metoda odhadu	χ^2	stupně volnosti df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko	ML	315,87	92	0,979	0,969	0,969	0,036	(0,032;0,040)	1,000
	FIML	389,12	92			0,966	0,038	(0,034;0,042)	1,000
	DWLS	297,10	92	0,990	0,990	0,990	0,034	(0,030;0,039)	1,000
Belgie	ML	285,78	92	0,978	0,968	0,968	0,036	(0,031;0,041)	1,000
	FIML	326,23	92			0,965	0,037	(0,032;0,041)	1,000
	DWLS	254,58	92	0,990	0,990	0,990	0,033	(0,028;0,038)	1,000
Česká republika	ML	251,87	92	0,971	0,958	0,966	0,039	(0,036;0,043)	1,000
	FIML	297,51	92			0,963	0,041	(0,035;0,046)	0,999
	DWLS	244,43	92	0,990	0,980	0,990	0,038	(0,033;0,044)	1,000
Německo	ML	479,80	92	0,978	0,968	0,960	0,039	(0,036;0,043)	1,000
	FIML	502,20	92			0,961	0,039	(0,036;0,042)	1,000
	DWLS	477,78	92	0,990	0,980	0,980	0,039	(0,036;0,043)	1,000
Dánsko	ML	353,89	92	0,968	0,953	0,951	0,046	(0,041;0,051)	0,913
	FIML	373,39	92			0,952	0,045	(0,041;0,050)	0,954
	DWLS	334,70	92	0,980	0,980	0,980	0,044	(0,039;0,049)	0,970
Španělsko	ML	358,29	92	0,967	0,951	0,948	0,047	(0,042;0,052)	0,839
	FIML	422,76	92			0,951	0,046	(0,041;0,050)	0,948
	DWLS	324,47	92	0,980	0,980	0,970	0,044	(0,039;0,049)	0,980
Finsko	ML	565,62	92	0,963	0,945	0,938	0,052	(0,048;0,056)	0,246
	FIML	588,51	92			0,937	0,052	(0,048;0,056)	0,205
	DWLS	539,15	92	0,980	0,970	0,970	0,050	(0,046;0,054)	0,460
Francie	ML	353,89	92	0,970	0,955	0,944	0,045	(0,040;0,050)	0,957
	FIML	365,89	92			0,944	0,045	(0,040;0,049)	0,969
	DWLS	303,96	92	0,990	0,980	0,980	0,040	(0,035;0,045)	1,000
Spojené království	ML	468,71	92	0,970	0,956	0,951	0,046	(0,042;0,050)	0,955
	FIML	477,48	92			0,950	0,046	(0,042;0,050)	0,963
	DWLS	444,17	92	0,990	0,980	0,970	0,044	(0,040;0,048)	0,990

Zdroj: ESS 2002; vlastní výpočty

Poznámka: V případě metody DWLS se jedná o Satorra-Bentler Scaled χ^2 . Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (přip. CFI \geq 0,92), RMSEA \leq 0,05 (přip. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Tabulka 20b Separátní analýza základního modelu (statistiky vhodnosti modelu) - srovnání výsledků tří metod odhadu parametrů modelu ML; FIML; DWLS

země	metoda odhadu	χ^2	stupně volnosti df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Řecko	ML	401,41	92	0,978	0,968	0,967	0,039	(0,035;0,043)	1,000
	FIML	446,84	92			0,967	0,039	(0,035;0,042)	1,000
	DWLS	387,75	92	0,990	0,980	0,980	0,038	(0,034;0,042)	1,000
Maďarsko	ML	265,34	92	0,976	0,965	0,968	0,037	(0,032;0,042)	1,000
	FIML	299,33	92			0,970	0,037	(0,032;0,041)	1,000
	DWLS	272,50	92	0,990	0,980	0,980	0,038	(0,033;0,043)	1,000
Švýcarsko	ML	369,91	92	0,974	0,962	0,956	0,041	(0,037;0,046)	1,000
	FIML	395,80	92			0,957	0,040	(0,036;0,044)	1,000
	DWLS	361,95	92	0,990	0,990	0,990	0,041	(0,036;0,045)	1,000
Irsko	ML	363,07	92	0,975	0,964	0,959	0,041	(0,036;0,045)	1,000
	FIML	372,08	92			0,962	0,039	(0,035;0,043)	1,000
	DWLS	334,97	92	0,990	0,980	0,980	0,039	(0,034;0,043)	1,000
Izrael	ML	333,50	92	0,980	0,970	0,958	0,036	(0,031;0,040)	1,000
	FIML	351,01	92			0,962	0,034	(0,030;0,037)	1,000
	DWLS	328,97	92	0,990	0,990	0,980	0,035	(0,031;0,039)	1,000
Itálie	ML	216,76	92	0,975	0,963	0,967	0,036	(0,030;0,042)	1,000
	FIML	228,15	92			0,969	0,035	(0,029;0,041)	1,000
	DWLS	211,13	92	0,990	0,980	0,990	0,035	(0,029;0,041)	1,000
Lucembursko	ML	205,66	92	0,974	0,962	0,961	0,035	(0,029;0,042)	1,000
	FIML	318,96	92			0,947	0,040	(0,035;0,045)	1,000
	DWLS	217,15	92	0,990	0,980	0,990	0,037	(0,031;0,043)	1,000

Zdroj: ESS 2002; vlastní výpočty

Poznámka: V případě metody DWLS se jedná o Satorra-Bentler Scaled χ^2 . Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Tabulka 20c Separátní analýza základního modelu (statistiky vhodnosti modelu) - srovnání výsledků tří metod odhadu parametrů modelu ML; FIML; DWLS

země	metoda odhadu	χ^2	stupně volnosti df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Nizozemí	ML	498,60	92	0,972	0,958	0,955	0,045	(0,041;0,049)	0,984
	FIML	530,95	92			0,954	0,045	(0,041;0,049)	0,987
	DWLS	409,34	92	0,990	0,980	0,980	0,040	(0,036;0,044)	1,000
Norsko	ML	520,70	92	0,969	0,954	0,935	0,048	(0,044;0,052)	0,747
	FIML	549,71	92			0,932	0,049	(0,045;0,053)	0,608
	DWLS	462,58	92	0,980	0,980	0,970	0,045	(0,041;0,049)	0,980
Polsko	ML	320,72	92	0,977	0,967	0,955	0,038	(0,033;0,042)	1,000
	FIML	356,11	92			0,960	0,037	(0,033;0,041)	1,000
	DWLS	288,00	92	0,990	0,980	0,980	0,035	(0,031;0,040)	1,000
Portugalsko	ML	403,66	92	0,958	0,938	0,926	0,054	(0,049;0,059)	0,114
	FIML	448,36	92			0,935	0,051	(0,046;0,055)	0,401
	DWLS	376,68	92	0,980	0,970	0,970	0,051	(0,046;0,057)	0,320
Švédsko	ML	513,56	92	0,964	0,947	0,939	0,050	(0,046;0,054)	0,472
	FIML	543,03	92			0,939	0,050	(0,046;0,054)	0,567
	DWLS	518,42	92	0,980	0,980	0,970	0,050	(0,046;0,055)	0,430
Slovinsko	ML	324,97	92	0,969	0,954	0,948	0,044	(0,039;0,049)	0,970
	FIML	344,57	92			0,949	0,043	(0,038;0,047)	0,995
	DWLS	311,29	92	0,990	0,980	0,980	0,043	(0,038;0,048)	0,990

Zdroj: ESS 2002; vlastní výpočty

Poznámka: V případě metody DWLS se jedná o Satorra-Bentler Scaled χ^2 . Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Z tabulek vyplývá, že metoda maximální věrohodnosti (ML) a metoda maximální věrohodnosti s úplnou informací (FIML) produkují velmi podobné výsledky a to i v případě zemí s větší mírou nekompletnosti dat, kde chybějící hodnoty (resp. případy) přesahují 20 % velikosti výběru (viz Lucembursko, Španělsko a Portugalsko). Odchytky v hodnotách statistik vhodnosti (CFI, RMSEA a 90 % interval spolehlivosti RMSEA) byly vždy až na úrovni druhého a třetího řádu čísla a i tak indexy pokaždé generovaly stejné závěry o vhodnosti modelu. Odhady standardizovaných faktorových zátěží byly více méně zcela totožné (viz Tabulka 21). Z tohoto je možné usoudit, že lze dále analyzovat data bez chybějících hodnot metodou odhadu maximální věrohodnosti (ML) bez větších obav.

Tabulka 21 Separátní analýza základního modelu realizovaná na agregovaném datovém souboru (odhad standardizovaných faktorových zátěží) - srovnání výsledků tří metod odhadu parametrů modelu ML; FIML; DWLS

Latentní proměnná	Proměnné	Metody odhadu		
		ML	FIML	DWLS
kolektivní sociální kapitál	MEET	0,46	0,47	0,44
	SOCACTIV	0,49	0,49	0,59
	FRIEND	0,47	0,49	0,41
	VOLUNORG	0,32	0,33	0,35
individuální sociální kapitál	HELPFUL	0,65	0,65	0,66
	FAIR	0,76	0,76	0,75
	TRUSTED	0,76	0,76	0,79
důvěra v instituce	PARLAM	0,86	0,86	0,87
	LEGAL	0,65	0,65	0,66
	POLIT	0,63	0,63	0,61
institucionální politické odcizení	VOTE	0,76	0,76	0,80
	CARE	0,79	0,79	0,84
	POLIT	0,27	0,27	0,29
individuální politické odcizení	OPINION	0,60	0,62	0,58
	POLACTIV	0,57	0,58	0,66
	DIFFICULT	0,56	0,57	0,57
	INTEREST	0,68	0,70	0,76

Zdroj: ESS 2002, spojený datový soubor 22 zemí; vlastní výpočty

Poznámka: Všechny parametry jsou statisticky významné na hladině 0,001 anebo více.

Dalším úskalím byla při testování modelu normalita dat. Na začátku analýzy bylo nutné zhodnotit, jaký typ proměnných se bude v modelu analyzovat, tzn. ordinální či kardinální proměnné. Běžným standardem je zařazovat do modelu kardinální proměnné a pro odhad parametrů modelu použít metodu maximální věrohodnosti (ML). V tomto případě je ale nezbytné splnit několik požadavků: a) vícerozměrné normální rozdělení dat⁷⁰ a b) dostatečně velký vzorek [Byrne 2010; Steenkamp, Baumgartner 1998; Vandenberg, Lance 2000].

Velmi často se ale v sociálních vědách pracuje s tzv. ordinálními proměnnými (např. několikabodové Lickertovy škály), která pak značí práci s kategoriálními daty. S těmito proměnnými se za určitých okolností může dále zacházet jako s kardinálními proměnnými, ale pokud je záměrem při odhadu parametrů modelu dále aplikovat metodu maximální věrohodnosti na korelační nebo kovarianční matici dat, je nutné podle Byrne [2010: 148] sledovat a) zda má proměnná dostatečný počet kategorií (tzn. tradičně více než čtyři či pět), b) normalitu dat a c) velikost výběru. Byrne shrnuje výsledky řady výzkumů a podotýká, že nenormalita dat je daleko závažnějším problémem než počet kategorií ordinální proměnné [potvrzuje také Gao, Mokhtarian, Johnston 2008; DiStefano 2002; Míndrilá 2010; Muthén a

⁷⁰ Je nutné tedy ověřit normální rozdělení každé položky zvlášť a poté normální rozdělení všech položek dohromady (tzn. vícerozměrné normální rozdělení).

Kaplan 1985] a že pokud budou mít data relativně normální rozdělení, stačí, aby měly položky čtyři a více kategorií, a lze použít metodu maximální věrohodnosti [Bentler, Chou 1987: 88].

Mnoho statistiků namítá, že takto postupovat nelze a že ordinální proměnné nespĺňují parametry spojitého měření [podrobněji Jöreskog 2002]. Proto není podle nich možné dále strukturní modelování zakládat na korelační matici, ale na jiné například polychoriální nebo polyseriální korelační matici [více Uebersax 2006a, 2006b].⁷¹ Také je podle některých statistiků vhodné použít jinou metodu odhadu parametrů. Jako nejlepší se ukazuje metoda diagonálně vážených nejmenších čtverců (DWLS) anebo její různé modifikace (např. RDWLS, WLSMV) [Flora, Curran 2004; Jöreskog 1990; Míndrilá 2010; Ollson 1979; Temme 2006]. Tyto metody odhadu není možné použít v programu AMOS, ale jen v programu LISREL anebo Mplus.

Zdá se, že v sociálních vědách je běžnou praxí následovat spíše doporučení první a kontrolovat počet kategorií, normalitu ordinálních dat a velikost výběru [např. Billiet, Welkenhuysen-Gybels 2004; Byrne 2008; Cieciuch, Davidov 2012; Meuleman, Davidov, Billiet 2009, atp.]. Počet respondentů (i počet respondentů po vyloučení chybějících hodnot) v analyzovaných 22 zemích se v testovaném základním modelu zdá být zcela dostatečný (tj. v každé zemi přesahuje 1 000 respondentů, což je obecně považováno za velký vzorek, viz Tabulka 20). V modelu mělo šest proměnných 5 kategorií odpovědí, jedna proměnná pouze 4 kategorie, jedna proměnná 7 kategorií a osm proměnných 11 kategorií (viz Tabulka 3, Příloha 1); tedy pouze u jedné proměnné by hrozilo riziko malého množství kategorií a i tento počet je těsně na hraně.

Při kontrole normality dat obecně platí, že neexistuje žádný konsensus ohledně akceptovatelné míry nenormality, ale lze v literatuře nalézt, že nejčastěji se jako limit udává hodnota 2 pro velikost šikmosti a hodnota 7 pro velikost špičatosti rozdělení jedné proměnné v modelu. Pokud šikmost a špičatost rozdělení dané proměnné nepřekročí tyto hodnoty, nebude ještě příliš odchýlené od normálního rozdělení [viz Byrne 2010: 103; Míndrilá 2010: 61]. Špičatost se navíc ve strukturním modelování považuje za větší problém než šikmost, proto se z hlediska vícerozměrného rozdělení proměnných sleduje vícerozměrná špičatost. Koeficient vícerozměrné špičatosti by neměl přesáhnout hodnotu 5 [Bentler 2005; Byrne 2010: 104], i když Gao, Mokhtarian, Johnston [2008] ukazují výzkumy, kde zjištěná nenormalita vícerozměrného rozdělení (koeficient vícerozměrné špičatosti dosahoval hodnoty až 40) zkreslovala pouze přibližně pět procent odhadovaných standardizovaných reziduí.

Výsledky analýzy jednorozměrného rozdělení testovaných položek ukázaly, že absolutní hodnota šikmosti proměnných nepřesáhla velikost dva (kromě Irska), navíc tuto odchylku způsobovala ve všech zemích jen jedna proměnná FRIEND. Všechny ostatní proměnné nepřesáhly z hlediska šikmosti ani hodnotu jedna (viz Tabulka 22). V případě špičatosti se objevily shodné výsledky – všechny proměnné kromě FRIEND nepřesáhly hodnotu jedna a ani špičatost rozdělení proměnné FRIEND nepřesáhla limit hodnoty sedm v žádné zemi. Z pohledu jednorozměrného rozdělení tedy lze konstatovat, že rozdělení jednotlivých proměnných je normální i přes jejich ordinální charakter. Normalita vícerozměrného rozdělení byla již více ohrožena, žádná hodnota koeficientu vícerozměrné špičatosti se nenacházela pod limitní hodnotou pět. Avšak téměř ve všech zemích daný koeficient nepřesáhl hodnotu 40 a ve všech zemích se jedná o kladnou vícerozměrnou špičatost, proto lze očekávat jen mírné zkreslení odhadu parametrů v modelu.

⁷¹ To jsou typy korelací mezi ordinálními proměnnými, resp. ordinálními a spojitými proměnnými. Mohou to být i tetrachorické korelace, které jsou mezi dvěma dichotomickými proměnnými.

Všechny předložené argumenty (velikost výběru, počet kategorií i normalita dat) potvrdily, že lze přistoupit v analýze k vybraným ordinálním proměnným jako ke kardinálním proměnným a aplikovat metodu maximální věrohodnosti na odhad parametrů modelu. Protože ale statistici radí se k danému problému neobracet zády, byla provedena předběžná separátní analýza základního modelu dvěma způsoby – 1) byla analyzována kovarianční matice dat odhadní metodou maximální věrohodnosti (ML) v programu AMOS ve všech zemích a 2) byla analyzována polychorická korelační matice dat odhadní metodou diagonálně vážených nejmenších čtverců (DWLS) v programu LISREL také ve všech zemích (viz příklad syntaxe v Příloze 3).

Tabulka 22 Kontrola normality dat – údaje šikmosti a špičatosti rozdělení proměnných v daných zemích, koeficient vícerozměrné špičatosti

země	Interval šikmosti	Interval špičatosti	Koeficient vícerozměrné špičatosti
Rakousko	(0,111; -1,989)	(-0,012; 4,970)	37,130
Belgie	(0,014; -1,573)	(0,025; 3,823)	35,714
Česká republika	(-0,042; -1,643)	(-0,055; 3,215)	24,993
Německo	(0,023; -1,593)	(-0,077; 3,639)	41,268
Dánsko	(0,006; -1,550)	(0,026; 3,832)	36,314
Španělsko	(0,006; -1,488)	(-0,068; 3,131)	32,812
Finsko	(0,039; -1,790)	(-0,006; 5,268)	39,153
Francie	(-0,016; -1,192)	(-0,124; 2,961)	35,533
Spojené království	(0,028; -1,528)	(0,095; 3,503)	34,108
Řecko	(0,018; -1,729)	(-0,094; 3,892)	34,687
Maďarsko	(-0,017; -1,195)	(0,120; -1,114)	17,014
Švýcarsko	(0,053; -1,357)	(-0,063; 3,029)	33,421
Irsko	(0,040; -2,130)	(-0,266; 6,858)	36,028
Izrael	(0,015; -1,825)	(0,041; 3,988)	29,771
Itálie	(-0,015; -0,934)	(-0,092; 1,140)	19,894
Lucembursko	(-0,008; -1,799)	(0,008; 3,633)	20,173
Nizozemí	(-0,032; -1,864)	(0,003; 5,997)	50,769
Norsko	(0,050; -1,402)	(0,008; 3,315)	43,860
Polsko	(-0,003; -1,234)	(-0,016; 1,759)	27,507
Portugalsko	(-0,036; -1,344)	(-0,010; 2,524)	22,617
Švédsko	(0,043; -1,296)	(-0,075; 2,530)	33,517
Slovinsko	(0,006; -1,488)	(-0,075; 2,698)	30,178

Zdroj: ESS 2002; vlastní výpočty (program AMOS)

Poznámka: Interval šikmosti či špičatosti znamená údaj nejmenší a největší šikmosti/špičatosti, které se mezi 16 položkami v dané zemi vyskytovaly. Koeficient vícerozměrné špičatosti je ve skutečnosti „Mardia’s normalized coefficient, jinak „Mardia’s multivariate kurtosis“.

Výsledky těchto dvou typů analýz byly dále mezi sebou porovnávány. Opět se ukázalo, že jsou si velice podobné (viz Tabulky 20). Odchyly v hodnotách statistik vhodnosti (CFI, RMSEA a 90 % interval spolehlivosti RMSEA) byly opakovaně až na úrovni druhého a třetího řádu čísla a indexy vyvozovaly shodné závěry o vhodnosti modelu; konkrétně velikosti indexů RMSEA se lišily jen o setiny. Zdálo se, že v některých případech má metoda maximální věrohodnosti tendenci vhodnost modelu podceňovat (např. u Portugalska, Švédska, Norska anebo Finska), index CFI u metody ML ukazoval na středně vhodný model, zatímco u metody DWLS na velmi dobrý model. Rozdíly ale nebyly nijak dramatické.

Odchyly v odhadech velikosti standardizovaných faktorových zátěží byly také minimální, jak ukazuje Tabulka 22. Pořadí proměnných podle velikosti faktorových zátěží na daném faktoru zůstává nezměněno. Největší výchylku ukazuje položka SOCACTIV. Závěrem lze konstatovat, že i srovnání výsledků analýz dvou odlišných metod odhadu parametrů (ML a DWLS) neukázalo, že je nutné kvůli ordinálnímu charakteru proměnných a mírně vychýlené

vícerozměrné špičatosti dat dále v analýze použít jinou odhadní metodu než maximální věrohodnost (ML).

6.2 Systematické chyby měření položek z postojových škál

Jak již bylo částečně řečeno, na základě teorie i výsledků již provedených výzkumů lze předpokládat, že některé testované položky vybraných postojových škál budou zatížené chybou měření, což by se také mělo projevit v testech modelů jednotlivých úrovní ekvivalence škál. Faktorů, které mohou ohrozit srovnatelnost postojových škál a jejich položek, je mnoho a přibývají s rostoucím počtem participujících zemí v šetření. Souhrnně se v metodologické literatuře tyto faktory nazývají chyby měření, které mají dvojí povahu a to systematickou a nesystematickou.⁷² V případě mezinárodní komparace se častěji zvažuje zkreslující vliv systematických chyb měření (bias), v nichž se koncentruje skupina rušivých faktorů, které vyplývají ze strukturálních a kulturních rozdílů mezi zeměmi. Tyto rušivé faktory pak ohrožují srovnatelnost měření mezi různými kulturami a mají za následek nesprávné závěry studií, neboť teoretický koncept nebo nástroj měření nemá stejný psychologický význam ve všech zahrnutých kulturních skupinách [Van de Vijver 1998: 43].⁷³ Navíc není ani možné odhalit nějaký univerzální zdroj zkreslení výsledků analýz, neboť ty jsou vždy závislé na konkrétních skupinách, které jsou do analýzy zařazeny. Van de Vijver a Leung [1997] identifikovali tři zdroje systematických chyb měření, které ohrožují ekvivalenci položek v mezinárodním výzkumu; konstruktovou (construct bias), metodologickou (method) a položkovou (item) chybu měření.

Konstruktová systematická chyba (construct bias) zamezuje komparaci teoretických jevů, protože nemají v různých kulturních skupinách stejný význam. Zdroje těchto chyb se ukrývají především v rozdílném kulturním chování (např. „být tradiční či moderní“ má jiný význam v západních zemích a jiný v asijských zemích) anebo nedostatečným pokrytím všech podstatných aspektů zkoumaného jevu měřicími položkami; některé mohou být v určitých situacích dokonce neadekvátní [Byrne, Watkins 2003: 157].

Metodologická systematická chyba (method bias) se týká všech zdrojů chyb vyrůstajících z metodologických aspektů výzkumu: výběru, nástroje měření a administrace dotazníků. Jedním ze zdrojů mohou být tedy a) odlišné kulturní charakteristiky výběrových souborů, které mají za následek jejich nesrovnatelnost. Dalším zdrojem je b) vliv formulace otázek/výroků a kategorií odpovědí v dotazníku, např. odlišná znalost logiky hodnoticích škál, které předložené výroky „známkuji“. Jiným významným zdrojem podobného rázu je specifický styl výběru odpovědí respondentem, který může mít různou podobu. Respondenti si mohou vybírat krajní body na škále měření anebo se jim vyhýbat (tj. response contraction bias), mohou systematicky uvádět souhlas s daným výrokiem a to bez ohledu na obsah dané položky (tj. acquiescence bias), anebo si volit takový škálový bod, který je zobrazuje v lepším světle (tj. sociální desirabilita).⁷⁴ Posledním zdrojem metodologické systematické chyby může být c) rozdílná administrace dotazníků nebo specifický vliv tazatelů na proces dotazování.

⁷² Systematická chyba je konstantní chyba, která se jednoduše definuje jako odchylka měření od pravdivé hodnoty a objevuje se při každé aplikaci daného designu výzkumu. Nesystematická chyba vypovídá o tom, že každé měření má svůj rozptyl (hodnoty jsou různé u jednotlivých respondentů), tato chyba se tedy liší u jednotlivých aplikací designu výzkumu [více Groves 2004].

⁷³ Ve statistické terminologii to znamená, že latentní proměnná je sice měřena stejnými indikátory, ale reprezentuje v každé skupině jiný teoretický koncept.

⁷⁴ Sociální desirabilitu by bylo možno definovat jako určitou potřebu sebe prezentace, která vede respondenty k zobrazování sama sebe v lepším světle, a souvisí s nadhodnocováním všeobecného mínění a sociálně akceptovatelného chování. V podstatě je otázka sociálně desirabilní tehdy, pokud odklání respondenty od variant odpovědí, které jsou považovány za společensky nežádoucí [Vinopal 2008: 47]. Tento jev je diskutován ze dvou úhlů pohledu, za prvé z hlediska fungování sociální desirability jako stylu odpovědí respondenta (response

Položková systematická chyba (item bias) způsobuje nejrůznější odchylky na úrovni položek v dotazníku, protože je spojena s odlišným kulturním či sociálním výkladem významu obsahu položek. Zkreslení v interpretaci položek nejvíce umocňuje překlad dotazníků do mnoha různých jazyků [Byrne, Watkins 2003: 157-158].

Robert, Lee a Chan [2006] shrnuli potenciální zdroje ohrožující statistickou srovnatelnost mezinárodních dat na základě kompilace existujících studií poněkud odlišně. Pokud se podle nich liší výběrové soubory buď v 1) kultuře, nebo 2) jazyku, 3) organizační (resp. pracovní) úrovni či 4) výzkumném kontextu, může dojít ke zkreslení výsledků komparativních analýz. Ekvivalenci položek tak definují jako kvalitu měření, která vzniká v interakci nástroje měření se specifickými charakteristikami výběrového souboru a s kontextem administrace dotazníku [2006: 67]. Jejich největším přínosem je ovšem zmapování toho, jak jednotlivé zdroje chyb ovlivňují výsledky ověřování ekvivalence v konfirmační faktorové analýze – resp. kdy dochází k identifikaci problematické položky v různých skupinách na základě odlišné velikosti faktorové zátěže či konstanty v regresních rovnicích.⁷⁵

Vysoká faktorová zátěž v regresní rovnici ukazuje na silný vztah mezi měřicí položkou a latentní proměnnou (tj. položka ji dobře měří)⁷⁶, nulová faktorová zátěž naznačuje absenci vztahu mezi měřicí položkou a latentní proměnnou [Robert, Lee, Chan 2006: 67]. Pokud faktorové zátěže položek nemají v různých skupinách srovnatelnou velikost, značí to odlišně silný vztah mezi měřicími položkami a latentními proměnnými v jednotlivých skupinách a pravděpodobně také měření jiného teoretického konceptu. Škála měření zkoumaných položek je stejně rozsáhlá a má stejnou jednotku, ale jednotlivé skóry (slovní či numerické) jsou pravděpodobně interpretovány poněkud jiným způsobem.

Naproti tomu velikosti konstant položek jsou rovnocenné, pokud respondenti v různých skupinách, kteří mají stejný postoj k teoretickému konceptu (tj. latentní proměnné), využívají rozsah měřicí škály stejným způsobem a měřený skór opravdu zachycuje jejich názor či hodnocení – výsledky tedy nezkrsluje metodologická systematická chyba měření [Robert, Lee, Chan 2006: 67-68]. Naopak rozdíly mezi velikostmi konstant v jednotlivých skupinách jsou důsledkem rozdílů mezi průměry přímo měřených položek, které jsou ale větší či menší než je očekáváno na základě velikostí faktorových zátěží a průměrů latentních proměnných; vstupuje do toho ještě další zkreslující faktor. Tudiž pokud některá položka měří latentní proměnnou - například míru mezilidské důvěry - velikost její konstanty $c = 2,7$ lze interpretovat tak, že skór latentní proměnné mezilidské důvěry je v průměru o 2,7 jednotek větší, než je skutečná míra mezilidské důvěry respondenta (resp. skór měřicí položky důvěry) [Steinmetz 2013: 8].

Prostřednictvím odlišně velkých faktorových zátěží (směrnic) a konstant položek se nejčastěji podle Roberta, Lee a Chana [2006: 69] manifestují následující zdroje zkreslení (viz Tabulka 23). První soubor zkreslujících faktorů vychází z kulturních rozdílů mezi

style), z druhé perspektivy se sociální desirabilita vnímá jako osobnostní psychosociální charakteristika [Johnson, Van de Vijver 2003: 195]. Některá témata jsou ve všech sociálních skupinách systematicky podhodnocována (např. kriminální jednání, užívání drog, konzumace alkoholu, některé druhy příjmů atp.), naopak některá jsou systematicky nadhodnocována (volební účast, šetření energií atd.) [blíže Tourangeau, Rips, Rasinski 2002]. I když se sociální desirabilita ukazuje jako univerzální koncept, který se objevuje ve všech sociálních skupinách, obsahuje kulturně specifické faktory a má různý vliv na respondenty v různých kulturách a sociálních skupinách.

⁷⁵ Problematický indikátor lze definovat jako indikátor, který není vhodný použít pro měření v dané skupině, neboť názor respondentů není vlivem chyby měření úspěšně přenesen do skóru přímo měřeného indikátoru a neodráží se v něm (např. ve skóru hodnocení výroku) [Steinmetz 2013: 10].

⁷⁶ Jinak řečeno – malé rozdíly mezi odpověďmi respondentů v měřicí položce korespondují s velkými rozdíly v postojích k danému teoretickému konceptu, tzn. s velkými rozdíly mezi skóry latentní proměnné [Robert, Lee, Chan 2006].

skupinami. Například může být daný teoretický koncept měřen etickými (kulturně univerzálními) položkami, ale obsah těchto položek není relevantní v každé zkoumané skupině a bylo by vhodnější použít emické položky. Porucha ekvivalence tohoto druhu se projevuje rozdílnými faktorovými zátěžemi položek, neboť dané měřicí položky nediskriminují mezi jednotlivými úrovněmi latentních proměnných (resp. mezi jejich skóry) rovnocenným způsobem.

Z kulturních odlišností mezi skupinami dále vychází rozdílný styl odpovědí respondentů, např. volba krajních kategorií odpovědí. Tato systematická chyba měření se projevuje odlišnou velikostí směrnic položek a někdy také odlišnou velikostí faktorových variancí. Na druhou stranu tendence respondentů k souhlasu s výrokem bez ohledu na jeho obsah a projevy sociální desirability vyúsťují v rozdílné velikosti konstant položek. Posledním typem chyb měření kulturního rázu je tzv. referenční efekt, který říká, že postoj vůči teoretickému konceptu anebo výběr kategorie na škále odpovědí je řízen názory blízké sociální skupiny, se kterou se respondent srovnává. Nesrovnává se však se skupinou, se kterou ho porovnává výzkumník v analýze, proto se tato systematická chyba měření odráží na velikosti směrnic [blíže Robert, Lee, Chan 2006: 70].

Tabulka 23 Shrnutí potenciálních zdrojů narušení ekvivalence položek na úrovni měření (tj. shody parametrů směrnic a konstant)

Potenciální zdroje zkreslení			
Ekvivalence faktorových zátěží položek / směrnice položky		Ekvivalence konstant položek / konstanta položky	
Etické položky v baterii otázek	Kultura	Systematická volba souhlasu s daným výrokem (tj. acquiescence bias)	Kultura
Tendence k volbě krajních (extrémních) kategorií odpovědí (tj. response style)		Tendence k volbě krajních (extrémních) kategorií odpovědí	
		Referenční (rámcový) efekt sociální skupiny	
Nesprávný překlad idiomů a metafor	Jazyk/ Překlad	Nesprávný překlad nespécifikovaných výrazů (intenzita, množství, frekvence činností, atp.)	Jazyk/ Překlad
Oslabené jazykové dovednosti respondenta, špatná interpretace		Oslabené jazykové dovednosti respondenta, špatná interpretace	
Kontext pracovní pozice	Organizace	Referenční (rámcový) efekt organizace, pracovního místa	Organizace
		Kontext pracovní pozice	
		Sociální desirabilita, vliv tazatele a atmosféry dotazování	Kontext odpovědi

Zdroj: podle Robert, Lee, Chan [2006: 69]

Druhý největší soubor potenciálních zdrojů zkreslení vyvstává z rozdílného jazyka šetřených skupin a problematického překladu. Kamenem úrazu jsou tu především idiomatické výrazy, metafory a kulturně specifická vyjádření, neboť jejich překlad nemůže téměř nikdy vyústit ve zcela konceptuálně ekvivalentní významy pojmů. Tento problém se následně projeví na různě velkých faktorových zátěžích položek, protože mají většinou různě silný vztah s danou latentní proměnnou a nemusí pokrývat všechny aspekty zkoumaného jevu. Z hlediska nesrovnatelné velikosti konstant položek se z jazykového hlediska jeví jako neproblematičtější slova s nespécifikovaným významem, často vyjadřující frekvenci či intenzitu činnosti (např. zřídka, někdy, občas, často), kvantitu, pravděpodobnost, hodnocení (např. dobrý, špatný) či vlastnost (např. tradiční, ambiciózní, cílevědomý, atp.) [Robert, Lee, Chan 2006: 70]. Tyto pojmy nemají většinou stabilní význam vzhledem k daným kulturám a zpravidla vyúsťují v neekvivalentní překlad, který nadhodnocuje či podhodnocuje danou kvantitu nebo hodnocení. Navíc se také ukazuje, že pojmy vyjadřující méně extrémní postoj

či kvantitu (např. spíše souhlasím, někdy, atp.) jsou pro respondenty více atraktivní a volí je jako svou odpověď častěji [více Chan 2000]. Z tohoto důvodu, i když obě položky stále odkazují ke stejné latentní proměnné, jsou tímto velikosti jejich konstant značně ovlivněny [Robert, Lee, Chan 2006: 70].

Třetí nejvýznamnější zdroj chyb měření vyvstává z tzv. kontextu odpovědi [2006: 72]. Efekty kontextu jsou souhrnným termínem pro zdroj vlivů na proces odpovídání respondenta a patří sem například téma výzkumu, uspořádání otázek v dotazníku, grafická úprava, způsob sběru dat či formát prezentace (verbální/vizuální), mentální schopnosti respondenta, myšlenkové pochody, emoce atd. [Vinopal 2009: 405]. Z celého tohoto výčtu zdůrazňují Robert, Lee a Chan [2006] především vliv sociálně desirabilních otázek a otázek s náročným obsahem (např. postoj ke složitým společenským otázkám), které způsobují vědomé či nevědomé posouvání odpovědí na škále měření mimo tzv. skutečnou hodnotu/skór, což má za následek rozdílné velikosti konstant daných položek.⁷⁷

V případě položek vybraných postojových škál lze nejpravděpodobněji očekávat silný vliv kulturních a jazykových rozdílů mezi zeměmi. Vzhledem k precizní práci multikulturních koordinačních týmů plánujících design výzkumu ESS by se zde nemělo jednat o nevhodné použití etických položek, což by mělo vliv na velikosti směrníc položek. Ale lze predikovat vliv rozdílného stylu odpovědí respondentů jako je volba krajních kategorií odpovědí (vliv na směrnice i konstanty), tendence respondentů k souhlasu s výrokem bez ohledu na jeho obsah, vliv sociálně desirabilních otázek a referenčního efektu (vliv na konstanty). Již z teorie je patrné, že obzvláště položky politického odcizení a důvěry v instituce tímto typem zkreslení trpí. Selingson [1980: 631-632] shrnuje kritiku, která některé položky politické efektivity provází už od 70. let. Například obsah některých položek se zdá být dvojznačný a není zcela jasné, která odpověď už značí pocit politického odcizení. Položka DIFFICULT [*Jak často se politika zdá být tak složitá, že vlastně nemůžete pochopit, co se děje?*] byla prohlášena za sociálně desirabilní, neboť kladná odpověď na ni staví respondenta do nepříjemné situace, kdy musí prohlásit, že něco nechápe nebo něčemu nerozumí. Většina kritik se zaměřuje také na formát odpovědí, který může svádět k tendenci souhlasit s danými výroky bez ohledu na obsah nebo k výběru neutrálních prostředních odpovědí.⁷⁸ To vše naznačuje, že v testech modelů ekvivalence se bude pravděpodobně vyskytovat problém se shodou konstant těchto položek.

Z hlediska rozdílného jazyka a překladu položek se zdají být postojové škály oproštěny od metafor či idiomatických výrazů, které mají vliv na velikosti směrníc položek. Na druhou stranu výskyt nespécifikovaných výrazů označujících frekvenci či intenzitu hodnocení je v testovaných položkách velmi častý (např. dost, jen trochu, zřídka, občas, někteří, přibližně, atp.), což může snadno způsobit problém s ekvivalencí konstant položek. Jak navíc vysvětlují Robert, Lee a Chan [2006: 76], jazykové rozdílnosti se také významně odráží na ekvivalenci variancí chyb měření, tudíž je možné předpokládat i narušení této striktní ekvivalence mezi zeměmi. Závěrem lze tedy souhrnně podotknout, že častější výskyt problému s ekvivalencí konstant testovaných položek sociálního kapitálu a politického odcizení se jeví jako pravděpodobnější než potíže s ekvivalencí směrníc položek mezi zkoumanými zeměmi.

⁷⁷ Robert, Lee a Chan [2006: 71-72] vysvětlují také řadu chyb měření vyvstávajících z vlivu pracovního zařazení respondenta, z členství v daných organizacích a z vlivu pracovních skupin, ale tyto zdroje chyb jsou pro tuto studii nevýznamné, neboť se nezabývá testováním položek z pracovní oblasti. Pro bližší seznámení se s tímto zdrojem chyb je doporučena jejich publikace.

⁷⁸ Lze připustit, že ESS se pokusilo tuto tendenci zmírnit přeformulováním výroků na otázky, ve kterých se již neudává míra souhlasu, ale jiný typ odpovědi.

7 SEPARÁTNÍ ANALÝZA SROVNATELNOSTI POSTOJOVÝCH ŠKÁL

Základní model, u něhož bylo předběžnými analýzami ověřeno, že je zcela vyhovující z hlediska identifikovatelnosti, chybějících údajů i normality dat, byl nejdříve testován v každé zemi zvlášť, bez toho, že by se v analýze kladly nějaké požadavky na shodu velikosti parametrů mezi zeměmi [doporučení Byrne 2004, 2008]. Tento krok je důležitý kvůli potvrzení skutečnosti, že model reprezentuje data z každé skupiny vhodným způsobem, tj. vyhovuje počet faktorů, vztahy mezi nimi jsou relevantní a všechny položky přispívají k vysvětlení daného faktoru dostatečně.

Datový soubor byl očištěn od chybějících hodnot (resp. případů) metodou „listwise deletion“ a jako metoda odhadu parametrů modelu byla použita metoda maximální věrohodnosti ML (vysvětlení viz výše). Tabulka 24 ukazuje výsledky separátní analýzy, tzn. statistiky vhodnosti základního modelu v každé zkoumané zemi zvlášť.

Z výstupů analýzy vyplývá, že základní model byl vyhovující v každé zkoumané zemi a reprezentoval data z výzkumu adekvátním způsobem. Daný závěr podle všech očekávání nepotvrdily jen χ^2 testy, jejichž výsledky jsou statisticky významné na hladině spolehlivosti 99 % (p value < 0,0001) s 92 stupni volnosti, a lze tak podle nich zamítnout nulovou hypotézu o shodě modelované (reprodukované) kovarianční matice se vstupní kovarianční maticí vypočítanou z reálných dat. Jak již ale bylo řečeno, χ^2 statistiky jsou velmi citlivé na velikost výběrového souboru a požadují téměř perfektní shodu jednotlivých matic [více Jöreskog, Sörbom 1993: kap. 4]. Z tohoto důvodu byla v této studii vhodnost modelu vyhodnocována na základě hodnot indexů GFI, AGFI, CFI, RMSEA a TLI, které posoudily základní model jako model velmi dobře reprezentující reálná data. Hodnoty GFI a AGFI neklesly téměř ve všech zemích pod hodnotu 0,95. Index CFI se nikde nenacházel pod hodnotou 0,92 a index RMSEA téměř nikdy nevystoupal nad hodnotu 0,05 (tj. středně dobrý výsledek). Model nejhůře reprezentuje data ve Finsku, Portugalsku a Švédsku a i tam výsledky ukazují na průměrně vhodný model.

Z toho vyplývá, že je možné dále předpokládat, že ve všech zemích byly danými škálovými položkami měřeny podobné teoretické koncepty - sociální kapitál a politické odcizení - a ty spolu velmi podobně vzájemně souvisí. Avšak ještě nebylo otestováno, zda je možné vůbec vzájemně statisticky porovnávat dané latentní proměnné či dokonce měřicí položky napříč zkoumanými zeměmi, tj. komparace dat v meritorní analýze ještě tudíž možná není.

Tabulka 24 Separátní analýza základního modelu - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko	315,87	92	3,433	0,979	0,969	0,969	0,036	(0,032;0,040)	1,000
Belgie	285,78	92	3,106	0,978	0,968	0,968	0,036	(0,031;0,041)	1,000
Česká republika	251,87	92	2,738	0,971	0,958	0,966	0,039	(0,036;0,043)	1,000
Německo	479,80	92	5,215	0,978	0,968	0,960	0,039	(0,036;0,043)	1,000
Dánsko	353,89	92	3,847	0,968	0,953	0,951	0,046	(0,041;0,051)	0,913
Španělsko	358,29	92	3,894	0,967	0,951	0,948	0,047	(0,042;0,052)	0,839
Finsko	565,62	92	6,148	0,963	0,945	0,938	0,052	(0,048;0,056)	0,246
Francie	353,89	92	3,847	0,970	0,955	0,944	0,045	(0,040;0,050)	0,957
Spojené království	468,71	92	5,095	0,970	0,956	0,951	0,046	(0,042;0,050)	0,955
Řecko	401,41	92	4,363	0,978	0,968	0,967	0,039	(0,035;0,043)	1,000
Maďarsko	265,34	92	2,884	0,976	0,965	0,968	0,037	(0,032;0,042)	1,000
Švýcarsko	369,91	92	4,021	0,974	0,962	0,956	0,041	(0,037;0,046)	1,000
Irsko	363,07	92	3,946	0,975	0,964	0,959	0,041	(0,036;0,045)	1,000
Izrael	333,50	92	3,625	0,980	0,970	0,958	0,036	(0,031;0,040)	1,000
Itálie	216,76	92	2,356	0,975	0,963	0,967	0,036	(0,030;0,042)	1,000
Lucembursko	205,66	92	2,235	0,974	0,962	0,961	0,035	(0,029;0,042)	1,000
Nizozemí	498,60	92	5,420	0,972	0,958	0,955	0,045	(0,041;0,049)	0,984
Norsko	520,70	92	5,660	0,969	0,954	0,935	0,048	(0,044;0,052)	0,747
Polsko	320,72	92	3,486	0,977	0,967	0,955	0,038	(0,033;0,042)	1,000
Portugalsko	403,66	92	4,388	0,958	0,938	0,926	0,054	(0,049;0,059)	0,114
Švédsko	513,56	92	5,582	0,964	0,947	0,939	0,050	(0,046;0,054)	0,472
Slovinsko	324,97	92	3,532	0,969	0,954	0,948	0,044	(0,039;0,049)	0,970

Zdroj: ESS 2002; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

8 SIMULTÁNNÍ ANALÝZA SROVNATELNOSTI POSTOJOVÝCH ŠKÁL

Simultánní analýza začíná testováním základního modelu ve všech zemích najednou. Tento model se už v jazyku strukturního modelování nenazývá základní model, ale konfigurální model, který má ale zcela totožnou podobu jako základní model. Proto se v prvním kroku simultánní analýzy testuje tzv. konfigurální ekvivalence. Nadále se sleduje, zda je konfigurální model totožný ve všech 22 zemích. V tuto chvíli také začíná hierarchické testování čím dál více restriktivních modelů, u nichž se kladou různé požadavky na shodu jednotlivých parametrů modelu. Konfigurální model je stále model bez jakýchkoli restrikcí – nepožaduje se v něm žádná shoda parametrů, všechny jsou volně odhadovány (tj. vypočítány) metodou maximální věrohodnosti. Konfigurální model má jednu základní statistickou úlohu a to sloužit jako první model (tj. Model 1), ke kterému lze porovnat druhý metrický model s restrikcemi na shodu všech faktorových zátěží anebo jakýkoli další více restriktivní model.

V této studii byly pro potřeby ověření srovnatelnosti postojových škál v simultánní analýze testovány modely hierarchicky v následujícím pořadí: 1) konfigurální ekvivalence, 2) metrická ekvivalence (popř. částečná metrická ekvivalence), 3) ekvivalence korelace chyb měření, 4) skalární ekvivalence (popř. částečná skalární ekvivalence), 5) ekvivalence kovariancí faktorů, 6) ekvivalence variancí faktorů (popř. jejich částečná ekvivalence) a 7) ekvivalence variancí chyb měření (viz Tabulka 25). Testování ekvivalence průměrů faktorů bylo vynecháno, průměry latentních proměnných byly rovnou vypočítány. V navržené struktuře testování jednotlivých modelů lze vysledovat logický směr ověřování nejprve ekvivalence měření v modelech měření a následně strukturní ekvivalence ve strukturním modelu.

Tabulka 25 Navržené pořadí testů modelů invariance

Separátní analýza	Teoretická konceptuální úroveň	Test vhodnosti základního modelu v každé zemi zvlášť, ověření základních předpokladů (chybějící hodnoty, normalita dat)
Simultánní analýza	Konceptuální úroveň	1. Konfigurální ekvivalence (test ekvivalence základního modelu ve všech zemích najednou; ověření počtu faktorů a velikostí parametrů)
	Úroveň měření	2. Metrická ekvivalence (test shody velikostí směrníc položek) 3. Ekvivalence korelace chyb měření (test shody korelací chyb měření) 4. Skalární ekvivalence (test shody velikostí konstant položek) 5. Ekvivalence variancí chyb měření (test shody variancí chyb měření – zde testováno až jako poslední model)
	Strukturální úroveň	6. Ekvivalence variancí faktorů (test shody velikostí variancí faktorů) 7. Ekvivalence kovariancí faktorů (test shody velikostí kovariancí faktorů)

Pokud separátní analýza základního modelu ukazuje, že model má podobnou faktorovou strukturu ve všech zemích, není důvod, proč by i *test konfigurálního modelu (Model 1)* neměl dopadnout příznivě. Ze shrnutí tabulky 26 je možné potvrdit, že hodnoty indexů CFI > 0,95 a RMSEA < 0,05 jsou zcela vyhovující a že data jsou konfigurálně ekvivalentní. Model s danou faktorovou strukturou adekvátně reprezentuje data ve všech zemích a koncepty jsou tudíž v jednotlivých zemích podobně chápány, interpretovány a i měřeny.

Tabulka 26 Statistiky vhodnosti pro testy ekvivalence: shrnutí / 22 zemí

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	8171,67 (2024)	-	0,954 (-)	0,009 (0,009;0,009)	0,940	-	přijat
Model 2 (Metrická ekvivalence)	10094,57 (2276)	1922,90*** (252)	0,942 (0,012)	0,010 (0,009;0,010)	0,933 (0,007)	Model 1	nepřijat
Model 2a (Částečná metrická ekvivalence)	9868,79 (2272)	1697,12*** (248)	0,944 (0,01)	0,010 (0,009;0,010)	0,934 (0,006)	Model 1	přijat
Model 2b (Ekvivalence korelace chyb měření)	10234,73 (2293)	365,94*** (21)	0,941 (0,003)	0,010 (0,009;0,010)	0,932 (0,002)	Model 2a	přijat
Model 3 (Skalární ekvivalence)	41685,24 (2629)	31450,52*** (336)	0,710 (0,231)	0,020 (0,020;0,020)	0,709 (0,223)	Model 2b	nepřijat
Model 3a (Částečná skalární ekvivalence)	22401,32 (2483)	12166,59*** (190)	0,852 (0,089)	0,015 (0,015;0,015)	0,843 (0,089)	Model 2b	nepřijat
Model 4 (Ekvivalence kovariancí faktorů)	11222,52 (2503)	987,80*** (210)	0,935 (0,006)	0,010 (0,010;0,010)	0,932 (0,00)	Model 2b	přijat
Model 5 (Ekvivalence variancí faktorů)	12999,38 (2608)	1776,86*** (105)	0,923 (0,012)	0,010 (0,010;0,011)	0,922 (0,01)	Model 4	nepřijat
Model 5a (Částečná ekvivalence variancí faktorů)	12553,18 (2606)	1330,66*** (103)	0,926 (0,009)	0,010 (0,010;0,010)	0,925 (0,007)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	21240,40 (2942)	8687,22*** (336)	0,836 (0,09)	0,013 (0,013;0,013)	0,877 (0,048)	Model 5a	nepřijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: $GFI \geq 0,95$, $AGFI \geq 0,95$, $CFI \geq 0,95$ (příp. $CFI \geq 0,92$), $RMSEA \leq 0,05$ (příp. $RMSEA \leq 0,08$), $PCLOSE \geq 0,5$; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Poznámka 2: Hodnoty χ^2 testů jsou tak vysoké, protože jsou součtem všech χ^2 testů z jednotlivých zemí.

Stále ale ještě ale není potvrzeno, že lze latentní a manifestní proměnné porovnávat v daných zemích statisticky. Je možné, že latentní proměnné individuálního politického odcizení a kolektivního sociálního kapitálu nejsou měřeny zcela ekvivalentně ve všech zemích, neboť velikosti standardizovaných faktorových zátěží těchto měřicích položek vykazují nejvyšší variabilitu napříč zeměmi (viz Tabulka 27-28). Stejně tak jsou značně proměnlivé i velikosti korelací faktorů individuálního a institucionálního odcizení s faktorem důvěry v instituce, tudíž i jejich vztahy nemusí být všude srovnatelné (viz Tabulka 29).

Proto bylo dále nutné klást na konfigurační model postupně různé restriktce ohledně shody různých parametrů. Takto byl vystavěn *metrický model (Model 2)*, kde bylo požadováno, aby si byly všechny faktorové zátěže (směrnice) ve všech 22 zemích rovny. Z tabulky 26 vyplývá, že plné metrické ekvivalence mezi všemi zeměmi dosaženo nebylo. Změna velikosti indexu CFI (Δ CFI) mezi modely 1 a 2 je větší než 0,01, což značí statisticky významnou odlišnost (v tomto konkrétním případě jen velmi malou, Δ CFI = 0,012). V této situaci je možné přistoupit k testování částečné metrické ekvivalence [více Byrne, Shavelson, Muthén 1989]. Pokud budou alespoň dva indikátory měřící danou latentní proměnnou plně ekvivalentní (tj. je splněna podmínka částečné ekvivalence), stále může dojít k validní komparaci daných manifestních či latentních proměnných. *Model částečné metrické ekvivalence* byl také testován (*Model 2a*) a při jeho komparaci s konfiguračním modelem (Model 1) byly vygenerovány dobré výsledky (Δ CFI \leq 0,01), což znamená, že tyto dva testované modely se od sebe statisticky významně neliší.

V případě testování částečné ekvivalence je nutné identifikovat ty směrnice nebo konstanty položek, které se mezi zeměmi příliš odlišují a činí daný model nesrovnatelný s tím předešlým, méně restriktivním modelem. Takovéto parametry se zjišťují na základě modifikačních indexů, které ukazují, o kolik se vylepší (resp. sníží) výsledky χ^2 testů a dalších statistik vhodnosti modelu. Také se sleduje změna samotného odhadu parametrů, protože už se dále neklade požadavek shody jeho velikosti s ostatními parametry v ostatních zemích, daný parametr se v modelu takzvaně uvolní a odhaduje se jeho skutečná velikost. Protože jsou modifikační indexy statisticky založeny na χ^2 testech a jsou citlivé na velikost souboru, sledují se jen modifikační indexy (M.I.) větší než 20 a odhady změny velikosti parametrů (Par Change) větší než absolutní hodnota čísla 0,1.⁷⁹ Chan [2000] navrhuje jiný limit pro modifikační indexy a to M.I. > 10, ale v této studii postačilo sledovat k dosažení vyšší úrovně ekvivalence modelu vyšší hodnoty modifikačních indexů.

⁷⁹ Modifikační indexy jsou ve skutečnosti χ^2 testy pro jednotlivé restriktce – jejich vysoká hodnota ukazuje, že požadavek na shodu daného parametru s ostatními parametry není v testovaném modelu správný. Někdy je vhodné sledovat modifikační indexy i v případě, že statistiky vhodnosti ukazují na velmi dobré modely a plnou ekvivalenci dat [podrobněji <http://essedunet.nsd.uib.no/cms/topics/immigration/2/.html>].

Tabulka 27 Standardizované faktorové zátěže (tzn. regresní váhy) v konfiguračním modelu (Model 1, bez restrikcí) pro jednotlivé země – koncept sociálního kapitálu

země	individuální sociální kapitál			kolektivní sociální kapitál			
	TRUSTED	FAIR	HELPFUL	FRIEND	VOLUNORG	MEET	SOCACTIV
Rakousko	0,714	0,804	0,706	0,527	0,304	0,482	0,434
Belgie	0,713	0,656	0,595	0,398	0,559	0,264	0,567
Česká republika	0,792	0,795	0,621	0,425	0,376	0,430	0,542
Německo	0,650	0,684	0,605	0,477	0,375	0,425	0,503
Dánsko	0,765	0,770	0,538	0,351	0,404	0,215	0,572
Španělsko	0,795	0,723	0,621	0,619	0,389	0,441	0,330
Finsko	0,719	0,712	0,603	0,411	0,417	0,332	0,587
Francie	0,658	0,592	0,531	0,526	0,353	0,419	0,411
Spojené království	0,709	0,757	0,643	0,590	0,327	0,361	0,401
Řecko	0,738	0,759	0,673	0,322	0,426	0,350	0,463
Maďarsko	0,752	0,684	0,684	0,713	0,292	0,441	0,449
Švýcarsko	0,700	0,695	0,561	0,557	0,325	0,374	0,431
Irsko	0,688	0,812	0,601	0,306	0,380	0,196	0,392
Izrael	0,711	0,800	0,511	0,571	0,349	0,305	0,344
Itálie	0,666	0,783	0,618	0,887	0,387	0,290	0,174
Lucembursko	0,702	0,642	0,453	0,521	0,234	0,298	0,445
Nizozemí	0,762	0,751	0,562	0,359	0,408	0,393	0,570
Norsko	0,704	0,699	0,488	0,328	0,339	0,351	0,661
Polsko	0,663	0,669	0,490	0,387	0,457	0,308	0,360
Portugalsko	0,782	0,655	0,548	0,526	0,309	0,498	0,293
Švédsko	0,723	0,711	0,534	0,294	0,434	0,239	0,593
Slovinsko	0,719	0,771	0,637	0,346	0,465	0,375	0,512
Variabilita	0,145	0,179	0,253	0,593	0,325	0,302	0,487

Zdroj: ESS 2002; vlastní výpočty

Tabulka 28 Standardizované faktorové zátěže (tzn. regresní váhy) v konfigurálním modelu (Model 1, bez restrikcí) pro jednotlivé země – koncept politického odcizení

země	individuální odcizení				institucionální odcizení			důvěra ve veřejné instituce		
	INTEREST	DIFFICULT	POLACTIV	OPINION	VOTE	CARE	POLIT	POLIT	PARLAM	LEGAL
Rakousko	0,698	0,628	0,518	0,604	0,749	0,814	0,240	0,594	0,894	0,630
Belgie	0,680	0,504	0,654	0,506	0,632	0,771	0,244	0,679	0,827	0,679
Česká republika	0,631	0,634	0,579	0,696	0,743	0,780	0,257	0,631	0,872	0,687
Německo	0,685	0,563	0,470	0,679	0,755	0,775	0,291	0,590	0,849	0,565
Dánsko	0,640	0,647	0,602	0,627	0,805	0,771	0,197	0,673	0,862	0,596
Španělsko	0,665	0,624	0,588	0,688	0,785	0,796	0,234	0,644	0,800	0,733
Finsko	0,669	0,513	0,564	0,599	0,716	0,778	0,254	0,657	0,870	0,636
Francie	0,711	0,586	0,584	0,551	0,723	0,681	0,200	0,698	0,778	0,659
Spojené království	0,734	0,526	0,603	0,451	0,703	0,794	0,159	0,759	0,836	0,606
Řecko	0,590	0,500	0,776	0,810	0,721	0,796	0,238	0,643	0,887	0,665
Maďarsko	0,679	0,556	0,399	0,688	0,768	0,698	0,269	0,606	0,803	0,785
Švýcarsko	0,707	0,618	0,582	0,616	0,640	0,766	0,247	0,603	0,833	0,689
Irsko	0,753	0,481	0,621	0,494	0,752	0,800	0,234	0,649	0,858	0,604
Izrael	0,493	0,509	0,485	0,653	0,742	0,676	0,248	0,604	0,916	0,540
Itálie	0,740	0,496	0,618	0,712	0,752	0,814	0,230	0,628	0,838	0,490
Lucembursko	0,643	0,552	0,543	0,505	0,705	0,693	0,270	0,605	0,888	0,670
Nizozemí	0,674	0,646	0,587	0,616	0,719	0,799	0,141	0,776	0,783	0,701
Norsko	0,637	0,551	0,691	0,407	0,725	0,748	0,293	0,590	0,845	0,586
Polsko	0,610	0,597	0,566	0,659	0,730	0,711	0,200	0,661	0,769	0,631
Portugalsko	0,695	0,599	0,652	0,635	0,747	0,826	0,210	0,648	0,811	0,681
Švédsko	0,699	0,518	0,606	0,508	0,707	0,775	0,299	0,641	0,841	0,642
Slovinsko	0,521	0,608	0,581	0,657	0,740	0,568	0,195	0,636	0,865	0,677
Variabilita	0,260	0,166	0,377	0,403	0,173	0,258	0,158	0,186	0,096	0,295

Zdroj: ESS 2002; vlastní výpočty

Tabulka 29 Korelace faktorů v konfigurálním modelu (Model 1, bez restrikcí) pro jednotlivé země

země	ITRUST - SOCNET	INSTDIS - INDIS	INDIS - ITRUST	INDIS - INSTRUST	INSTDIS - INSTRUST	INDIS - SOCNET	ITRUST - INSTRUST	INSTRUST - SOCNET	INSTDIS - SOCNET	INSTDIS - ITRUST
Rakousko	0,334	0,401	-0,179	0,207	0,488	-0,250	-0,291	-0,176	-0,256	-0,293
Belgie	0,314	0,467	-0,172	0,337	0,642	-0,267	-0,551	-0,311	-0,291	-0,441
Česká republika	0,238	0,392	-0,221	0,183	0,601	-0,230	-0,455	-0,181	-0,206	-0,406
Německo	0,330	0,315	-0,163	0,205	0,674	-0,236	-0,436	-0,137	-0,235	-0,405
Dánsko	0,278	0,351	-0,118	0,388	0,649	-0,177	-0,387	-0,238	-0,337	-0,275
Španělsko	0,264	0,480	-0,186	0,128	0,445	-0,163	-0,327	-0,066	-0,056	-0,333
Finsko	0,336	0,400	-0,002	0,276	0,665	-0,267	-0,565	-0,314	-0,329	-0,404
Francie	0,449	0,451	-0,175	0,324	0,681	-0,267	-0,489	-0,306	-0,295	-0,334
Spojené království	0,323	0,365	-0,152	0,242	0,697	-0,107	-0,469	-0,211	-0,235	-0,400
Řecko	0,189	0,350	-0,108	0,068	0,434	-0,574	-0,264	-0,158	-0,334	-0,302
Maďarsko	0,311	0,414	-0,222	0,284	0,554	-0,224	-0,482	-0,122	-0,126	-0,389
Švýcarsko	0,365	0,351	-0,208	0,282	0,632	-0,144	-0,479	-0,223	-0,206	-0,486
Irsko	0,278	0,444	-0,167	0,359	0,599	-0,315	-0,339	-0,225	-0,213	-0,267
Izrael	0,347	0,286	-0,170	0,037	0,410	-0,175	-0,236	-0,101	-0,219	-0,364
Itálie	0,318	0,520	-0,231	0,230	0,493	-0,159	-0,395	-0,098	-0,117	-0,266
Lucembursko	0,380	0,317	-0,225	0,310	0,496	-0,284	-0,437	-0,246	-0,206	-0,358
Nizozemí	0,278	0,433	-0,200	0,366	0,635	-0,271	-0,558	-0,243	-0,248	-0,427
Norsko	0,264	0,380	-0,051	0,335	0,645	-0,249	-0,480	-0,229	-0,238	-0,343
Polsko	0,384	0,308	-0,187	0,198	0,541	-0,260	-0,364	-0,384	-0,316	-0,362
Portugalsko	0,197	0,390	-0,189	0,218	0,475	-0,168	-0,394	-0,162	-0,201	-0,312
Švédsko	0,192	0,415	-0,208	0,357	0,662	-0,328	-0,538	-0,172	-0,303	-0,431
Slovinsko	0,311	0,312	-0,115	0,261	0,460	-0,323	-0,478	-0,263	-0,249	-0,317
Variabilita	0,26	0,23	0,23	0,35	0,29	0,47	0,33	0,32	0,28	0,17

Zdroj: ESS 2002; vlastní výpočty

Poznámka: ITRUST (individuální sociální kapitál), INSTRUST (důvěra v instituce), SOCNET (kolektivní sociální kapitál), INDIS (individuální politické odcizení), INSTDIS (institucionální politické odcizení).

V částečně metrickém modelu (Model 2a) musely být na základě nejvyšších modifikačních indexů postupně uvolněny faktorové zátěže položek DIFFICULT a OPINION v Řecku a položek POLACTIV a OPINION v Norsku – jedná se v obou případech o latentní proměnnou individuální politické odcizení (viz Tabulka 30). Modifikační indexy spojené s těmito položkami ale nebyly příliš vysoké (M.I. se pohybovaly mezi 48–51), ale odhadované změny velikosti parametrů významně přesahovaly hlavně v Norsku hodnotu 0,1 (Par Change od 0,20 do 0,45). V ostatních 20 zemích byla potvrzena plně metrická ekvivalence dat.⁸⁰

Plná metrická ekvivalence dat značí, že obsah každé položky je interpretován podobným způsobem ve všech skupinách a že respondenti z různých skupin na otázky odpovídají stejným způsobem. Navíc také potvrzuje, že měřicí položky mají k daným latentním proměnným stejně silný vztah ve všech skupinách. Tyto závěry platí pro všechny země kromě Řecka a Norska, které se při simultánní komparaci všech šetřených zemích poněkud vychýlily. Je patrné, že v Řecku a Norsku nemají některé položky měřicí individuální politické odcizení stejně silný vztah k dané latentní proměnné jako je tomu u jiných zemí. Z hodnot uvolněných faktorových zátěží vyplývá, že v případě Řecka je faktorová zátěž položky OPINION vyšší než faktorové zátěže téže položky v ostatních zemích, což značí, že přeložená verze položky v Řecku koresponduje s latentní proměnnou lépe než v jiných zemích, podobně je tomu v Norsku u položky POLACTIV. Naopak přeložená položka OPINION do norštiny má menší faktorovou zátěž než je tomu v jiných zemích a má tak méně silný vztah s faktorem individuálního odcizení; podobně jako položka DIFFICULT v Řecku.

Tabulka 30 Odhady nestandardizovaných faktorových zátěží v metrickém modelu a částečně metrickém modelu / 22 zemí; velikosti uvolněných faktorových zátěží / GR, NO

Latentní proměnná	Měřicí položka	Plně metrický model	Částečně metrický model	GR	NO
individuální odcizení	INTEREST	1,000	1,000		
	DIFFICULT	1,132	1,158	0,900	
	POLACTIV	1,344	1,330		1,933
	OPINION	1,126	1,117	1,366	0,807
institucionální odcizení	VOTE	1,000	1,000		
	CARE	1,064	1,064		
	POLIT	0,738	0,715		
důvěra v instituce	POLIT	0,725	0,728		
	PARLAM	1,000	1,000		
	LEGAL	0,794	0,795		
individuální sociální kapitál	TRUSTED	1,026	1,026		
	FAIR	1,000	1,000		
	HELPFUL	0,816	0,816		
kolektivní sociální kapitál	FRIEND	1,416	1,414		
	VOLUNORG	1,940	1,938		
	MEET	1,000	1,000		
	SOCACTIV	0,789	0,790		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných faktorových zátěží (směrnic) ve srovnání s plně metrickým modelem, červená barva vyšší hodnoty směrnic.

⁸⁰ V některých případech je nezbytné na základě prověření velikosti modifikačních indexů uvolnit v modelu faktorové zátěže referenčních položek. Protože není možné uvolnit faktorovou zátěž referenční položky (už je totiž zafixovaná ve všech zemích na jednotku), mělo by v původním modelu dojít k jejich přehodnocení a volbě nových (pokud možno invariantních) referenčních položek. Pak je zapotřebí nový model opět otestovat.

[více <http://essedunet.nsd.uib.no/cms/topics/immigration/2/.html>]

Škála měření zkoumaných položek je stejně rozsáhlá a má stejnou jednotku, ale jednotlivé skóry odpovědí jsou v Řecku a Norsku pravděpodobně interpretovány poněkud jiným způsobem. Na vině může být ztížený překlad významu anglické fráze „make your mind up“ nebo „political issues“ vyskytujících se ve znění položek OPINION a POLACTIV do řečtiny a norštiny anebo obecná tendence respondentů z Řecka a Norska vybírat v tomto případě krajní kategorie odpovědí, resp. jejich specifický styl odpovědí.

Z ověření částečné metrické ekvivalence vybraných položek ve všech zemích plyne jedna podstatná informace týkající se dalšího možného statistického porovnávání těchto dat. Ve všech zemích je možné statisticky validně porovnávat vztahy zkoumaných latentních proměnných s jinými proměnnými z datového souboru, u kterých byla také ověřena metrická ekvivalence, a to i v Řecku a Norsku, protože byla ve faktoru individuálního odcizení dodržena alespoň částečná metrická ekvivalence. V Řecku i Norsku byl ověřen také dopad částečné metrické ekvivalence na velikosti korelací částečně ekvivalentní latentní proměnné individuálního politického odcizení s ostatními latentními proměnnými: 1) institucionálním odcizením, 2) důvěrou v instituce, 3) individuálním sociálním kapitálem a 4) kolektivním sociálním kapitálem. Výsledky korelací těchto latentních proměnných pro plně metrický model v Řecku jsou 0,38; 0,1; -0,12; -0,55 popořadě a pro částečně metrický model 0,36; 0,08; -0,12; -0,54 popořadě. V Norsku jsou výsledky velikostí korelací pro plnou metrickou ekvivalenci následující 0,36; 0,32; -0,05; -0,15 popořadě - a pro částečnou ekvivalenci 0,39; 0,33; -0,05; -0,18 popořadě. Rozdíly jsou minimální, a proto je vhodné položky narušující plnou metrickou ekvivalenci v modelu ponechat a ne je z měření v těchto zemích vyřadit.

V rámci testování ekvivalence v modelu měření je v případě potřeby možné otestovat navíc shodu korelací chyb měření. Již bylo řečeno, proč je korelace chyb měření položek MEET a SOCACTIV v testovaném modelu nezbytná. Proto bylo i dále otestováno, zda jsou korelace chyb měření shodné ve všech zemích. Výsledky zanesené v Tabulce 26 ukazují, že *model ekvivalence korelací chyb měření (Model 2b)* a model částečné metrické ekvivalence (Model 2a) nebyly statisticky významně odlišné ($\Delta CFI < 0,01$) a lze tak potvrdit, že korelace chyb měření těchto dvou položek je mezi zeměmi srovnatelná a že se obě položky pravděpodobně obsahově překrývají (tj. korelují) v každé zkoumané zemi.

V hierarchickém pořadí pak následoval *test skalární ekvivalence (Model 3)*. V případě, že již byla identifikována částečná metrická ekvivalence, není nutné testovat plnou skalární ekvivalenci a lze okamžitě uvolnit konstanty položek, které nesplnily plnou metrickou ekvivalenci. V tomto případě bylo ale zajímavé vysledovat, nakolik je plně skalární model (Model 3) shodný s modelem částečné metrické ekvivalence (Model 2a), resp. modelem ekvivalence korelace chyb měření (Model 2b).⁸¹ Ukázalo se, že plně skalární model je statisticky významně odlišný od předchozího modelu (pokles indexu CFI byl opravdu podstatný, $\Delta CFI = 0,231$). Muselo být v některých případech volně odhadováno více než dvě konstanty měřicích položek v dané latentní proměnné, a proto nebyly v některých skupinách splněny ani požadavky částečné skalární ekvivalence měření. Jediné dvě latentní proměnné sociálního kapitálu jsou částečně srovnatelné pouze v deseti zemích střední a západní Evropy (včetně Izraele); to jsou nejlepší výsledky, kterých bylo možno dílem určitých kompromisů dosáhnout (viz Tabulka 31).

⁸¹ Restrikce se na sebe v hierarchicky seřazených modelech v podstatě nabalují, tudíž model ekvivalence korelace chyb měření v sobě obsahuje jak požadavky na shodu téměř všech (ekvivalentních) faktorových zátěží (viz částečná metrická ekvivalence), tak i požadavek na shodu korelací chyb měření ve všech zemích. Pokud se v modelech takto restrikce přidávají k sobě, porovnávají se modely v hierarchickém pořadí (tzn. více restriktivní model vůči méně restriktivnímu).

Z toho vyplývá, že pokud by bylo výzkumným záměrem přímo statisticky porovnávat průměry daných 16 položek mezi 22 zeměmi najednou (např. průměrnou důvěru respondentů k parlamentu dané země), nedošlo by k validní komparaci, neboť položky nejsou skalárně ekvivalentní. Navíc není ani možné validně srovnávat průměry latentních proměnných důvěry v instituce a politického odcizení ve všech zemích, protože k tomu je nutné splnit alespoň parametry částečné skalární ekvivalence položek. Z výsledků tak vyplývá, že rozdíly mezi průměry některých měřicích položek jsou způsobené ještě něčím jiným než rozdíly mezi průměry latentních proměnných, resp. rozdílnými názory na teoretické koncepty, a že se to odvíjí od příslušnosti respondenta k dané zemi. Jinak řečeno všude neplatí, že respondent, který vykazuje například nízkou důvěru v politické instituce, odpovídá také v dotazníku, že osobně nedůvěřuje buď politikům, nebo Českému parlamentu či právnímu systému – průměr měřicích položek by pak byl také velmi nízký. Protože bylo do analýzy zařazeno příliš velké množství kulturně a jazykově odlišných zemí, pravděpodobně došlo ke kombinaci různých zdrojů chyb měření (např. stylu odpovědí respondentů, sociálně desirabilních otázek, referenčního efektu, překladu nespecifikovaných výrazů frekvence a hodnocení výroků, atp.), které způsobují odchylky ve velikostech konstant položek, i když ty stále měří stejný koncept.

Tabulka 31 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / AT, BE, CZ, DE, FR, UK, CH, IE, IL, LU

Latentní proměnné	Individuální sociální kapitál			Kolektivní sociální kapitál			
	TRUSTED	FAIR	HELPFUL	FRIEND	VOLUNORG	MEET	SOCACTIV
Skalární model	5,004	5,663	4,948	8,562	4,579	5,079	2,779
Částečně skalární model	5,067	5,705	4,924	8,521	4,722	5,121	2,754
Rakousko		5,450			4,204		
Belgie			4,519	8,347	5,128		
Česká republika			4,425		3,323	4,545	
Německo	4,635				4,079	4,840	
Francie	4,585			8,123			2,909
Spojen.králov.			5,452		3,600		
Švýcarsko	5,390				4,853		
Irsko			5,746	9,001			
Izrael			4,574		5,505	5,336	
Lucembursko			4,499	8,762	6,857		

Zdroj: ESS 2002; vlastní výpočty

Poznámka 1: Model částečné skalární ekvivalence: $\chi^2 = 4757,24$ ($\Delta \chi^2 = 633,03$); $df = 1074$ ($\Delta df = 37$); $CFI = 0,94$ ($\Delta CFI = 0,01$); $RMSEA = 0,014$; 90 % C.I. $RMSEA = (0,014; 0,014)$; $TLI = 0,933$ ($\Delta TLI = 0,009$)

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

V případě faktorů sociálního kapitálu je ale možné jejich průměry porovnávat mezi deseti evropskými zeměmi, neboť zde částečná skalární ekvivalence prokázána byla. K porovnávání průměrů latentních proměnných se zde přistupuje podle doporučení Byrne, Shavelsona a Muthéna [1989] a Vandenberg a Lance [2000], že 1) postačí dvě plně ekvivalentní položky a že 2) problematické položky nesmí ve faktoru převažovat. Steinmetz [2013] ale v současnosti upozorňuje svou studií na to, že už jedna jediná vychýlená konstanta položky způsobuje velký problém při statistickém hodnocení významnosti rozdílů mezi průměry latentních proměnných. Proto jsou ve faktoru, kde dochází k výskytu poloviny problematických položek, dané země barevně odlišené a výsledky srovnání relativizované (viz Tabulka 32).⁸²

⁸² Zatím bez zvažování velikosti rozdílů mezi uvolněnými konstantami, protože ověření dopadu vychýlených konstant na velikost a rozdíly latentních průměrů by si žádalo další hlubší studii, kterou není možné z důvodu rozsahu a zaměření zařadit do této práce.

Když se pomocí strukturní analýzy průměrů (MACS) vypočítávají průměry latentních proměnných, musí být průměr dané latentní proměnné zvolen v jedné zemi jako referenční hodnota (klasicky: průměr = 0) a proti ní se vypočítávají všechny ostatní průměry dané latentní proměnné ve všech dalších zemích. V tomto případě byla jako referenční hodnota zvolena každá latentní proměnná v Rakousku, a proto má vždy průměr 0,00. Všechny ostatní se proti ní porovnávají a určuje se pořadí daných zemí podle hodnocení daného konceptu. Pro srovnání je v tabulce 32 uveden i jednoduchý strukturní průměr vypočítaný přímo ze skóre položek měřících daný faktor. Z údajů je patrné, že strukturní průměr generuje méně jasné pořadí zemí než latentní průměr, protože na základě testu ANOVY nejsou mezi sebou jednotlivé průměry v řadě zemí vůbec statisticky významně odlišné, tj. ukazují, že země se v daných postojích neliší. Ve všech případech rozhoduje o pořadí zemí spíše latentní průměr. Proto se také doporučuje v případě částečné skalární ekvivalence pracovat namísto hrubých skóre měřících položek s faktorovými skóre patřičných latentních proměnných. Reeskens a Hooghe [2008] také dokumentují, že lze takto dojít k více validním, přesným a preciznějším výsledkům při komparaci dat.

Tabulka 32 Srovnání strukturních a latentních průměr /AT, BE, CZ, DE, FR, UK, CH, IE, IL, LU

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	5,23	0,046	3/4/5/6/7	0,00	x	3/4/5/6
Belgie	4,98	0,044	4/5/6/7/8/9	-0,196**	0,063	7
Česká republika	4,50*	0,058	8/9/10	-0,678**	0,070	10
Německo	5,10	0,033	3/4/5/6/7/8/9	-0,091	0,056	3/4/5/6
Francie	4,85	0,045	5/6/7/8/9/10	-0,366**	0,067	9
Spojen.království	5,35	0,039	3/4/5/6	0,106	0,060	3/4/5/6
Švýcarsko	5,77	0,038	1/2	0,591**	0,061	1
Irsko	5,81*	0,045	1/2	0,554**	0,062	2
Izrael	4,84	0,044	4/5/6/7/8/9/10	-0,359**	0,061	8
Lucembursko	5,14	0,056	3/4/5/6/7/8/9	0,02	0,075	3/4/5/6
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	5,22	0,026	3/4/5/6/7	0,00	x	5/6/7
Belgie	5,29	0,030	3/4/5/6/7	-0,009	0,025	5/6/7
Česká republika	4,59*	0,035	10	-0,418**	0,031	10
Německo	5,03	0,021	8/9	-0,126**	0,022	9
Francie	5,27	0,029	3/4/5/6/7	0,029	0,026	5/6/7
Spojen.království	5,02	0,025	8/9	-0,107**	0,024	8
Švýcarsko	5,36	0,024	2/3/4/5/6/7	0,055**	0,024	4
Irsko	5,36	0,024	2/3/4/5/6/7	0,109**	0,024	3
Izrael	5,52	0,026	2/3/4	0,133**	0,024	2
Lucembursko	5,89*	0,036	1	0,338**	0,032	1

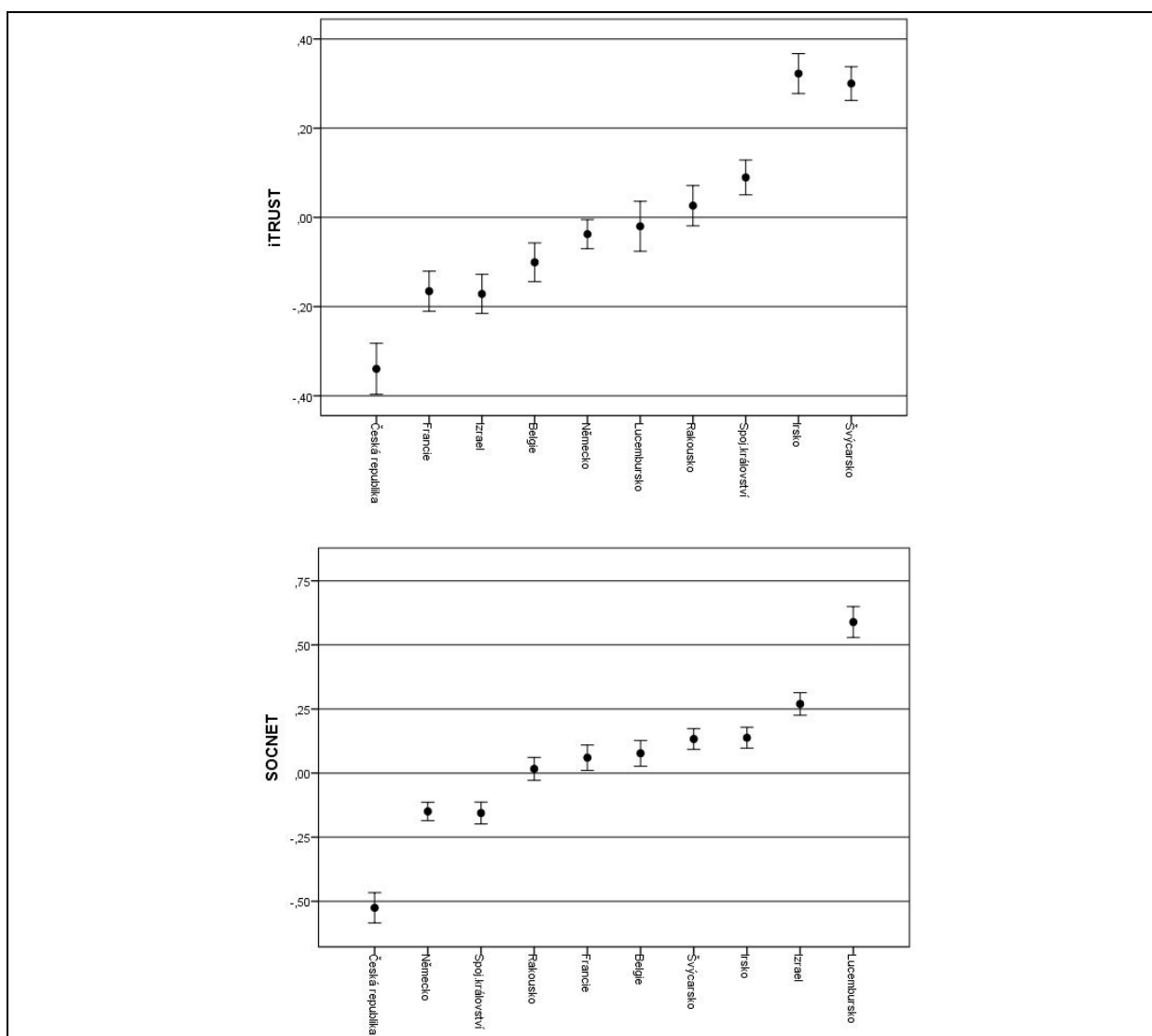
Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturní průměr je vypočten jako průměr skóre hodnocení měřících položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4/...), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 2: Některé země jsou modře zvýrazněny, protože vychýlené konstanty poloviny položek mohou zkreslovat rozhodnutí o statistické významnosti rozdílů mezi jejich latentními průměry.

V analýze průměrů individuálního sociálního kapitálu vyniká Švýcarsko a Irsko s nejvyšší mírou mezilidské důvěry (viz Tabulka 32 nebo Graf 1). Naopak patrně nejhorší výsledky vykazují Izrael, Francie a Česká republika. Pokud porovnáme reálnou velikost strukturálních průměrů, lze říci, že v zemích střední a západní Evropy (a v Izraeli) lidé manifestují spíše průměrnou důvěru v lidi s rozptylem (4,5–5,8 na škále od 0 do 10). Z hlediska kolektivního sociálního kapitálu se na nejvyšších příčkách umísťuje Izrael, Irsko a Lucembursko a na těch nejnižších Německo a Česká republika. Na základě reálných hodnot strukturálního průměru se jedná spíše o nadprůměrný kolektivní kapitál v evropských zemích s rozptylem (4,6–5,9 na škále od 0 do 8). Zajímavá je nevyvážená situace v Izraeli s poměrně vysokým kolektivním sociálním kapitálem (tj. hodně neformálních sociálních sítí), ale nízkým individuálním sociálním kapitálem (tj. nízká mezilidská důvěra). Je ale otázka, zda se v těchto údajích odráží skutečné kulturní rozdíly, nebo spíše systematické chyby měření. Není znám skutečný dopad nestejně velkých konstant položek VOLUNORG a MEET, které jsou v Izraeli (o 0,4 až 1) vyšší než je standardní průměr v evropských zemích (viz Tabulka 31), na velikost a statistickou odlišnost latentního průměru, proto je vhodné interpretovat rozdíly v pořadí zemí obezřetně.

Graf 1 Srovnání průměrů faktorů individuálního (ITRUST) a kolektivního (SOCNET) sociálního kapitálu a jejich 95 % intervaly spolehlivosti napříč vybranými zeměmi



Prostřednictvím testů dalších modelů už byla ověřována strukturní ekvivalence ve strukturní části modelu. Tento typ analýzy je nezbytný v případě zájmu o ověření konstruktové validity měřicího nástroje a soustředí se výhradně na vztahy mezi latentními proměnnými. Jako další byl tedy testován *model ekvivalence kovariancí faktorů (Model 4)*. Výsledky ukazují, že kovariance mezi latentními proměnnými jsou ve všech zemích totožné a že latentní proměnné mezi sebou tudíž vykazují shodné vztahy ($\Delta CFI < 0,01$). Například, že vysoké individuální a institucionální politické odcizení je vzájemně propojeno s nízkým individuálním a kolektivním sociálním kapitálem – tedy slabou mezilidskou důvěrou a slabými sociálními vazbami. Stejně tak se i nízká úroveň sociálního kapitálu odráží na klesající důvěře ve veřejné instituce, a naopak vysoký individuální sociální kapitál podporuje vysoký kolektivní sociální kapitál. Podobně spolu úzce souvisí i individuální a institucionální politické odcizení – a tyto závěry platí pro všechny země.

Test *modelu ekvivalence variancí latentních proměnných (Model 5)* odhalil, že variance faktorů nejsou všude shodné ($\Delta CFI = 0,012$) a na základě modifikačních indexů, které byly značně vysoké (M.I. se pohybují od 155 do 287) vyplynulo, že v Řecku a Izraeli nejsou variance faktoru důvěry ve veřejné instituce srovnatelné s variancemi téhož faktoru v ostatních zemích. Z toho vyplývá, že rozsah skóre této latentní proměnné v Řecku a Izraeli se liší od rozsahu skóre téže latentní proměnné v jiných skupinách. Navíc to naznačuje výskyt zkreslujícího stylu odpovědí respondentů v Řecku a Izraeli, kteří častěji volí krajní kategorie odpovědí v případě položek měřících tento faktor.

Jako poslední byl testován *model ekvivalence variancí chyb měření (Model 6)*, který je obecně považován za velmi restriktivní a je nezbytný pouze, pokud je cílem vypočítat a srovnávat položkovou reliabilitu. Z analýzy tohoto modelu lze vyvodit, že chyby měření jednotlivých položek nejsou srovnatelné napříč 22 zeměmi ($\Delta CFI > 0,01$) a není možné ani naplnit požadavky částečné ekvivalence chyb měření, což je při tomto počtu zemí zcela očekávatelný výsledek.

Simultánní analýza ekvivalence 16 měřících položek v pěti faktorech ve 22 zemích odhalila, že data jsou konfiguračně ekvivalentní a tudíž, že ve všech zemích byly sledované sociální jevy chápány a interpretovány velmi podobným způsobem a dané faktory měřeny srovnatelně. Tato úroveň ekvivalence je ale spíše teoretického rázu – s jistotou lze usoudit, že se laicky řečeno ve všech zemích hovořilo o jedné a té samé věci. Z praktického hlediska je ale také důležité vědět, zda je možné mezi sebou dané měřicí položky a latentní proměnné napříč všemi zeměmi statisticky porovnávat. Protože bylo v datech dosaženo alespoň částečné metrické ekvivalence, ze statistického hlediska lze validně zkoumat vztahy mezi šetřenými latentními proměnnými a dalšími metricky ekvivalentními proměnnými z výzkumu ve všech zemích (např. použitím regresní a korelační analýzy, atp.). Nebylo však dosaženo ani částečné skalární ekvivalence, tudíž dále není možné mezi zeměmi porovnávat průměry měřících položek a latentních proměnných (např. testovat rozdíl mezi průměry v ANOVĚ či T-testech). Dokonce se doporučuje i v regresní analýze používat faktorové skóre namísto jednotlivých měřících položek, u kterých nebyla skalární ekvivalence prokázána; výsledky jsou tak více validní [viz Reeskens, Hooghe 2008].

Z hlediska konstruktové validity postojových škál je možné z výsledků analýzy usoudit na poměrně stabilní a validní měření – bylo dosaženo částečné metrické ekvivalence dat, také plné ekvivalence kovariancí faktorů a částečné ekvivalence variancí faktorů. Model měření a strukturní model je napříč zeměmi značně rigidní a nebyly identifikovány významnější odchylky v modelech způsobené kulturním prostředím. Opět to značí, že vztahy mezi jednotlivými teoretickými koncepty jako sociální kapitál a politické odcizení jsou pevné a velmi podobné ve všech šetřených zemích. Konceptuálně se tedy ve výzkumu zkoumala jedna problematika, všude stejně vnímaná a podobně hodnocená.

9 SIMULTÁNNÍ ANALÝZA KULTURNĚ HOMOGENNÍCH SKUPIN ZEMÍ

Komparace měřicích položek, které byly v dotaznících překládány do různých jazyků a interpretovány různými kulturními skupinami, je vždy poměrně složitá. Klíčovým faktem vyplývajícím z prezentovaných výsledků simultánní analýzy ekvivalence je to, že není možné porovnávat *všechny položky* na určité statistické úrovni *ve všech 22 zemích*. Navzdory kvalitní metodologické přípravě výzkumu plné ekvivalence zkoumaných postojových škál dosaženo nebylo, i když je to v rámci analýzy mezinárodních dat často předpokládáno. Pokud se ale do analýzy vybere takto rozsáhlá skupina zemí a potažmo i takto komplexní model, není většinou pravděpodobné, že by byla v datech nalezena plná skalární (tj. silná) ekvivalence anebo dokonce plná ekvivalence chyb měření (tj. striktní ekvivalence).

Avšak - sice za cenu jistých objetí – tuto situaci je možné vyřešit jiným způsobem. Protože lze odvodit, že na silnou ekvivalenci položek bude mít pravděpodobně největší zkreslující vliv kultura, jazyk země a překlad položek, v případě neuspokojivých výsledků ekvivalence škál je vhodné si buď skupinu zemí omezit podle logického klíče, anebo zmenšit soubor testovaných měřicích proměnných; či obojí zároveň. Současně se tak mohou odkrýt i konkrétní příčiny selhání celkové mezinárodní ekvivalence těchto škál. Proto byla ve studii skupina 22 zemí dále rozdělena na menší kulturně kompaktnější celky. Takto vzniklo sedm skupin zemí, které byly nadále analyzovány samostatně. Z dalších analýz byl vyloučen Izrael, neboť ho není možné ani kulturně, ani jazykově a ani historicky zařadit do žádné skupiny zemí. Model zůstal prozatím stejně komplexní, protože hlavním cílem studie je ověřit ekvivalenci měřicích položek zvolených postojových škál a případně odhalit ty, které jsou systematicky zkreslené.

9.1 Simultánní analýza zemí střední a jihovýchodní Evropy



První analyzovanou skupinou jsou země, které se řadí územně i historicky do střední a jihovýchodní Evropy - Česká republika, Polsko, Slovinsko, Maďarsko a Německo, Rakousko, Lichtenštejnsko a Švýcarsko. První skupinu zemí k sobě lingvisticky pojí západoslovanské jazyky a druhou skupinu společný základ západogermánských jazyků. Jedinou výjimku zde tvoří Maďarsko, kde se hovoří ugrickým jazykem z ugrofinské skupiny jazyků. I když dané země velmi silně spojují společné historické kořeny, v testech ekvivalence měření se ukázala pouze úroveň

metrické ekvivalence. Kvůli snaze o identifikaci ještě silnějších úrovní ekvivalence položek (např. skalární anebo ekvivalence variancí chyb měření) byla daná skupina zemí rozdělena ještě na dva menší soubory, které si jsou z jazykového hlediska ještě bližší.

9.1.1 Česká republika, Maďarsko, Polsko a Slovinsko

Česká republika a Polsko územně náleží do střední Evropy a jejich oficiální jazyky se řadí do západoslovanské rodiny jazyků. Maďarština patří do jiné jazykové rodiny než slovanské jazyky – k tzv. ugrofinským jazykům, i přesto lze ale tyto země považovat za kulturně a historicky si blízké. Čtvrtou vybranou zemí je Slovinsko, které územně spadá do jihovýchodní části Evropy, ale jeho oficiálním jazykem je slovinština, jež spadá do větve jihoslovanských jazyků – ne tak lingvisticky vzdálených od západoslovanských jazyků [Lewis et al. 2013].

Tabulka 33 Statistika vhodnosti pro testy ekvivalence: shrnutí / CZ, PL, SI, HU

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1162,91 (368)	-	0,960 (-)	0,020 (0,018;0,021)	0,948 (-)	-	přijat
Model 2 (Metrická ekvivalence)	1350,45 (404)	187,53*** (36)	0,952 (0,008)	0,021 (0,019;0,022)	0,943 (0,005)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	1364,91 (407)	14,46** (3)	0,951 (0,001)	0,021 (0,019;0,022)	0,943 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	3042,23 (455)	1677,33*** (48)	0,869 (0,082)	0,032 (0,031;0,033)	0,862 (0,081)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1585,59 (422)	220,68*** (15)	0,941 (0,01)	0,022 (0,021;0,023)	0,933 (0,01)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1689,62 (452)	104,03*** (30)	0,937 (0,004)	0,022 (0,021;0,023)	0,934 (0,001)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1805,71 (467)	116,09*** (15)	0,932 (0,005)	0,023 (0,022;0,024)	0,930 (0,004)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	2400,18 (515)	594,48*** (48)	0,905 (0,027)	0,026 (0,025;0,027)	0,911 (0,019)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	2047,37 (507)	241,66*** (40)	0,922 (0,1)	0,023 (0,022;0,024)	0,926 (0,004)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (přip. CFI \geq 0,92), RMSEA \leq 0,05 (přip. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Všechny tyto skutečnosti hovoří pro to ověřit, zda bude mít tento kulturní a jazykový soulad (resp. nesoulad) vliv na úroveň srovnatelnosti měřících položek testovaných postojových škál. Postup ověřování modelu byl obdobný jako u předešlé analýzy s jedinou výjimkou, že již nebyl testován základní model v každé zemi zvlášť, protože tyto výsledky jsou už zaznamenané v Tabulce 24. Lze tedy rovnou přistoupit k simultánní analýze konfiguračního modelu (viz Tabulka 33).⁸³ V tabulce 33 jsou také zaneseny výstupy analýzy ekvivalence měření a strukturní ekvivalence modelu ve čtyřech zkoumaných zemích (CZ, HU, PL, SI). Opět se prokázalo, že data jsou konfiguračně a metricky ekvivalentní a že korelace chyb měření jsou také v daných zemích totožně velké (viz také výsledky z Tabulky 26). Stejně tak se potvrdila ekvivalence kovariancí a variancí všech faktorů ve všech zařazených zemích.

Zajímavější jsou ale výsledky skalární ekvivalence položek, která ani v těchto kulturně si podobných zemích nebyla naplněna (viz Tabulka 33). Rozdíl mezi metrickým a plně skalárním modelem je signifikantně významný, neboť rozdíl velikostí indexu CFI je značně vysoký (Δ CFI = 0,082). Částečná skalární ekvivalence potvrzena v datech byla, ale platí pouze pro latentní proměnné individuálního a kolektivního sociálního kapitálu; všechny ostatní konstanty položek měřících politické odcizení a důvěru ve veřejné instituce musely být na základě posouzení modifikačních indexů (M.I. > 20) uvolněny. Největší modifikační indexy přesahující ve většině případů hodnotu 100, které korespondovaly také s vyššími odhady změn velikosti parametrů (Par Change od 0,24 do 1,11), se týkaly položek VOLUNORG v Maďarsku a Slovinsku, OPINION v Maďarsku a HELPFUL v Polsku.

Tabulka 34 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / CZ, PL, SI, HU

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	CZ	PL	SI	HU
individuální odcizení	INTEREST	2,649		2,769	2,624	2,676	2,511
	DIFFICULT	3,153		3,240	3,201	3,097	3,012
	POLACTIV	3,951		4,054	3,814	3,823	4,143
	OPINION	2,944		3,082	2,988	3,060	2,607
institucionální odcizení	VOTE	3,968		4,021	4,037	3,969	3,716
	CARE	3,941		3,944	3,969	4,068	3,707
důvěra v instituce	POLIT	6,895		6,891	7,159	6,937	6,089
	PARLAM	6,124		6,424	6,451	5,925	4,973
	LEGAL	5,934		6,288	6,270	5,698	4,898
individuální sociální kapitál	TRUSTED	4,003	4,056				
	FAIR	4,699	4,748				
	HELPFUL	3,839		3,803	3,317	4,240	4,214
kolektivní sociální kapitál	FRIEND	8,233	8,205				
	VOLUNORG	3,466	3,362			4,606	2,429
	MEET	4,379	4,424				4,065
	SOCACTIV	2,613	2,588				

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Stejně tak může mít zkreslující vliv sociálně desirabilní povaha položky INTEREST dotazující se na zájem respondenta o politiku s tím, že v občanské společnosti je zájem o politiku společensky oceňován obzvláště ve vzdělanějších vrstvách společnosti, kam může

⁸³ Nebylo by nutné ani opakovat test konfiguračního a metrického modelu, neboť už v předešlé analýze bylo prokázáno, že tyto úrovně ekvivalence jsou potvrzeny v datech ze všech zemí. Z důvodů přehlednosti byly výsledky těchto testů opět v tabulkách zaznamenány s tím, že ve shrnutí závěrů následujících analýz se klade důraz především na vyhodnocení skalární ekvivalence a ekvivalence chyb měření.

být tazatel hypoteticky respondenty zařazován. Stejný případ mohl nastat i při dotazování položky DIFFICULT, kdy je respondent vystaven obtížné situaci vypovědět, že něčemu v oblasti politiky případně nerozumí. Nejpravděpodobnější vliv bude mít ale možná zdroj zkreslení pocházející z jazyka a překladu dotazníku, neboť obzvláště otázky na politické odcizení obsahují nespecifikované výrazy množství či intenzity jako je „dost, jen trochu“ (INTEREST), „zřídka, občas“ (DIFFICULT), „pravděpodobně“ (POLACTIV), „ani těžké, ani lehké“ (OPINION), „někteří, většina“ (VOTE, CARE), u nichž může překlad podhodnocovat nebo nadhodnocovat udané množství, nebo dochází k výběru atraktivnější, nenáročné střední varianty odpovědi.

V případě faktoru individuálního sociálního kapitálu se jako nejproblematictější ukazuje položka HELPFUL, jejíž konstanta musela být uvolněna v každé zemi a volně odhadnuta, tj. její velikost nemohla být zafixována jako shodná ve všech zemích. Konstanty této položky jsou vyšší v Maďarsku a Slovinsku a nižší v České republice a Polsku. U položek VOLUNORG a MEET je tomu naopak, jejich konstanty jsou v Maďarsku menší. U položky HELPFUL se může jednat o problém s překladem anglické fráze „try to be helpful“, u položky VOLUNORG o mírně sociálně desirabilní charakter při posuzování důležitosti dobrovolných organizací pro život respondenta. Obecně se pak spíše asi jedná o vliv referenčního efektu, kdy jsou dané otázky interpretovány v souvislosti s postoji blízké sociální skupiny.

Test modelu ekvivalence variancí chyb měření nepotvrdil plnou shodu chyb měření daných proměnných ve čtyřech zemích, ale byla ověřena částečná ekvivalence variancí chyb měření ve faktoru individuálního a kolektivního sociálního kapitálu a individuálního politického odcizení (viz Tabulka 35). Nejvíce se odchyluje Česká republika, u které se zdá, že položky měřící institucionální odcizení a důvěru ve veřejné instituce mají menší chyby měření a jsou tudíž více reliabilní, což částečně potvrzuje i předchozí analýza reliability.

Tabulka 35 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / CZ, PL, SI, HU

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	CZ	PL	SI	HU
individuální odcizení	INTEREST	0,410	0,376			0,513	
	DIFFICULT	0,765	0,768				
	POLACTIV	1,107	1,107				
	OPINION	0,520	0,518				
institucionální odcizení	VOTE	0,397	0,431	0,302			
	CARE	0,427	0,454	0,304			
důvěra v instituce	POLIT	1,744	1,846	1,322			
	PARLAM	1,727	1,704	1,126			2,292
	LEGAL	3,089	3,102				
individuální sociální kapitál	TRUSTED	2,636	2,638				
	FAIR	2,763	2,761				
	HELPFUL	3,428	3,428				
kolektivní sociální kapitál	FRIEND	3,085	2,886			2,239	4,714
	VOLUNORG	6,303	6,227				
	MEET	2,221	2,250				
	SOCACTIV	0,675	0,683				

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Protože z hlediska skalární ekvivalence i velikosti chyb měření způsobuje největší problém Maďarsko, je možné ho ze skupiny zemí vyloučit a testovat ekvivalenci položek jen ve třech následujících zemích – České republice, Polsku a Slovinsku (viz Tabulka 37). Analýza v těchto zemích ukazuje, že položky jsou opět jen částečně skalárně ekvivalentní, jejich nesrovnatelné konstanty způsobily příliš velký pokles indexu CFI mezi modely (Δ CFI = 0,041). Tentokrát ale nebylo zapotřebí uvolňovat konstanty více jak jedné či dvou měřicích položek v jednom faktoru najednou, byla tak identifikována částečná skalární ekvivalence ve všech latentních proměnných. Modifikační indexy vztahující se k uvolněným položkám jsou ale stále poměrně vysoké (M.I. od 52 do 113, Par Change od 0,41 do 0,81).

Zdá se, že po vyřazení Maďarska z testovaných skupin problémy se srovnatelností konstant položek měřicích politické odcizení a důvěru v instituce až na drobné výjimky téměř vymizely. Z tohoto důvodu je možné, že zdroje chyb měření těchto položek jsou spíše jazykové a překladatelské povahy - maďarština jako jediný jazyk vybočuje ze slovanských jazyků - a budou se opakovat i u jiných zkoumaných skupin zemí. Nejzajímavější výsledky jsou opakovaně spojené s měřicími položkami HELPFUL a VOLUNORG, které je stále zapotřebí uvolňovat ve všech zemích najednou – nemají shodně velké konstanty s tím, že pro položky měřené ve Slovinsku jsou většinou podstatně vyšší než pro ty v České republice či Polsku (viz Tabulka 36).

Příčinou vychýlených konstant položek v případě Slovinska může být také kromě referenčního efektu a vlivu specifických stylů odpovědi respondentů problém s tzv. oslabenou jazykovou dovedností respondentů [Robert, Lee, Chan 2006: 71]. Ten nastává ve chvíli, kdy jsou o vyplnění dotazníku požádáni respondenti, kteří jazyk, ve kterém je předložen dotazník, úplně dokonale neovládají. Ve Slovinsku byly distribuovány dotazníky ve slovinštině, ale některé tamní národnostní menšiny hovoří italsky a maďarsky [Lewis et al. 2013], a i tito respondenti mohli být součástí šetření.

Tabulka 36 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / CZ, PL, SI

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	CZ	PL	SI
individuální odcizení	INTEREST	2,679	2,679			
	DIFFICULT	3,18	3,177			
	POLACTIV	3,889	3,837	4,001		
	OPINION	3,035	3,034			
institucionální odcizení	VOTE	4,013	4,006			
	CARE	3,984	3,938			4,095
důvěra v instituce	POLIT	7,023	7,089	6,813		
	PARLAM	6,319	6,29			
	LEGAL	6,137	6,197			5,907
individuální sociální kapitál	TRUSTED	3,99	4,014			
	FAIR	4,719	4,747			
	HELPFUL	3,724		3,81	3,315	4,213
kolektivní sociální kapitál	FRIEND	8,292	8,193			8,495
	VOLUNORG	3,748		3,126	3,534	4,644
	MEET	4,454	4,454			
	SOCACTIV	2,62	2,624			

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Tabulka 37 Statistiky vhodnosti pro testy ekvivalence: shrnutí / CZ, PL, SI

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	897,58 (276)	-	0,957 (-)	0,023 (0,022;0,025)	0,944 (-)	-	přijat
Model 2 (Metrická ekvivalence)	943,42 (332)	45,84** (24)	0,955 (0,002)	0,023 (0,021;0,024)	0,946 (0,002)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	948,42 (302)	5,00* (2)	0,954 (0,001)	0,023 (0,022;0,025)	0,946 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1585,00 (334)	636,59*** (32)	0,913 (0,041)	0,030 (0,029;0,032)	0,906 (0,04)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1122,73 (325)	174,31*** (23)	0,944 (0,01)	0,024 (0,023;0,026)	0,938 (0,008)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1190,78 (345)	68,05*** (20)	0,941 (0,003)	0,024 (0,023;0,026)	0,939 (0,00)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1250,84 (355)	60,06*** (10)	0,938 (0,003)	0,025 (0,023;0,026)	0,937 (0,002)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	1621,48 (387)	370,64*** (32)	0,914 (0,024)	0,028 (0,026;0,029)	0,920 (0,017)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	1401,64 (379)	150,80*** (24)	0,929 (0,009)	0,025 (0,024;0,027)	0,932 (0,005)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

V takovéto situaci dochází k obdobné variantě zkusení jako při překladu měřicího nástroje, neboť respondenti nemusí kvůli omezené jazykové znalosti zcela shodně interpretovat význam položek, obzvláště idiomů, frází a nespecifikovaných pojmů označujících množství či frekvenci.⁸⁴ Výsledky testování skalární ekvivalence potvrzují, že průměry měřicích položek by stále ještě neměly být přímo porovnávány, ale je možná statistická komparace průměrů všech latentních proměnných v těchto třech zemích a faktorů sociálního kapitálu i v Maďarsku (viz Tabulka 38).

Tabulka 38 Srovnání strukturních a latentních průměrů / CZ, PL, SI, HU

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Česká republika	4,5	0,058	1/2/3	0,638**	0,076	1
Polsko	3,87**	0,042	4	0,00	x	4
Slovinsko	4,33	0,056	1/2/3	0,400**	0,075	3
Maďarsko	4,35	0,052	1/2/3	0,438**	0,073	2
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Česká republika	4,59	0,035	2/3	-0,038	0,031	2/3
Polsko	4,64	0,027	2/3	0,00	x	2/3
Slovinsko	5,14**	0,032	1	0,349**	0,033	1
Maďarsko	4,07**	0,036	4	-0,387**	0,035	4
(ne)důvěra ve veřejné instituce						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Česká republika	6,42	0,056	1/2/3	-0,268**	0,077	2
Polsko	6,67**	0,042	1	0,00	x	1
Slovinsko	6,16**	0,056	2	-0,598**	0,079	3
individuální politické odcizení						
země	Strukturní průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Česká republika	3,27	0,22	1/2/3	0,075**	0,021	1
Polsko	3,17	0,018	1/2/3	0,00	x	2/3
Slovinsko	3,15	0,021	1/2/3	-0,01	0,021	2/3
institucionální politické odcizení						
země	Strukturní průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Česká republika	4,89	0,031	1/2/3	-0,096**	0,030	3
Polsko	5,08	0,024	1/2/3	0,00	x	1/2
Slovinsko	4,98	0,029	1/2/3	-0,002	0,029	1/2

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturní průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů, resp. MACS. Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 2: Maďarsko je modře zvýrazněno, protože vychýlené konstanty poloviny položek mohou zkusovat rozhodnutí o statistické významnosti rozdílů mezi latentními průměry.

⁸⁴ Stejná situace platí i pro Českou republiku, kde je částečně úředním jazykem na celém území slovenština a polština a kde některé národnostní menšiny hovoří bulharsky, chorvatsky, maďarsky, německy, romsky, rusky, ukrajinsky, srbsky, atp. V Polsku jsou pomocnými úředními jazyky němčina, běloruština, kašubština a litevština, a početnou národnostní menšinou jsou Slezané [Lewis et al. 2013].

Z analýzy průměrů lze usoudit, že Česká republika vykazuje ze všech zemí nejvyšší míru individuálního sociálního kapitálu a také individuálního politického odcizení, Slovinsko nejvyšší míru kolektivního sociálního kapitálu a Polsko zase nejnižší míru důvěry ve veřejné instituce a jeden z nejvyšších průměrů institucionálního odcizení. Avšak z hlediska posouzení reálné velikosti (strukturních) průměrů se stále jedná ve všech zemích spíše o podprůměrnou míru sociálního kapitálu (hodnocení mezilidské důvěry cca 4,5 na škále od 0 do 10 v České republice, v Polsku cca 3,9), spíše vyšší nedůvěru ve veřejné instituce (hodnocení cca 6,5) a nadprůměrné politické odcizení. Hodnocení Maďarska je do tabulky zaneseno pouze v případě sociálního kapitálu, neboť podle výsledků testu skalární ekvivalence není možné validně interpretovat komparaci průměrů žádného jiného faktoru z maďarského modelu s ostatními zeměmi. Na tomto příkladu lze vidět, že průměry latentních proměnných je možné vzájemně porovnávat, ale většinou jen v určitém omezeném souboru zemí.

Je patrné, že i v případě výběru kulturně celistvého souboru zemí, kdy se kulturní zdroj chyb měření cíleně minimalizuje, nemusí měřicí položky ještě stále vykazovat vysokou úroveň skalární ekvivalence. Zajímavá je skutečnost, že po vyřazení Maďarska byly podmínky alespoň částečné skalární ekvivalence ve všech latentních proměnných dodrženy, což by ukazovalo na mírnou převahu důležitosti příbuzného jazyka respondentů nad podobnou kulturou či stejným historickým zázemím šetřených zemí; a implicitně tak na silnější zdroj chyb měření pramenících z rozdílného jazyka a překladu dotazníků.

V těchto třech zemích byl testován ještě jeden poslední model ekvivalence variancí chyb měření (Model 6 a 7). Opět byla potvrzena částečná ekvivalence chyb měření všude kromě latentní proměnné institucionálního politického odcizení, kde jsou chyby měření jednotlivých položek příliš rozdílné a musely být z restrikcí uvolněny. Česká republika opakovaně ukazuje měření s menšími chybami a tudíž více reliabilními položkami a Slovinsko naopak položky s vyššími chybami měření (viz Tabulka 39).

Tabulka 39 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / CZ, PL, SI

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	CZ	PL	SI
individuální odcizení	INTEREST	0,418	0,370			0,519
	DIFFICULT	0,759	0,762			
	POLACTIV	1,045	1,045			
	OPINION	0,539	0,536			
institucionální odcizení	VOTE	0,398		0,308	0,395	0,512
	CARE	0,412	0,419	0,302		
důvěra v instituce	POLIT	1,657		1,251	1,546	2,156
	PARLAM	1,551	1,558			
	LEGAL	3,251	3,246			
individuální sociální kapitál	TRUSTED	2,722	2,702			
	FAIR	2,642	2,927	1,837		
	HELPFUL	3,614	3,624			
kolektivní sociální kapitál	FRIEND	2,877	2,877			
	VOLUNORG	5,968	5,871			
	MEET	2,192	2,191			
	SOCACTIV	0,684	0,684			

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

9.1.2 Německo, Rakousko a Švýcarsko

Německo, Rakousko i Švýcarsko spadají také do střední části Evropy s tím, že ve všech těchto zemích se oficiálně hovoří němčinou (resp. západogermánským jazykem). Jen Švýcarsko je zajímavým vícejazyčným státem, kde hovoří většina Švýcarů sice německy, ale nemalá menšina také francouzsky, italsky anebo rétorománsky; což více méně kopíruje rozložení sousedních států [Lewis et al. 2013]. Lze tedy v těchto zemích předpokládat značnou podobnost ve vnímání měřicích položek z dotazníku, avšak je možné, že Švýcarsko bude nepatrně vybočovat.

Výsledky analýzy modelů ekvivalence ukazují, že i v těchto třech značně podobných státech položky nedosáhly ani plné skalární ekvivalence, ani plné ekvivalence variancí chyb měření. Jak rozdílné konstanty položek (Δ CFI = 0,075), tak i variance chyb měření (Δ CFI = 0,027) měly za následek příliš velké rozdíly ve velikostech indexu CFI mezi jednotlivými modely (viz Tabulka 41).

Test částečné skalární ekvivalence modelu odhalil, že položky měřící faktor institucionálního odcizení a faktor důvěry v instituce jsou jen metricky ekvivalentní, protože jejich konstanty jsou v Německu a Rakousku výrazně vyšší než ve Švýcarsku. Nelze tedy validně interpretovat rozdíly mezi těmito faktory v jednotlivých zemích na základě jejich průměrů, ale pouze jejich vztahy s ostatními proměnnými. Modifikační indexy, na základě nichž musely být uvolněny konstanty těchto položek, byly poměrně vysoké (M.I. od 20 do 106, Par Change od 0,2 do 0,3).

Z faktoru individuálního odcizení musely být v modelu uvolněny konstanty položek POLACTIV a OPINION (M.I. od 31 do 92) ve všech třech zemích. Naopak faktor individuálního sociálního kapitálu je v daných zemích plně skalárně ekvivalentní a ve faktoru kolektivního sociálního kapitálu musely být z restrikce shody uvolněny konstanty položek FRIEND (M.I. = 24), VOLUNORG (M.I. = 51) anebo MEET (M.I. = 32) podle specifické situace dané země (viz Tabulka 40).

Tabulka 40 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / AT, DE, CH

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	AT	DE	CH
individuální odcizení	INTEREST	2,224	2,231			
	DIFFICULT	2,906	2,912			
	POLACTIV	3,613		3,868	3,471	3,554
	OPINION	2,638		2,565	2,599	2,801
institucionální odcizení	VOTE	3,786		3,998	3,828	3,490
	CARE	3,661		3,763	3,833	3,286
důvěra v instituce	POLIT	6,069		6,365	6,524	5,117
	PARLAM	5,007		4,799	5,597	4,307
	LEGAL	4,131		3,876	4,382	3,893
individuální sociální kapitál	TRUSTED	5,069	5,196			
	FAIR	5,804	5,849			
	HELPFUL	5,062	5,097			
kolektivní sociální kapitál	FRIEND	8,536	8,486	8,727		
	VOLUNORG	4,368	4,166			4,873
	MEET	5,016	5,138		4,855	
	SOCACTIV	2,774	2,776			

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Tabulka 41 Statistiky vhodnosti pro testy ekvivalence: shrnutí / AT, DE, CH

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1165,59 (276)	-	0,962 (-)	0,022 (0,021;0,024)	0,950 (-)	-	přijat
Model 2 (Metrická ekvivalence)	1263,76 (300)	98,17*** (24)	0,958 (0,004)	0,022 (0,021;0,024)	0,950 (0,00)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	1278,75 (302)	14,99** (2)	0,958 (0,00)	0,022 (0,021;0,024)	0,950 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	3053,12 (334)	1774,37*** (32)	0,883 (0,075)	0,036 (0,034;0,037)	0,874 (0,076)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1476,33 (315)	197,58*** (13)	0,950 (0,008)	0,024 (0,023;0,025)	0,943 (0,007)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1548,31 (335)	71,98*** (20)	0,948 (0,002)	0,024 (0,023;0,025)	0,944 (0,001)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1770,49 (345)	222,18*** (10)	0,939 (0,009)	0,025 (0,024;0,027)	0,936 (0,008)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	2421,36 (377)	650,87*** (32)	0,912 (0,027)	0,029 (0,028;0,030)	0,916 (0,02)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	2026,08 (372)	255,58*** (27)	0,929 (0,01)	0,026 (0,025;0,027)	0,931 (0,005)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Opět se zde opakuje scénář, kdy nejsou v daných zemích srovnatelné konstanty položek politického odcizení a důvěry v instituce. Švýcarsko bylo otestováno i ve skupině s jinými zeměmi (např. AT–CH, DE–CH, FR–BE–LU–CH), které jsou kulturně či jazykové sprízněné, avšak nikdy nebyla nalezena plná skalární ekvivalence všech položek, ani částečná skalární ekvivalence položek ve všech faktorech. Jediné latentní proměnné, které se zdají být alespoň částečně porovnatelné, jsou faktory individuálního a kolektivního sociálního kapitálu. V případě faktorů politického odcizení a důvěry v instituce nebylo v těchto zemích nikdy dosaženo ani částečné skalární ekvivalence položek.

Důvody mohou být stejné jako v případě předešlé komparace středoevropských zemí se Slovinskem a to je oslabená jazyková dovednost respondentů ve Švýcarsku, kde část obyvatel hovoří německy a jiná část francouzsky, italsky či rétorománsky. Ve všech vlnách šetření ESS byl tento problém ve Švýcarsku řešen nabídkou vícejazyčné varianty dotazníku a to v němčině, francouzštině a italštině. Přesto se právě položky měřící politické odcizení, které obsahují špatně interpretovatelné výrazy množství, míry a frekvence, liší ve velikosti svých konstant. Tudíž jazykové a překladatelské problémy jsou patrné i při komparaci těchto zemí a ve srovnání s referenčním efektem a vlivem specifického stylu či kontextu odpovědí, které mají spíše kulturní povahu, se jeví jako nejpravděpodobnější. Ovšem tento předpoklad by bylo nutné opět otestovat.

Z hlediska chyb měření si položky v daných zemích nestojí špatně. Rozdíl mezi modelem plně a částečné ekvivalence variancí chyb měření nebyl příliš velký ($\Delta CFI = 0,017$). Faktory sociálního kapitálu a institucionálního odcizení dokonce dosáhly plné ekvivalence chyb měření (viz Tabulka 42). V případě individuálního odcizení musely být uvolněny chyby měření položky OPINION ve všech zemích (M.I. od 50 do 74) a ve faktoru důvěry v instituce chyby měření položek POLIT v Rakousku (M.I. = 102) a PARLAM a LEGAL ve Švýcarsku (M.I. = 64, 46). V Rakousku lze podle velikostí uvolněných chyb měření usoudit, že chyby měření daných položek jsou větší než ve zbylých dvou zemích a že položky jsou proto méně reliabilní. Ve Švýcarsku je situace opačná a položky tam dosahují větší míry reliability.

Tabulka 42 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / AT, DE, CH

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	AT	DE	CH
individuální odcizení	INTEREST	0,364	0,364			
	DIFFICULT	0,711	0,708			
	POLACTIV	1,236	1,239			
	OPINION	0,561		0,748	0,424	0,574
institucionální odcizení	VOTE	0,412	0,411			
	CARE	0,356	0,357			
důvěra v instituce	POLIT	1,739		2,423		
	PARLAM	1,317	1,598			0,824
	LEGAL	3,363	3,640			2,583
individuální sociální kapitál	TRUSTED	2,836	2,836			
	FAIR	2,264	2,265			
	HELPFUL	2,767	2,755			
kolektivní sociální kapitál	FRIEND	2,050	2,050			
	VOLUNORG	6,947	6,945			
	MEET	1,534	1,534			
	SOCATIV	0,673	0,673			

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Protože se Švýcarsko z jazykového hlediska mírně odlišuje, bylo z další analýzy vyloučeno. Ekvivalence položek byla tedy dále otestována pouze v Německu a Rakousku. Výsledky testu modelů ekvivalence konečně ukazují, že měřicí položky jsou částečně skalárně ekvivalentní a částečně ekvivalentní z hlediska chyb měření tentokrát ve všech faktorech v modelu (viz Tabulka 44).

Ve faktoru individuálního a institucionálního odcizení musely být při testu z restrikcí vyloučeny v daných zemích konstanty položek POLACTIV (M.I. = 76, 59) a VOTE (M.I. = 65, 45), z faktoru důvěry v instituce konstanta položky PARLAM (M.I. = 75, 45) a z faktoru individuálního sociálního kapitálu konstanta položky FAIR (M.I. = 55, 40). Latentní proměnná kolektivního sociálního kapitálu se ukazuje být v těchto dvou zemích jako plně skalárně ekvivalentní (viz Tabulka 43).

Tabulka 43 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / AT, DE

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	AT	DE
individuální odcizení	INTEREST	2,216	2,221		
	DIFFICULT	2,899	2,902		
	POLACTIV	3,640		3,878	3,460
	OPINION	2,586	2,585		
institucionální odcizení	VOTE	3,903		4,044	3,814
	CARE	3,808	3,816		
důvěra v instituce	POLIT	6,506	6,496		
	PARLAM	5,364		5,011	5,543
	LEGAL	4,245	4,205		
individuální sociální kapitál	TRUSTED	4,826	4,800		
	FAIR	5,664		5,344	5,873
	HELPFUL	4,973	4,955		
kolektivní sociální kapitál	FRIEND	8,553	8,555		
	VOLUNORG	4,142	4,144		
	MEET	4,936	4,937		
	SOCACTIV	2,779	2,780		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Je opět zřejmé, že vyřazení země, která se jazykově lišila a ve které bylo nezbytné v terénním šetření využít překladu dotazníku do několika jazyků, významně vylepšilo celkovou srovnatelnost jednotlivých položek. Potvrdilo se, že v Německu a Rakousku lze dojít k validní komparaci sebraných dat a to jak na základě ověřování vztahů faktorů s jinými proměnnými, tak i při vzájemném srovnávání průměrů zkoumaných faktorů. Jako stále skalárně neekvivalentní se ukazují čtyři specifické položky, u nichž je podle mého názoru vhodné hledat zdroj zkreslení v kontextu odpovědi, neboť kulturní a jazykové zdroje zkreslení položek byly specifickým výběrem zemí eliminovány na minimum. I když stále zde zůstává aktuální otázka vlivu oficiálních národnostních menšin v Rakousku, které hovoří turecky a chorvatsky, a jiných regionálních jazyků jako je slovinština a maďarština [Lewis et al. 2013]. I tyto drobné odchylky mohou být příčinou oslabené jazykové dovednosti některých respondentů, kteří vyplňovali v Rakousku pouze německou verzi dotazníku.

Tabulka 44 Statistiky vhodnosti pro testy ekvivalence: shrnutí / AT, DE

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	795,67 (184)	-	0,964 (-)	0,027 (0,025;0,029)	0,953 (-)	-	přijat
Model 2 (Metrická ekvivalence)	818,82 (196)	23,15** (12)	0,963 (0,001)	0,026 (0,024;0,028)	0,955 (0,002)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	824,97 (197)	6,15** (1)	0,963 (0,00)	0,026 (0,024;0,028)	0,955 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1426,46 (213)	601,49*** (16)	0,928 (0,035)	0,035 (0,033;0,037)	0,919 (0,036)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	955,28 (209)	160,31*** (12)	0,956 (0,007)	0,028 (0,026;0,030)	0,949 (0,006)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	983,57 (219)	28,29** (10)	0,955 (0,001)	0,027 (0,026;0,029)	0,951 (0,002)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1118,09 (224)	134,51*** (5)	0,947 (0,008)	0,029 (0,028;0,031)	0,943 (0,008)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	1421,46 (240)	303,38*** (16)	0,930 (0,017)	0,033 (0,031;0,034)	0,930 (0,013)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	1207,32 (238)	89,23*** (14)	0,943 (0,004)	0,030 (0,028;0,031)	0,942 (0,001)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Neboť byly splněny všechny požadavky částečné skalární ekvivalence položek ve všech faktorech, lze dále porovnávat jejich průměry v Německu a Rakousku mezi sebou. V případě Švýcarska lze na základě výsledků analýz porovnávat pouze faktory sociálního kapitálu a individuálního politického odcizení. Strukturní a latentní průměry generují velmi podobné pořadí zemí podle hodnocení daných konceptů (viz Tabulka 45).

Tabulka 45 Srovnání strukturních a latentních průměrů / AT, DE, CH

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	5,23	0,046	2/3	0,092	0,058	2/3
Německo	5,1**	0,033	2/3	0,00	x	2/3
Švýcarsko	5,77**	0,038	1	0,687**	0,053	1
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	5,22	0,026	1/2	0,169**	0,029	2
Německo	5,03**	0,021	3	0,00	x	3
Švýcarsko	5,36	0,024	1/2	0,191**	0,027	1
(ne)důvěra ve veřejné instituce						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	5,06**	0,046	2	-0,718**	0,070	2
Německo	5,53**	0,036	1	0,00	x	1
individuální politické odcizení						
země	Strukturní průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	2,91*	0,018	1/2	0,095**	0,021	2
Německo	2,78*	0,014	2/3	0,00	x	3
Švýcarsko	2,89	0,018	1/2	0,101**	0,021	1
institucionální politické odcizení						
země	Strukturní průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Rakousko	4,74	0,027	1/2	0,072**	0,024	1
Německo	4,75	0,022	1/2	0,00	x	2

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturní průměr je vypočten jako průměr skóreů hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 % a * na hladině 90 %.

Z hlediska individuálního sociálního kapitálu se zdá, že ve Švýcarsku lidé nejlépe hodnotí svou mezilidskou důvěru; Rakousko i Německo si vedou velmi podobně, ale o něco hůře než Švýcarsko. Všechny země se však nacházejí přibližně v průměrném hodnocení mezilidské důvěry. Při hodnocení sociálních vztahů a sítí Rakousko a Švýcarsko předčilo Německo a vykazují nadprůměrnou intenzitu sociálních kontaktů. Z hlediska nedůvěry v instituce naopak získává přední místo Německo s nadprůměrným hodnocením. V hodnocení politického odcizení se umísťuje na prvním místě Rakousko opět s nadprůměrnými čísly.

Model ekvivalence variancí chyb měření potvrdil v Německu a Rakousku částečnou ekvivalenci variancí chyb měření položek ve faktorech individuálního odcizení a důvěry v instituce a plnou ekvivalenci chyb měření položek v ostatních latentních proměnných.

V Rakousku chyby měření opět nabývají vyšších hodnot a položky OPINION a POLIT jsou tudíž méně reliabilní než v Německu (viz Tabulka 46).

Tabulka 46 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / AT, DE

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	AT	DE
individuální odcizení	INTEREST	0,379	0,379		
	DIFFICULT	0,730	0,726		
	POLACTIV	1,242	1,245		
	OPINION	0,549		0,740	0,416
institucionální odcizení	VOTE	0,374	0,372		
	CARE	0,347	0,349		
důvěra v instituce	POLIT	1,884		2,461	1,492
	PARLAM	1,373	1,373		
	LEGAL	3,732	3,731		
individuální sociální kapitál	TRUSTED	3,076	3,077		
	FAIR	2,257	2,257		
	HELPFUL	2,746	2,745		
kolektivní sociální kapitál	FRIEND	2,234	2,234		
	VOLUNORG	7,197	7,195		
	MEET	1,599	1,599		
	SOCACTIV	0,688	0,689		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

9.2 Simultánní analýza zemí jižní a západní Evropy



Do druhé skupiny byly vybrány kulturně si blízké země z jižní a západní Evropy, které mají stejné lingvistické základy románských jazyků. Proto, i když sem Francie územně a i z hlediska moderního politického a ekonomického vývoje nepatří (spadá spíše do západní Evropy), je do této skupiny kvůli jazykové příbuznosti zařazena i ona. Jediná řečtina nepatří mezi románské jazyky vzniklé z latiny, ale řadí se mezi samostatnou vývojovou větev indoevropských jazyků [Lewis et al. 2013]. Pokud byly předešlé předpoklady o silném vlivu podobných jazykových kořenů zemí na srovnatelnost dat správné, i v této skupině zemí by se mělo Řecko

z hlediska ekvivalence položek více odchylovat.

Výsledky analýzy ekvivalence položek tento předpoklad potvrdily. V předběžném testu se ukázalo, že pokud bude do analýzy zahrnuto i Řecko, model bude dokonce jen pouze částečně metrický (Δ CFI = 0,013) a skalární ekvivalence se nepotvrdí vůbec (Δ CFI = 0,236; viz Tabulka 47). Data z Řecka neprokázala plnou metrickou ekvivalenci položek individuálního odcizení ani v souhrnné analýze všech 22 zemí, a ani v této menší kulturně kohezivní skupině, uvolněna by musela být směrnice položky DIFFICULT (M.I. = 23) a položky FRIEND (M.I. = 14).

Tabulka 47 Statistiky vhodnosti pro testy ekvivalence: shrnutí / ES, PT, IT, GR, FR

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90% C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1734,05 (460)	-	0,953 (-)	0,020 (0,019;0,021)	0,939 (-)	-	přijat
Model 2 (Metrická ekvivalence)	2164,44 (508)	430,39*** (48)	0,939 (0,013)	0,021 (0,020;0,022)	0,928 (0,011)	Model 1	nepřijat
Model 3 (Skalární ekvivalence – všechny země)	8687,323 (572)	6522,89*** (64)	0,703 (0,236)	0,044 (0,044;0,045)	0,688 (0,24)	Model 2	nepřijat
Model 3a (Skalární ekvivalence – bez GR)	4170,48 (556)	2006,04*** (48)	0,868 (0,071)	0,030 (0,029;0,031)	0,857 (0,071)	Model 2	nepřijat
Model 3a (Skalární ekvivalence – bez ES)	7842,84 (556)	5678,41*** (48)	0,733 (0,206)	0,043 (0,042;0,043)	0,712 (0,216)	Model 2	nepřijat
Model 3a (Skalární ekvivalence – bez PT)	7842,69 (556)	5678,25*** (48)	0,733 (0,206)	0,043 (0,042;0,043)	0,712 (0,216)	Model 2	nepřijat
Model 3a (Skalární ekvivalence – bez IT)	7804,45 (556)	5640,01*** (48)	0,735 (0,208)	0,043 (0,042;0,043)	0,714 (0,214)	Model 2	nepřijat
Model 3a (Skalární ekvivalence – bez FR)	7424,22 (556)	5259,78*** (48)	0,748 (0,191)	0,041 (0,041;0,042)	0,729 (0,199)	Model 2	nepřijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: $GFI \geq 0,95$, $AGFI \geq 0,95$, $CFI \geq 0,95$ (příp. $CFI \geq 0,92$), $RMSEA \leq 0,05$ (příp. $RMSEA \leq 0,08$), $PCLOSE \geq 0,5$; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Tabulka 48 Statistiky vhodnosti pro testy ekvivalence: shrnutí / ES, PT, IT, FR (bez GR)

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90% C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1332,59 (368)	-	0,946 (-)	0,023 (0,022;0,024)	0,930 (-)	-	přijat
Model 2 (Metrická ekvivalence)	1511,27 (404)	178,68*** (36)	0,938 (0,008)	0,024 (0,022;0,025)	0,926 (0,004)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	1515,65 (407)	4,39 (3)	0,938 (0,00)	0,023 (0,022;0,025)	0,927 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	3490,21 (455)	1974,56*** (48)	0,830 (0,108)	0,037 (0,036;0,038)	0,821 (0,106)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1691,21 (424)	175,56*** (17)	0,929 (0,009)	0,025 (0,023;0,026)	0,920 (0,007)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1788,83 (454)	97,62*** (30)	0,925 (0,004)	0,024 (0,023;0,026)	0,921 (0,001)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1879,36 (469)	90,53*** (15)	0,921 (0,004)	0,025 (0,023;0,026)	0,919 (0,002)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	2395,09 (517)	515,736*** (48)	0,895 (0,026)	0,027 (0,026;0,028)	0,902 (0,017)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	2081,48 (509)	202,12*** (40)	0,912 (0,01)	0,025 (0,024;0,026)	0,917 (0,002)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Pro vyřazení Řecka hovořily i statistiky vhodnosti skalárního modelu, který byl otestován postupně pětkrát; v každém pokusu byly uvolněny všechny konstanty položek z jedné země. χ^2 testy generovaly nejnižší hodnotu při uvolnění parametrů z Řecka, stejně tak indexy CFI a RMSEA v tomto případě mnohem lépe hodnotily daný model. Z následující analýzy bylo proto vyřazeno Řecko. Analýza ekvivalence měřících položek ve Španělsku, Francii, Portugalsku a Itálii potom vyprodukovala údaje, které potvrdily konfigurální a plnou metrickou ekvivalenci, stejně tak plnou ekvivalenci korelace chyb měření a kovariancí a variancí faktorů v daných zemích. Výsledky se shodují se závěry analýzy předešlé kulturní skupiny zemí ze střední a jihovýchodní Evropy.

Na druhou stranu opět nebyla prokázána ani plná skalární ekvivalence (Δ CFI = 0,108), ani plná ekvivalence variancí chyb měření (Δ CFI = 0,026). Požadavky částečné skalární ekvivalence splňují pouze tři latentní proměnné postojových škál politického odcizení a institucionální důvěry, u nichž nedošlo na základě posouzení velikosti modifikačních indexů k uvolnění více jak jedné či dvou položek v jednom faktoru. V latentní proměnné individuálního odcizení musely být postupně uvolněny konstanty položky OPINION ve všech zemích (M.I. od 30 do 130) a položky INTEREST ve Španělsku a Francii (M.I. = 135, 63). V latentní proměnné důvěry ve veřejné instituce musela být zbavena restrikce položka LEGAL ve Španělsku a Itálii (M.I. = 71, 128) a PARLAM ve Francii (M.I. = 66). Ve zbylých dvou latentních proměnných sociálního kapitálu muselo být uvolněno z restrikce více proměnných a nelze tak porovnávat jejich průměry navzájem mezi těmito čtyřmi zeměmi (viz Tabulka 49).

Tabulka 49 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / ES, PT, IT, FR

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	ES	PT	IT	FR
individuální odcizení	INTEREST	2,829	2,812	2,990			2,699
	DIFFICULT	3,235	3,238				
	POLACTIV	4,103	4,108				
	OPINION	3,387		3,122	3,315	3,543	3,545
institucionální odcizení	VOTE	4,051	4,108				3,909
	CARE	3,875			4,017		
důvěra v instituce	POLIT	6,604	6,619				
	PARLAM	5,389	5,248				5,756
	LEGAL	5,369	5,410	5,757		4,533	
individuální sociální kapitál	TRUSTED	4,553		4,891	4,160	4,530	4,398
	FAIR	5,219		5,240	5,263	4,551	5,571
	HELPFUL	4,232		4,441	3,938	4,023	4,361
kolektivní sociální kapitál	FRIEND	8,198		8,374	8,493	7,705	8,091
	VOLUNORG	5,576		5,629	5,544	6,433	4,807
	MEET	5,262		5,287	5,709	4,835	5,192
	SOCACTIV	2,676		2,612	2,665	2,361	2,907

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

V případě zemí z jižní Evropy se objevuje zajímavá skutečnost a to problém s ekvivalencí konstant položek měřících sociální kapitál. Ve Španělsku jsou konstanty těchto položek významně vyšší, v Itálii naopak podstatně nižší oproti Španělsku a ve zbylých zemích se velikost konstant nahodile mění. Příčiny tohoto problému lze vidět jak v kulturních, tak i jazykových zdrojích chyb měření. Jazyky sice mají stejný lingvistický - románský - základ, ale v řadě ohledů se stále podstatně liší. Například ve Španělsku se v jednotlivých

autonomních společenstvích hovoří kromě španělštiny také regionálními úředními jazyky, proto je od roku 2002 dotazník ESS distribuován ve Španělsku jak ve španělštině, tak i v katalánštině (druhý nejrozšířenější jazyk), a od roku 2006 také v baskičtině a galicijštině.

Stejně tak je tomu s kulturním zázemím a vývojem daných zemí. Španělsko a Portugalsko jsou si v tomto ohledu nejbližší, neboť jako státy vykrytalizovaly ze společného křesťanského království. A i když mají Itálie a Francie v řadě věcí mnoho společného s historií Španělska a Portugalska, prošly jiným historickým vývojem a v dnešní moderní době už Francie očividně vybočuje a z hlediska ekonomické vyspělosti je zařazována do oblasti západní Evropy. Tyto rozdílnosti ale přesto nenarušují interpretaci položek měřících politické odcizení a důvěru v instituce; překlad významu pojmů označujících nespécifikované množství, intenzitu a frekvenci tedy evidentně není v románských jazycích problematický. Potíže se srovnatelností dat začínají až u položek měřících sociální kapitál, což je proti ostatním zemím nezvyklé. Roli tady může hrát silnější referenční efekt, neboť respondent může častěji srovnávat své postoje s blízkými sociálními skupinami, nebo efekt specifického stylu odpovědí (např. volba krajních kategorií škály, sociální desirabilita, atp.).

Co se týče ekvivalence chyb měření, problém není nijak vážný – kromě faktoru institucionálního odcizení v Portugalsku dosáhly všechny ostatní položky alespoň částečné ekvivalence chyb měření (viz Tabulka 50). Z velikostí uvolněných parametrů chyb měření lze usoudit, že v Portugalsku mají položky CARE (M.I. = 35) a VOTE (M.I. = 28) na rozdíl od Francie i ostatních zemí menší chyby měření a jsou tudíž více reliabilní. Naopak v Itálii má položka LEGAL (M.I. = 25) a SOCACTIV (M.I. = 27) výrazně vyšší chybu měření než v ostatních zemích, stejně tak ve Španělsku položka POLIT (M.I. = 24) a SOCACTIV (M.I. = 11).

Tabulka 50 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / ES, PT, IT, FR

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	ES	PT	IT	FR
individuální odcizení	INTEREST	0,421	0,420				
	DIFFICULT	0,930	0,931				
	POLACTIV	0,894	0,894				
	OPINION	0,677	0,677				
institucionální odcizení	VOTE	0,401	0,347		0,304		0,59
	CARE	0,381	0,405		0,257		
důvěra v instituce	POLIT	1,726	1,557	2,175			
	PARLAM	1,885	1,892				
	LEGAL	3,216	2,982			4,071	
individuální sociální kapitál	TRUSTED	2,401	2,386				
	FAIR	2,609	2,339				3,248
	HELPFUL	3,246	3,249				
kolektivní sociální kapitál	FRIEND	1,905	1,904				
	VOLUNORG	6,204	6,213				
	MEET	2,056	2,055				
	SOCACTIV	0,746	0,623	0,845		0,921	

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Lze si dále klást otázku, zda by položky nesplnily požadavky skalární ekvivalence při ještě větším zúžení testované skupiny zemí. Měření v Itálii a Francii se v analýze projevilo jako

pravděpodobná překážka plné skalární ekvivalence, v těchto zemích se objevovaly nejvyšší modifikační indexy požadující uvolnění konstanty daných položek a slibující zlepšení vhodnosti skalárního modelu. Navíc je italština i francouzština z hlediska jazyka jinou větví románských jazyků než španělština a portugalština⁸⁵, proto by bylo zajímavé zjistit, zda se po vyřazení Itálie a Francie zlepší úroveň skalární ekvivalence položek.

Výsledky ilustruje Tabulka 52. I tentokrát byl plně skalární model příliš odlišný od částečně skalárního modelu (Δ CFI = 0,036), avšak tentokrát splňují požadavky částečné skalární ekvivalence všechny latentní proměnné, protože u nich nebyly uvolněny více jak jedna či dvě konstanty proměnných zároveň (viz Tabulka 51). V latentní proměnné individuálního odcizení musely být na základě modifikačních indexů zbaveny restriktce položky INTEREST (M.I. = 36, 51) a OPINION (M.I. = 19, 21), ve faktoru institucionálního odcizení položka CARE, ve faktoru důvěry v instituce položka LEGAL, ve faktoru individuálního sociálního kapitálu položka TRUSTED (M.I. = 30, 37) a ve faktoru kolektivního sociálního kapitálu položka MEET (M.I. = 20, 22). Navíc všechny položky vykazují ve Španělsku a Portugalsku zcela shodné variance chyb měření a tudíž i srovnatelnou položkovou reliabilitu.

Opět se potvrdilo, že vyřazením dvou zemí (Francie a Itálie), které se od Španělska a Portugalska mírně kulturně i jazykově odchylovaly, přineslo určité zlepšení v rámci širší možné komparace dat. Jak analyzované vztahy, tak i srovnávání průměrů latentních proměnných může být dále validně interpretováno v rámci těchto dvou zemí. Ostatní položky, které nejsou ani v těchto dvou zemích plně skalárně ekvivalentní, budou pravděpodobně vystaveny jiným než kulturním zdrojům zkreslení dat, např. vlivu kontextu odpovědi a možná i jazykovým zdrojům chyb měření kvůli početným a silným regionálním úředním jazykům ve Španělsku.

Tabulka 51 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / ES, PT

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	ES	PT
individuální odcizení	INTEREST	2,919		3,016	2,768
	DIFFICULT	3,244	3,238		
	POLACTIV	4,156	4,151		
	OPINION	3,237		3,149	3,318
institucionální odcizení	VOTE	4,181	4,177		
	CARE	4,029		3,973	4,071
důvěra v instituce	POLIT	6,866	6,864		
	PARLAM	5,396	5,399		
	LEGAL	5,758		5,904	5,608
individuální sociální kapitál	TRUSTED	4,476		4,749	4,156
	FAIR	5,171	5,181		
	HELPFUL	4,130	4,137		
kolektivní sociální kapitál	FRIEND	8,421	8,422		
	VOLUNORG	5,588	5,586		
	MEET	5,483		5,296	5,692
	SOCACTIV	2,638	2,639		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

⁸⁵ Italština patří mezi tzv. východorománské jazyky. Španělština a portugalština se sice nacházejí ve stejné západorománské větvi jazyků jako francouzština, ale francouzština patří mezi tzv. gallo-románské jazyky a španělština a portugalština mezi ibersko-románské jazyky [Lewis et al. 2013].

Tabulka 52 Statistiky vhodnosti pro testy ekvivalence: shrnutí / ES, PT

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90% C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	761,95 (184)	-	0,938 (-)	0,036 (0,033;0,038)	0,920 (-)	-	přijat
Model 2 (Metrická ekvivalence)	792,82 (196)	30,87** (12)	0,936 (0,002)	0,035 (0,032;0,038)	0,922 (0,002)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	793,08 (197)	0,26 (1)	0,936 (0,00)	0,035 (0,032;0,037)	0,923 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1150,64 (213)	357,56*** (16)	0,900 (0,036)	0,042 (0,040;0,044)	0,887 (0,036)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	901,79 (207)	108,71*** (10)	0,926 (0,01)	0,037 (0,034;0,039)	0,914 (0,009)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	929,34 (217)	27,55** (10)	0,924 (0,002)	0,036 (0,034;0,039)	0,916 (0,002)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	957,54 (222)	28,20*** (5)	0,922 (0,002)	0,037 (0,034;0,039)	0,915 (0,001)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	1065,21 (238)	107,68*** (16)	0,912 (0,01)	0,037 (0,035;0,040)	0,911 (0,004)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Protože byla potvrzena alespoň částečná skalární ekvivalence všech faktorů, bylo možné přistoupit k porovnání jejich průměrů ve Španělsku a Portugalsku. V případě jejich komparace s Francií a Portugalskem mohou být strukturně analyzovány pouze faktory politického odcizení a důvěry v instituce. Údaje v tabulce 53, která tuto komparaci zaznamenává, opět odkrývají to, že strukturní a latentní průměry negenerují vždy stejné pořadí zemí. V případě kolektivního sociálního kapitálu, důvěry v instituce a politického odcizení není strukturní průměr schopen jasně stanovit pořadí zemí a srovnat je tak mezi sebou. Latentní průměry určují pořadí zemí jistěji, a proto se také lépe interpretují a vhodněji slouží srovnávacím analýzám.

Tabulka 53 Srovnání strukturních a latentních průměrů / ES, PT, IT, FR

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Španělsko	4,86**	0,050	1	0,534**	0,066	1
Portugalsko	4,32**	0,049	2	0,00	x	2
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Španělsko	5,47	0,031	1/2	-0,158**	0,044	2
Portugalsko	5,59	0,032	1/2	0,00	x	1
(ne)důvěra ve veřejné instituce						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Španělsko	5,8	0,055	1/2/3	-0,433**	0,084	3
Portugalsko	6,19**	0,054	1/2	0,00	x	1
Itálie	5,34**	0,055	3/4	-0,731**	0,084	4
Francie	5,7	0,051	2/3/4	-0,408**	0,080	2
individuální politické odcizení						
země	Strukturní průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Španělsko	3,41	0,022	1/2/3/4	0,113**	0,030	1
Portugalsko	3,34	0,024	1/2/3/4	0,00	x	2/3/4
Itálie	3,37	0,025	1/2/3/4	0,05	0,033	2/3/4
Francie	3,4	0,021	1/2/3/4	0,05	0,030	2/3/4
institucionální politické odcizení						
země	Strukturní průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Španělsko	4,86	0,032	2/3/4	-0,185**	0,032	2
Portugalsko	5,17**	0,030	1	0,00	x	1
Itálie	4,74	0,034	2/3/4	-0,275**	0,035	3
Francie	4,64	0,030	2/3/4	-0,407**	0,032	4

Zdroj: ESS 2002; vlastní výpočty

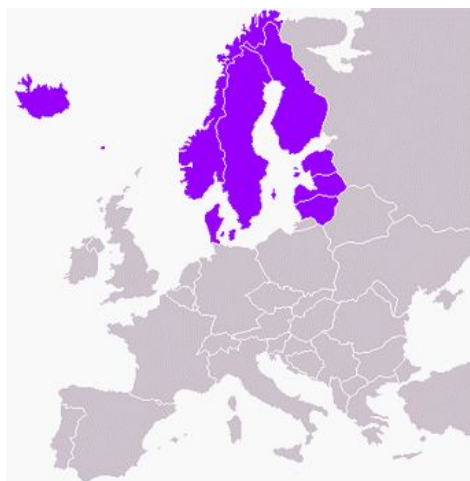
Poznámka: Strukturní průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 2: Španělsko a Francie jsou modře zvýrazněny, protože vychýlené konstanty poloviny položek mohou zkreslovat rozhodnutí o statistické významnosti rozdílů mezi latentními průměry.

Ve Španělsku je patrná vyšší míra individuálního sociálního kapitálu a individuálního politického odcizení než v Portugalsku. V Portugalsku zase vyšší míra kolektivního

sociálního kapitálu, nedůvěry ve veřejné instituce a větší institucionální odcizení než ve Španělsku (obdobný příklad jako při srovnání České republiky a Polska). Údaje ve Francii a Itálii ukazují na nepatrně větší důvěru ve veřejné instituce, ale v případě politického odcizení jsou velmi podobné hodnocení Španělska. Z hlediska reálných hodnot ukazatelů se jedná o průměrné až nadprůměrné hodnocení individuálního sociálního kapitálu (hodnocení cca 5 na škále od 0 - 10), vyšší míru nedůvěry ve veřejné instituce a nadprůměrné politické odcizení.

9.3 Simultánní analýza zemí severní Evropy



Dánské, norské a švédské království leží v tzv. Skandinávii, která označuje oblast v severní Evropě se společným kulturním, historickým a etno-lingvistickým dědictvím. Někdy se za součást Skandinávie v rozšířené definici považuje i Finsko nebo Island, ale pak už by se mělo hovořit spíše o tzv. severských zemích.

Dánština, norština a švédština spadají pod tzv. severogermánské jazyky, avšak finština už nespadá do větve indoevropských jazyků, ale do úplně jiné uralské rodiny jazyků, konkrétně baltofinských jazyků [Lewis et al. 2013]. Opět bylo možné si ověřit,

zda má tento rozkol mezi kulturou a jazykem stejný vliv na srovnatelnost měřicích položek ve zkoumaném modelu, jaký byl potvrzen v předešlých skupinách zemí.

Závěry z analýzy ekvivalence naznačují, že v těchto čtyřech zemích jsou data konfiguračně i metricky ekvivalentní, korelace chyb měření jsou shodné ve všech zemích a faktory spolu i podobně kovariují. Na druhou stranu nebyla opakovaně potvrzena ani plná skalární ekvivalence, ani plná ekvivalence variancí chyb měření (viz Tabulka 54).

Latentní proměnné politického odcizení a důvěry v instituce nedosáhly ani částečné skalární ekvivalence ($\Delta CFI = 0,076$), bylo nutné uvolnit podle modifikačních indexů (M.I. od 23 do 154) více jak dvě proměnné v daných faktorech. Latentní proměnné sociálního kapitálu se v severských zemích ukazují jako částečně ekvivalentní a tudíž srovnatelné. Ve faktoru kolektivního sociálního kapitálu byly uvolněny položky VOLUNORG v Dánsku a Finsku (M.I. = 31, 22) a MEET ve Finsku a Norsku (M.I. = 39, 123). Nejvíce konstant položek muselo být zbaveno restriktce shody ve Finsku, kde také nabývaly významně vyšších hodnot než v ostatních zemích; v Dánsku zase naopak nejnižších hodnot (viz Tabulka 55).

Výsledky jsou velice podobné údajům z komparace zemí střední Evropy, kde položky měřící faktory politického odcizení dosahovaly také jen stěží úroveň částečné skalární ekvivalence. Opět jsou zde pravděpodobně na vině jazykové a kulturní odlišnosti zkoumaných severských zemích, i když byly cíleným výběrem zemí co nejvíce eliminovány. Ve Finsku se hovoří v dnešní době dvěma úředními jazyky a to finštinou a švédštinou, protože Finsko bylo celá staletí součástí Švédského království; z tohoto důvodu byly také dotazníky ESS v této zemi nabízeny ve švédské i finské verzi. Finština je ale lingvisticky zcela jiný jazyk řadící se do rodiny uralských jazyků, kterým dnes ve Finsku hovoří většina obyvatel. Pro menšinu obyvatel je mateřským jazykem stále švédština, přičemž některé další menšiny Finů hovoří také ruštinou, estonštinou nebo v Laponsku dokonce sámským jazykem [Lewis et al. 2013].

Tabulka 54 Statistiky vhodnosti pro testy ekvivalence: shrnutí / DK, FI, NO, SE

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1953,78 (368)	-	0,940 (-)	0,025 (0,024;0,026)	0,922 (-)	-	přiját
Model 2 (Metrická ekvivalence)	2129,82 (404)	176,04*** (36)	0,935 (0,005)	0,025 (0,023;0,026)	0,923 (0,00)	Model 1	přiját
Model 2a (Ekvivalence korelace chyb měření)	2130,40 (407)	0,58 (3)	0,935 (0,00)	0,024 (0,023;0,025)	0,923 (0,00)	Model 2	přiját
Model 3 (Skalární ekvivalence)	4189,71 (455)	2059,31*** (48)	0,859 (0,076)	0,034 (0,033;0,035)	0,851 (0,072)	Model 2a	nepřiját
Model 3a (Částečná skalární ekvivalence)	2395,15 (421)	264,76*** (14)	0,925 (0,01)	0,026 (0,025;0,027)	0,915 (0,008)	Model 2a	přiját
Model 4 (Ekvivalence kovariancí faktorů)	2495,54 (451)	100,39*** (30)	0,923 (0,002)	0,025 (0,024;0,026)	0,918 (0,00)	Model 3a	přiját
Model 5 (Ekvivalence variancí faktorů)	2583,89 (466)	88,35*** (15)	0,920 (0,003)	0,025 (0,024;0,026)	0,918 (0,00)	Model 4	přiját
Model 6 (Ekvivalence variancí chyb měření)	3195,42 (514)	611,53*** (48)	0,899 (0,021)	0,027 (0,026;0,028)	0,905 (0,013)	Model 5	nepřiját
Model 7 (Částečná ekvivalence variancí chyb měření)	2897,14 (509)	313,25*** (43)	0,910 (0,01)	0,025 (0,024;0,026)	0,915 (0,003)	Model 5	přiját

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Tabulka 55 Odhady konstant ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / DK, FI, NO, SE

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	DK	FI	NO	SE
individuální odcizení	INTEREST	2,423		2,247	2,549	2,464	2,341
	DIFFICULT	3,073		2,947	3,422	2,979	2,916
	POLACTIV	3,307		2,695	3,779	3,218	3,212
	OPINION	2,889		2,506	2,921	2,920	3,104
institucionální odcizení	VOTE	3,202		3,119	3,408	3,091	3,149
	CARE	3,079		3,077	3,235	3,027	2,966
důvěra v instituce	POLIT	5,146		4,622	5,138	5,434	5,117
	PARLAM	4,120		3,908	4,126	4,306	3,901
	LEGAL	3,459		2,943	3,204	3,678	3,817
individuální sociální kapitál	TRUSTED	6,519	6,545				
	FAIR	6,939	7,001				
	HELPFUL	5,913	6,019		5,721		
kolektivní sociální kapitál	FRIEND	8,781	8,809				
	VOLUNORG	4,215	4,226	3,794	4,431		
	MEET	5,426	5,341		5,219	5,745	
	SOCACTIV	2,901	2,901				

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Norsko je zase v dnešní době známé tím, že významná část jeho obyvatel má zahraniční původ. Velmi početnými přistěhovaleckými skupinami jsou kromě Švédů také Poláci, Litevci, Iráčané a Somálci [<http://www.migrationinformation.org>]. Tudiž kromě norštiny a sámštiny hovoří řada dalších Norů jiným rodným jazykem, a proto dotazník pouze v norštině již možná v dnešní době pro potřeby mezinárodní komparace Norska s Evropou nestačí.⁸⁶ Ve Švédsku najdeme podobné jazykové odlišnosti, z menšin jsou zde nejpočetnější vedle Finů, Dánů a Sámů přistěhovalci z Íránu, Íráku a Turecka. Méně kritická imigrační situace platí pro Dánsko s dánštinou jako hlavním úředním jazykem.

Všechny tyto jazykové zvláštnosti mohou být důvodem toho, proč se určité měřicí položky v těchto zemích natolik vychylují při komparaci konstant; vše hovoří pro oslabenou jazykovou dovednost některých respondentů hovořících jedním z menšinových jazyků a pro značnou obtížnost překladu a nepřesnou interpretaci nespécifických výrazů obsažených v daných otázkách.

Z hlediska testování modelu ekvivalence variancí chyb modelu lze usoudit na poměrně dobré výsledky. Ve všech faktorech bylo dosaženo částečné ekvivalence chyb měření s tím, že nejvíce chyb měření muselo být opět uvolněno ve Finsku – jedná se o položky DIFFICULT (M.I. = 43), CARE (M.I. = 35) a SOCACTIV (M.I. = 94). Na základě jejich velikostí je možné zhodnotit reliabilitu položek měření, které se ve Finsku jeví jako nejméně reliabilní (vzhledem k ostatním zemím) – jejich chyby měření jsou vyšší (viz Tabulka 56).

⁸⁶ Stejný případ se začíná objevovat i v Německu, kde má přistěhovalecký původ přibližně 17 % obyvatel Německa – nejvýraznější skupinou s jinou státní příslušností začínají být v Německu Turci, Italové a Jugoslávci [<http://www.migrationinformation.org>].

Tabulka 56 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / DK, FI, NO, SE

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	DK	FI	NO	SE
individuální odcizení	INTEREST	0,326	0,325				
	DIFFICULT	0,800	0,728		0,996		
	POLACTIV	1,095	1,094				
	OPINION	0,694	0,696				
institucionální odcizení	VOTE	0,516	0,52				
	CARE	0,436	0,447	0,594	0,317		
důvěra v instituce	POLIT	1,238	1,236				
	PARLAM	1,304	1,303				
	LEGAL	2,873	2,973				
individuální sociální kapitál	TRUSTED	1,943	1,943				
	FAIR	1,649	1,649				
	HELPFUL	2,935	2,935				
kolektivní sociální kapitál	FRIEND	1,647	1,644				
	VOLUNORG	6,716	6,709				
	MEET	1,584	1,585				
	SOCACTIV	0,518	0,512		0,749	0,347	

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

I v tomto případě bylo možné přistoupit k ještě většímu zúžení zkoumané skupiny zemí a pokusit se nalézt plnou skalární ekvivalenci položek. Z předešlého testu skalárního modelu vyplývá, že nejproblematictější zemí je Finsko, bylo zde zapotřebí uvolnit nejvíce konstant položek. Avšak v průběhu další analýzy modelů se ukazuje, že ani za přítomnosti Dánska nelze dosáhnout alespoň částečné skalární ekvivalence položek ve všech latentních proměnných. Proto z dalšího testovaného souboru skupin nebylo vyřazeno jen Finsko, ale i Dánsko.

Testy ekvivalence položek v Norsku a Švédsku konečně potvrzují částečnou skalární ekvivalenci ve všech faktorech zahrnutých do modelu, některé konstanty ještě stále způsobovaly ve skalárním modelu příliš velký rozdíl v hodnotách indexu CFI (Δ CFI = 0,027, viz Tabulka 57). Z faktoru individuálního odcizení byla uvolněna na základě modifikačních indexů konstanta položky OPINION (M.I. = 27, 32), z faktoru důvěry v instituce položka LEGAL (M.I. = 16, 17), z faktoru individuálního sociálního kapitálu položka TRUSTED (M.I. = 19, 24) a z faktoru kolektivního sociálního kapitálu MEET (M.I. = 49, 59). Latentní proměnná institucionálního odcizení je dokonce plně skalárně ekvivalentní (viz Tabulka 58). Testy modelů dále odhalily, že všechny variance chyb měření těchto položek jsou v Norsku a Švédsku shodné a že jejich položková reliabilita je tudíž srovnatelná, což znamená, že byla v datech identifikována tzv. striktní ekvivalence.

Opětovně se dokazuje, že skalární ekvivalence je mnohem více citlivá na kulturní i jazykové rozdíly mezi zeměmi než metrická ekvivalence. Po vyloučení Finska (jazykově nejvychýlenější země) a Dánska (kulturně a historicky nejvzdálenější země) byla identifikována částečná skalární ekvivalence ve všech latentních proměnných. Ale i přes jistou kulturní a jazykovou spřízněnost Norska a Švédska zůstávají čtyři měřicí položky stále skalárně nesrovnatelné. Důvody mohou být opět mnohé od kontextu odpovědi po oslabenou jazykovou dovednost početných národnostních menšin žijících v těchto zemích, ale jisté je to, že ztěžují interpretaci daných otázek a nevhodně posouvají hodnocení respondentů na škále odpovědí, i když dané položky stále měří jeden a ten samý koncept.

Tabulka 57 Statistiky vhodnosti pro testy ekvivalence: shrnutí / NO, SE

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1034,27 (184)	-	0,937 (-)	0,035 (0,033;0,037)	0,918 (-)	-	přijat
Model 2 (Metrická ekvivalence)	1078,45 (196)	44,18*** (12)	0,935 (0,002)	0,034 (0,032;0,036)	0,920 (0,002)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	1078,79 (197)	0,34 (1)	0,935 (0,00)	0,034 (0,032;0,036)	0,921 (0,001)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1454,46 (213)	375,67*** (16)	0,908 (0,027)	0,039 (0,037;0,041)	0,897 (0,024)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1199,32 (209)	120,54*** (12)	0,927 (0,008)	0,035 (0,033;0,037)	0,916 (0,005)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1232,23 (219)	32,91*** (10)	0,925 (0,002)	0,035 (0,033;0,037)	0,918 (0,002)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1241,40 (224)	9,17 (5)	0,925 (0,00)	0,034 (0,033;0,036)	0,919 (0,00)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	1349,93 (240)	108,53*** (16)	0,918 (0,007)	0,035 (0,033;0,037)	0,918 (0,001)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Tabulka 58 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / NO, SE

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	NO	SE
individuální odcizení	INTEREST	2,407	2,409		
	DIFFICULT	2,951	2,952		
	POLACTIV	3,213	3,216		
	OPINION	3,004		2,900	3,130
institucionální odcizení	VOTE	3,131	3,133		
	CARE	3,010	3,012		
důvěra v instituce	POLIT	5,319	5,325		
	PARLAM	4,158	4,162		
	LEGAL	3,784		3,608	3,972
individuální sociální kapitál	TRUSTED	6,402		6,567	6,186
	FAIR	6,842	6,839		
	HELPFUL	6,015	6,011		
kolektivní sociální kapitál	FRIEND	8,694	8,694		
	VOLUNORG	4,237	4,238		
	MEET	5,556		5,760	5,318
	SOCACTIV	2,931	2,936		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Pravděpodobně se opravdu jedná o ideální souhru kulturních a jazykových podobností mezi zeměmi, která je tolik podstatná pro dosažení plné skalární ekvivalence položek. Například Finsko má společné ugrofinské jazykové kořeny pouze s Maďarskem, proto byla otestována ekvivalence položek i v těchto dvou zemích. I přes podobný lingvistický základ jazyka byla v daných zemích prokázána pouze konfigurální a metrická ekvivalence dat, ale částečná skalární ekvivalence položek se neobjevila ani u jednoho faktoru. Tudíž nejsou podstatné jen společné jazykové kořeny daných skupin či zemí, ale i kulturní kontext života respondentů, který je utvářen historicky proměnlivými sociálními, ekonomickými a politickými vlivy.

Srovnání strukturálních a latentních průměrů mezi testovanými zeměmi severní Evropy opět odhaluje skutečnost, že strukturální průměry nejsou mezi sebou ve vybraných zemích statisticky významně odlišné a že podle nich tedy není možné určit pořadí zemí v rámci hodnocení daných konceptů. Latentní průměry nejsou schopny určit pořadí Norska a Švédska pouze v rámci hodnocení politického odcizení.

Co se týče konceptu sociálního kapitálu, Norsko a Dánsko vykazuje obecně vyšší míru individuálního sociálního kapitálu než Švédsko a Finsko, resp. více mezilidské důvěry. V případě sociálních vztahů si nejhůře vede Finsko. Na druhou stranu je také patrné, že v Norsku je vyšší míra nedůvěry ve veřejné instituce než ve Švédsku. Podle velikostí strukturálních průměrů lze soudit, že míra politického odcizení je v těchto zemích spíše průměrná (cca 2,9 na škále hodnocení u individuálního odcizení, cca 3,8 u institucionálního odcizení). Navíc na základě statisticky nevýznamných rozdílů mezi průměry lze usoudit, že Norsko a Švédsko jsou v rámci hodnocení sociálního kapitálu a politického odcizení velice podobné země.

Tabulka 59 Srovnání strukturních a latentních průměrů / NO, SE, DK, FI

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Norsko	6,53	0,033	1/2/3/3	0,305**	0,052	2
Švédsko	6,27**	0,038	2/3/4	0,00	x	4
Dánsko	6,85**	0,042	1/2	0,667**	0,058	1
Finsko	6,33	0,035	2/3/4	0,142**	0,052	3
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Norsko	5,44	0,022	1/2/3/4	0,04**	0,019	1
Švédsko	5,26	0,025	1/2/3/4	0,00	x	2/3
Dánsko	5,28	0,027	1/2/3/4	0,017	0,021	2/3
Finsko	5,27	0,024	1/2/3/4	-0,069**	0,021	4
(ne)důvěra ve veřejné instituce						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Norsko	4,46	0,039	1/2	0,206**	0,067	1
Švédsko	4,38	0,044	1/2	0,00	x	2
individuální politické odcizení						
země	Strukturní průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Norsko	2,89	0,016	1/2	0,017	0,020	1/2
Švédsko	2,91	0,018	1/2	0,00	x	1/2
institucionální politické odcizení						
země	Strukturní průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Norsko	3,84	0,025	1/2	-0,022	0,029	1/2
Švédsko	3,80	0,028	1/2	0,00	x	1/2

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturní průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 2: Finsko je modře zvýrazněno, protože vychýlené konstanty poloviny položek mohou zkreslovat rozhodnutí o statistické významnosti rozdílů mezi latentními průměry.

9.4 Simultánní analýza zemí západní Evropy



Ostrovní část západní Evropy tvoří Spojené království a Irsko, pevninskou část pak Francie, Nizozemsko, Belgie a Lucembursko. Tyto země mají společnou kulturu i historii, přesto je bylo vhodné rozdělit ještě do menších skupin i na základě jazykové podobnosti.

V první skupině byla otestována ekvivalence položek ve Spojeném království, Irsku a Nizozemsku. Angličtina, skotština i nizozemština se řadí do tzv. zápogermánských jazyků. Jako kontrastní země zde vystupovalo Irsko, neboť irština, kterou se (menšinově) hovoří v Irské republice a Severním Irsku, se řadí mezi keltské jazyky; stejně tak i skotská gaelština [Lewis et al. 2013].

Ve druhé skupině byly otestovány modely v zemích Beneluxu (Belgii, Nizozemsku a Lucembursku), neboť tyto země pojí hospodářské a ekonomické zájmy a řadí se mezi nejvyspělejší v Evropě. Jednalo se v tomto případě tedy spíše o test významu kulturní shody než jazykové příbuznosti. Ve třetí skupině – Francii, Belgii a Lucembursku – byly položky analyzovány kvůli velmi blízké jazykové shodě daných zemí. V Belgickém království jsou oficiálními jazyky nizozemština, francouzština a němčina, v Lucembursku je to pak lucemburština, francouzština a němčina.

9.4.1 Spojené království, Irsko a Nizozemsko

Jako první skupina testovaných zemí byla vybrána kombinace Spojeného království Velké Británie a Severního Irsku, Irské republiky a Nizozemska. V tomto souboru zemí dochází k zajímavému mixu stejných kulturních základů s odlišnými jazykovými kořeny zápogermánských jazyků (angličtina, skotština i nizozemština) a keltských jazyků (irština), které navíc nejsou jasně územně ohraničené; některá nářečí pocházející z keltských jazyků najdeme i ve Spojeném království (např. v Severním Irsku a Skotsku). Souhrnné výsledky analýzy jsou zaznamenány v Tabulce 60.

Stejně jako v předchozích analýzách byla v datech identifikována konfigurální a plně metrická ekvivalence. Lze tedy dále v těchto třech zemích bez obav testovat vztahy latentních proměnných s dalšími metricky ekvivalentními proměnnými z výzkumu. Navíc se také potvrdila konstruktová validita testovaného modelu, neboť ve všech zemích jsou vzájemné vztahy latentních proměnných podobné a srovnatelné. Plná skalární ekvivalence opět nebyla potvrzena ($\Delta CFI = 0,069$) a stejně tak ani ekvivalence chyb měření ($\Delta CFI = 0,031$).

Částečnou skalární ekvivalenci vykazují, tak jako tomu bylo i v předešlých skupinách zemí, pouze faktory sociálního kapitálu. U latentních proměnných politického odcizení bylo nutné odstranit podmínku shody všech konstant měřicích položek; ty nejsou tedy nadále v analyzovaných zemích srovnatelné na základě průměrů (viz Tabulka 61). V rámci individuálního sociálního kapitálu bylo nezbytné na základě modifikačních indexů uvolnit konstantu jen jedné položky HELPFUL v Nizozemsku (M.I. = 114). Ve faktoru kolektivního sociálního kapitálu muselo dojít ke zrušení restriktce u položky VOLUNORG ve všech zemích (M.I. od 76 do 137) a položky FRIEND v Nizozemsku (M.I. = 120).

Tabulka 60 Statistiky vhodnosti pro testy ekvivalence: shrnutí / UK, IE, NL

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	1330,38 (276)	-	0,955 (-)	0,025 (0,024;0,027)	0,941 (-)	-	přijat
Model 2 (Metrická ekvivalence)	1501,41 (300)	171,03*** (24)	0,948 (0,007)	0,026 (0,025;0,027)	0,938 (0,003)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	1545,13 (302)	43,73*** (2)	0,946 (0,002)	0,026 (0,025;0,028)	0,936 (0,002)	Model 2	přijat
Model 3 (Skalární ekvivalence)	3185,48 (334)	1640,34*** (32)	0,877 (0,069)	0,038 (0,037;0,039)	0,868 (0,068)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1796,83 (312)	251,70*** (10)	0,936 (0,01)	0,028 (0,027;0,029)	0,926 (0,01)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1867,48 (332)	70,65*** (20)	0,934 (0,002)	0,028 (0,027;0,029)	0,928 (0,002)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1995,45 (342)	127,97*** (10)	0,929 (0,005)	0,029 (0,027;0,030)	0,925 (0,003)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	2745,78 (374)	750,34 (32)	0,898 (0,031)	0,033 (0,032;0,034)	0,902 (0,023)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	2201,41 (365)	205,99*** (23)	0,921 (0,008)	0,029 (0,028;0,030)	0,922 (0,003)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Velikosti konstant položek měřících faktory politického odcizení a důvěry v instituce jsou ve Spojeném království a Irské republice výrazně vyšší než je tomu v Nizozemsku, proto se Nizozemsko v této úrovni analýzy tolik odlišuje. Příčina se pravděpodobně skrývá opět v kombinaci kulturních a jazykových odlišnostech, neboť Spojené království a Irsko mají jednak společný (i když bouřlivý) historický vývoj, a také společný jazykový mix angličtiny (germánského jazyka) a irštiny a skotštiny (keltského jazyka). To vše z nich tvoří silnou dvojici vystupující v opozici proti menší zemi s hlavním úředním jazykem nizozemštiny patřící do Nizozemského království v pevninské části Evropy, které bylo v historii ohrožováno především Francií. Může zde tedy hrát roli jak specifický styl a kontext odpovědí respondentů, tak i obtížně přeložitelné výrazy v otázkách. Položka VOLUNORG z faktoru kolektivního sociálního kapitálu může být zatížena lehčím sociálně desirabilním charakterem, který se potom odráží na srovnatelnosti dat.

Tabulka 61 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / UK, IE, NL

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	UK	IE	NL
individuální odcizení	INTEREST	2,432		2,492	2,638	2,266
	DIFFICULT	3,098		3,201	3,085	3,022
	POLACTIV	3,787		3,667	3,841	3,848
	OPINION	2,958		2,858	2,895	3,055
institucionální odcizení	VOTE	3,632		3,690	3,804	3,488
	CARE	3,411		3,414	3,548	3,339
důvěra v instituce	POLIT	5,737		6,024	6,267	5,270
	PARLAM	5,132		5,155	5,579	4,904
	LEGAL	4,789		4,823	4,883	4,724
individuální sociální kapitál	TRUSTED	5,439	5,465			
	FAIR	5,923	5,947			
	HELPFUL	5,492	5,755			5,129
kolektivní sociální kapitál	FRIEND	8,657	8,830			8,324
	VOLUNORG	4,493		3,724	4,599	4,956
	MEET	5,169	5,169			
	SOCACTIV	2,781	2,781			

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Variance chyb měření nesplňovaly požadavek ani částečné rovnosti kvůli jejich příliš rozdílné velikosti u proměnných především ve faktoru důvěry ve veřejné instituce – položky POLIT a LEGAL (M.I. od 41 do 59; M.I. od 55 do 87) musely být uvolněny ve všech zemích (viz Tabulka 62). Všechny ostatní faktory splňují požadavky alespoň částečné ekvivalence variance chyb měření. Největší velikost chyb měření ve faktoru důvěry v instituce a individuálního sociálního kapitálu lze přitom nalézt u Irska. V Nizozemsku se naopak ukazují menší chyby měření v rámci faktoru individuálního odcizení. Zdá se proto, že měřicí položky v Nizozemsku vykazují nejvyšší míru reliability v porovnání s ostatními zeměmi.

Tabulka 62 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / UK, IE, NL

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	UK	IE	NL
individuální odcizení	INTEREST	0,360	0,419			0,281
	DIFFICULT	0,869	0,859			
	POLACTIV	1,101	1,106			
	OPINION	0,802	0,890			0,637
institucionální odcizení	VOTE	0,476	0,477			
	CARE	0,367	0,365			
důvěra v instituce	POLIT	1,236		1,186	1,719	0,859
	PARLAM	1,686	1,527		2,250	
	LEGAL	3,348		3,604	4,067	2,528
individuální sociální kapitál	TRUSTED	2,499	2,146		3,287	
	FAIR	1,749	1,770			
	HELPFUL	2,830	2,825			
kolektivní sociální kapitál	FRIEND	2,029	2,015			
	VOLUNORG	6,949	6,978			
	MEET	1,751	1,746			
	SOCACTIV	0,665	0,747		0,509	

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Vyřazení Nizozemska jako kulturně i jazykově lehce vychýlené země v dané skupině z další analýzy se opět ukázalo jako prospěšné pro možnosti srovnávání dat mezi Spojeným královstvím a Irskem. Dokonce v tomto případě dosáhly téměř všechny měřicí položky plně skalární ekvivalence, což je poměrně ojedinělý případ (viz Tabulka 63).

Tabulka 63 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / UK, IE

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	UK	IE
individuální odcizení	INTEREST	2,565	2,565		
	DIFFICULT	3,156	3,156		
	POLACTIV	3,759	3,760		
	OPINION	2,880	2,881		
institucionální odcizení	VOTE	3,759	3,765		
	CARE	3,494	3,501		
důvěra v instituce	POLIT	6,194	6,208		
	PARLAM	5,398	5,411		
	LEGAL	4,900	4,909		
individuální sociální kapitál	TRUSTED	5,248	5,233		
	FAIR	5,768	5,754		
	HELPFUL	5,643	5,631		
kolektivní sociální kapitál	FRIEND	8,817		8,577	9,007
	VOLUNORG	4,119		3,621	4,616
	MEET	5,106	5,092		
	SOCACTIV	2,778	2,771		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Tabulka 64 Statistiky vhodnosti pro testy ekvivalence: shrnutí / UK, IE

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	831,78 (184)	-	0,955 (-)	0,031 (0,029;0,033)	0,941 (-)	-	přijat
Model 2 (Metrická ekvivalence)	860,62 (196)	28,84** (12)	0,953 (0,002)	0,030 (0,028;0,032)	0,943 (0,002)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	870,37 (197)	9,75** (1)	0,952 (0,001)	0,030 (0,028;0,032)	0,942 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1196,52 (213)	326,15*** (16)	0,931 (0,021)	0,035 (0,033;0,037)	0,922 (0,02)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1031,91 (211)	161,53*** (14)	0,942 (0,01)	0,032 (0,030;0,034)	0,934 (0,008)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1059,29 (221)	27,38** (10)	0,941 (0,001)	0,032 (0,030;0,034)	0,936 (0,002)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1109,33 (226)	50,04*** (5)	0,938 (0,003)	0,032 (0,030;0,034)	0,934 (0,00)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	1345,70 (242)	236,37*** (16)	0,923 (0,015)	0,035 (0,033;0,037)	0,923 (0,011)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	1233,76 (240)	124,43*** (14)	0,930 (0,008)	0,033 (0,031;0,035)	0,930 (0,004)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Zbývající dvě položky FRIEND a VOLUNORG z kolektivního sociálního kapitálu nemusí být stále srovnatelné díky referenčnímu efektu nebo specifickému kontextu odpovědi. Nezanedbatelný vliv může mít v tomto případě také jazyková bariéra početných národních skupin, jako jsou Angličané, Skotové, Velšané a Irové, které jsou v dnešní době doplňovány o řady imigrantských skupin neevropského původu ve Spojeném království (z Indie, Pákistánu, Bangladěše, Afriky, atd.) a evropského původu v Irsku (z Polska, Litvy, Lotyšska, Německa, atp.) [http://www.migrationinformation.org]. Je pravděpodobné, že i drobné jazykové nesrovnalosti se významně odráží na velikostech konstant položek, které se pak stávají hůře srovnatelnými.

Protože se potvrdila částečná skalární ekvivalence položek, bylo možné přistoupit ke komparaci strukturních a latentních průměrů všech proměnných ve Spojeném království a Irsku a průměrů faktorů sociálního kapitálu i v Nizozemí. Ze shrnutí tabulky 65 je možné usoudit, že výsledky analýzy strukturních a latentních průměrů jsou nyní zcela totožné, ale že někdy stále není možné jasně rozhodnout o pořadí zemí podle hodnocení daných konceptů.

Tabulka 65 Srovnání strukturních a latentních průměrů / UK, IE, NL

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Spojené království	5,35**	0,039	3	-0,467**	0,064	3
Irsko	5,81**	0,045	1/2	0,00	x	1/2
Nizozemsko	5,73	0,034	1/2	0,034	0,061	1/2
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Spojené království	5,02**	0,025	3	-0,178**	0,023	3
Irsko	5,36**	0,024	1/2	0,00	x	1/2
Nizozemsko	5,35	0,023	1/2	-0,028	0,017	1/2
(ne)důvěra ve veřejné instituce						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Spojené království	5,5	0,044	1/2	-0,094	0,074	1/2
Irsko	5,55	0,049	1/2	0,00	x	1/2
individuální politické odcizení						
země	Strukturní průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Spojené království	3,08	0,018	1/2	-0,069	0,026	1/2
Irsko	3,11	0,019	1/2	0,00	x	1/2
institucionální politické odcizení						
země	Strukturní průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Spojené království	4,48	0,027	1/2	-0,045	0,028	1/2
Irsko	4,52	0,029	1/2	0,00	x	1/2

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturní průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 2: Země jsou ve faktoru kolektivního sociálního kapitálu modře zvýrazněny, protože vychýlené konstanty poloviny položek mohou zkreslovat rozhodnutí o statistické významnosti rozdílů mezi latentními průměry.

U politického odcizení a důvěry v instituce není kvůli statisticky nevýznamným rozdílům mezi průměry možné určit, zda se ve svých postojích vůbec dané země liší. Tudíž důvěra ve veřejné instituce se jeví ve Spojeném království a Irsku jako průměrná a shodná (cca hodnota 5,5 na škále od 0 do 10) a politické odcizení jako lehce nadprůměrné. Co se týče sociálního kapitálu, jak mezilidská důvěra, tak sociální vazby se odkrývají v Irsku jako nepatrně silnější než ve Spojeném Království a v hodnocení spíše nadprůměrné; s výsledky Irska korespondují i údaje z Nizozemska.

Rozdílné chyby měření je možné identifikovat ve Spojeném království a Irsku pouze u položky POLIT (M.I. = 20, 22) ve faktoru důvěry v instituce a položky SOCACTIV (M.I. = 31, 34) ve faktoru kolektivního sociálního kapitálu s tím, že se ukazuje, že Irsko má spíše větší chyby měření a v tomto případě menší reliabilitu daných položek (viz Tabulka 66). Všechny ostatní položky dosáhly v těchto zemích tzv. striktní ekvivalence a jejich chyby měření a potažmo reliabilita jsou totožné.

Tabulka 66 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / UK, IE

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	UK	IE
individuální odcizení	INTEREST	0,361	0,361		
	DIFFICULT	0,987	0,986		
	POLACTIV	1,121	1,121		
	OPINION	0,923	0,923		
institucionální odcizení	VOTE	0,495	0,497		
	CARE	0,383	0,381		
důvěra v instituce	POLIT	1,471		1,148	1,803
	PARLAM	1,714	1,721		
	LEGAL	3,871	3,866		
individuální sociální kapitál	TRUSTED	2,825	2,826		
	FAIR	1,845	1,844		
	HELPFUL	2,916	2,916		
kolektivní sociální kapitál	FRIEND	1,862	1,848		
	VOLUNORG	7,479	7,485		
	MEET	1,844	1,841		
	SOCACTIV	0,675		0,808	0,549

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

9.4.2 Benelux – Nizozemsko, Belgie a Lucembursko

Nizozemsko, které bylo vyřazeno z předešlé analýzy Spojeného království a Irsku, lze kulturně zařadit i do skupiny zemí Beneluxu. Belgie a Lucembursko byly z historického hlediska dříve součástí Nizozemského království, od kterého se v 19. století oddělily, a v současné době spolu tyto země tvoří volné společenství založené na principu hospodářské unie. I když oficiálními jazyky Belgie a Lucemburska jsou francouzština a němčina (a nizozemština v Belgii), Nizozemsko se k nim územně tradičně řadí, proto bylo zařazeno i do této skupiny zemí. Výsledky analýzy jejich srovnatelnosti ukazuje tabulka 67.

Tabulka 67 Statistiky vhodnosti pro testy ekvivalence: shrnutí / NL, BE, LU

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	990,03 (276)	-	0,960 (-)	0,023 (0,022;0,025)	0,948 (-)	-	přijat
Model 2 (Metrická ekvivalence)	1102,37 (300)	112,34*** (24)	0,955 (0,005)	0,024 (0,022;0,025)	0,946 (0,00)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	1103,28 (302)	0,91 (2)	0,955 (0,00)	0,023 (0,022;0,025)	0,947 (0,001)	Model 2	přijat
Model 3 (Skalární ekvivalence)	2251,19 (334)	1147,92*** (32)	0,893 (0,062)	0,034 (0,033;0,036)	0,884 (0,063)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1288,63 (316)	185,36*** (14)	0,946 (0,009)	0,025 (0,024;0,027)	0,938 (0,009)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1320,31 (336)	31,68** (20)	0,945 (0,00)	0,025 (0,023;0,026)	0,941 (0,003)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1380,87 (346)	60,55*** (10)	0,942 (0,003)	0,025 (0,024;0,026)	0,940 (0,001)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	2292,22 (378)	911,36*** (32)	0,893 (0,049)	0,032 (0,031;0,034)	0,898 (0,042)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	1570,77 (366)	189,90*** (20)	0,933 (0,009)	0,026 (0,025;0,027)	0,934 (0,006)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015.

U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

I ve spojení Nizozemska s Belgií a Lucemburskem nedosáhly měřicí položky ani plné skalární ekvivalence (Δ CFI = 0,062), ani plné ekvivalence chyb měření (Δ CFI = 0,049). Částečně skalární se ukázaly být pouze položky faktorů politického odcizení a důvěry v instituce. V latentních proměnných sociálního kapitálu musely být uvolněny konstanty většiny položek, podobně jako ve skupině zemí z jižní Evropy. V případě individuálního politického odcizení musela být uvolněna na základě modifikačních indexů jen jedna položka INTEREST (M.I. od 19 do 62) a ve faktoru důvěry v instituce položka LEGAL (M.I. od 114 do 131), viz Tabulka 68. Konstanty položek měřících sociální kapitál jsou v Nizozemsku vyšší než v Belgii či Lucembursku, naopak konstanty položek měřící politické odcizení jsou nižší. Nizozemsko je tedy tou zemí, která se v tomto případě opět vychyluje.

Tabulka 68 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / NL, BE, LU

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	NL	BE	LU
individuální odcizení	INTEREST	2,391		2,280	2,534	2,497
	DIFFICULT	3,029	3,026			
	POLACTIV	3,867	3,868			
	OPINION	3,100	3,098			
institucionální odcizení	VOTE	3,573	3,572			
	CARE	3,378	3,378			
důvěra v instituce	POLIT	5,326	5,309			
	PARLAM	4,797	4,770			
	LEGAL	4,784		4,698	5,349	4,026
individuální sociální kapitál	TRUSTED	5,386		5,691	5,013	5,187
	FAIR	5,909		6,128	5,743	5,501
	HELPFUL	4,895		5,233	4,545	4,470
kolektivní sociální kapitál	FRIEND	8,421		8,338	8,345	8,756
	VOLUNORG	5,326		4,974	5,158	6,832
	MEET	5,208		5,266	5,186	5,043
	SOCACTIV	2,744		2,764	2,675	2,804

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Zajímavá je také skutečnost, že v souboru zemí Nizozemsko, Spojené království a Irsko dosáhly částečné skalární ekvivalence pouze položky měřící koncept sociálního kapitálu, a v tomto spojení zemí - Nizozemska s Belgií a Lucemburskem - se naopak odhaluje částečná skalární ekvivalence položek ve zbývajících faktorech politického odcizení a důvěry v instituce. Nizozemsko je takto poněkud dichotomicky rozděleno a lze v jeho případě srovnávat dané koncepty podle průměrů ve dvou kulturně a jazykově odlišných skupinách.⁸⁷ Je velmi obtížné identifikovat skutečný zdroj zkreslení rozdílné interpretace škálových bodů v Nizozemsku a vysvětlit, proč vlastně dochází k tomu, že Nizozemci vnímají a interpretují měřicí škály položek sociálního kapitálu shodně jako respondenti ze Spojeného království a Irska, a na druhou stranu, proč interpretují škály měření otázek politického odcizení a institucionální důvěry podobně jako respondenti z Belgie a Lucemburska.

Na jednu stranu pojí Nizozemsko s ostrovní částí Evropy jednak společné západogermánské lingvistické kořeny nizozemštiny a angličtiny, navíc velké množství Nizozemců hovoří také

⁸⁷ Nizozemsko bylo také otestováno ve skupině společně jen s Belgií, ale lepší výsledky skalární ekvivalence vygenerovány nebyly. V těchto dvou zemích data neukázala částečnou skalární ekvivalenci ve všech faktorech – některé latentní proměnné se tak stále objevovaly jako nesrovnatelné z hlediska faktorových průměrů.

anglicky bez větších problémů, a jednak poměrně velká ekonomická vyspělost. Tyto společné prvky ale nebrání výskytu problémů se srovnatelností konstant položek měřících politické odcizení. Příčinou mohou být jak odlišnosti v historických a kulturních souvislostech, tak i v rozdílné imigrační politice v Nizozemsku, kde nejsou přistěhovalecké skupiny zdaleka tak početné jako ve Spojeném království. Menší množství imigrantů hovořících jiným rodným jazykem současně snižuje potenciální jazykové bariéry při vyplňování dotazníku.

Na druhou stranu je Nizozemsko pevně historicky i kulturně spojeno s Belgií a Lucemburskem. Jedním z úředních jazyků je navíc v Belgii i nizozemština, ve které byly kromě francouzštiny distribuovány také dotazníky ESS. V Belgii a Lucembursku se sice hovoří i francouzsky a německy, ale většina Nizozemců těmito jazyky hovoří též. Přistěhovalecké skupiny nejsou v těchto zemích nijak početné, většinou se jedná o Italy, Francouze, Němce, Portugaly a Turky [http://www.migrationinformation.org]. Jediný problém souvisí s regionálními variantami nizozemštiny, francouzštiny a němčiny, které vykazují drobné rozdíly ve slovní zásobě a sémantice.⁸⁸ Bez hlubší studie ale není nemožné jednoznačně říci, proč dochází k rozdílné interpretaci měřících škál těchto položek tak specifickým způsobem, jakým se to děje v případě Nizozemska, a vlivem jakých zdrojů zkreslení k tomu vlastně dochází.

V testu modelu ekvivalence variancí chyb měření se odhalilo, že v těchto třech zemích nemá většina měřících položek podobné chyby měření a položky nedosahují tzv. striktní ekvivalence. Polovina položek má v Belgii a Nizozemsku významně vyšší chyby měření než je tomu v Lucembursku a ty musely být z restrikcí uvolněny; jsou tedy méně reliabilní. Ve faktoru individuálního sociálního kapitálu musely být z restrikcí ve všech zemích na základě modifikačních indexů uvolněny položky FAIR (M.I. od 95 do 126) a HELPFUL (M.I. od 45 do 63), viz Tabulka 69.

Tabulka 69 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / NL, BE, LU

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	NL	BE	LU
individuální odcizení	INTEREST	0,395		0,449	0,535	0,304
	DIFFICULT	0,829	0,821			
	POLACTIV	1,113	0,982		1,672	
	OPINION	0,721	0,651		0,968	
institucionální odcizení	VOTE	0,517	0,519			
	CARE	0,396	0,353		0,559	
důvěra v instituce	POLIT	1,238		1,362	1,797	0,886
	PARLAM	1,476	1,491			
	LEGAL	2,850	2,837			
individuální sociální kapitál	TRUSTED	2,259	2,211			
	FAIR	2,390		2,674	3,943	1,514
	HELPFUL	3,317		3,527	4,618	2,621
kolektivní sociální kapitál	FRIEND	2,314	2,313			
	VOLUNORG	6,675	6,660			
	MEET	1,865	1,740		2,354	
	SOCACTIV	0,705	0,707			

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

⁸⁸ Jedná se např. o vlámsčinu – belgická varianta nizozemštiny, belgickou francouzštinu a lucemburštinu, která vznikla jako kombinace francouzštiny a němčiny [Lewis et al. 2013].

Protože se Nizozemsko opět prokázalo v dané skupině zemí jako pouze částečně skalárně ekvivalentní, bylo možné ho z další analýzy vyřadit a otestovat ekvivalenci položek jen v Belgii a Lucembursku. Výsledky v Tabulce 71 ukazují, že v datech byla identifikována opakovaně pouze částečná skalární ekvivalence (Δ CFI = 0,065), ale tentokrát ve všech faktorech, jejichž průměry jsou tak v Belgii a Lucembursku vzájemně porovnatelné. Navíc položky měřící faktor institucionálního a individuálního politického odcizení se v Belgii a Lucembursku ukazují jako plně skalárně ekvivalentní (viz Tabulka 70). Z faktoru důvěry v instituce musela být na základě modifikačních indexů uvolněna z restrikce konstanta položky LEGAL (M.I. = 97, 159) a z faktorů sociálního kapitálu konstanty položek FAIR, FRIEND a VOLUNORG.

Tabulka 70 Odhady konstant ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / BE, LU

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	BE	LU
individuální odcizení	INTEREST	2,543	2,544		
	DIFFICULT	3,039	3,040		
	POLACTIV	3,896	3,899		
	OPINION	3,149	3,151		
institucionální odcizení	VOTE	3,679	3,684		
	CARE	3,436	3,442		
důvěra v instituce	POLIT	5,450	5,456		
	PARLAM	4,776	4,752		
	LEGAL	4,911		5,412	4,065
individuální sociální kapitál	TRUSTED	5,044	5,040		
	FAIR	5,643		5,725	5,431
	HELPFUL	4,480	4,495		
kolektivní sociální kapitál	FRIEND	8,495		8,348	8,725
	VOLUNORG	5,730		5,163	6,771
	MEET	5,138	5,132		
	SOCACTIV	2,722	2,711		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Vyřazení Nizozemska mělo opět určitý vliv na vylepšení skalární ekvivalence položek, neboť v Belgii a Lucembursku lze v tuto chvíli komparovat průměry všech latentních proměnných. Řada měřících položek je dokonce plně skalárně ekvivalentní, lze tudíž přistoupit i k porovnávání jejich průměrů. V případě kolektivního sociálního kapitálu jsou v Lucembursku konstanty položek FRIEND a VOLUNORG stále vyšší než v Belgii, u měření mezilidské a institucionální důvěry je tomu naopak. Tyto drobné rozdíly mezi Belgií a Lucemburskem mohou být způsobeny regionálními odlišnostmi ve francouzštině a němčině nebo vlivem celkového specifického stylu odpovědi.

Tabulka 71 Statistiky vhodnosti pro testy ekvivalence: shrnutí / BE, LU

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	491,45 (184)	-	0,965 (-)	0,025 (0,022;0,028)	0,955 (-)	-	přijat
Model 2 (Metrická ekvivalence)	545,07 (196)	53,62*** (12)	0,961 (0,004)	0,026 (0,023;0,029)	0,952 (0,003)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	545,12 (197)	0,05 (1)	0,961 (0,00)	0,026 (0,023;0,028)	0,952 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1138,74 (213)	593,62*** (16)	0,896 (0,065)	0,041 (0,038;0,043)	0,882 (0,07)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	644,52 (209)	99,4*** (12)	0,951 (0,01)	0,028 (0,026;0,031)	0,944 (0,008)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	656,98 (219)	12,46 (10)	0,951 (0,00)	0,028 (0,025;0,030)	0,946 (0,00)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	685,85 (224)	28,87*** (5)	0,948 (0,003)	0,028 (0,026;0,030)	0,944 (0,002)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	944,46 (240)	258,61*** (16)	0,921 (0,027)	0,033 (0,031;0,036)	0,921 (0,023)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	781,37 (237)	95,52*** (13)	0,939 (0,009)	0,030 (0,027;0,032)	0,938 (0,006)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

V případě testu modelu ekvivalence variancí chyb měření výsledky zpochybnily plnou ekvivalenci, ale potvrdily částečnou ekvivalenci chyb měření. Musely být na základě modifikačních indexů uvolněny parametry položek POLACTIV (M.I. = 39, 64), OPINION (M.I. = 13, 21) a FAIR (M.I. = 13, 21) s tím, že lze podle velikosti volně odhadnutých variancí chyb měření usoudit, že položky měřící kolektivní sociální kapitál mají v Lucembursku vyšší chybu měření než položky v Belgii a jsou proto méně reliabilní, naopak položky měřící individuální sociální kapitál a důvěru v instituce jsou v Lucembursku více reliabilní.

Tabulka 72 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / BE, LU

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	BE	LU
individuální odcizení	INTEREST	0,447	0,443		
	DIFFICULT	0,941	0,944		
	POLACTIV	1,149		0,848	1,643
	OPINION	0,817		0,709	1,006
institucionální odcizení	VOTE	0,574	0,575		
	CARE	0,448	0,447		
důvěra v instituce	POLIT	1,523	1,524		
	PARLAM	1,344	1,341		
	LEGAL	3,210	3,212		
individuální sociální kapitál	TRUSTED	2,515	2,523		
	FAIR	3,086		2,606	3,902
	HELPFUL	3,874	3,862		
kolektivní sociální kapitál	FRIEND	2,373	2,375		
	VOLUNORG	7,215	7,200		
	MEET	2,134	2,136		
	SOCACTIV	0,781	0,782		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

9.4.3 Francie, Belgie a Lucembursko

Francie, Belgie a Lucembursko vzájemně pojí jedna důležitá věc a to je oficiální úřední jazyk francouzština a územní příslušnost k západní Evropě. Je tedy velká pravděpodobnost, že v těchto zemích také nalezneme kvůli jejich společnému kulturnímu a jazykovému základu vzájemně srovnatelné položky.⁸⁹ Shrnující výsledky ukazuje tabulka 73. Údaje potvrzují konfigurační a metrickou ekvivalenci a totožné kovariance mezi latentními proměnnými. V těchto třech zemích ale opět nebyla identifikována plná skalární ekvivalence. Částečně skalární jsou proměnné měřící faktory sociálního kapitálu.

⁸⁹ V Belgii i Lucembursku je jedním z oficiálních jazyků také němčina, proto byly analyzovány modely ekvivalence i v souboru zemí s Německem a Rakouskem. Ukázalo se, že není možné v tomto spojení prokázat částečnou skalární ekvivalenci proměnných ve všech faktorech – pouze v latentních proměnných sociálního kapitálu. V obou zemích je pak nutné uvolnit konstanty položek HELPFUL, VOLUNORG a FRIEND. V případě Lucemburska je navíc nutné uvolnit z restrikce více než polovinu variancí chyb měření, které jsou příliš vysoké.

Tabulka 73 Statistiky vhodnosti pro testy ekvivalence: shrnutí / FR, BE, LU

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	845,33 (276)	-	0,958 (-)	0,023 (0,021;0,024)	0,945 (-)	-	přijat
Model 2 (Metrická ekvivalence)	968,96 (300)	123,63*** (24)	0,951 (0,007)	0,023 (0,022;0,025)	0,941 (0,004)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	972,48 (302)	3,52 (2)	0,951 (0,00)	0,023 (0,022;0,025)	0,941 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	2088,70 (334)	1116,23*** (32)	0,871 (0,08)	0,036 (0,035;0,037)	0,861 (0,08)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	1070,31 (310)	97,83*** (8)	0,944 (0,007)	0,025 (0,023;0,026)	0,935 (0,006)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	1099,42 (330)	29,11* (20)	0,943 (0,001)	0,024 (0,022;0,026)	0,938 (0,003)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	1140,90 (340)	41,48*** (10)	0,941 (0,00)	0,024 (0,023;0,026)	0,937 (0,001)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	1336,48 (354)	195,58*** (14)	0,928 (0,013)	0,026 (0,025;0,028)	0,926 (0,011)	Model 5	nepřijat
Model 7 (Částečná ekvivalence variancí chyb měření)	1293,96 (368)	153,05*** (28)	0,932 (0,009)	0,025 (0,023;0,026)	0,933 (0,004)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

V individuálním sociálním kapitálu musela být uvolněna na základě modifikačních indexů konstanta položky TRUSTED (M.I. od 15 do 22) a v kolektivním sociálním kapitálu konstanty položek VOLUNORG (M.I. od 33 do 150) a SOCACTIV (M.I. od 37 do 56). V případě politického odcizení a faktoru důvěry v instituce musela být zbavena restriktce shody většina konstant proměnných a průměry těchto latentních proměnných tedy nejsou validně vzájemně srovnatelné (viz Tabulka 74).

Tabulka 74 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / FR, BE, LU

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	FR	BE	LU
individuální odcizení	INTEREST	2,594		2,681	2,589	2,479
	DIFFICULT	3,118		3,256	3,077	2,991
	POLACTIV	3,975		4,109	3,945	3,844
	OPINION	3,287		3,527	3,218	3,028
institucionální odcizení	VOTE	3,735		3,828	3,738	3,617
	CARE	3,539		3,718	3,484	3,405
důvěra v instituce	POLIT	5,792		6,347	5,686	5,103
	PARLAM	5,062		5,534	4,973	4,438
	LEGAL	4,998		5,163	5,581	3,822
individuální sociální kapitál	TRUSTED	4,832		4,492	4,882	5,252
	FAIR	5,622	5,624			
	HELPFUL	4,456	4,459			
kolektivní sociální kapitál	FRIEND	8,367	8,361			
	VOLUNORG	5,365		4,876	5,097	6,795
	MEET	5,161	5,158			
	SOCACTIV	2,804		2,924	2,646	2,802

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Konstanty položek měřících faktory politického odcizení a institucionální důvěry jsou ve Francii výrazně vyšší než v Belgii či Lucembursku, přičemž největší rozdíly se objevují při komparaci Francie s Lucemburskem. Příčiny se mohou skrývat v kulturních i jazykových rozdílnostech. Lucembursko je vyspělým státem s vysokou životní úrovní, nízkou zaměstnaností a menším množstvím menšin pocházejících převážně z Evropské unie. Francouzština se zde nejčastěji užívá v politických kruzích a ve státní sféře. Němčina je vyučována na základních školách a je nejčastěji používána v médiích. Lucemburština je naproti tomu místním dialektem, kterým se hovoří doma, a někdy v ní vychází i tisk a knihy [Lewis et al. 2013]. Francie je sice také ekonomicky vyspělým státem v západní Evropě, ale nevyhýbají se jí v dnešní době mnohé sociální problémy, velmi často spojované s početnými národnostními menšinami z Evropy, Afriky i Číny a jejich obrovskou etnickou různorodostí. Všechny tyto zdroje chyb měření poté ovlivňují volbu kategorií na škále měření a způsobují různé systematické odchylky (např. volbu krajních hodnot, tendenci k souhlasu bez ohledu na obsah položky, větší atraktivitu méně extrémní varianty odpovědi, atp), obzvláště v případě takových proměnných, které obsahují nespecifikované výrazy množství, intenzity a frekvence, jako je tomu u faktorů politického odcizení.

Z předchozích analýz skalární ekvivalence vyplývá, že je již bylo možné porovnat průměry všech faktorů mezi Belgií a Lucemburskem, navíc byla možná komparace ještě s Nizozemskem ve faktorech politického odcizení a důvěry v instituce, a s Francií v případě

faktorů sociálního kapitálu.⁹⁰ Výsledky komparace jsou shrnuté v Tabulce 75. Srovnání strukturálních a latentních průměrů v těchto zemích odkrývá, že Lucembursko má vyšší míru mezilidské důvěry a sociálních vazeb než Belgie a Francie. V hodnotách reálných čísel strukturálního průměru se ale stále jedná o průměrné hodnocení těchto sociálních vazeb a důvěry (hodnocení cca 5–6). Naopak v Belgii lze najít vyšší míru nedůvěry ve veřejné instituce a politického odcizení (opět se ukazuje průměrné až nadprůměrné obecné hodnocení), nejnižší míra politického odcizení se objevuje v Nizozemsku.

Tabulka 75 Srovnání strukturálních a latentních průměrů / BE, LU, FR, NL

individuální sociální kapitál						
země	Strukturální průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Lucembursko	5,14	0,056	1/2/3	0,219**	0,069	1
Belgie	4,98	0,044	1/2/3	0,00	x	2
Francie	4,85	0,045	1/2/3	-0,162**	0,062	3
kolektivní sociální kapitál						
země	Strukturální průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Lucembursko	5,89**	0,036	1	0,225**	0,030	1
Belgie	5,29	0,030	2/3	0,00	x	2/3
Francie	5,27	0,029	2/3	0,004	0,014	2/3
(ne)důvěra ve veřejné instituce						
země	Strukturální průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Lucembursko	4,42**	0,060	3	-0,866**	0,080	3
Belgie	5,42**	0,049	1	0,00	x	1
Nizozemsko	4,86**	0,037	2	-0,455**	0,073	2
individuální politické odcizení						
země	Strukturální průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Lucembursko	3,07	0,026	1/2/3	-0,120**	0,030	2
Belgie	3,21**	0,019	1/2	0,00	x	1
Nizozemsko	3,03**	0,016	2/3	-0,191**	0,027	3
institucionální politické odcizení						
země	Strukturální průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Lucembursko	4,02	0,036	2/3	-0,086**	0,034	2
Belgie	4,31**	0,028	1	0,00	x	1
Nizozemsko	3,97	0,023	2/3	-0,22**	0,030	3

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturální průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturální analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturální průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 2: Země jsou ve faktoru kolektivního sociálního kapitálu modře zvýrazněny, protože vychýlené konstanty poloviny položek mohou zkreslovat rozhodnutí o statistické významnosti rozdílů mezi latentními průměry.

⁹⁰ Nastává zde obdobný případ jako v případě komparace Nizozemska se 1) Spojeným královstvím a Irskem a 2) Belgií a Lucemburskem, kdy v prvním případě byly srovnatelné jen faktory sociálního kapitálu a ve druhém případě jen faktory politického odcizení a institucionální důvěry. V této situaci je srovnatelná Belgie a Lucembursko 1) s Nizozemskem (politické odcizení a důvěra v instituce) a 2) s Francií (sociální kapitál). Příčiny tohoto zajímavého jevu by bylo nutné dále podrobně prozkoumat.

Testy ekvivalence variancí chyb měření vygenerovaly o něco příznivější údaje. Plné ekvivalence chyb měření dosahují položky ze všech třech zemí ve faktoru důvěry v instituce a individuálního sociálního kapitálu. Ostatní latentní proměnné jsou částečně ekvivalentní z hlediska chyb měření. Musely být uvolněny variance chyb měření na základě modifikačních indexů u položky POLACTIV (M.I. = 63), OPINION (M.I. = 28) a FAIR v Lucembursku (M.I. = 16) a položky SOCACTIV ve Francii (M.I. = 38). V Lucembursku se stejně jako v předešlých analýzách ukazuje, že chyby měření položek měřících politické odcizení a sociálního kapitálu jsou vyšší než v ostatních zemích (tj. dané proměnné jsou méně reliabilní), proto musely být tyto parametry v modelu opět uvolněny a volně odhadnuty (viz Tabulka 76).

Tabulka 76 Nestandardizované odhady variancí chyb měření modelu ekvivalence variancí chyb měření; velikosti uvolněných variancí chyb měření / FR, BE, LU

Latentní proměnná	Měřicí položka	Model ekvivalence chyb měření	Model částečné ekvivalence chyb měření	FR	BE	LU
individuální odcizení	INTEREST	0,446	0,434			
	DIFFICULT	0,926	0,924			
	POLACTIV	1,068	0,922			1,679
	OPINION	0,784	0,718			0,998
institucionální odcizení	VOTE	0,568	0,566			
	CARE	0,449	0,459			
důvěra v instituce	POLIT	1,444	1,453			
	PARLAM	1,599	1,595			
	LEGAL	3,240	3,231			
individuální sociální kapitál	TRUSTED	2,633	2,639			
	FAIR	3,142	2,926			3,920
	HELPFUL	3,857	3,845			
kolektivní sociální kapitál	FRIEND	2,462	2,418			
	VOLUNORG	7,048	7,030			
	MEET	2,012	2,026			
	SOCACTIV	0,655	0,855	0,548		

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Přestože výsledky testů byly příznivé, největší překážky se objevovaly u vzájemné srovnatelnosti položek z Francie a Lucemburska. Proto bylo z testované skupiny vyřazeno ještě Lucembursko a je dále zkoumána srovnatelnost položek pouze ve Francii a Belgii.⁹¹ Tabulka 77 zachycuje analýzu modelů ekvivalence testovaných v těchto zemích. Zde se ukazuje, že jediný model nemohl být přijat kvůli horším výsledkům statistik vhodnosti ($\Delta CFI > 0,01$) a to konkrétně model plné skalární ekvivalence.

⁹¹ Vztah Belgie a Lucemburska se již prokázal.

Tabulka 77 Statistiky vhodnosti pro testy ekvivalence: shrnutí / FR, BE

Model	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	Komparace	Rozhodnutí
Model 1 (Konfigurální ekvivalence)	639,68 (184)	-	0,957 (-)	0,028 (0,026;0,031)	0,944 (-)	-	přijat
Model 2 (Metrická ekvivalence)	722,42 (196)	82,74*** (12)	0,951 (0,006)	0,030 (0,027;0,032)	0,940 (0,004)	Model 1	přijat
Model 2a (Ekvivalence korelace chyb měření)	725,22 (197)	2,8* (1)	0,951 (0,00)	0,030 (0,027;0,032)	0,940 (0,00)	Model 2	přijat
Model 3 (Skalární ekvivalence)	1150,36 (213)	425,14*** (16)	0,912 (0,039)	0,038 (0,036;0,040)	0,901 (0,039)	Model 2a	nepřijat
Model 3a (Částečná skalární ekvivalence)	844,58 (209)	119,37*** (12)	0,941 (0,01)	0,032 (0,029;0,034)	0,932 (0,008)	Model 2a	přijat
Model 4 (Ekvivalence kovariancí faktorů)	860,36 (219)	15,77 (10)	0,940 (0,001)	0,031 (0,029;0,033)	0,934 (0,00)	Model 3a	přijat
Model 5 (Ekvivalence variancí faktorů)	886,13 (224)	25,77*** (5)	0,938 (0,002)	0,031 (0,029;0,033)	0,934 (0,00)	Model 4	přijat
Model 6 (Ekvivalence variancí chyb měření)	984,46 (240)	98,33*** (16)	0,930 (0,008)	0,032 (0,030;0,034)	0,930 (0,004)	Model 5	přijat

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Detailnější pohled na odhady parametrů modelu částečné skalární ekvivalence položek odhaluje, že všechny latentní proměnné splňují podmínky částečné skalární ekvivalence a položky faktoru institucionálního odcizení jsou dokonce plně skalárně ekvivalentní (viz Tabulka 78). Z faktoru individuálního odcizení musela být zbavena restrikce na základě modifikačních indexů pouze jedna položka OPINION (M.I. = 25, 21), ve faktoru důvěry v instituce položka LEGAL (M.I. = 73, 64) a v latentních proměnných sociálního kapitálu položky TRUSTED (M.I. = 10, 17) a SOCACTIV (M.I. = 42, 53). Navíc statistiky vhodnosti posledního modelu ekvivalence variancí chyb měření ověřují, že všechny položky mají podobnou velikost chyb měření ($\Delta CFI \leq 0,01$) a splňují tak nejpřísnější striktní ekvivalenci, tj. jsou podobně reliabilní.

Tabulka 78 Odhady konstant položek ve skalárním modelu; velikosti uvolněných konstant v částečně skalárním modelu / FR, BE

Latentní proměnná	Měřicí položka	Skalární model	Částečně skalární model	FR	BE
individuální odcizení	INTEREST	2,635	2,635		
	DIFFICULT	3,162	3,162		
	POLACTIV	4,024	4,024		
	OPINION	3,362		3,483	3,261
institucionální odcizení	VOTE	3,786	3,784		
	CARE	3,596	3,594		
důvěra v instituce	POLIT	6,007	6,006		
	PARLAM	5,238	5,244		
	LEGAL	5,397		4,940	5,801
individuální sociální kapitál	TRUSTED	4,677		4,535	4,802
	FAIR	5,609	5,614		
	HELPFUL	4,418	4,422		
kolektivní sociální kapitál	FRIEND	8,239	8,224		
	VOLUNORG	4,978	4,959		
	MEET	5,196	5,180		
	SOCACTIV	2,801		2,937	2,618

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Vyřazení Lucemburska mělo tedy příznivý vliv na srovnatelnost položek mezi Francií a Belgií, neboť ty jsou porovnatelné ve všech latentních proměnných. Tyto čtyři konstanty, které vykazovaly rozdílné velikosti, mohou zkreslovat jazykové odlišnosti francouzštiny a belgické francouzštiny mírně se lišící slovní zásobou i sémantikou. Dále to také může být skutečnost, že více než polovina Belgičanů hovoří vlámsky (belgickou nizozemštinou), a proto byly také dotazníky v této zemi distribuovány ve dvou jazycích. V neposlední řadě zde může hrát roli kontext odpovědi nebo specifický styl odpovědí respondentů v daných zemích.

Na základě částečné skalární ekvivalence položek je také dále možné ve Francii a Belgii navzájem porovnávat průměry všech latentních proměnných. Neboť část tohoto srovnání byla naznačena již výše (viz Tabulka 75), lze se v tuto chvíli již zaměřit pouze na komparaci faktorů politického odcizení a důvěry v instituci, které mezi Belgií a Francií ještě porovnávané nebyly.⁹² Strukturální průměry opět nebyly schopny v případě některých

⁹² Položky Belgie, Francie a Lucemburska byly vzájemně (částečně skalárně) srovnatelné pouze v konceptu sociálního kapitálu. Nicméně po vygenerování výsledků samostatné analýzy ekvivalence modelů ve skupině 1) Francie a Belgie a 2) Belgie a Lucemburska se ukázalo, že (pouze) v těchto dvojicích jsou měřicí indikátory ekvivalentní a srovnatelné ve všech faktorech. Do porovnávání průměrů latentních položek tak nelze zahrnout

faktorů jednoznačně rozhodnout pořadí zemí při hodnocení daného konceptu, ale jinak produkují relativně shodné pořadí jako latentní průměry. Co se týče míry důvěry v instituce a politického odcizení, lze usoudit, že Francie v těchto konceptech předčí Belgii a obecně vykazuje spíše nadprůměrnou nedůvěru ve veřejné instituce (hodnocení cca 5,7) a také nadprůměrné politické odcizení (hodnocení cca 3,3 a 4,6). Z relativního srovnání pořadí zemí v tabulkách 75 a 79 také vyplývá, že pokud je na tom Francie hůře z hlediska hodnocení (ne)důvěry v instituce a politického odcizení, předstihne kromě Belgie také Lucembursko a Nizozemsko.

Tabulka 79 Srovnání strukturních a latentních průměrů / FR, BE

individuální sociální kapitál						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Francie	4,85	0,045	1/2	-0,153**	0,061	2
Belgie	4,98	0,044	1/2	0,00	x	1
kolektivní sociální kapitál						
země	Strukturní průměr (0 – 8)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Francie	5,27	0,029	1/2	0,043	0,024	1/2
Belgie	5,29	0,030	1/2	0,00	x	1/2
(ne)důvěra ve veřejné instituce						
země	Strukturní průměr (0 – 10)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Francie	5,7**	0,051	1	0,421**	0,073	1
Belgie	5,42**	0,049	2	0,00	x	2
individuální politické odcizení						
země	Strukturní průměr (1 – 5)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Francie	3,4	0,021	1/2	0,169**	0,027	1
Belgie	3,21	0,019	1/2	0,00	x	2
institucionální politické odcizení						
země	Strukturní průměr (1 – 7)			Latentní průměr		
	průměr	S.E.	pořadí	průměr	S.E.	pořadí
Francie	4,64**	0,030	1	0,199**	0,030	1
Belgie	4,31**	0,028	2	0,00	x	2

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Strukturní průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi (test ANOVA), latentní průměr je výstupem strukturní analýzy průměrů (resp. MACS). Pořadí, které není jasně určené (1/2/3/4), je v tabulce zaneseno proto, že latentní průměry anebo strukturní průměry v daných zemích nebyly statisticky významně odlišné, proto se o jejich pořadí nedá jednoznačně rozhodnout. ** značí naopak statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

9.5 Shrnutí výsledků simultánní analýzy dat

Z hlediska tradičních cílů komparativních výzkumů – srovnávat co nejvíce zemí mezi sebou – se zdají být prezentované výsledky na první pohled poněkud překvapivé, možná i lehce skličující. Překvapivé být mohou, ale skličující jistojistě ne. Prvním výborným výsledkem je to, že měřicí položky testovaných postojových škál z dotazníkového šetření první vlny ESS (2002) prokázaly ve všech zemích konfigurální a přinejmenším částečnou metrickou ekvivalenci. Jediné země, které se od parametrů plně metrické ekvivalence vychylovaly, bylo

všechny tři země, ale je možné provést relativní srovnávání pořadí zemí vůči jedné referenční (v tomto případě Belgii, viz ve výkladu dále).

Řecko a Norsko, ale i tyto země nakonec prokázaly alespoň částečnou metrickou ekvivalenci dat. Nalezení tzv. slabé (a silné) ekvivalence [Horn, McArdle 1992; Meredith 1993] značí velmi kvalitně naplánovaný design výzkumu jak z hlediska sběru dat, tak i formulace multilingvních dotazníků. Dosažená konfigurační ekvivalence dat dosvědčuje, že jsou teoretické jevy v každé skupině chápány a měřeny stejně. Metrická ekvivalence dat, znamená, že je obsah každé položky a každé kategorie (slovní i numerické) měřicí škály vnímán, chápán a interpretován velmi podobným způsobem ve všech skupinách respondentů. Tyto úrovně ekvivalence byly u daných položek v první vlně šetření ESS zjištěny. Nejenže je tak vhodné tyto měřicí položky i v dalších šetřeních reprodukovat z hlediska formulace i překladu, zjištěné skutečnosti především jasně určují, jaké statistické analýzy lze v rámci meritorní analýzy těchto položek použít při komparaci dat.

Jestliže je v datech ověřena konfigurační ekvivalence, je potvrzeno, že zkoumané teoretické jevy v participujících zemích vůbec existují a že respondenti z těchto zemí chápou jejich význam podobně, respektive je dokážou hodnotit na základě předložených měřicích položek. Pokud je navíc v testech ekvivalence položek aplikována konfirmační faktorová analýza, konfigurační ekvivalence dat dále značí, že testovaný model adekvátně reprezentuje data ve všech skupinách.⁹³ Konfigurační ekvivalence už ale neposkytuje další informace o tom, jak je možné s daty dále komparativně pracovat. Je naprosto nezbytnou podmínkou pro jakékoli komparativní analýzy, ale většinou je nutné pokračovat v dalších testech ekvivalence postojových škál.

Z metrické ekvivalence je už pak možné vyvodit více užitečných informací o tom, jak přistoupit ke komparaci dat. Ověřená metrická ekvivalence v datech značí, že je možné statisticky porovnávat vztahy testovaných latentních proměnných (tj. postojových škál) s jinými také metricky ekvivalentními proměnnými z dotazníku mezi zeměmi. Je tedy možné a nanejvýš vhodné analyzovat tyto vztahy pomocí korelační a případně regresní analýzy. Neboť je právě toto hlavním cílem většiny komparativních analýz, je metrická ekvivalence položek pro další statistické záměry většinou zcela dostačující. I v případě prokázání plné metrické ekvivalence položek je ale stále vhodné využívat v analýzách celou baterii otázek (resp. latentní proměnnou či faktor, faktorové skóry, atp.) a ne jen samostatné měřicí položky. Pokud je ale výzkumným cílem srovnat průměry manifestních či latentních proměnných mezi sebou, je dalším nezbytným krokem test skalární ekvivalence položek.

Test skalární ekvivalence položek postojových škál sociálního kapitálu a politického odcizení vyloučil jakoukoli komparaci průměrů 16 manifestních proměnných mezi 22 zeměmi. Navíc není ani možné validně srovnávat průměry latentních proměnných důvěry v instituce a politického odcizení ve všech zemích, protože položky těchto faktorů neprokázaly ani částečnou skalární ekvivalenci. V případě faktorů sociálního kapitálu je ale možné jejich průměry porovnávat mezi deseti evropskými zeměmi, neboť zde částečná skalární ekvivalence položek prokázána byla.

Skalární ekvivalence je už poměrně striktní požadavek na srovnatelnost položek, které musí v tomto ohledu prokázat především svou „statistickou srovnatelnost“; v případě použití konfirmační faktorové analýzy se jedná o srovnatelnost konstant položek v regresní rovnici dané proměnné. Ukazuje se, že tyto přísné parametry skalární ekvivalence jsou na rozdíl od metrické ekvivalence jen zřídka kdy naplněny i v mírně jazykově odlišných zemích, natož pak v zemích s jinými historickými a/nebo kulturními kořeny. Protože do analýzy bylo zařazeno příliš velké množství kulturně a jazykově odlišných zemí, pravděpodobně došlo ke kombinaci různých zdrojů chyb měření (např. stylu odpovědí respondentů, sociálně

⁹³ Pokud testy navíc potvrdí ve všech skupinách zemí ekvivalenci variancí a kovariancí latentních proměnných, je navíc i ověřena konstruktová validita tohoto měřicí nástroje (tj. modelu).

desirabilních otázek, referenčního efektu, překladu nespecifikovaných výrazů frekvence a hodnocení výroků, atp.), které způsobují odchylky ve velikostech konstant položek, i když ty stále měří stejný koncept.

Proto je adekvátním rozhodnutím v případě nedostatečné skalární ekvivalence dat rozdělit v analýze země do několika kulturně a jazykově homogennějších skupin, kde šance na nalezení vysoké úrovně skalární ekvivalence anebo dokonce ekvivalence chyb měření (tzv. striktní ekvivalence) významně vzrůstá. Výsledky dalších simultánních analýz tyto předpoklady jen potvrdily. Země byly rozděleny do několika skupin podle geografické polohy, historických souvislostí a především jazykové shody. Do každé skupiny byla navíc vždy zařazena jedna anebo více zemí, které měly s ostatními zeměmi značně velkou míru historické a kulturní shody, ale rozdílné jazykové kořeny. Tyto země sloužily jako diferencující prvek, který měl ověřit, zda by bylo možné posoudit míru vlivu kultury a jazyka na vzájemnou srovnatelnost měřicích položek.

Ukazuje se, že vliv stejného dorozumivacího jazyka v některých situacích zcela předčil vliv podobné kultury vybraných zemí. V případě první skupiny zemí ze střední Evropy (CZ, PL, SI, HU) byl koncept politického odcizení částečně skalárně ekvivalentní až po vyloučení Maďarska z analýzy. Stejně výsledky se objevují i v případě vyloučení Švýcarska, Itálie a Francie, Dánska a Finska, Nizozemska a Lucemburska z daných homogenních skupin zemí (viz Tabulka 80). Vždy se jedná o země, kde se hovoří buď zcela odlišným jazykem (např. maďarština vs. slovanské jazyky, finština vs. severogermánské jazyky) anebo ne tak podobným jazykem (např. francouzština a italština vs. španělština, dánština vs. norština, nizozemština vs. angličtina). V případě kulturně odlišných zemí (např. Francie – Anglie, Španělsko – Norsko atp.) se ani o částečné skalární ekvivalenci položek daných postojových škál nelze bavit – nebylo jí vůbec možné dosáhnout.

Na druhou stranu pouhá shoda jazykových kořenů nestačí, což dosvědčuje test ekvivalence položek mezi Maďarskem a Finskem, jejichž jazyky patří do společné rodiny uralských jazyků. Ani v těchto dvou zemích nebyla prokázána částečná skalární ekvivalence a nelze v nich porovnávat průměry testovaných proměnných. Potvrzuje se takto, že až rovnováha mezi dostatečnou kulturní a jazykovou blízkostí zemí vytvoří ideální podmínky pro dosažení vyšší úrovně skalární ekvivalence dat. Proto země se společným silným kulturním a současně také jazykovým základem vykazují částečnou nebo plnou skalární ekvivalenci *ve všech latentních proměnných*: Česká republika, Polsko a Slovinsko; Rakousko a Německo; Španělsko a Portugalsko; Norsko a Švédsko; Spojené království a Irsko; Belgie a Lucembursko a Francie a Belgie.⁹⁴ Příčiny drobných rozdílů ve velikostech některých konstant položek v těchto stejnorodých dvojicích zemí jsou už pak pravděpodobně dílem kontextu odpovědi a/nebo slabší jazykové dovednosti některých národnostních menšin hovořících buď specifickými nářečími či zcela jiným rodným jazykem. Závěry plynoucí z údajů v analýzách potvrzují nejen smysluplnost komparativní analýzy průměrů položek především v kulturně si podobných zemích (viz Graf 2, Příloha 4), ale také důležitost precizního překladu otázek, který očividně určuje další možnosti statistických analýz měřicích položek.

Testy skalární ekvivalence položek zkoumaných postojových škál dále podle předpokladů ukázaly, že položky měřící latentní proměnné sociálního kapitálu jsou méně problematické z hlediska komparace než položky měřící politické odcizení a důvěru v instituce, což se zcela shoduje s výsledky jiných výzkumů a studií [Balch 1974; Seligson 1980; Reeskens a Hooghe [2008]; van der Veld a Saris [2009]; Blasius, Thiesen 2000]. Měřicí položky politického

⁹⁴ Výsledky korespondují i se zjištěními Davidova a De Beuckelaera [2010], kteří testovali 16 párů zemí se stejným a 30 párů s odlišným dorozumivacím jazykem. Autoři v této metodologické stati potvrzují, že země s podobnými jazyky vykazují větší míru skalární ekvivalence položek.

odcizení jsou pravděpodobně skutečně zkresleny metodologickou systematickou chybou měření (konkrétně stylem odpovědí, sociální desirabilitou a tendencí k souhlasu bez ohledu na obsah položky), což je předmětem diskuzí již řadu let. V této studii se ukázal být kamenem úrazu ale především problematický překlad vágních kvantifikátorů množství a intenzity jako je „dost, jen trochu“ (INTEREST), „zřídka, občas“ (DIFFICULT), „pravděpodobně“ (POLACTIV), „ani těžké, ani lehké“ (OPINION), „někteří, většina“ (VOTE, CARE), které jsou pro položky měřící postojovou škálu politického odcizení velmi typické a pravděpodobně tak výrazně komplikují jejich kulturní srovnatelnost. Všechny tyto závěry by ale bylo vhodné dále detailněji otestovat nejlépe kognitivními rozhovory, teorií odpovědi na položku, ANOVA testy nebo hierarchickým strukturním modelováním.

Tabulka 80 Metrická a skalární ekvivalence v datech v jednotlivých skupinách zemí

	Koncept	sociální kapitál		politické odcizení		
	Latentní proměnná	individuální sociální kapitál	kolektivní sociální kapitál	důvěra v instituce	individuální odcizení	institucionální odcizení
Vyloučená země	Skupina zemí	Úroveň ekvivalence dat				
	CZ, PL, SI, HU	✓✓	✓✓	✓	✓	✓
HU	CZ, PL, SI	✓✓	✓✓	✓✓	✓✓	✓✓
	AT, DE, CH	✓✓	✓✓	✓	✓	✓
CH	AT, DE	✓✓	✓✓✓	✓✓	✓✓	✓✓
	ES, PT, IT, FR	✓	✓	✓✓	✓✓	✓✓
IT, FR	ES, PT	✓✓	✓✓	✓✓	✓✓	✓✓
	DK, FI, NO, SE	✓✓	✓✓	✓	✓	✓
DK, FI	NO, SE	✓✓	✓✓	✓✓	✓✓	✓✓✓
	GB, IE, NL	✓✓	✓✓	✓	✓	✓
NL	GB, IE	✓✓✓	✓✓	✓✓✓	✓✓✓	✓✓✓
	NL, BE, LU	✓	✓	✓✓	✓✓	✓✓
NL	BE, LU	✓✓	✓✓	✓✓	✓✓✓	✓✓✓
	FR, BE, LU	✓✓	✓✓	✓	✓	✓
LU	FR, BE	✓✓	✓✓	✓✓	✓✓	✓✓✓

Zdroj: ESS (2002); vlastní výpočty

Poznámka: ✓ značí metrickou ekvivalenci dat, ✓✓ znamenají dosažení částečné skalární ekvivalence položek ve faktoru, ✓✓✓ v oranžovém poli znamenají dosažení plné skalární ekvivalence položek ve faktorech.

Z hlediska analýzy ekvivalence jednotlivých položek postojových škál lze ještě zkonfrontovat výsledky základních statistických analýz - korelační analýzy, analýzy reliability a faktorové analýzy – s výsledky strukturního modelování. Základní analýzy ukázaly na pět nejproblematictějších položek, jejichž koeficienty korelace, Cronbachovo alfa a faktorové zátěže z hlediska velikosti mezi zeměmi značně kolísají, tj. pohybují se od 0,2 do 0,3 bodu

rozdílu mezi velikostmi jednotlivých koeficientů – byly to: HELPFUL, SOCACTIV, LEGAL, OPINION a VOTE (viz Tabulka 81).

Strukturní modelování v rámci analýzy metrické ekvivalence dat tyto závěry nepotvrzuje, téměř všechny položky se v daných zemích ukázaly být plně metricky ekvivalentní a tudíž srovnatelné. To, co ale metrický test identifikoval, je problém s komparací Řecka a Norska. Základní analýzy tyto země také vyčlenily jako odchýlené, ale nelze z těchto výsledků vyvodit, jak lze s těmito zeměmi dále zacházet v průběhu komparace; tyto detailní informace poskytuje až strukturní modelování. Test skalární ekvivalence položek provedený ve všech zemích najednou ukázal, že položky nejsou ani částečně skalárně ekvivalentní a že tak lze validně interpretovat vztahy latentních proměnných s ostatními metricky ekvivalentními proměnnými, ale ne už průměry měřicích položek, ani průměry latentních proměnných mezi všemi zeměmi zároveň.

Až simultánní analýza skalární ekvivalence v relativně homogenních skupinách zemí závěry základních analýz částečně potvrzuje, ale zároveň ukazuje, že situace je v jednotlivých skupinách zemí poměrně specifická. Strukturní modelování jako nejproblematictější odhalilo položky HELPFUL, VOLUNORG, LEGAL, OPINION a VOTE, jejichž konstanty se vychylovaly nejčastěji a musely být při testech modelu skalární ekvivalence uvolněny (viz Tabulka 82). Shoda výsledků skalárního testu ve strukturním modelování a základních analýz (resp. explorativní faktorové analýzy) je tedy poměrně velká. Komplikace je ale v tom, že výsledky základních analýz jsou poměrně povrchní a nenaznačují 1) jaký důsledek má tato vychýlená položka pro komparaci celé latentní proměnné, 2) jak specifická je tato situace pro různé skupiny zemí, 3) jak se odráží tato odchylka položky na celkové komparaci dat, resp. nenaznačuje, jaký statistický nástroj by bylo vhodné v případě takto vychýlené položky dále při meritorní analýze dat použít a 4) zda by bylo vhodné některé položky či země z komparativní analýzy přímo vyřadit. Lze tedy pravděpodobně identifikovat problematické položky už ze základních analýz, ale nelze říci, jak je tedy v daném souboru zemí vlastně možné postojové škály komparovat dál.

Tabulka 81 Analýza problematických položek z hlediska srovnatelnosti; porovnání výsledků čtyř analýz

Typ analýzy	Problematická položka na základě variability koeficientů					Problematická země
	Individuální sociální kapitál	Kolektivní sociální kapitál	Důvěra ve veřejné instituce	Individuální politické odcizení	Institucionální politické odcizení	
Korelační analýza	FAIR HELPFUL	SOCACTIV MEET	LEGAL	OPINION POLACTIV	POLIT CARE	Lucembursko Norsko, Izrael
Analýza reliability	TRUSTED	IMFRDS	PARLAM	OPINION	VOTE	Lucembursko Norsko, Izrael
EFA	HELPFUL	SOCACTIV FRIEND	LEGAL	OPINION	VOTE	Irsko, Izrael Řecko
SEM	HELPFUL	VOLUNORG	LEGAL	OPINION	VOTE	Řecko, Norsko a další

Poznámka: EFA (explorativní faktorová analýza), SEM (strukturní modelování)

Analýza chyb měření testovaných položek slouží spíše jako doplňkový druh analýzy reliability položek. Její výsledky někdy korespondují s tím, jaká země musí být ze skupiny vyřazena - většinou mají tyto země položky s většími chybami měření, viz např. Maďarsko, Itálie, Francie, Dánsko, Finsko v Tabulce 83, ale nemusí tomu tak nezbytně být. Výjimkou je kupříkladu Irsko, které vykazovalo mnohem větší míru chyb měření než Nizozemsko, přesto se ukázalo, že se Spojeným královstvím tvoří ideálnější dvojici z hlediska komparace. Vždy tedy záleží na zvolené skupině zemí, mezi kterými jsou data komparována.

Tabulka 82 Uvolněné konstanty neekvivalentních položek při testování částečné skalární ekvivalence v daných skupinách zemí

Skupiny zemí		CZ, PL, SI, HU	CZ, PL, SI	AT, DE, CH	AT, DE	ES, PT, IT, FR	ES, PT	DK, FI, NO, SE	NO, SE	GB, IE, NL	GB, IE	NL, BE, LU	BE, LU	FR, BE, LU	FR, BR	počet uvolnění
Latentní proměnná	Měřicí položka															
individuální odcizení	INTEREST	all				ES,FR	all	all		all		all		all		23
	DIFFICULT	all						all		all				all		14
	POLACTIV	all	CZ	all	all			all		all				all		20
	OPINION	all		all		all	all	all	all	all				all	all	27
institucionální odcizení	VOTE	all		all	all	FR		all		all				all		23
	CARE	all	SI	all		PT	all	all		all				all		21
důvěra v institute	POLIT	all	CZ	all				all		all				all		18
	PARLAM	all		all	all	FR		all		all				all		20
	LEGAL	all		all		ES, IT	all	all	all	all		all	all	all	all	30
individuální sociální kapitál	TRUSTED					all	all		all			all		all	all	16
	FAIR				all	all						all	all			11
	HELPFUL	all	all			all		FI		NL		all				18
kolektivní sociální kapitál	FRIEND		SI	AT		all				NL	all	all	all			14
	VOLUNORG	SI, HU	all	CH		all		DK, FI		all	all	all	all	all		25
	MEET	HU		DE		all	all	FI, NO	all			all				15
	SOCATIV					all						all		all	all	12

Zdroj: ESS 2002; vlastní výpočty

Poznámka: Šedě jsou označeny homogenní skupiny zemí, ve kterých bylo možné po vyloučení problematické země identifikovat zlepšení úrovně ekvivalence a nalézt alespoň částečnou skalární ekvivalenci položek ve všech faktorech.

Poznámka 2: Zkratka ‚all‘ znamená uvolnění konstant dané položky ve všech zemích ze zkoumané skupiny.

Tabulka 83 Uvolněné variance chyb měření při testu modelu ekvivalence chyb měření (tzv. striktní ekvivalence) – srovnání ve skupinách

země		CZ	PL	SI	HU	AT	DE	CH	ES	PT	IT	FR	DK	FI	NO	SE	GB	IE	NL	BE	LU	FR	
Latentní proměnná	Měřicí položka																						
individuální odcizení	INTEREST			x															x				
	DIFFICULT													x									
	POLACTIV																					x	
	OPINION					x	x	x											x			x	
institucionální odcizení	VOTE	x								x		x											
	CARE	x								x			x	x									
důvěra v instituce	POLIT	x				x			x								x	x	x				
	PARLAM	x			x			x											x				
	LEGAL							x			x						x	x	x				
individuální sociální kapitál	TRUSTED																		x				
	FAIR	x										x										x	
	HELPFUL																						
kolektivní sociální kapitál	FRIEND			x	x																		
	VOLUNORG																						
	MEET																						
	SOCACTIV								x		x			x	x		x	x					x

Zdroj: ESS 2002; vlastní výpočty

Poznámka: V zeleném poli jsou označeny ty chyby měření, které se ukázaly být menší než chyby měření položky v jiných zemích z určené skupiny (tzn. vyšší reliabilitu). Naopak červené pole označuje chyby měření, které byly podstatně vyšší a značí nižší reliabilitu položky v dané zemi (pouze ale ve srovnání se zeměmi z určené skupiny). V šedém poli se nacházejí země, které byly ze skupiny později vyřazeny.

10 SIMULTÁNNÍ ČASOVÁ ANALÝZA SROVNATELNOSTI ŠKÁL DŮVĚRY

Horn a McArdle [1992: 117] definují neměnnost měření jako situaci, kdy operacionalizace konstruktů vyústí v měření zcela totožných znaků a to i za různých okolností studia daného fenoménu. Ekvivalence položek se tedy nemusí testovat jen mezi kulturně, historicky, etnicky či jazykově odlišnými nebo příbuznými skupinami, ale i mezi různými časovými vlnami sběru dat, kde se zjišťuje srovnatelnost položek v čase. Řada mezinárodních výzkumů se realizuje opakovaně již několik desítek let a komparace položek v rámci jedné země měřených v různých časových vlnách šetření je jedním z dalších cílů mezinárodních, ale i národních komparativních analýz. Mezinárodní data tak mohou poskytnout každé zemi zvlášť hlubší pohled do struktury její společnosti a společenského dění a teoreticky umožňují porovnávat stejné teoretické koncepty po řadu let.

I v této studii bylo zajímavé nahlédnout do časové srovnatelnosti vybraných položek. Data z mezinárodního šetření ESS nabízejí již šest vln šetření, které postihují posledních deset let vývoje ve vybraných společnostech. Do hlavního dotazníku, který zůstává poměrně stabilní, bylo v dalších letech ze zkoumaných 16 položek avšak zařazeno již jen několik z nich (viz Tabulka 4). Konkrétně se jedná o položky měřící latentní proměnnou důvěry ve veřejné instituce a individuálního sociálního kapitálu, resp. faktoru mezilidské důvěry, které jsou neměnnou součástí hlavního dotazníku po celých deset let. I z hlediska participujících zemí došlo v průběhu času k obměně - některé země se daný rok výzkumu neúčastnily a některé do výzkumu naopak přibyly (viz Tabulka 84). Do analýzy ekvivalence je tak možné zařadit již pouze šest položek i nadále testovaných ve vybraných 22 zemích.⁹⁵

Tabulka 84 Přehled vybraných zemí participujících v jednotlivých vlnách výzkumu ESS

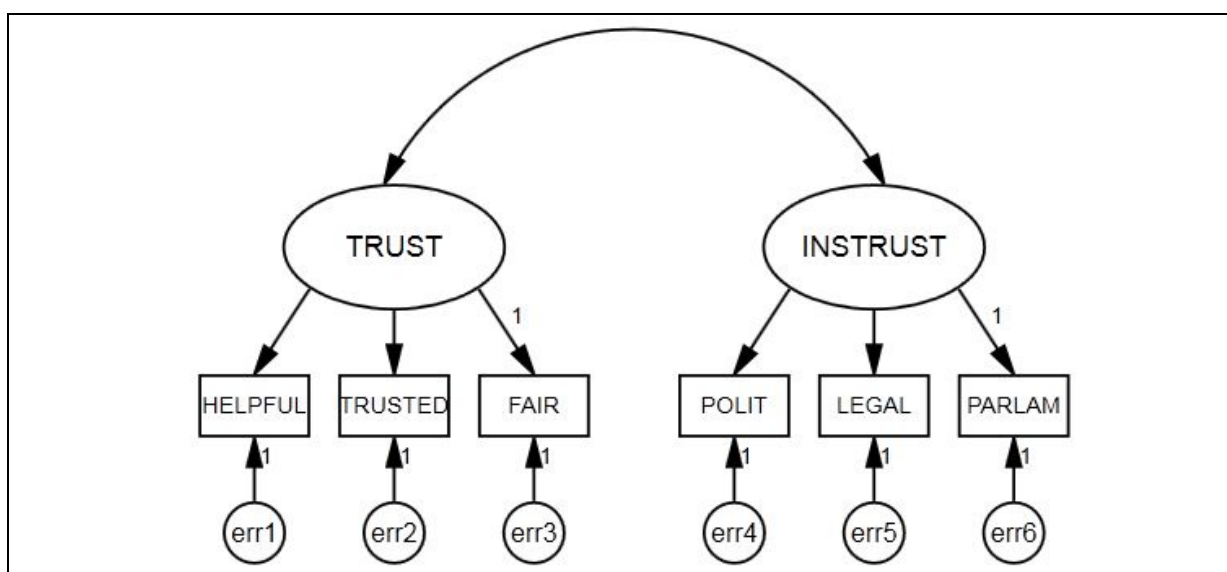
zkratka	země	První vlna	Druhá vlna	Třetí vlna	Čtvrtá vlna	Pátá vlna	Šestá vlna
		(2002)	(2004)	(2006)	(2008)	(2010)	(2012)
AT	Rakousko	x	x	x	x		
BE	Belgie	x	x	x	x	x	x
CZ	Česká republika	x	x		x	x	x
DE	Německo	x	x	x	x	x	x
DK	Dánsko	x	x	x	x	x	x
ES	Španělsko	x	x	x	x	x	x
FI	Finsko	x	x	x	x	x	x
FR	Francie	x	x	x	x	x	
UK	Spoj. království	x	x	x	x	x	x
GR	Řecko	x	x		x	x	
HU	Maďarsko	x	x	x	x	x	
CH	Švýcarsko	x	x	x	x	x	x
IE	Irsko	x	x	x	x	x	x
IL	Izrael	x			x	x	x
IT	Itálie	x	x				
LU	Lucembursko	x	x				
NL	Nizozemí	x	x	x	x	x	x
NO	Norsko	x	x	x	x	x	x
PO	Polsko	x	x	x	x	x	x
PT	Portugalsko	x	x	x	x	x	x
SE	Švédsko	x	x	x	x	x	x
SI	Slovinsko	x	x	x	x	x	x

Poznámka: Šedé pole značí absenci dané země v udaném roce šetření ESS; modrá pole znamenají, že ještě nebylo možné v lednu 2014 získat nové datové soubory šesté vlny ESS (2012).

⁹⁵ Ve studii byly dále analyzovány jen země, které od roku 2002 více méně pravidelně na výzkumu ESS participují (viz 22 zemí v Tabulce 84). Země, které se do šetření přidaly postupně ve druhé nebo třetí vlně, nejsou do této časové analýzy ekvivalence položek zařazeny.

Pro potřeby časové analýzy ekvivalence vybraných postojoyých škál byl vystaven nový model, který tentokrát zachycuje vztahy pouze mezi dvěma latentními proměnnými – důvěrou v instituce a mezilidskou důvěrou (viz Diagram 9). Dané faktory měří totožné tři položky jako v předešlé analýze ekvivalence (jejich znění a odpovědi viz Tabulka 3, Příloha 1), jako referenční proměnná je opět zvolena v latentní proměnné mezilidské důvěry položka FAIR a ve faktoru důvěry v instituce položka PARLAM. V rámci předběžného testování nového modelu byly zmapovány počty respondentů z každé vlny šetření v každé zemi a počty chybějících případů včetně podílu chybějících hodnot u jednotlivých proměnných, stejně tak došlo k ověření normality dat. Výsledky se nijak neodchylovaly od předešlých závěrů vyplývajících z předběžné analýzy základního modelu a podmínek strukturního modelování (viz kapitola 6).

Diagram 9 **Strukturní model konceptu důvěry ve veřejné instituce a mezilidské důvěry – simultánní analýza v čase**



Poznámka: Vysvětlení ke zkratkám položek a faktorů nabízí souhrnně Tabulka 3 v Příloze 1.

Z hlediska chybějících případů dosáhly všechny další vlny šetření ESS mnohem lepších výsledků než první vlna výzkumu z roku 2002. Žádná země již neměla míru chybějících případů vyšší než 20 % (viz Tabulka 85). V druhé vlně šetření se stále ve vůdčí pozici drželo Lucembursko (16 %), v dalších letech se pohybovalo na hranici lehce nad deset procent pouze Portugalsko a Španělsko. A v páté a šesté vlně šetření už žádná země nepřesáhla míru deseti procent chybějících případů z velikosti výběrového souboru. Podíl chybějících hodnot u jednotlivých proměnných od roku 2002 nepřesáhl hranici dvě procenta. Počet respondentů byl po vyloučení chybějících případů stále dostačující – nejmenší soubory tradičně přesahují velikost 1 200 respondentů.

Co se týče normality analyzovaných dat, výsledky jednorozměrného rozdělení ukázaly, že absolutní hodnota šikmosti proměnných nepřesáhla v žádné zemi a v žádné další vlně velikost 2 a v případě špičatosti hodnotu 7. Většina položek se pohybovala v intervalu od -1 do 1, lehčí vychýlení se objevilo opět u položky FRIEND a ve Finsku. Normalita vícerozměrného rozdělení (resp. koeficient vícerozměrné špičatosti) se sice nenacházela pod limitní hodnotou 5, ale ani nepřesáhla hodnotu 40. Jednotlivé položky mají v tomto modelu vždy na hodnotící škále 11 kategorií odpovědí. I vzhledem k uspokojivým výsledkům předchozích testů byla nadále i v časové analýze aplikována metoda listwise deleted pro vyřazení chybějících případů ze souborů a parametry modelu byly odhadovány metodou maximální věrohodnosti (ML) v programu AMOS 21.0.

Tabulka 85 Kontrola chybějících dat (missing values) a přehled počtu respondentů v jednotlivých zemích v šesti vlnách šetření ESS

země	První vlna		Druhá vlna		Třetí vlna		Čtvrtá vlna		Pátá vlna		Šestá vlna	
	ESS (2002)		ESS (2004)		ESS (2006)		ESS (2008)		ESS (2010)		ESS (2012)	
	počet respondentů	chybějící hodnoty	počet respondentů	chybějící hodnoty	počet respondentů	chybějící hodnoty	počet respondentů	chybějící hodnoty	počet respondentů	chybějící hodnoty	počet respondentů	chybějící hodnoty
Rakousko	2257	6,42 %	2256	6,87 %	2405	8,57 %	2255	6,30 %				
Belgie	1899	8,95 %	1778	3,09 %	1798	2,00 %	1760	1,76 %	1704	1,70 %	1869	1,23 %
Česká republika	1360	9,04 %	3026	8,79 %			2018	3,47 %	2386	2,98 %	2009	6,07 %
Německo	2919	3,49 %	2870	4,32 %	2916	4,05 %	2751	3,42 %	3031	4,35 %	2958	2,77 %
Dánsko	1506	3,85 %	1487	4,98 %	1505	4,19 %	1610	2,55 %	1576	3,55 %	1650	3,03 %
Španělsko	1729	12,03 %	1663	7,40 %	1876	7,62 %	2576	11,34 %	1885	5,25 %	1889	5,51 %
Finsko	2000	1,95 %	2022	2,47 %	1896	1,90 %	2195	1,64 %	1878	1,70 %	2197	1,27 %
Francie	1503	2,86 %	1806	2,33 %	1986	2,82 %	2073	2,41 %	1728	1,74 %		
Spojené království	2052	2,68 %	1897	4,43 %	2394	4,51 %	2352	3,32 %	2422	6,98 %	2286	6,78 %
Řecko	2566	5,85 %	2406	3,41 %			2072	1,83 %	2715	2,14 %		
Maďarsko	1685	8,43 %	1498	7,28 %	1518	9,42 %	1544	5,96 %	1561	6,34 %		
Švýcarsko	2040	6,47 %	2141	8,17 %	1804	6,71 %	1819	9,51 %	1506	7,90 %	1493	7,90 %
Irsko	2046	7,48 %	2286	6,39 %	1800	9,17 %	1764	2,27 %	2576	7,38 %	2628	5,90 %
Izrael	2499	7,60 %					2490	9,52 %	2294	9,76 %	2508	7,58 %
Itálie	1207	5,72 %	1529	2,94 %								
Lucembursko	1552	21,91 %	1635	16,33 %								
Nizozemí	2364	3,72 %	1881	3,51 %	1889	3,28 %	1778	2,36 %	1829	2,84 %	1845	2,38 %
Norsko	2036	1,03 %	1760	1,31 %	1750	0,91 %	1549	1,87 %	1548	1,81 %	1624	1,79 %
Polsko	2110	8,96 %	1716	6,93 %	1721	7,55 %	1619	8,89 %	1751	9,25 %	1898	7,85 %
Portugalsko	1511	12,44 %	2052	7,12 %	2222	11,48 %	2367	10,22 %	2150	8,51 %	2151	4,56 %
Švédsko	1999	5,15 %	1948	4,62 %	1927	5,40 %	1830	3,66 %	1497	4,81 %	1847	4,71 %
Slovinsko	1519	7,50 %	1442	7,49 %	1476	9,15 %	1286	7,23 %	1403	7,41 %	1257	5,25 %

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty

Poznámka 1: Šedé pole značí absenci dané země v udaném roce šetření ESS; k modře označeným zemím nebylo ještě možné v lednu 2014 získat nové datové soubory za šestou vlnu šetření ESS (2012).

Tento model (viz Diagram 9) bylo možné nadále pokládat za tzv. základní model. Základní model byl nejdříve testován v každé zemi zvlášť a ve všech vlnách, bez toho, že by se v analýze kladly nějaké požadavky na shodu velikosti parametrů ve zkoumaných zemích (viz princip separátní analýzy). Z výstupů analýzy vyplývá, že základní model je vyhovující v každé zkoumané zemi a v každé vlně šetření od roku 2002 do roku 2012 a reprezentuje data z výzkumu adekvátním způsobem (viz Tabulky 86–91, Příloha 5). Daný závěr podle všech očekávání nepotvrzují jen χ^2 testy, jejichž výsledky jsou statisticky významné na hladině spolehlivosti 99 % (p value < 0,0001) s 8 stupni volnosti, ale praktické indexy vhodnosti modelu (např. GFI, AGFI, CFI, RMSEA) ukazují na velmi kvalitní model. Velikost indexu CFI neklesá v žádné vlně šetření pod hodnotu 0,97, většinou se pohybuje mezi 0,98 – 0,99. Model nejhůře reprezentuje data v první vlně šetření v Lucembursku a Norsku, ve druhé vlně ve Spojeném království a Portugalsku, ve třetí vlně šetření ve Finsku a Švýcarsku, ve čtvrté vlně ve Finsku a Dánsku, v páté vlně opět ve Finsku a Irsku a nakonec v šesté vlně v Dánsku, Švýcarsku a Norsku.

Protože již nebylo nadále záměrem otestovat srovnatelnost položek v jednotlivých vlnách šetření napříč všemi vybranými zeměmi (to bylo ukázáno pouze na datech z první vlny šetření), ale provést simultánní analýzu srovnatelnosti položek *v čase* v rámci *jedné země*, musely se změnit určité parametry pro výběr zkoumaných skupin. Skupiny, které jsou zde základem několika výběrové konfirmační faktorové analýzy, nejsou tentokrát rozdílné země, ale pouze data z jedné země sebraná v různých časových vlnách šetření. Takto lze prokázat, zda je daná položka (resp. latentní proměnná) měřena v jedné a té samé zemi stále stejným způsobem a zda je stále stejně chápána, vnímána a interpretována. Je to velmi důležitá informace pro ty, kteří mají zájem sledovat a statisticky testovat jeden sociální jev (např. mezilidskou důvěru) v konkrétní zemi po určitý časový úsek. Simultánní časová analýza je proto provedena v každé zemi zvlášť.

Řada sekvenčních kroků testování ekvivalence zůstává stejná jako v případě simultánní analýzy mezi zeměmi. Základem je 1) konfigurální ekvivalence, další úroveň je 2) plná nebo částečná metrická ekvivalence, a 3) skalární ekvivalence položek, 4) lze otestovat ekvivalenci kovariancí faktorů, 5) ekvivalenci variancí faktoru a 6) ekvivalenci variancí chyb měření. Neboť testování položek probíhá vždy v rámci jedné země, základním předpokladem bylo, že testy modelů prokážou nejvyšší úroveň skalární a striktní ekvivalence bez větších problémů. Nedochozí zde k žádnému kulturnímu, historickému, ani jazykové střetu, pouze k postupnému vývoji dané společnosti a názorových konstelací v čase.

Tento předpoklad byl potvrzen alespoň v rámci metrické ekvivalence. V každé zemi je v datech ze všech vln šetření prokázána plná konfigurální ekvivalence (index CFI neklesá pod hodnotu 0,98 a ve většině zemí přesahuje hodnotu 0,99) a také plná metrická ekvivalence (viz Tabulka 92). Tato úroveň srovnatelnosti položek umožňuje dále statisticky porovnávat vztahy mezi dvěma zkoumanými latentními proměnnými anebo latentní proměnnou a jinou metricky ekvivalentní proměnnou z výzkumu (např. korelační či regresní analýzou) v datech *z dané země ve všech vlnách šetření*. V tabulce 92 lze nalézt podrobný přehled velikostí (nestandardizovaných) faktorových zátěží jednotlivých měřicích proměnných z modelu metrické ekvivalence; resp. sílu vztahu měřicích položek k danému faktoru, která byla potvrzena v průběhu let jako více méně neměnná.

Tabulka 92 Výsledky testu modelu plné metrické ekvivalence (Model 2) – porovnání s konfiguračním modelem (Model 1)

země	Faktorové zátěže v metrickém modelu				Statistiky vhodnosti modelu				
	HELPFUL	TRUSTED	POLIT	LEGAL	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)
Rakousko	0,955	0,945	0,823	0,739	287,96 (44)	75,18 (12)***	0,987 (0,004)	0,026 (0,023;0,028)	0,983 (0,00)
Belgie	0,829	1,216	0,963	0,922	273,96 (68)	31,61 (20)***	0,989 (0,00)	0,017 (0,015;0,019)	0,985 (0,00)
Česká rep.	0,830	1,094	0,900	0,880	257,28 (56)	65,61 (16)***	0,993 (0,002)	0,019 (0,017;0,021)	0,990 (0,00)
Německo	0,920	1,046	0,859	0,753	425,71 (68)	48,38 (20)***	0,988 (0,001)	0,018 (0,016;0,019)	0,984 (0,005)
Dánsko	0,811	1,118	0,846	0,629	368,86 (68)	64,33 (20)***	0,981 (0,003)	0,022 (0,020;0,024)	0,975 (0,00)
Španělsko	0,844	1,165	0,899	0,939	211,83 (68)	80,20 (20)***	0,991 (0,004)	0,014 (0,012;0,016)	0,989 (0,002)
Finsko	0,898	1,100	0,908	0,701	541,33 (68)	50,78 (20)***	0,980 (0,001)	0,024 (0,022;0,026)	0,974 (0,00)
Francie	0,864	1,082	0,946	0,841	199,15 (56)	22,66 (16)***	0,989 (0,00)	0,017 (0,014;0,020)	0,985 (0,005)
Spoj. král.	0,875	1,058	0,914	0,740	325,97 (68)	55,91 (20)***	0,990 (0,001)	0,017 (0,015;0,019)	0,986 (0,00)
Řecko	0,860	0,990	0,785	0,833	255,99 (44)	106,09 (12)***	0,990 (0,004)	0,023 (0,020;0,025)	0,986 (0,004)
Maďarsko	0,928	1,046	0,794	0,879	182,32 (56)	39,07 (16)***	0,992 (0,001)	0,018 (0,015;0,021)	0,989 (0,002)
Švýcarsko	0,845	1,022	0,896	0,894	337,46 (68)	37,53 (20)***	0,983 (0,001)	0,020 (0,018;0,022)	0,978 (0,00)
Irsko	0,888	0,971	0,878	0,691	283,40 (68)	67,58 (20)***	0,990 (0,002)	0,016 (0,014;0,018)	0,987 (0,00)
Izrael	0,759	0,980	0,753	0,686	198,43 (44)	75,46 (12)***	0,989 (0,004)	0,020 (0,017;0,023)	0,985 (0,003)
Itálie	0,762	0,878	0,849	0,672	56,41 (20)	4,04 (4)	0,991 (0,00)	0,026 (0,018;0,035)	0,987 (0,003)
Lucembur.	0,736	1,013	0,892	0,861	59,77 (20)	4,09 (4)	0,990 (0,00)	0,028 (0,020;0,036)	0,986 (0,00)
Nizozemí	0,792	1,137	0,963	0,829	421,18 (68)	86,29 (20)***	0,985 (0,003)	0,022 (0,020;0,024)	0,980 (0,00)
Norsko	0,737	1,071	0,825	0,686	355,03 (68)	27,73 (20)***	0,983 (0,001)	0,020 (0,018;0,023)	0,978 (0,00)
Polsko	0,865	1,012	0,813	0,830	163,30 (68)	42,14 (20)***	0,994 (0,001)	0,012 (0,010;0,014)	0,992 (0,00)
Portugalsko	0,944	1,135	0,752	0,813	365,23 (68)	124,77 (20)***	0,985 (0,005)	0,020 (0,018;0,022)	0,980 (0,001)
Švédsko	0,789	1,047	0,914	0,771	312,88 (68)	32,52 (20)***	0,988 (0,00)	0,019 (0,016;0,021)	0,984 (0,00)
Slovinsko	0,781	0,981	0,838	0,942	127,03 (68)	36,78 (20)***	0,996 (0,001)	0,011 (0,008;0,013)	0,995 (0,00)

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Poznámka 2: Položky FAIR a PARLAM jsou referenční a jejich faktorová zátěž je ve všech zemích zafixována na jednotku.

Tabulka 93 Výsledky testu modelu skalární ekvivalence (Model 3) – porovnání s metrickým modelem (Model 2)

země	Statistiky vhodnosti modelu					Rozhodnutí o ekvivalenci (E)
	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	
Rakousko	426,01 (61)	138,06 (17)***	0,981 (0,006)	0,026 (0,024;0,029)	0,982 (0,001)	částečná E
Belgie	461,99 (92)	188,03 (24)***	0,980 (0,009)	0,020 (0,018;0,021)	0,980 (0,00)	částečná E
Česká republika	548,39 (75)	291,11 (19)***	0,983 (0,01)	0,025 (0,023;0,027)	0,983 (0,007)	částečná E
Německo	693,43 (98)	267,71 (30)***	0,980 (0,008)	0,019 (0,018;0,020)	0,982 (0,002)	plná E
Dánsko	561,9 (93)	193,04 (25)***	0,971 (0,01)	0,024 (0,022;0,026)	0,972 (0,003)	částečná E
Španělsko	281,14 (89)	69,31 (21)***	0,988 (0,003)	0,015 (0,013;0,017)	0,987 (0,002)	částečná E
Finsko	798,39 (98)	257,06 (30)***	0,971 (0,009)	0,024 (0,023;0,026)	0,973 (0,001)	plná E
Francie	345,69 (80)	146,54 (24)***	0,979 (0,01)	0,019 (0,017;0,021)	0,980 (0,005)	plná E
Spojené království	545,58 (96)	219,60 (28)***	0,982 (0,008)	0,019 (0,018;0,021)	0,983 (0,003)	částečná E
Řecko	334,46 (53)	78,46 (9)***	0,987 (0,003)	0,024 (0,021;0,026)	0,985 (0,001)	částečná E
Maďarsko	349,32 (72)	167,00 (16)***	0,982 (0,01)	0,023 (0,021;0,026)	0,981 (0,008)	částečná E
Švýcarsko	520,58 (98)	183,13 (30)***	0,974 (0,009)	0,021 (0,019;0,023)	0,981 (0,001)	plná E
Irsko	504,20 (86)	220,81 (18)***	0,981 (0,009)	0,020 (0,018;0,022)	0,980 (0,007)	částečná E
Izrael	301,61 (55)	103,18 (11)***	0,982 (0,007)	0,022 (0,020;0,025)	0,981 (0,004)	částečná E
Itálie	92,61 (25)	36,2 (5)***	0,984 (0,007)	0,032 (0,025;0,039)	0,981 (0,006)	částečná E
Lucembursko	110,84 (26)	51,07 (6)***	0,980 (0,01)	0,036 (0,029;0,043)	0,977 (0,009)	plná E
Nizozemí	668,95 (96)	247,76 (28)***	0,975 (0,01)	0,023 (0,021;0,025)	0,977 (0,003)	částečná E
Norsko	488,35 (93)	133,32 (25)***	0,977 (0,006)	0,021 (0,019;0,022)	0,978 (0,00)	částečná E
Polsko	322,41 (85)	159,11 (17)***	0,985 (0,009)	0,017 (0,015;0,019)	0,984 (0,006)	částečná E
Portugalsko	564,88 (86)	199,65 (18)***	0,975 (0,01)	0,022 (0,020;0,024)	0,974 (0,006)	částečná E
Švédsko	544,24 (94)	231,37 (26)***	0,978 (0,01)	0,021 (0,020;0,023)	0,979 (0,005)	částečná E
Slovinsko	261,27 (89)	134,23 (21)***	0,989 (0,007)	0,016 (0,014;0,018)	0,989 (0,006)	částečná E

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Mnohem překvapivější jsou ale výsledky testování skalárního modelu v jednotlivých zemích. Prvním předpokladem bylo, že dosáhnout plné skalární ekvivalence položek v dané zemi nebude příliš obtížné – položky totiž nejsou vystaveny možnému systematickému zkreslení jinými historickými a kulturními kořeny nebo odlišnými rodnými jazyky respondentů. Ukazuje se ale, že plné skalární ekvivalence všech šesti položek v obou faktorech bylo dosaženo pouze v pěti zemích – Německu, Francii, Švýcarsku, Finsku a Lucembursku. Ve všech ostatních zemích muselo dojít na základě modifikačních indexů k uvolnění restrikce shody konstant alespoň u jedné položky. Obecně se zdá, že mnohem stabilnější je z hlediska komparace faktor mezilidské důvěry. Jeho položky nebyly srovnatelné v čase pouze v první vlně šetření v Izraeli, ve druhé vlně v Irsku, první a druhé vlně šetření v Polsku a třetí vlně šetření v Portugalsku. Nejčastěji musely být uvolněny konstanty položek HELPFUL a TRUSTED, ale z přehledu je zřejmé, že se jedná spíše o individuální případy (viz Tabulka 94).

Problém systematictějšího rázu je spatřován spíše v položkách faktoru důvěry ve veřejné instituce. Daný faktor, respektive jeho měřicí položky nejsou ani částečně skalárně ekvivalentní alespoň v jedné vlně šetření ve čtyřech zemích – České republice, Dánsku, Švédsku a Izraeli, a dokonce zcela nesrovnatelné v čase v devíti dalších zemích – např. v Polsku, Irsku, Řecku, Španělsku, atp. Průměry těchto položek (PARLAM, POLIT a LEGAL) anebo této latentní proměnné (INSTRUST) tak nejsou vzájemně porovnatelné v dané zemi někdy i ve více než dvou vlnách šetření a sledování vývoje hodnocení institucionální důvěry na základě některých statistických ukazatelů je tak prakticky nemožné. Největší problém z hlediska srovnatelnosti vykazuje položka LEGAL (44 případů vynechání) a PARLAM (37 případů vynechání). V některých zemích dochází ke specifické situaci, kdy musely být uvolněny konstanty jedné jediné nesrovnatelné položky ve všech vlnách šetření, například v České republice, Polsku a Slovinsku u položky LEGAL.

Tabulka 94 Přehled položek s uvolněnými konstantami v částečně skalárním modelu v jednotlivých zemích a vlnách šetření/1. část

země	První vlna	Druhá vlna	Třetí vlna	Čtvrtá vlna	Pátá vlna	Šestá vlna
	ESS (2002)	ESS (2004)	ESS (2006)	ESS (2008)	ESS (2010)	ESS (2012)
Německo						
Francie						
Švýcarsko						
Finsko						
Lucembursko						
Spojené království	PARLAM					LEGAL
Itálie	LEGAL	LEGAL				
Rakousko				LEGAL		
Nizozemí	LEGAL	PARLAM				
Norsko	LEGAL				POLIT	PARLAM
						LEGAL
						POLIT
Česká republika	LEGAL	LEGAL		LEGAL	LEGAL	LEGAL
	POLIT					
Belgie	PARLAM			PARLAM	LEGAL	
	POLIT				POLIT	
					PARLAM	
Dánsko	LEGAL			PARLAM		LEGAL
				LEGAL		POLIT
Portugalsko	HELPFUL	PARLAM	FAIR	PARLAM	PARLAM	PARLAM
	PARLAM	POLIT	TRUSTED			LEGAL
	POLIT		PARLAM			POLIT

Tabulka 94 Přehled položek s uvolněnými konstantami v částečně skalárním modelu v jednotlivých zemích a vlnách šetření/2. část

země	První vlna	Druhá vlna	Třetí vlna	Čtvrtá vlna	Pátá vlna	Šestá vlna
Švédsko	LEGAL	POLIT		TRUSTED		TRUSTED
Španělsko	LEGAL	HELPFUL	LEGAL	LEGAL	HELPFUL	PARLAM
	POLIT	POLIT		PARLAM	PARLAM	POLIT
		LEGAL			POLIT	LEGAL
Slovensko	LEGAL	LEGAL	LEGAL	LEGAL	PARLAM	PARLAM
					LEGAL	LEGAL
					POLIT	POLIT
Maďarsko	PARLAM	PARLAM	PARLAM	PARLAM	PARLAM	
	LEGAL			HELPFUL		
	POLIT			POLIT		
Řecko	PARLAM	PARLAM		PARLAM	PARLAM	
	LEGAL	LEGAL		LEGAL	LEGAL	
	POLIT	POLIT		POLIT	POLIT	
Irsko		PARLAM	PARLAM		PARLAM	PARLAM
		LEGAL	LEGAL			
		POLIT	POLIT			
		HELPFUL				
		TRUSTED				
Izrael	TRUSTED				PARLAM	HELPFUL
	HELPFUL					
	PARLAM					
	LEGAL					
Polsko	FAIR	TRUSTED	PARLAM		LEGAL	
	TRUSTED	HELPFUL				
	HELPFUL					
	PARLAM	PARLAM	LEGAL			
	LEGAL	LEGAL				
POLIT						

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Zelená barva značí nižší hodnoty uvolněných parametrů položek ve srovnání s plně ekvivalentním modelem, červená barva vyšší hodnoty parametrů.

Poznámka 2: Šedá pole značí absenci země ve výzkumu, k modře označeným zemím nebylo ještě možné v lednu 2014 získat nové datové soubory za šestou vlnu šetření ESS (2012).

Příčiny tohoto neočekávaného zjištění mohou být různé, ale zcela jistě potvrzují vysokou citlivost skalární ekvivalence jak vůči rozdílným historickým a jazykovým kořenům daných zemí, tak i vůči drobným změnám ve vnímání obsahu dotazníkových položek a jejich měřících škál způsobených vlivem postojových změn ve společnosti. Vzhledem k eliminaci kulturních i jazykových rozdílů v časové analýze jedné země se jako nejpravděpodobnější zkreslující faktor jeví podle rozboru Roberta, Lee a Chana [2006: 69] kontext odpovědi, referenční efekt porovnávání odpovědi s názory blízké skupiny, sociálně desirabilní charakter položek, vliv specifických stylů odpovědí respondentů a možný výskyt jazykových bariér přistěhovaleckých skupin v některých zemích, které se v průběhu času vzhledem k imigrační politice Evropské unie stále zvětšují. Nebo se také může jednat o proměnu interpretace a vnímání rozsahu měřící škály důvěry vyvolanou sociálními, politickými a ekonomickými zvraty, které se v celé Evropě v posledních deseti letech odehrávají. Jaký zdroj zkreslení je ale relevantní pro jakou zemi a proč se daná zjištění nevztahují k otázkám mezilidské důvěry, ale

zejména k položkám institucionální důvěry, je další výzkumnou otázkou, na kterou není možné v této studii odpovědět vyčerpávajícím způsobem.

Možnosti porovnávání strukturních a latentních průměrů jsou ilustrovány v Tabulkách 95 (faktor mezilidské důvěry) a 96 (faktor důvěry v instituce). Oranžová pole značí, že v dané vlně šetření není možné ani strukturní, ani latentní průměr porovnávat s ostatními průměry faktorů z jiných vln šetření, neboť nebyla prokázána částečná skalární ekvivalence příslušných položek. V případě strukturních průměrů je uveden pouze statisticky významný rozdíl mezi nejmenší a největší velikostí průměru, ostatní kombinace nebylo možné do tabulky zahrnout. Proměnlivost hodnocení mezilidské důvěry či důvěry v instituce v *dané zemi* byla určena na základě statisticky významných rozdílů mezi průměry latentní proměnné v daných vlnách šetření. Pro každou zemi je situace značně specifická.

V některých případech si je hodnocení důvěry v jednotlivých vlnách šetření v dané zemi velmi podobné (tj. latentní proměnné generují stejná pořadí země v jednotlivých vlnách), v dané zemi pravděpodobně tedy nedošlo v čase k žádné změně, např. situace Finska, Francie a Dánska v hodnocení mezilidské důvěry. Porovnávání průměrů v Rakousku, Řecku, Nizozemí a Slovinsku zase odhalilo postupné zvyšování důvěry respondentů v lodi obecně. Postupné zhoršování mezilidské důvěry lze nalézt v Irsku, Švédsku a Portugalsku. V každé ze jmenovaných zemí je ještě možné posoudit velikost strukturních průměrů, které vesměs značí průměrnou míru mezilidské důvěry v daných zemích až na drobné výjimky, např. nápadně vyšší míru důvěry v severských zemích – Dánsku, Finsku, Norsku a Švédsku - a překvapivě značně nízkou v Řecku a Polsku.

Pomineme-li mnohem horší srovnatelnost hodnocení míry důvěry ve veřejné instituce v jednotlivých zemích, ukazuje se, že kromě stále optimistického Finska, Švýcarska a Švédska většina zemí manifestuje spíše zhoršující se důvěru v politické instituce (např. Polsko, Česká republika) anebo minimálně stabilně nízké či kolísavé hodnocení. Z hlediska velikosti strukturních průměrů se dlouhodobě nachází v nadprůměrné úrovni hodnocení pouze Dánsko, Finsko, Švýcarsko a Nizozemí. Nejnížší hodnocení míry důvěry ve veřejné instituci dnes vykazuje Česká republika, Polsko a Portugalsko (viz Tabulka 96).

V dalším kroku byla otestována shoda modelu ekvivalence kovariancí faktorů v každé zemi zvlášť. Výsledky testů odhalily naprostou shodu vztahu mezi faktory napříč všemi vlnami šetření v každé zemi (tzn. konstruktovou validitu modelu). Model ekvivalence kovariancí faktorů byl porovnáván se statistikami vhodnosti skalárního (nebo částečně skalárního) modelu a v řadě případů se téměř zcela shodoval s předchozím modelem skalární ekvivalence ($\Delta CFI \leq 0,01$). V tabulce 97 je obsaženo také shrnutí velikostí kovariancí mezi faktory mezilidské a institucionální důvěry z modelu ekvivalence kovariancí faktorů (Model 5), které jsou značně vysoké. Mezilidská důvěra a důvěra v instituce se tedy v každé zemi stabilně po deset let vzájemně silně ovlivňují.

Dalším překvapením byly výstupy testu striktní ekvivalence položek. Model ekvivalence variancí chyb měření prokázal svou shodu s předchozím modelem ve většině případů ($\Delta CFI \leq 0,01$) a bylo tak ve 13 zemích potvrzeno, že položky měřící dané faktory mají ve všech vlnách šetření stejně velké chyby měření a tím pádem stejnou míru reliability daných položek (viz Tabulka 98). Pouze v devíti zemích bylo nutné v jedné nebo více vlnách šetření uvolnit na základě modifikačních indexů variance chyb měření. Nejčastěji se jednalo o položku POLIT. V řadě zemí šlo v průběhu času o vylepšení míry reliability dané položky a snížení chyby měření oproti předešlým vlnám výzkumu. Například v případě čtvrté a páté vlny šetření v Řecku došlo ke snížení variancí chyb měření oproti první vlně u čtyř položek (viz Tabulka 99). Většinou se jedná o výskyt větších chyb měření položek v prvních vlnách šetření a jejich snížení a zlepšení reliability těchto položek v pozdějších vlnách šetření v dané zemi, což by mohlo být odrazem postupného zkvalitňování designu výzkumu ESS a sběru dat.

Tabulka 95 Porovnání změn strukturálních a latentních průměrů v dané zemi v průběhu šesti vln šetření ESS (2002 – 2012) – mezilidská důvěra

země	ESS (2002)		ESS (2004)		ESS (2006)		ESS (2008)		ESS (2010)		ESS (2012)	
	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr
Rakousko	5,23*	0,00 (4)	5,39	0,18*(2)	5,36	0,15*(3)	5,54*	0,31*(1)				
Belgie	4,96	0,00 (5)	4,95*	-0,01 (5)	5,14	0,16* (4)	5,27*	0,28*(1)	5,23	0,24*(2)	5,16	0,20* (3)
Česká republika	4,46	0,00 (3)	4,44*	-0,03 (3)			4,74*	0,28*(1)	4,64	0,15*(2)	4,62	0,09 (3)
Německo	5,09	0,00 (4)	5,08*	-0,01 (4)	5,14	0,04 (4)	5,26	0,17* (2)	5,17	0,09* (3)	5,31*	0,22* (1)
Dánsko	6,82	0,00 (1)	6,7	-0,13* (3)	5,14*	0,03 (1)	6,81	-0,04 (1)	6,75	-0,11* (2)	6,83*	-0,01 (1)
Španělsko	4,87	0,00 (4)	4,78*	-0,06 (4)	5,02	0,16*(3)	4,91	0,03 (4)	5,09	0,18*(2)	5,11*	0,21* (1)
Finsko	6,33	0,00 (2)	6,34	0,01 (2)	6,44	0,09 (1)	6,32	-0,03 (2)	6,36	0,01 (2)	6,42	0,07 (2)
Francie	4,83	0,00 (1)	4,92	0,08 (1)	4,92	0,08 (1)	4,91	0,07 (1)	4,86	0,02 (1)		
Spoj. království	5,35*	0,00 (5)	5,45	0,09 (5)	5,64	0,29*(2)	5,53	0,18*(4)	5,59	0,24*(3)	5,66*	0,31*(1)
Řecko	3,45	0,00 (4)	3,62	0,17*(3)			3,65	0,18*(2)	3,74	0,32*(1)		
Maďarsko	4,31	0,00 (3)	4,21*	-0,1 (3)	4,47	0,16*(2)	4,35	0,04 (3)	4,67	0,36*(1)		
Švýcarsko	5,75*	0,00(6)	5,89	0,15*(3)	5,95	0,20*(1)	5,90	0,14*(4)	5,87	0,13*(5)	5,95*	0,17*(2)
Irsko	5,80	0,00 (1)			5,68	-0,12*(2)	5,84	0,03 (1)	5,46*	-0,35*(4)	5,53	-0,28*(3)
Izrael							5,21	0,00 (1)	5,11	-0,14* (2)	5,24	-0,02 (1)
Itálie	4,43*	0,00 (1)	4,19*	-0,23*(2)								
Lucembursko	5,13	0,00 (1)	5,15	-0,04 (1)								
Nizozemí	5,73*	0,00 (4)	5,81	0,07 (4)	5,77	0,05 (4)	5,90	0,16*(3)	6,04*	0,30*(1)	5,98	0,24* (2)
Norsko	6,53	0,00 (2)	6,51	-0,01 (2)	6,64	0,13*(1)	6,52	-0,01 (2)	6,58	0,05 (2)	6,59	0,05 (2)
Polsko					4,13*	0,00 (4)	4,28	0,32*(2)	4,45*	0,49* (1)	4,28	0,31* (3)
Portugalsko	4,31*	0,00 (1)	4,18	-0,13*(2)			4,12	-0,21* (4)	4,19	-0,13* (3)	4,05*	-0,26* (5)
Švédsko	6,26	0,00 (3)	6,17*	-0,1* (4)	6,32	0,05 (3)	6,38	0,12*(2)	6,40*	0,14* (1)	6,19	-0,09 (3)
Slovinsko	4,30	0,00 (5)	4,44	0,14* (4)	4,50	0,19*(3)	4,73	0,43*(2)	4,28*	-0,06 (5)	4,87*	0,56* (1)

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Čísla v závorkách za latentním průměrem značí pořadí, oranžová pole značí, že položky daného faktoru nebyly ve vyznačeném časovém úseku ani částečně skalární a nemůže být porovnáván průměr tohoto faktoru s ostatními.

Poznámka 2: Strukturální průměr je vypočten jako průměr skóre hodnocení měřicích položek daného faktoru za danou zemi, latentní průměr je výstupem analýzy strukturálního modelování (resp. MGCFA). * značí statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 3: Šedá pole značí absenci země ve výzkumu, k modře označeným zemím nebylo ještě možné v lednu 2014 získat nové datové soubory za šestou vlnu šetření ESS (2012).

Tabulka 96 Porovnání změn strukturálních a latentních průměrů v dané zemi v průběhu šesti vln šetření ESS (2002 – 2012) – důvěra v instituce

země	ESS (2002)		ESS (2004)		ESS (2006)		ESS (2008)		ESS (2010)		ESS (2012)	
	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr	Strukt. průměr	Latentní průměr
Rakousko	4,89*	0,00 (1)	4,63	-0,3* (4)	4,76	-0,17* (2)	4,62*	-0,26* (3)				
Belgie			5,41	0,00 (1)	5,25	-0,20* (3)	5,47	0,10 (1)			5,25*	-0,19* (2)
Česká republika			6,83*	0,00 (1)			6,69	-0,08 (1)	6,67	-0,11* (2)	6,75	-0,01 (1)
Německo	5,54	0,00 (3)	5,75*	0,25* (2)	5,75	0,27* (1)	5,44	-0,14* (4)	5,66	0,17* (3)	5,23*	-0,36* (5)
Dánsko	3,73	0,00 (1)	3,61*	-0,15* (2)	5,75*	-0,25* (3)			3,91	-0,34* (4)		
Španělsko												
Finsko	4,22	0,00 (2)	4,06	-0,18* (4)	3,99*	-0,21* (6)	4,01	-0,19* (5)	4,42*	0,34* (1)	4,07	-0,13* (3)
Francie	4,30	0,00 (4)	4,18	0,14* (3)	4,16	0,21* (2)	4,33*	0,02 (4)	4,08*	0,31* (1)		
Spoj. království	5,51*	0,00 (6)	5,75	0,33* (3)	5,83*	0,43* (1)	5,69	0,29* (4)	5,77	0,41* (2)	5,56	0,22* (5)
Řecko												
Maďarsko			6,42	0,00(2)	6,56*	-0,19*(3)			6,00*	0,53*(1)		
Švýcarsko	4,38	0,00 (2)	4,55*	0,22* (1)	4,38	0,02 (2)	4,35	-0,01 (2)	4,30	-0,07 (2)	4,03*	-0,37* (3)
Irsko	5,56	0,00 (2)					6,01	0,6* (1)				
Izrael							6,00	0,00 (1)			5,71	-0,34* (2)
Itálie	4,67	0,00 (2)	4,26	0,41* (1)								
Lucembursko	5,58	0,00 (1)	5,67	-0,14* (2)								
Nizozemí	4,86	0,00 (2)	5,10*	0,31* (1)	4,65	-0,19* (3)	4,47*	-0,39* (6)	4,52	-0,33* (5)	4,56	-0,24* (4)
Norsko	4,46	0,00 (2)	4,66*	0,28* (1)	4,45	0,05 (2)	4,27	-0,14* (3)	4,05	-0,41* (4)		
Polsko							6,96	0,00 (1)	6,55*	-0,76* (3)	7,07*	-0,20* (2)
Portugalsko					6,51	0,00 (3)	6,74	0,27* (2)	7,20	0,84* (1)		
Švédsko	4,40	0,00 (4)	4,88*	0,57* (1)	4,61	0,27* (2)	4,50	0,14* (3)	4,04	-0,37* (5)	4,35*	-0,03 (4)
Slovinsko	6,21	0,00 (1)	6,30	0,04 (1)	6,13	-0,11 (1)	5,96*	-0,31* (2)				

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Čísla v závorkách za strukturálním a latentním průměrem značí pořadí, oranžová pole značí, že položky daného faktoru nebyly ve vyznačeném časovém úseku ani částečně skalární a nemůže být porovnáván průměr tohoto faktoru s ostatními.

Poznámka 2: Strukturální průměr je vypočten jako průměr skóre hodnocení měřících položek daného faktoru za danou zemi, latentní průměr je výstupem analýzy strukturálního modelování (resp. MGCFA). * značí statisticky významnou odlišnost daného průměru od ostatních průměrů na hladině 95 %.

Poznámka 3: Velikost strukturálního průměru směřuje od nejlepšího hodnocení (1) k nejhoršímu hodnocení (10) politických institucí.

Poznámka 4: Šedá pole značí absenci země ve výzkumu, k modře označeným zemím nebylo ještě možné v lednu 2014 získat nové datové soubory za šestou vlnu šetření ESS (2012).

Tabulka 97 Výsledky testu modelu ekvivalence kovariancí faktorů (Model 4) – porovnání se skalárním modelem (Model 3)

země	Kovariance faktorů v Modelu 4	Statistiky vhodnosti modelu				
		χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)
Rakousko	-1,601	518,67 (64)	92,66 (3)***	0,977 (0,004)	0,029 (0,027;0,031)	0,978 (0,004)
Belgie	-1,310	471,928 (97)	9,94 (5)**	0,980 (0,00)	0,019 (0,018;0,021)	0,981 (0,00)
Česká republika	-1,456	559,18 (79)	10,79 (4)**	0,983 (0,00)	0,024 (0,023;0,026)	0,984 (0,001)
Německo	-1,490	713,59 (103)	20,17 (5)***	0,979 (0,001)	0,019 (0,018;0,020)	0,983 (0,001)
Dánsko	-1,172	567,81 (98)	5,91 (5)	0,971 (0,00)	0,023 (0,021;0,025)	0,973 (0,001)
Španělsko	-0,983	295,21 (89)	14,07 (0)**	0,988 (0,00)	0,015 (0,013;0,017)	0,988 (0,001)
Finsko	-1,288	805,19 (102)	6,80 (4)	0,971 (0,00)	0,024 (0,022;0,026)	0,974 (0,00)
Francie	-1,233	354,18 (84)	8,49 (4)	0,978 (0,001)	0,019 (0,017;0,021)	0,981 (0,00)
Spojené království	-1,386	562,99 (101)	17,42 (5)***	0,982 (0,00)	0,019 (0,017;0,020)	0,984 (0,00)
Řecko	-1,003	355,32 (56)	20,85 (3)***	0,986 (0,001)	0,024 (0,021;0,026)	0,985 (0,00)
Maďarsko	-1,541	363,22 (76)	13,9 (4)***	0,981 (0,001)	0,023 (0,021;0,025)	0,981 (0,00)
Švýcarsko	-1,013	529,19 (103)	8,61 (5)	0,974 (0,00)	0,020 (0,019;0,022)	0,977 (0,00)
Irsko	-1,229	520,33 (91)	16,13 (5)***	0,980 (0,001)	0,020 (0,018;0,021)	0,981 (0,00)
Izrael	-1,338	323,85 (58)	22,24 (3)***	0,981 (0,001)	0,023 (0,020;0,025)	0,980 (0,001)
Itálie	-1,459	92,66 (26)	0,05 (1)	0,984 (0,00)	0,031 (0,025;0,038)	0,982 (0,00)
Lucembursko	-1,469	110,88 (27)	0,04 (1)	0,980 (0,00)	0,035 (0,028;0,042)	0,978 (0,00)
Nizozemí	-1,237	691,33 (101)	22,38 (5)***	0,975 (0,00)	0,023 (0,021;0,024)	0,977 (0,00)
Norsko	-1,159	499,99 (98)	11,65 (5)***	0,977 (0,00)	0,020 (0,018;0,022)	0,979 (0,00)
Polsko	-1,265	344,50 (90)	22,08 (5)***	0,984 (0,001)	0,017 (0,015;0,019)	0,984 (0,00)
Portugalsko	-1,013	615,03 (91)	50,15 (5)***	0,973 (0,002)	0,023 (0,021;0,024)	0,973 (0,001)
Švédsko	-1,410	554,54 (99)	10,30 (5)***	0,977 (0,001)	0,021 (0,019;0,023)	0,979 (0,00)
Slovinsko	-1,726	267,29 (94)	6,02 (5)	0,989 (0,00)	0,015 (0,013;0,018)	0,989 (0,00)

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Tabulka 98 Výsledky testu modelu ekvivalence variancí chyb měření (Model 6) – porovnání s modelem ekvivalence variancí faktorů (Model 5)

země	Statistiky vhodnosti modelu					Rozhodnutí o ekviv. (E)
	χ^2 (df)	$\Delta \chi^2$ (Δ df)	CFI (Δ CFI)	RMSEA (90 % C. I.)	TLI (Δ TLI)	
Rakousko	735,29 (87)	184,77 (17)***	0,967 (0,008)	0,030 (0,028;0,032)	0,977 (0,002)	částečná E
Belgie	696,30 (137)	197,59 (30)***	0,969 (0,01)	0,020 (0,018;0,021)	0,980 (0,00)	plná E
Česká republika	1027,33 (111)	240,84 (24)***	0,967 (0,008)	0,029 (0,027;0,030)	0,978 (0,00)	plná E
Německo	847,34 (143)	100,03 (30)***	0,976 (0,003)	0,017 (0,016;0,018)	0,985 (0,002)	plná E
Dánsko	763,05 (137)	163,84 (29)***	0,961 (0,009)	0,023 (0,021;0,024)	0,975 (0,00)	částečná E
Španělsko	579,54 (129)	154,01 (30)***	0,973 (0,008)	0,018 (0,017;0,020)	0,981 (0,001)	plná E
Finsko	963,20 (142)	119,54 (30)***	0,966 (0,003)	0,022 (0,021;0,023)	0,978 (0,00)	plná E
Francie	455,85 (116)	74,17 (24)***	0,973 (0,004)	0,018 (0,016;0,020)	0,982 (0,001)	plná E
Spojené království	725,74 (141)	127,72 (30)***	0,977 (0,004)	0,018 (0,017;0,019)	0,985 (0,00)	plná E
Řecko	629,83 (73)	185,74 (11)***	0,974 (0,008)	0,028 (0,026;0,031)	0,979 (0,004)	částečná E
Maďarsko	548,48 (107)	134,65 (23)***	0,971 (0,007)	0,024 (0,022;0,026)	0,980 (0,001)	částečná E
Švýcarsko	654,47 (143)	97,30 (30)***	0,969 (0,004)	0,019 (0,017;0,020)	0,980 (0,00)	plná E
Irsko	801,61 (129)	215,64 (28)***	0,969 (0,009)	0,021 (0,019;0,022)	0,979 (0,001)	částečná E
Izrael	483,84 (80)	118,78 (16)***	0,971 (0,007)	0,024 (0,022;0,026)	0,978 (0,002)	částečná E
Itálie	115,02 (34)	18,89 (6)***	0,981 (0,003)	0,030 (0,024;0,036)	0,983 (0,00)	plná E
Lucembursko	134,64 (35)	18,34 (6)***	0,976 (0,003)	0,033 (0,027;0,039)	0,980 (0,00)	plná E
Nizozemí	891,96 (141)	157,84 (30)***	0,968 (0,005)	0,022 (0,020;0,023)	0,979 (0,00)	plná E
Norsko	708,72 (135)	190,36 (27)***	0,967 (0,009)	0,020 (0,019;0,022)	0,978 (0,002)	částečná E
Polsko	537,94 (130)	193,44 (40)***	0,974 (0,008)	0,018 (0,016;0,019)	0,982 (0,002)	plná E
Portugalsko	905,48 (130)	209,82 (29)***	0,960 (0,009)	0,023 (0,021;0,024)	0,972 (0,001)	částečná E
Švédsko	801,35 (139)	197,28 (30)***	0,967 (0,008)	0,021 (0,020;0,023)	0,979 (0,001)	plná E
Slovinsko	425,76 (128)	150,78 (24)***	0,981 (0,008)	0,017 (0,016;0,019)	0,987 (0,004)	částečná E

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty - každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA. Změna ve velikosti indexů (Δ) nesmí přesáhnout u CFI 0,01, u RMSEA 0,015. U $\Delta \chi^2$ značí * statistickou významnost na hladině 90 %, ** na hladině 95 % a *** na hladině 99 %.

Tabulka 99 Uvolněné variance chyb měření při testu modelu ekvivalence chyb měření (tzv. striktní ekvivalence) – srovnání položek v dané zemi v čase

země		AT4	DK2	GR1	GR4	GR5	HU1	IE2	IE5	IL1	IL5	NO1	NO2	PT3	SI1	SI3	SI5	SI6
Latentní proměnná	Měřicí položka																	
důvěra v instituce	POLIT			x		x										x	x	x
	PARLAM						x			x							x	
	LEGAL				x							x			x		x	
mezilidská důvěra	TRUSTED		x	x	x							x	x			x		
	FAIR	x		x	x			x	x					x				
	HELPFUL										x							

Zdroj: ESS 2002, ESS 2004, ESS 2006, ESS 2008, ESS 2010, ESS 2012; vlastní výpočty

Poznámka: V zeleném poli jsou označeny ty chyby měření, které se ukázaly být menší (a reliabilnější) než chyby měření položky z jiné vlny šetření v dané zemi. Naopak červené pole označuje chyby měření, které byly podstatně vyšší a značí nižší reliabilitu položky v časovém úseku (pouze ale ve srovnání v čase v rámci jedné země).

Poznámka 2: Číslo za zkratkou země značí číslo vlny šetření (1 = první vlna šetření; 6 = šestá vlna šetření).

ZÁVĚR

Tato práce pojednala téma srovnatelnosti postojových škál v mezinárodních výzkumech především z metodologického hlediska. Jako úvod do tématu měl čtenář nejprve šanci seznámit se s historií komparativních výzkumů, s konceptem ekvivalence a s její základní kategorizací na teoretické úrovni a úrovni měření, přičemž tento počin je třeba chápat jako jednu z řady možností, jak jednotlivé druhy ekvivalence vykládat a jak chápat jejich vzájemné postavení. První tři teoretičtější laděné kapitoly měly posloužit jednak ke sblížení se specifickou terminologií a následně k udržení logické linie při výkladu technik určených k testování srovnatelnosti škál a jejich aplikaci na konkrétní dvě postojové škály. Případným zájemcům snad budou v této části práce užitečnými také odkazy na empirické studie, které vysvětlují aplikaci těchto techniky při testování ekvivalence položek na různých mezinárodních datech a další bohatě citovaná zahraniční i česká odborná literatura.

V další části práce se čtenář již dostal k hlavnímu výzkumnému cíli této studie, kterým bylo posoudit mezinárodní srovnatelnost jedněch z nejznámějších, často v mezinárodních výzkumech využívaných postojových škál sociálního kapitálu a politického odcizení, tak jak jsou měřeny v Evropském sociálním výzkumu (2002). Na základě bohatých empirických materiálů uvedených v této studii lze jejich srovnatelnost zhodnotit následujícím způsobem. Postojové škály byly v datech ověřeny jako konfigurálně a metricky ekvivalentní. Je tedy zřejmé, že koncepty sociálního kapitálu a politického odcizení jsou ve všech zemích chápány podobně a že jejich měřicí položky jsou interpretovány stejným způsobem. Skalární ekvivalence těchto škál ve všech zemích potvrzena nebyla, je tedy pravděpodobné, že rozsah měřicí škály není vnímán respondenty z různých kultur stejně a je odlišně využíván, a že jednotlivé slovní i numerické body hodnocení položky v postojové škále mají pro respondenty také různý význam. Proto i když respondenti zastávají více méně shodný názor a teoretický koncept hodnotí stejně, jejich naměřený skór nemusí být stejný. Protože bylo do analýzy zařazeno příliš velké množství kulturně a jazykově odlišných zemí, pravděpodobně došlo ke kombinaci různých zdrojů chyb měření, které způsobily odchylky ve velikostech konstant položek, i když ty stále měří stejný koncept.

Na základě toho bylo jednoznačně vyvozeno, že jednotlivé úrovně ekvivalence dat na úrovni měření jsou různě citlivé na jemné kulturní a jazykové odlišnosti mezi zkoumanými zeměmi a na kontext dlouhodobého vývoje názorů a postojů ve společnosti. Konfigurální ekvivalence je vůči zmíněným odlišnostem obecně nejodolnější. Metrická ekvivalence je nejvíce narušena zřejmými kulturními rozdíly mezi zeměmi a odlišnými interpretacemi měřicích položek. Skalární ekvivalence je potom ze všech typů ekvivalencí nejvíce citlivá na kombinaci jak kulturních, tak i jazykových odlišností a to i těch velmi jemných, a navíc ji ovlivňuje i vývoj společenských postojů k daným konceptům – tedy určitý historický kontext, jak ukazuje simultánní časová analýza srovnatelnosti škál důvěry.

To vše jsou důležité poznatky proto, že dosažená úroveň ekvivalence postojových škál dále určuje, jakým způsobem lze s daty komparativně nakládat. Implikace daných závěrů do praxe je potom taková, že ve všech participujících zemích v ESS (2002) není problém vzájemně porovnávat vztahy mezi těmito dvěma postojovými škálami či vztahy těchto postojových škál s jinou další metricky ekvivalentní proměnnou z datového souboru, jednak prostřednictvím korelační či regresní analýzy, tak i strukturního modelování. Tento typ meritorních analýz je na mezinárodní data aplikován také nejčastěji, takže zde se žádný zásadní problém neobjevuje. Komplikace by nastala, až když by bylo výzkumným záměrem porovnat i průměry jednotlivých postojových škál nebo jejich měřicích položek mezi zeměmi, protože výsledky skalární ekvivalence ukazují, že by taková komparace nebyla validní a že to tudíž provést nelze. Pokud by byl ale o takovýto druh srovnávací analýzy průměrů opravdový zájem, podrobnější simultánní analýza položek ukazuje, že v určitých malých skupinách zemí

je možné průměry postojových škál vzájemně srovnávat. Většinou se jedná o homogenní skupiny zemí jak z hlediska geografie, kultury a historie, tak i jazykových kořenů, protože více méně jakákoli větší jazyková odchylka znemožňuje komparovat průměry postojových škál.

Rámeček 5 Homogenní skupiny zemí ze simultánní analýzy s nejvyšší úrovní ekvivalence postojových škál

- 1) střední a jihovýchodní Evropa: Česká republika, Polsko a Slovinsko
- 2) střední Evropa: Německo a Rakousko
- 3) jižní Evropa: Španělsko a Portugalsko
- 4) severní Evropa: Norsko a Švédsko
- 5) západní Evropa: Spojené království a Irsko
- 6) západní Evropa: Belgie a Lucembursko
- 7) západní Evropa: Belgie a Francie

Je ale vhodné upozornit na to, že interpretace skalární ekvivalence v sobě skýtá jistá nebezpečí. Není samozřejmě nezbytné jakkoli omezovat sběr dat, mezinárodní výzkum lze realizovat v různě odlišných zemích po celém světě, analýzu dat je ovšem zapotřebí tomuto faktu přizpůsobit a porovnávat průměry postojových škál spíše jen ve značně homogenních skupinách zemí. Obecně nízká (popř. žádná) úroveň skalární ekvivalence ještě není v komparativním výzkumu absolutní bariérou, protože zmíněná analýza průměrů jak postojových škál, tak i jejich měřicích položek, je při analýze mezinárodních dat využívána často ve spojení s dalšími typy analýz anebo je možné průměry nahradit jinými složitějšími statistickými ukazateli. Mezinárodní data se zpravidla používají pro testování poměrně komplexních hypotéz a pro tyto cíle je samotná analýza průměrů značně elementární a bez potřebného kontextu. Na což se také snažili upozornit mnozí sociální vědci v 90. letech 20. století prostřednictvím kritiky namířené proti vágním srovnávacím tabulkám masově se objevujících v řadě sekundárních analýz mezinárodních dat.

Testy skalární ekvivalence položek zkoumaných postojových škál také podle předpokladů ukázaly, že položky měřící latentní proměnné sociálního kapitálu jsou méně problematické z hlediska komparace než položky měřící politické odcizení a důvěru v instituce, což se zcela shoduje s výsledky jiných výzkumů a studií. Strukturní modelování jako nejproblematičtější odhalilo v postojové škále sociálního kapitálu položky HELPFUL, VOLUNORG a LEGAL, a v postojové škále politického odcizení položky OPINION a VOTE.

Rámeček 6 Znění položek nejvíce zatížených systematickou chybou měření

HELPFUL: „Řekl(a) byste, že se lidé většinou snaží pomáhat druhým, nebo že se většinou starají o sebe?“

VOLUNORG: „Jak důležité jsou ve Vašem životě následující věci?: dobrovolné organizace“

LEGAL: „Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: právní systém“

OPINION: „Jak je pro Vás těžké či lehké udělat si názor na politické otázky?“

VOTE: „Řekl(a) byste, že politici se zajímají jen o získávání hlasů voličů spíše než o názory lidí?“

Tyto položky jsou pravděpodobně zkresleny metodologickou systematickou chybou měření. V této studii lze komplikace s mezinárodní komparací připsat kromě stylu odpovědí, sociální desirability a referenčního efektu především problematickému překladu vágních kvantifikátorů množství a intenzity jako je například slovní vyjádření „ani těžké, ani lehké“ (OPINION) a „někteří, většina“ (VOTE), které pravděpodobně výrazně ztěžují jejich kulturní srovnatelnost. U položky HELPFUL se může jednat o problém s překladem anglické fráze „try to be helpful“. V případě nalezení takovýchto problematických položek lze postupovat

různým způsobem od vyloučení dané položky z postojové škály až po úplné zamítnutí jakékoli komparace škál mezi zeměmi [viz Poortinga 1989: 749-751]. Ale van de Vijver [2003a: 145] upozorňuje na to, že dosahování validity v šetření a odstraňování položek zatížených systematickou chybou měření nesmí být jen slepým empiricismem vědců. Pokusy o získání maximální ekvivalence dat mohou někdy selhat a občas skutečnost, že nemůžeme data srovnávat, ukazuje na přítomnost zajímavých kulturních rozdílů, jimž by měla být věnována stejná pozornost jako budování ekvivalence samotné.

Ze všech výše uvedených důvodů je ale nicméně stejně nezbytné otestovat a následně zhodnotit ekvivalenci jakékoli postojové škály, kterou máme v úmyslu srovnávat mezi zeměmi. Nelze se z hlediska srovnatelnosti spoléhat jen na kvalitně odvedenou práci výzkumného týmu, který daný výzkum systematicky plánuje a připravuje. To, že jsou postojové škály pečlivě formulovány, testovány v pilotážích kvůli validitě a reliabilitě a jejich měřicí položky odborně překládány do různých jazyků a testovány v průběhu kognitivních rozhovorů, ještě zdaleka neznamená, že budou škály plně srovnatelné mezi zeměmi i v sebraných datech. Tuto skutečnost je nezbytně opakovaně ověřovat s každou novou vlnou daného šetření. Je to jedna z nevýhod tohoto aspektu kvality výzkumu, neboť nelze definitivně stanovit – jako je tomu v případě určení validity a reliability měření v pilotážích a experimentálních šetřeních – že dané postojové škály měřené položkami jsou srovnatelné. Úroveň srovnatelnosti škál je značně závislá na konkrétních datech, se kterými se pracuje.

V této práci použita technika strukturního modelování, příp. její další modifikace – tj. hierarchické strukturní modelování, je nejenom nejčastěji využívanou metodou při posuzování úrovně srovnatelnosti již vyvinutých, robustních postojových škál, které se testují na velkých vzorcích populace a znovu se používají v dalších vlnách výzkumu. Tato metoda je navíc vhodná k testování nových, ještě nezavedených škál, neboť dokáže pracovat i s poměrně malými vzorky dat z pilotáží a předvýzkumů (cca 100 respondentů) a identifikovat jejich problematické části a místo výskytu (tzn. skupinu/zemi) případných potíží. Další oblast jejího budoucího metodologického využití by mohl být i nakonec jakýkoli národní výzkum, který zkoumá společenské postoje v různých sociálních skupinách a má za cíl je v těchto skupinách také vzájemně porovnávat. I tehdy je vhodné ověřit, zda tyto sociální skupiny interpretují teoretický koncept i měřicí položky v dotazníku podobným způsobem.

LITERATURA

- Abramson, P. R. 1983. *Political Attitudes in America. Formation and Change*. New York: W. H. Freeman & Company.
- Albright, J. J., H. M. Park. 2009. „Confirmatory Factor Analysis Using Amos, LISREL, Mplus, and SAS/STAT CALIS.“ *Working Paper. The University Information Technology Services (UITS)* [online]. *Center for Statistical and Mathematical Computing, Indiana University* [cit. 14. 1. 2012]. Dostupné z: <<http://rt.uits.iu.edu/visualization/analytics/docs/cfa-docs/cfa.pdf> >.
- Almond, G. A., S. Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton, NJ: Princeton University Press.
- Anýžová, P. 2013. „Ekvivalence položek v mezinárodních datech: základní vymezení a možnosti analýzy.“ *Data a výzkum – SDA Info* 7 (1): 29–56.
- Arbuckle, J. L. 2012. *IBM SPSS® Amos™ 21 User's Guide*. Amos Development Corporation. Crawfordville, FL.
- Arbuckle, J. L. 1996. „Full information estimation in the presence of incomplete data.“ Pp. 243-277 in G. A. Marcoulides, R. E. Schumacker (Eds.). *Advanced structural equation modeling: Issues and techniques*. Mahwah, NJ: Erlbaum.
- Arts, W., L. Halman. 1999. „New Directions in Quantitative Comparative Sociology – An Introduction.“ *International Journal of Comparative Sociology* 40 (1): 1–12.
- Ægisdóttir, S., L. H. Gerstein, D. C. Çinarbas. 2008. „Methodological Issues in Cross-Cultural Counseling Research: Equivalence, Bias, and Translations.“ *The Counseling Psychologist* 36 (2): 188–219.
- Balch, G. 1974. „Multiple Indicators in Survey Research: The Concept „Sense of Political Efficacy“.“ *Political Metodology* 1 (1): 1–43.
- Bentler, P. M. 2005. *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software.
- Bentler, P. M., C-P Chou. 1987. „Practical issues in structural modeling.“ *Sociological Methods & Research* 16 (1): 78–117.
- Berry, J. W., Y. H. Poortinga, J. Pandey (Eds.). 1997. *Handbook of Cross-Culture Psychology: Theory and Methods*. Boston: Allyn and Bacon.
- Bielby, W. T., R. M. Hauser. 1977. „Structural equation models.“ *Annual Review of Sociology* 3 (1): 137-161.
- Billiet, J. 2003. „Cross-cultural equivalence with structural equation modeling.“ Pp. 247-263 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross cultural survey methods*. New Jersey: John Wiley & Sons.
- Billiet, J., J. Welkenhuysen-Gybels. 2004. „Assessing Cross-national Construct Equivalence in the ESS: The Case of Six Items.“ *Conference Paper. International conference on social science methodology, edition 6* [online]. Amsterdam, August 17/20 2004 [cit. 02. 11. 2010]. Dostupné z: <<https://lirias.kuleuven.be/bitstream/123456789/84246/1/DA+2004-40.pdf>>.
- Bhalla, G., L. Y. S. Lin. 1987. „Cross-Cultural Marketing Research: A Discussion of Equivalence Issues and Measurement Strategies.“ *Psychology & Marketing* 4 (4): 275-285.
- Blasius, J., V. Thiessen. 2000. „Methodological Artifacts in Measures of Political Efficacy and Trust: A Multiple Correspondence Analysis.“ *Political Analysis* 9 (1): 1-20.
- Blasius, J., V. Thiessen. 2006. „Assessing Data Quality and Construct Comparability in Cross-national Surveys.“ *European Sociological Review* 22 (3): 229–242.

- Blair, J., L. Piccinino. 2005. „The Development and Testing of Instruments for Cross-Culture and Multi-Cultural Surveys.“ Pp. 13-30 in J. H. P. Hoffmeyer-Zlotnic, J. A. Harkness (Eds.). *Methodological Aspects in Cross- National Research*. Mannheim: Zuma Nachrichten 11.
- Borg, I. 1998. „A Facet-Theoretical Approach to Item Equivalency.“ Pp. 145-158 in J. A. Harkness (Ed.). *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA-Nachrichten Spezial 3.
- Bourdieu, P. 1986. „The Forms of Capital.“ Pp. 241–258 in J. G. Richardson (Ed.). *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood.
- Braun, M., J. A. Harkness. 2005. „Text and Context: Challenges to Comparability in Survey Questions.“ Pp. 95–109 in J. H. P. Hoffmeyer-Zlotnic, J. A. Harkness (Eds.). *Methodological Aspects in Cross- National Research*. Mannheim: Zuma Nachrichten 11.
- Braun, M., T. P. Johnson. 2010. „An Illustrative Review of Techniques for Detecting Inequivalences.“ Pp. 375–393 in J. A. Harkness et al. (Eds.). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Braun, M., J. Scott. 1998. „Multidimensional Scaling and Equivalence: Is having a Job the Same as Working?“ Pp. 129-144 in J. A. Harkness (Ed.). *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA-Nachrichten Spezial 3.
- Brennan, R. L. 2001. *Generalizability theory*. New York: Springer-Verlag.
- Brislin, R. W. 1976. „Comparative Research Methodology: Cross-cultural Studies.“ *International Journal of Psychology* 11 (3): 215–229.
- Byrne, B. M. 1988. „The Self Description Questionnaire III: Testing for equivalent factorial validity across ability.“ *Educational and Psychological Measurement* 48 (2): 397–406.
- Byrne, B. M. 1998. *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.
- Byrne, B. M. 2004. „Testing for multigroup invariance using AMOS Graphics: A road less traveled.“ *Structural Equation Modeling: A Multidisciplinary Journal* 11 (2): 272–300.
- Byrne, B. M. 2006. *Structural equation modeling with EQS: Basic concepts, applications, and programming. Second edition*. Mahwah, NJ: Erlbaum.
- Byrne, B. M. 2008. „Testing for multigroup equivalence of a measuring instrument: A walk through the process.“ *Psicothema* 20 (4): 872-882.
- Byrne, B. M. 2010. *Structural equation modeling with AMOS: Basic concepts, applications, and programming. Second edition*. New York, NY: Taylor & Francis Group.
- Byrne, B. M. 2012. *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Taylor and Francis Group.
- Byrne, B. M., R. J. Shavelson, B. Muthén. 1989. „Testing for the Equivalence of Factor Covariance and Mean Structures: The Issue of Partial Measurement Invariance.“ *Psychological Bulletin* 105 (3): 456-466.
- Byrne, B. M., D. Watkins. 2003. „The Issue Of Measurement Invariance Revisited.“ *Journal of Cross-Cultural Psychology* 34 (2): 155–175.
- Campbell, A., G. Gurin, W. E. Miller. 1954. *The Voter Decides*. Evanston, IL: Row, Peterson.
- Campbell, D. T., D. W. Fiske. 1959. „Convergent and Discriminant Validation by the Multimethod-Multitrait Matrix.“ *Psychological Bulletin* 56 (2): 833–853.
- Cieciuch, J., E. Davidov. 2012. „A comparison of the invariance properties of the PVQ-40 and the PVQ-21 to measure human values across German and Polish Samples.“ *Survey Research Methods* 6 (1): 37-48.
- Citrin, J., H. McClosky, M. J. Shanks, P. M. Sniderman. 1975. „Personal and Political Sources of Political Alientation.“ *British Journal of Political Science* 5 (1): 1-31.
- Coleman, J. 1988. „Social Capital and the Creation of Human Capital.“ *American Journal of Sociology* 94 (supplement): 94–120.

- Craig, S. C., R. G. Niemi, G. E. Silver. 1990. „Political Efficacy and Trust: A report on the NES Pilot Study Items.“ *Political Behavior* 12 (3): 289–314.
- Davidov, E. 2010. „Testing for comparability of human values across countries and time with the third round of European Social Survey.“ *International Journal of Comparative Sociology* 51 (3): 171–191.
- Davidov, E. 2012. „Comparability of measurement instruments across countries or time points.“ Presentation [Online]. University of Tartu, 4.10.2012 [cit. 22.1.2014]. Dostupné z:
<http://www.ssi.ut.ee/sites/default/files/ssi/davidov_lecture2.pdf>.
- Davidov, E., A. De Beuckelaer. 2010. „How Harmful are Survey Translations? A Test with Schwartz’s Human Values Instrument.“ *International Journal of Public Opinion Research* 22 (4): 485–510.
- Davidov, E., H. Dülmer, E. Schlüter, P. Schmidt, B. Meuleman. 2012. „Using a multilevel structural equation modeling approach to explain cross-cultural measurement noninvariance.“ *Journal of Cross-Cultural Psychology* 43 (4): 558–575.
- Davidov, E., P. Schmidt. 2007. „Values and attitudes towards immigrants: A comparison between Germany and Israel.“ Prezentace, Workshop: *Immigration, public opinion and governmental policies*, Pompeu Fabra University, Barcelona, March 2007.
- Davidson, A. R., J. Jaccard, H. C. Triandis, M. L. Morales, R. Diaz-Guerrero. 1976. „Cross-cultural model testing: toward a solution of the etic-emic dilemma.“ *International Journal of Psychology* 11 (1): 1–13.
- Devins, G. M., M. Beiser, R. Dion, L. G. Pelletier, R. Gary Edwards. 1997. „Cross-Cultural Measurements of Psychological Well-Being: The Psychometric Equivalence of Cantonese, Vietnamese, and Laotian Translations of the Affect Balance Scale.“ *American Journal of Public Health* 87 (5): 794–799.
- Diehl, S., R. Terlutter (Eds.). 2006. *International advertising and communication. Current insights and empirical findings*. Wiesbaden: GWV.
- DiStefano, C. 2002. „The Impact of Categorization With Confirmatory Factor Analysis.“ *Structural Equation Modeling: A Multidisciplinary Journal* 9 (3): 327–346.
- Donahue, B. H. 2006. *The effect of partial measurement invariance on prediction*. (Unpublished PhD Dissertation). Athens: University of Georgia.
- Dorer, B. 2012. *Round 6 Translation Guidelines*. Mannheim: European Social Survey, GESIS.
- Eurobarometer. 2005. *Social Capital* [Online]. Special Eurobarometer 223. Wave 62.2 – TNS Opinion & Social [cit. 22. 1. 2014]. Dostupné z:
<http://ec.europa.eu/public_opinion/archives/ebs/ebs_223_en.pdf>.
- European Social Survey. 2010. *ESS Round 5 Translation Guidelines*. Mannheim, European Social Survey GESIS.
- European Social Survey. 2012. *Sampling for the European Social Survey Round VI: Principles and Requirements*. Mannheim: European Social Survey, GESIS.
- Ferragina, E. (Ed.). 2012. *Social Capital in Europe: A Comparative Regional Analysis*. Cheltenham: Edward Elgar Pub.
- Flora, D. B., P. J. Curran. 2004. „An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis With Ordinal Data.“ *Psychological Methods* 9 (4): 466–491.
- Fontaine, J. R. J. 2003. „Multidimensional scaling.“ Pp. 235–246 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Gao, S., P. L. Mokhtarian, R. A. Johnston. 2008. „Nonnormality of Data in Structural Equation Models.“ *Journal of the Transportation Research Board* 2082 (1): 116–124.

- Goerman, P. L., R. A. Caspar. 2010. „Managing the Cognitive Pretesting of Multilingual Survey Instruments: A Case Study of Pretesting of the U.S. Census Bureau Bilingual Spanish/English Questionnaire.“ Pp. 75–90 in J. A. Harkness et al. (Eds.). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Groves, R. M. 2004. *Survey Errors and Survey Costs*. New York: Wiley.
- Halman, L., R. Luijkx. 2006. „Social capital in contemporary Europe: evidence from the European Social Survey.“ *Portuguese Journal of Social Science* 5 (1): 65-90.
- Halperin, E., R. Sulitzeanu-Kenan. 2010. „Making a Difference: Political Efficacy and Policy Preferences Polarization.“ [online]. *Working Paper for Social Science Research Network (SSRN)* [cit. 20.1.2013]. Dostupné z: <<http://ssrn.com/abstract=1719564>> nebo <<http://dx.doi.org/10.2139/ssrn.1719564>>.
- Harris, M. 1976. „History and significance of the emic/etic distinction.“ *Annual Review of Anthropology* 5 (1): 329–350.
- Harkness, J. A. 2003. „Questionnaire translation.“ Pp. 35–56 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Harkness, J. A. (Ed.). 1998. *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA-Nachrichten Spezial 3.
- Harkness, J. A., M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. Ph Mohler, B-E Pennell, T. W. Smith (Eds.). 2010. *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Harkness, J. A., P. Ph. Mohler, F. J. R. Van de Vijver. 2003. „Comparative Research.“ Pp. 3–16 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Harkness J. A., A. Schoua-Glusberg. 1998. „Questionnaires in Translation.“ Pp. 87-126 in J. A. Harkness (Ed.). *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA-Nachrichten Spezial 3.
- Harkness, J. A., A. Viilar, B. Edwards. 2010. „Translation, Adaptation, and Design.“ Pp. 117-140 in J. A. Harkness et al. (Eds.). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Harkness, J. A., F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). 2003. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Hayes, B. C., C. S. Bean. 1993. „Political Efficacy: A Comparative Study of the United States, West Germany, Great Britain and Australia.“ *European Journal of Political Research* 23 (2): 261–280.
- Hendl, J. 2006. *Přehled statistických metod zpracování dat. Analýza a metaanalýza dat*. Praha: Portál.
- Herkenrath, M. 2002. „Quantitative Cross-National Analysis as a Research Tool in the Sociology of Developing Countries: A Critical Examination.“ *Current Sociology* [online] 50 (4): 517–530 [cit. 24. 10. 2009]. Dostupné z: <<http://csi.sagepub.com/cgi/content/abstract/50/4/517>>.
- Hoffmeyer-Zlotnic, J. H. P., J. A. Harkness (Eds.). 2005. *Methodological Aspects in Cross-National Research*. Mannheim: Zuma Nachrichten 11.
- Horn, J. L., J. J. McArdle. 1992. „A practical and theoretical guide to measurement invariance in aging research.“ *Experimental Aging Research* 18 (3): 117-144.
- Hox, J. J., T. M. Bechger. 1998. „An Introduction to Structural Equation Modeling.“ *Family Science Review* 11: 354–373.
- Hox, J. J., E. D. de Leeuw, M. J. S. Brinkhuis. 2010. „Analysis Models for Comparative Surveys.“ Pp. 395–418 in J. A. Harkness et al. (Eds.). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.

- Hsueh, K-H, L. R. Phillips, W.Y. Cheng, S. J. Fulton Picot. 2005. „Assessing Cross-Cultural Equivalence Through Confirmatory Factor Analysis.“ *Western Journal of Nursing Research* 27 (6): 755–771.
- Hui, C. H., H. C. Triandis. 1985. „Measurement in Cross-cultural psychology: A review and comparison of strategies.“ *Journal of Cross-Cultural Psychology* 16 (2): 131-152.
- Chan D. 2000. „Detection of differential item functioning on the Kirton Adaptation-Innovation Inventory using multiple-group mean and covariance structure analyses.“ *Multivariate Behavioral Research* 35 (2): 169–199.
- Cheung, G. W., R. B. Rensvold. 1999. „Testing factorial invariance across groups: A reconceptualization and proposed new method.“ *Journal of Management* 25 (1): 1-27.
- Cheung, G. W., R. B. Rensvold. 2002. „Evaluating goodness-of-fit indexes for testing measurement invariance.“ *Structural Equation Modeling: A Multidisciplinary Journal* 9 (2): 233–255.
- Christiansen, B., H. Hyman, R. Rommetveit. 1951. *Cross-National Social Research*. Oslo: Institut for Social Research.
- Jelínek, M., P. Květoň, D. Vobořil. 2011. *Testování v psychologii: Teorie odpovědi na položku a počítačové adaptivní chování*. Praha: Grada Publishing.
- Jeřábek, H., P. Soukup (Eds.) 2008. *Advanced Lazarsfeldian methodology*. Praha: Karolinum.
- Johnson, T. P. 1998. „Approaches to Equivalence in Cross-Cultural and Cross-National Survey Research.“ Pp. 1–40 in J. A. Harkness (Ed.). *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA-Nachrichten Spezial 3.
- Johnson, T. P., F. J. R. Van de Vijver. 2003. „Social desirability in cross-cultural research.“ Pp. 195 – 204 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler. *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Jöreskog, K. G. 1971. „Simultaneous Factor Analysis in Several Populations.“ *Psychometrika* 36 (4): 409–426.
- Jöreskog, K. G. 1990. „New developments in LISREL: Analysis of ordinal variables using polychoric correlations and weighted least squares.“ *Quality and Quantity* 24 (4): 387–404.
- Jöreskog, K. G. 2002. *Structural Equation Modeling with Ordinal Variables using LISREL*. [online]. [cit. 14.1.2012]. Dostupné z: <www.ssicentral.com/lisrel/techdocs/ordinal.pdf>.
- Jöreskog, K. G., D. Sörbom. 1993. *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software International.
- Jowell, R.. 1998. „How comparative is comparative research?“ Centre For Research Into Elections And Social Trends *CREST. Paper No 66* [online]. *Maison Francaise*, Oxford: June 26/27 1997 [cit. 27. 10. 2009]. Dostupné z: <<http://www.crest.ox.ac.uk/papers.htm>>.
- Jowell, R., R. C. Fitzgerald, G. Eva (Eds.). 2007. *Measuring Attitudes Cross-Nationally*. London: Sage Publications.
- Kankaras, M., G. Moors. 2009. „Measurement Equivalence in Solidarity Attitudes in Europe: Insights from a Multi-Group Latent-Class Factor Approach.“ *International Sociology* 24 (4): 557–579.
- Karp, J. A., S. A. Banducci. 2008. „Political Efficacy and Participation in Twenty-Seven Democracies: How Electoral Systems Shape Political Behaviour.“ *British Journal of Political Science* 38 (2): 311–334.
- Kline, R. B. 1998. *Principles and practice of structural equation modeling*. New York: Guilford.
- Kohn, M. L. (Ed.). 1989. *Cross-National Research in Sociology*. Newbury Park, CA: Sage.
- Kostecký, T., D. Čermák. 2003. „Výběrová šetření a analýza agregátních dat – diskuze na téma použitelnosti různých přístupů v komparativních analýzách politického chování.“ *Sociologický časopis* 39 (4): 529–550.

- Kreidl, M. 2004. „Metody měření reliability a validity.“ [online] *Socioweb* 2004/2: 6-8. [cit. 14.12.2012]. Dostupné z: <www.socioweb.cz/index.php?disp=teorie&shw=153&lst=106>.
- Krejčí, J.. 2006. „Mezinárodní sociálněvědní komparativní výzkum a Česká republika: Přehled výzkumů a dostupných dat.“ *Sociologický časopis* 42 (1): 149–173.
- Krejčí, J. 2008. *Kvalita sociálněvědních výběrových šetření v České republice*. Praha: Sociologické nakladatelství.
- Lewis, M. P., G. F. Simons, Ch. D. Fennig (Eds.). 2013. *Ethnologue: Languages of the World, Seventeenth edition*. [online]. Dallas, Texas: SIL International. [cit. 30. 8. 2013]. Dostupné z: <<http://www.ethnologue.com>>.
- Lievens, F., F. Anseel, M. H. Harris, J. Eisenberg. 2007. „Measurement Invariance of the Pay Satisfaction Questionnaire Across Three Countries.“ *Educational and Psychological Measurement* 67 (6): 1042–1051.
- Lievesley, D. 2001. „Making a Difference: A Role for the Responsible International Statistician?“ *Journal of the Royal Statistical Society* 50 (4): 367–406.
- Linek, L. 2010. *Zrození snu? Struktura a dynamika postojů k politickému režimu a jeho institucím a jejich důsledky*. Praha: SLON.
- Linek, L. 2011. „Institucionální odcizení a individuální odcizení od politiky.“ *Data a výzkum - SDA Info* 5 (1): 7-34.
- Little, R. J. A., D. B. Rubin. 1987. *Statistical analysis with missing data*. New York: Wiley.
- Lonner, W. J. 1985. „Issues in testing assessment in cross-cultural counseling.“ *The counseling Psychologist* 13: 599–614.
- Loury, G. C. 1977. „A Dynamic Theory of Racial Income Differences.“ Pp. 153-186 in P. A. Wallace, A. M. LaMond (Eds.). *Women, Minorities, and Employment Discrimination*. Lexington, MA: Heath.
- Madden, Ch. 2004. „Making Cross-country Comparisons of Culture Statistics: Problems and Solutions (International Comparisons of Culture Sectors: An Exploratory Investigation).“ [online] *IFACCA, Australian Council for the Arts*. Sydney, 2004 [cit. 04. 10. 2009]. Dostupné z: <www.australiacouncil.gov.au/>.
- Madsen, D. 1978. „A Structural Approach to the Explanation of Political Efficacy Levels Under Democratic Regimes.“ *American Journal of Political Science* 22 (4): 867-883.
- Marcoulides, G. A., R. E. Schumacker (Eds.). *Advanced structural equation modeling: Issues and techniques*. Mahwah, NJ: Erlbaum.
- Marsh, R. M. 1967. *Comparative Sociology. A codification of Cross-Societal Analysis*. New York: Harcourt, Brace & World, Inc.
- Matějů, P.. 1989. „Metoda strukturního modelování: Přehled základních problémů.“ *Sociologický časopis* 25 (4): 399–418.
- Matějů, P. 1993. „Determinanty ekonomického úspěchu v první fázi postkomunistické transformace. Česká republika 1989–1992.“ *Sociologický časopis* 29 (3): 341–366.
- Matějů, P. 2002. „Social Capital. Problems of Its Conceptualization and Measurement in Transforming Societies.“ [online]. *Conference Paper 6. Social Capital: The Challenge of International Measurement*. An international conference convened by the Organisation for Economic Co-operation and Development and the United Kingdom Office for National Statistics. London, UK, 25–27 September 2002 [cit. 22. 1. 2009]. Dostupné z: <http://www.stratif.cz/attachments/doc97/Soc_cap_mateju_london.pdf>.
- Matějů, P., A. Vitásková. 2006. „Trust and Mutually Beneficial Exchanges: Measuring Social Capital for Comparative Analyses.“ *Sociologický časopis/Czech Sociological Review* 42 (3): 493–516.
- May, H. 2006. „A Multilevel Bayesian Item Response Theory Method for Scaling Socioeconomic Status in International Studies of Education.“ *Journal of Educational and Behavioral Statistics* 31 (1): 63–79.

- Meade, A. W., E. C. Johnson, Ph. W. Braddy. 2006. „The Utility of Alternative Fit indices in Tests of Measurement Invariance.“ [online]. *Paper presented at the Antal Academy of Management Conference*. Atlanta, GA [cit. 20. 2. 2013]. Dostupné z: <<http://www4.ncsu.edu/~awmeade/Links/Papers/AFIsMI%28AoM06%29.pdf>>
- Meade, A. W., G. J. Lautenschlager. 2004. „Comparison of Item Response Theory and Confirmatory Factor Analytic Methodologies for Establishing Measurement Equivalence/Invariance.“ *Organizational Research Methods* 7 (4): 361–388.
- Meredith, W. 1993. „Measurement Invariance, Factor Analysis and Factorial Invariance.“ *Psychometrika* 58 (4): 525–543.
- Meuleman, B., E. Davidov, J. Billiet. 2009. „Changing attitudes toward immigration in Europe, 2002–2007: A dynamic group conflict theory approach.“ *Social Science Research* 38 (2): 352–365.
- Meulemann, H. (Ed.). 2008. *Social capital in Europe: similarity of countries and diversity of people? Multi-level analyses of the European Social Survey 2002*. Leiden/Boston: Brill.
- Milfont, T. L., R. Fisher. 2010. „Testing measurement invariance across group: Applications in cross-cultural research.“ *International Journal of Psychological Research* 3 (1): 111–121.
- Mîndrilă, D. 2010. „Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: a comparison of estimation bias with ordinal and multivariate non-normal data.“ *International Journal of Digital Society* 1 (1): 60–66.
- Mohler, P. Ph., T. P. Johnson. 2010. „Equivalence, Comparability, and Methodological Progress.“ Pp. 17–29 in J. A. Harkness et al. (Eds). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Montero, J. R., R. Gunther, M. Torcal. 1997. „Democracy in Spain: Legitimacy, Discontent, and Disaffection.“ *Studies in Comparative International Development* 32 (3): 124–160.
- Možný, I. 1991. *Proč tak snadno ... Někteří důvody sametové revoluce*. Praha: Sociologické nakladatelství.
- Mullen, M. R. 1995. „Diagnostic Measurement Equivalence in Cross-National Research.“ *Journal of International Business Studies* 26 (3): 573–596.
- Mushquash, Ch., B. P. O'Connor. 2006. „SPSS and SAS programs for generalizability theory analyses.“ *Behavior Research Methods* 38 (3): 542–547.
- Muthén, B., A. Christoffersson. 1981. „Simultaneous factor analysis of dichotomous variables in several groups.“ *Psychometrika* 46 (4): 407–419.
- Muthén, B., D. Kaplan. 1985. „A comparison of some methodologies for the factor analysis of non-normal Likert variables.“ *British Journal of Mathematical and Statistical Psychology* 38 (2): 171–189.
- Olsson, U. 1979. „Maximum likelihood estimation of the polychord correlation coefficient.“ *Psychometrika* 44: 443–460.
- Øyen, E. (Ed.). 1990. *Comparative Methodology: Theory and Practice in International Social Research*. London: Sage.
- Pan, Y., A. Landreth, H. Park, M. Hinsdale-Shouse, A. Schoua-Glusberg. 2010. „Cognitive Interviewing in Non-English languages: A Cross-Cultural Perspective.“ Pp. 91–113 in J. A. Harkness et al. (Eds). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Parts, E. 2008. „Indicators of social capital in the European Union.“ Working Paper [online]. IAREG 2008/02. University of Tartu. [cit 22. 1. 2014]. Dostupné z: <http://www.iareg.org/fileadmin/iareg/media/papers/WP2_02.pdf>.
- Payne, G. 1973. „Comparative sociology: some problems of theory and method.“ *British Journal of Sociology* 24 (1): 13–30.
- Paxton, P. 2002. „Social Capital and Democracy: An Interdependent Relationship.“ *American Sociological Review* 67 (2): 254–277.

- Plečtitá-Vlachová, K., F. Kalvas. 2005. „The European Social Survey (ESS): o projektu a dostupnosti dat.“ *SDA Info* VII (2): 14–16.
- Poortinga, Y. H. 1989. „Equivalence of cross-cultural data: An overview of basic issues.“ *International Journal of Psychology* 24 (6): 737-756.
- Przeworski, A., H. Teune. 1970. *The Logic of Comparative Social Inquiry*. New York: John Wiley & Sons.
- Putnam, R. 1995. „Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America.“ *PS: Political Science and Politics* 28 (4): 664–683.
- Putnam, R. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon and Schuster.
- Putnam, R. D., R. Leonardi, R. Y. Nanetti. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.
- Raju, N. S., L. J. Laffitte, B. M. Byrne. 2002. „Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory.“ *Journal of Applied Psychology* 87 (3): 517–529.
- Reeskens, T., M. Hooghe. 2008. „Cross-cultural measurement equivalence of generalized trust. Evidence from the European Social Survey (2002 and 2004).“ *Social Indicators Research* 85 (3): 515-532.
- Revilla, M., W. Saris. 2011. „How to evaluate the cross-cultural equivalence of single items.“ (RECSM, UPF) Mini conference: *Measurement invariance: Methods, problems, and further directions*. Zurich 15/16. 7. 2011.
- Richardson, J. G. (Ed.). 1986. *Handbook of Theory and Research for the Sociology of Education*. New York: Greenwood.
- Robert, Ch., W. C. Lee, K-Y Chan. 2006. „An empirical analysis of measurement equivalence with the INDCOL measure of individualism and collectivism: implications for valid cross-cultural inference.“ *Personnel Psychology* 59: 65-99.
- Rudá, E. 2012. „Teorie odpovědi na položku a její aplikace v sociologii.“ Nепublikovaná diplomová práce. Praha: FF UK.
- Řehák, J. 1998a. „Kvalita dat I: Klasický model měření reliability a jeho praktický aplikační výzkum“ *Sociologický časopis* 34: 51-60.
- Řehák, J. 1998b. „Kvalita dat II: Přístupy ohodnocování výzkumných instrumentů založené na modelování kovariančních struktur.“ *Sociologický časopis* 34 (2): 195–204.
- Saris, W. E. 2003. „Multitrait-Multimethod studies.“ Pp. 265–274 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Saris, W. E., I. Gallhofer. 2007. „Can questions travel successfully?“ Pp. 53–79 in R. Jowell et al. (Eds.). *Measuring Attitudes Cross-Nationally*. London: Sage Publications.
- Saris, W. E., I. Gallhofer. (n.d.). Report on the MTMM experiments in the pilot studies and proposals for round 1 of the ESS. European Social Survey.
- Saris, W. E., A. Satorra, W. van der Veld. 2009. „Testing Structural Equation Models or Detection of Misspecifications?“ *Structural Equation Modeling: A Multidisciplinary Journal* 16 (4): 561-582.
- Seligson, M. A. 1980. „A Problem-Solving Approach to Measuring Political Efficacy.“ *Social Science Quarterly* 60 (4): 630–642.
- Scheuch, E. K. 1993. „The Cross-culture Use of Sample Surveys: Problems of Comparability.“ *Historical Social Research* 18 (2): 104–138.
- Schmitt, N., G. Kuljanin. 2008. „Measurement invariance: Review of practise and implications.“ *Human Resource Management Review* 18 (4): 210-222.
- Schubert, J. 2010. „Klasická testová teorie reliability v metodologii výběrových šetření.“ *Data a výzkum - SDA Info* 4 (2): 105-122.

- Schwarz, N., D. Oyserman, E. Peytcheva. 2010. „Cognition, Communication, and Culture: Implications for the Survey Response Process.“ Pp. 177–190 in J. A. Harkness et al. (Eds.). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.
- Singh, J. 1995. „Measurement issues in cross-national research.“ *Journal of International Business Studies* 26 (3): 597–619.
- Smith, T. W. 2003. „Developing Comparable Questions in Cross-national Survey.“ Pp. 69–91 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross cultural survey methods*. New Jersey: John Wiley & Sons.
- Smith, P. B., R. Fischer. 2008. „Acquiescence, Extreme Response Bias and Culture: A Multilevel Analysis.“ Pp. 285–314 in J. R. Fons Van de Vijver, D. van Hemert, Y. H. Poortinga (Eds.). *Multilevel analysis of individuals and cultures*. New York: Taylor & Francis Group.
- Soukup, P. 2006a. „Proč užívat hierarchické lineární modely?“ *Sociologický časopis/Czech Sociological Review* 42 (5): 987–1012.
- Soukup, P. 2006b. „Čím větší, tím lepší (aneb mýty o reliabilitě).“ [online]. *Socioweb* 2006/7: 11–12. [cit. 14. 12. 2012]. Dostupné z: <<http://www.socioweb.cz/index.php?disp=teorie&shw=242&lst=112>>.
- Stachová, J., J. Bernard, D. Čermák. 2009. *Sociální kapitál v České republice a v mezinárodním srovnání*. Praha: Sociologický ústav AV ČR.
- Steenkamp, J-B. E. M., H. Baumgartner. 1998. „Assessing Measurement Invariance in Cross-National Consumer Research.“ *Journal of Consumer Research* 25 (1): 78–90.
- Steinmetz, H. 2013. Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 9 (1): 1–12.
- Stout, W. 2001. „Nonparametric Item response Theory: A Maturing and Applicable Measurement Modeling Approach.“ *Applied Psychological Measurement* 25 (3): 300–306.
- Stout, W. 2002. „Psychometrics: From Practice to Theory and Back (15 years of Nonparametric Multidimensional IRT, DIF/Test Equity, and Skills Diagnostic Assessment).“ *Psychometrika* 67 (4): 485–518.
- Straus, M. A. 1969. „Phenomenal Identity and Conceptual Equivalence of Measurement in Cross-National Comparative Research.“ *Journal of Marriage and the Family* 31 (2): 233–241.
- Šafr, J., M. Sedláčková. 2006. *Sociální kapitál: Koncepty, teorie a metody měření*. Praha: Sociologický ústav AV ČR.
- Temme, D. 2006. „Assessing measurement invariance of ordinal indicators in cross-national research.“ Pp. 455–472 in S. Diehl, R. Terlutter (Eds.). *International advertising and communication. Current insights and empirical findings*. Wiesbaden: GWV.
- Thomassen, J. (n.d.). „Opinions about political issues.“ European Social Survey. ESS Core Questionnaire Development [online]. Chapter 5, Part II: People’s underlying value orientations [cit. 14. 1. 2013]. Dostupné z: <http://www.europeansocialsurvey.org/index.php?option=com_docman&task=cat_view&gid=83&Itemid=80>.
- Tourangeau R., L. J. Rips, K. A. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Traxlerová, M., L. Rabušic. 2008. „Jak měřit bezmocnost: předběžné výsledky.“ *Data a výzkum – SDA Info* 2 (1): 7–29.
- Uebersax, J. S. 2006a. „Factor analysis and SEM with tetrachoric and polychoric correlations.“ [online]. *Statistical Methods for Rater Agreement web site*. [cit. 14. 1. 2012]. Dostupné z: <<http://john-uebersax.com/stat/sem.htm>>.

- Uebersax, J. S. 2006b. „The tetrachoric and polychoric correlation coefficients.“ [online]. *Statistical Methods for Rater Agreement web site*. [cit. 14. 1. 2012]. Dostupné z: <<http://john-uebersax.com/stat/tetra.htm>>.
- Urbánek, T. 2000. *Strukturální modelování v psychologii*. Brno: Psychologický ústav AV ČR a Nakladatelství Pavel Křepela.
- Urbánek, T., M. Šimeček. 2002. „Teorie odpovědi na položku.“ *Československá psychologie: Časopis pro psychologickou teorii a praxi* [online] 45 (5): 428-440 [cit. 14. 12. 2012]. Dostupný z: <www.psychodiagnostika.cz/urbanek.pdf>.
- Vandenberg, R. J. 2002. „Toward a further understanding of and improvement in measurement invariance methods and procedures.“ *Organizational Research Methods* 5 (2): 139-158.
- Vandenberg, R. J., Ch. E. Lance. 2000. „A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research.“ *Organizational Research Methods* 3 (1): 4–70.
- Van Deth, J. W. 2003. „Using published survey data.“ Pp. 291–309 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Van de Vijver, F. J. R. 1998. „Towards a Theory of Bias and Equivalence.“ Pp. 41–65 in J. A. Harkness (Ed.). *Cross-Cultural Survey Equivalence*. Mannheim: ZUMA-Nachrichten Spezial 3.
- Van de Vijver, F. J. R. 2003a. „Bias and equivalence: Cross-cultural perspectives.“ Pp. 143-155 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Van de Vijver, F. J. R. 2003b. „Bias and substantive analysis.“ Pp. 207–233 in J. A. Harkness, F. J. R. Van de Vijver, P. Ph. Mohler (Eds.). *Cross-cultural survey methods*. New Jersey: John Wiley & Sons.
- Van de Vijver, F. J. R., D. A. van Hemert, Y. H. Poortinga (Eds.). 2008. *Multilevel analysis of individuals and cultures*. New York: Taylor & Francis Group.
- Van de Vijver, F. J. R., K. Leung. 1997. „Methods and Data Analysis of Comparative Research.“ Pp. 257–301 in J. W. Berry, Y. H. Poortinga, J. Pandey (Eds.). *Handbook of Cross-Culture Psychology: Theory and Methods*. Boston: Allyn and Bacon.
- Van de Vijver, F. J. R., Y. H. Poortinga. 2002. „Structural Equivalence in Multilevel Research.“ *Journal of Cross-Cultural Psychology* 33 (2): 141–156.
- Van der Veld, W., W. E. Saris. 2009. „Causes of Generalized Social Trust: An Innovative Cross-National Evaluation.“ [online]. RECSM Working Paper 8. [cit. 22. 1. 2014]. Dostupné z: <<http://www.upf.edu/survey/working/working.html>>.
- Van Hemert, D. A., F. J. R. van de Vijver, Y. H. Poortinga, J. Georgas. 2002. „Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries.“ *Personality and Individual Differences* 33 (8): 1229–1249.
- Vinopal, J. 2008. *Kognitivní přístupy v metodologii výzkumných šetření: metoda okamžité validizace*. Praha: Sociologický ústav AV ČR.
- Vinopal, J. 2009. „Situace standardizovaného dotazování z hlediska kognitivních přístupů.“ *Sociologický časopis/Czech Sociological Review* 45 (2): 397–420.
- Wallace, P. A., A. M. LaMond (Eds.). 1977. *Women, Minorities, and Employment Discrimination*. Lexington, MA: Heath.
- Wang, M., S. S. Russell. 2005. „Measurement Equivalence of the Job Descriptive Index Across Chinese and American Workers: Results from Confirmatory Factor Analysis and Item Response Theory.“ *Educational and Psychological Measurement* 65 (4): 709–732.

- Welkenhuysen-Gybels, J. G. J., F. J. R. van de Vijver. 2001. „A comparison of methods for the evaluation of construct equivalence in a multigroup setting.“ [online]. Unpublished. *Proceedings of the Annual Meeting of the American Statistical Association* [cit. 22.1.2014]. Dostupné z: <<http://www.amstat.org/sections/srms/Proceedings/y2001/Proceed/00106.pdf>>.
- Woehr, D. J., J. P. Meriac. 2010. „Using Polytomous Item Response Theory to Examine Differential Item and Test Functioning: The Case of Work Ethic.“ Pp. 419-433 in J. A. Harkness et al. (Eds). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. New Jersey: John Wiley & Sons.

Internetové odkazy:

www.csdiworkshop.org
www.issp.org
www.europeansocialsurvey.org
www.capstan.be/content/home.html
www.sqp.nl
www.europeansocialsurvey.org/data/
<http://essedunet.nsd.uib.no/cms/topics/immigration/2/.html>
www.ethnologue.com
www.migrationinformation.org
<http://www.amstat.org/sections/srms/Proceedings/>

PŘÍLOHY

Příloha 1: znění vybraných otázek

Tabulka 3 Plné znění testovaných položek, jejich kódy v dotazníku ESS (první vlna, 2002) a kategorie odpovědí v českém a anglickém jazyce.

Kód otázky	Znění otázky	Kategorie odpovědí
Individuální odcizení (individual disaffection)		
INTEREST (B1)	Jak moc se zajímáte o politiku? Zajímáte se o ni:	1 - velmi
		2 - dost
		3 - jen trochu
		4 - vůbec ne
	How interested would you say you are in politics?	1 - very interested
		2 - quite interested
		3 - hardly interested
		4 - not at all interested
DIFFICULT (B2)	Jak často se politika zdá být tak složitá, že vlastně nemůžete pochopit, co se děje?	1 - nikdy
		2 - zřídka
		3 - občas
		4 - pravidelně
		5 - často
	How often does politics seem so complicated that you can't really understand what is going on?	1 - never
		2 - seldom
		3 - occasionally
		4 - regularly
		5 - frequently
POLACTIV (B3) (recod ⁹⁶)	Myslíte, že byste mohl(a) hrát aktivní roli ve skupině, která se zabývá politickými otázkami?	1 – určitě ne
		2 – pravděpodobně ne
		3 – nejsem si tím jistý/á
		4 – pravděpodobně ano
		5 – určitě ano
	Do you think that you could take an active role in a group involved with political issues?	1 – definitely not
		2 – probably not
		3 – not sure either way
		4 – probably
		5 - definitely
OPINION (B4) (recod)	Jak je pro Vás těžké či lehké udělat si názor na politické otázky?	1 – velmi těžké
		2 – těžké
		3 – ani těžké, ani lehké
		4 – lehké
		5 – velmi lehké
	How difficult or easy do you find it to make your mind up about political issues?	1 – very difficult
		2 - difficult
		3 – neither difficult nor easy
		4 – easy
		5 – very easy

⁹⁶ Recod označuje, že původní pořadí kategorií odpovědí bylo pro potřeby analýzy transformováno opačně (resp. zrcadlově obráceně) ve všech datových souborech ESS, se kterými se pracovalo, tak aby si všechny kategorie odpovědí logicky odpovídaly.

Institucionální odcizení		
VOTE (B6) (recod)	Řekl(a) byste, že politici se zajímají jen o získávání hlasů voličů spíše než o názory lidí?	1 – téměř všichni politici se zajímají jen o hlasy voličů
		2 – většina politiků se zajímá jen o hlasy voličů
		3 – někteří politici se zajímají jen o hlasy voličů
		4 – většina politiků se zajímá o názory lidí
		5 – téměř všichni politici se zajímají o názory lidí
	Would you say that politicians are just interested in getting people's votes rather than in people's opinions?	1 – nearly all politicians are just interested in votes
		2 – most politicians are just interested in votes
		3 – some politicians are just interested in votes, others aren't
		4 – most politicians are interested in people's opinions
		5 – nearly all politicians are interested in people's opinions
CARE (B5) (recod)	Myslíte si, že se politici obecně starají o to, co si lidé jako Vy myslí?	1 – téměř žádní politici se nestarají o to, co si lidé jako já myslí
		2 – velmi málo politiků se stará
		3 – někteří politici se starají
		4 – mnoho politiků se stará
		5 – většina politiků se stará o to, co si lidé jako já myslí
	Do you think that politicians in general care what people like you think?	1 – hardly any politicians care what people like me think
		2 – very few care
		3 – some care
		4 – many care
		5 – most politicians care what people like me think
POLIT (B10) (recod)	Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: politici	0 znamená, že instituci vůbec nedůvěřujete, 10 znamená, že jí naprosto důvěřujete.
	Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: politicians	0 means you do not trust an institution at all, and 10 means you have complete trust.

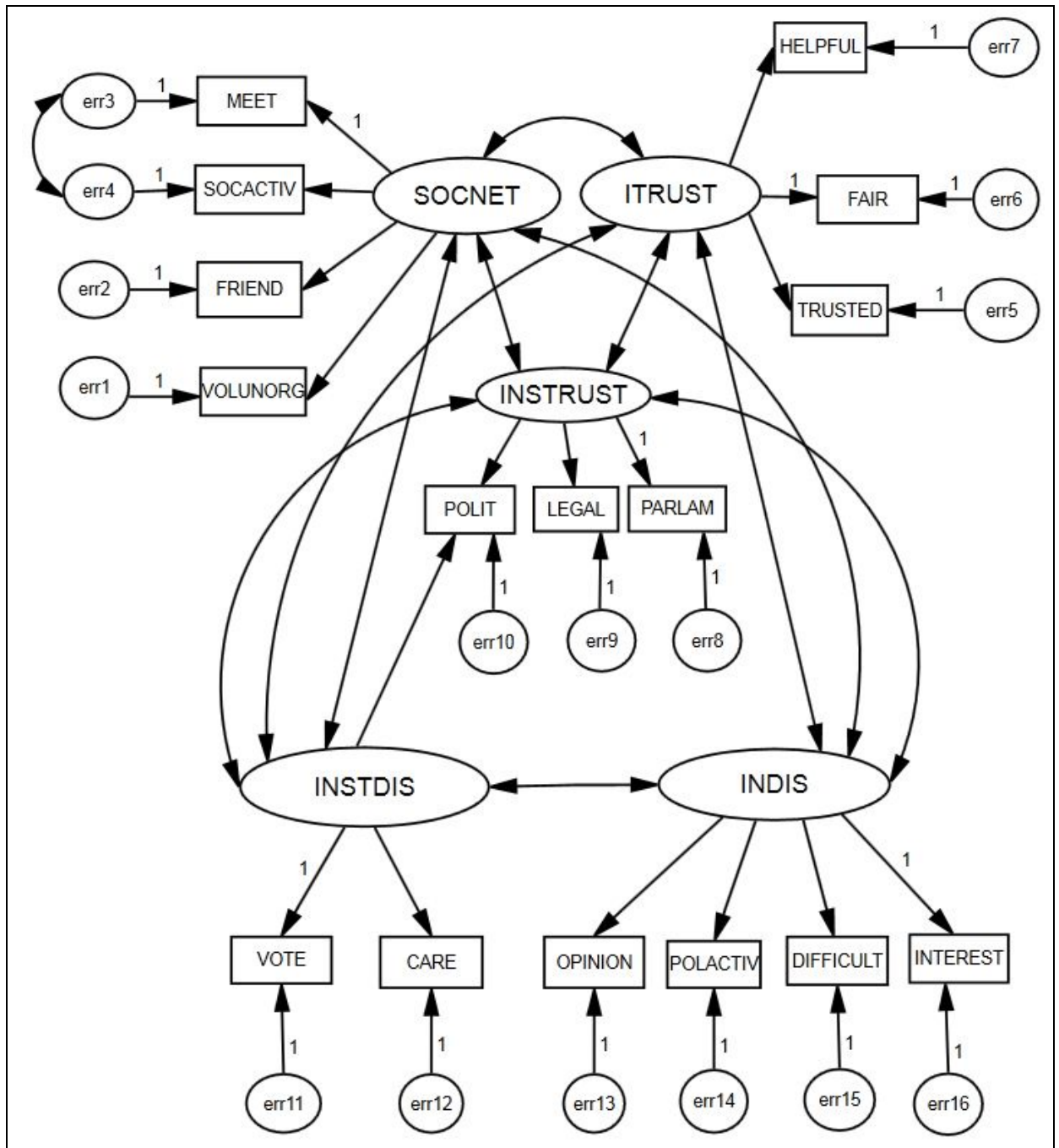
Důvěra v instituce		
POLIT (B10) (recod)	Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: politici	0 znamená, že instituci vůbec nedůvěřujete, 10 znamená, že jí naprosto důvěřujete.
	Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: politicians	0 means you do not trust an institution at all, and 10 means you have complete trust.
PARLAM (B7) (recod)	Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: Český parlament	0 znamená, že instituci vůbec nedůvěřujete, 10 znamená, že jí naprosto důvěřujete.
	Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: parliament	0 means you do not trust an institution at all, and 10 means you have complete trust.
LEGAL (B8) (recod)	Řekněte mi prosím s použitím stupnice 0 - 10 jak moc Vy osobně důvěřujete institucím, které budu číst nahlas: právní systém	0 znamená, že instituci vůbec nedůvěřujete, 10 znamená, že jí naprosto důvěřujete.
	Please tell me on a score of 0-10 how much you personally trust each of the institutions I read out: the legal system	0 means you do not trust an institution at all, and 10 means you have complete trust.

Individuální sociální kapitál (sociální důvěra)		
TRUSTED (A8)	Obecně vzato, řekl(a) byste, že se většině lidí dá důvěřovat, nebo že člověk nemůže být při jednání s lidmi nikdy dost opatrný?	0 – člověk nemůže být nikdy dost opatrný
		10 – většině lidí se dá důvěřovat
	Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?	0 – you can't be too careful
		10 – most people can be trusted
FAIR (A9)	Myslíte si, že by se Vás většina lidí snažila podvést, pokud by měli tu možnost, nebo by se snažili být poctiví?	0 – většina lidí by se mě snažila podvést
		10 – většina lidí by se snažila být poctivá
	Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?	0 – most people would try to take advantage of me
		10 – most people would try to be fair
HELPFUL (A10)	Řekl(a) byste, že se lidé většinou snaží pomáhat druhým, nebo že se většinou starají o sebe?	0 – lidé se většinou starají o sebe
		10 – lidé se většinou snaží pomáhat
	Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves?	0 – people mostly look out for themselves
		10 – people mostly try to be helpful

Kolektivní sociální kapitál (sociální sítě)		
FRIEND (E14)	Jak důležité jsou ve Vašem životě následující věci?: přátelé	0 – naprosto nedůležité
		10 – naprosto důležité
	How important are friends in your life?	0 – extremely unimportant
		10 – extremely important
VOLUNORG (E19)	Jak důležité jsou ve Vašem životě následující věci?: dobrovolné organizace	0 – naprosto nedůležité
		10 – naprosto důležité
	How important are voluntary organizations in your life?	0 – extremely unimportant
		10 – extremely important
MEET (C2)	Jak často se ve svém volném čase společensky scházíte s přáteli, příbuznými či kolegy?	1 – nikdy
		7 – každý den
	How often do you meet socially with friends, relatives or work colleagues?	1 – never
		7 – every day
SOACTIV (C4)	Když se srovnáte s jinými lidmi Vašeho věku, jak často se ve svém volném čase účastníte společenských aktivit?	1 – mnohem méně než většina lidí mého věku
		2 – méně než většina lidí mého věku
		3 – přibližně stejně jako lidé mého věku
		4 – více než většina lidí mého věku
		5 – mnohem více než většina lidí mého věku
	Compared to other people of your age, how often would you say you take part in social activities?	1 – much less than most
		2 – less than most
		3 – about the same
		4 – more than most
		5 – much more than most

Příloha 2: diagram strukturního modelu škál sociálního kapitálu a politického odcizení

Diagram 8 Strukturní model postojových škál sociálního kapitálu a politického odcizení



Příloha 3: syntax pro výpočet polychoriální korelační matice a test základního modelu metodou odhadu DWLS v LISRELU

Příklad syntaxu pro výpočet polychoriální korelační matice (PRELIS) a test základního modelu metodou odhadu DWLS pro Českou republiku v LISRELU (SIMPLIS output)

```
SY='C:\DATA\MODEL FINAL\SI\ESS1_MULTIGROUP_CZ.PSF'  
OU MA=PM SM=maticeCZ.pm AC=maticeCZ.ac
```

```
Observed variables: TRUSTED FAIR HELPFUL INTEREST DIFFICULT POLACTIV  
OPINION CARE VOTE PARLAM LEGAL POLIT MEET SOCACTIV FRIEND VOLUNORG  
Correlation matrix from File maticeCZ.pm
```

```
Asymptotic Covariance Matrix from File maticeCZ.ac
```

```
Sample size: 1307
```

```
Latent variables:
```

```
SOCNET ITRUST INSTRUST INSTDIS INDIS
```

```
Equations:
```

```
MEET=1*SOCNET
```

```
FRIEND=SOCNET
```

```
VOLUNORG=SOCNET
```

```
SOCACTIV=SOCNET
```

```
FAIR=1*ITRUST
```

```
TRUSTED=ITRUST
```

```
HELPFUL=ITRUST
```

```
PARLAM=1*INSTRUST
```

```
LEGAL=INSTRUST
```

```
POLIT=INSTRUST
```

```
VOTE=1*INSTDIS
```

```
CARE=INSTDIS
```

```
POLIT=INSTDIS
```

```
INTEREST=1*INDIS
```

```
DIFFICULT=INDIS
```

```
POLACTIV=INDIS
```

```
OPINION=INDIS
```

```
LET THE ERRORS BETWEEN MEET AND SOCACTIV CORRELATE
```

```
Options: ME=DW
```

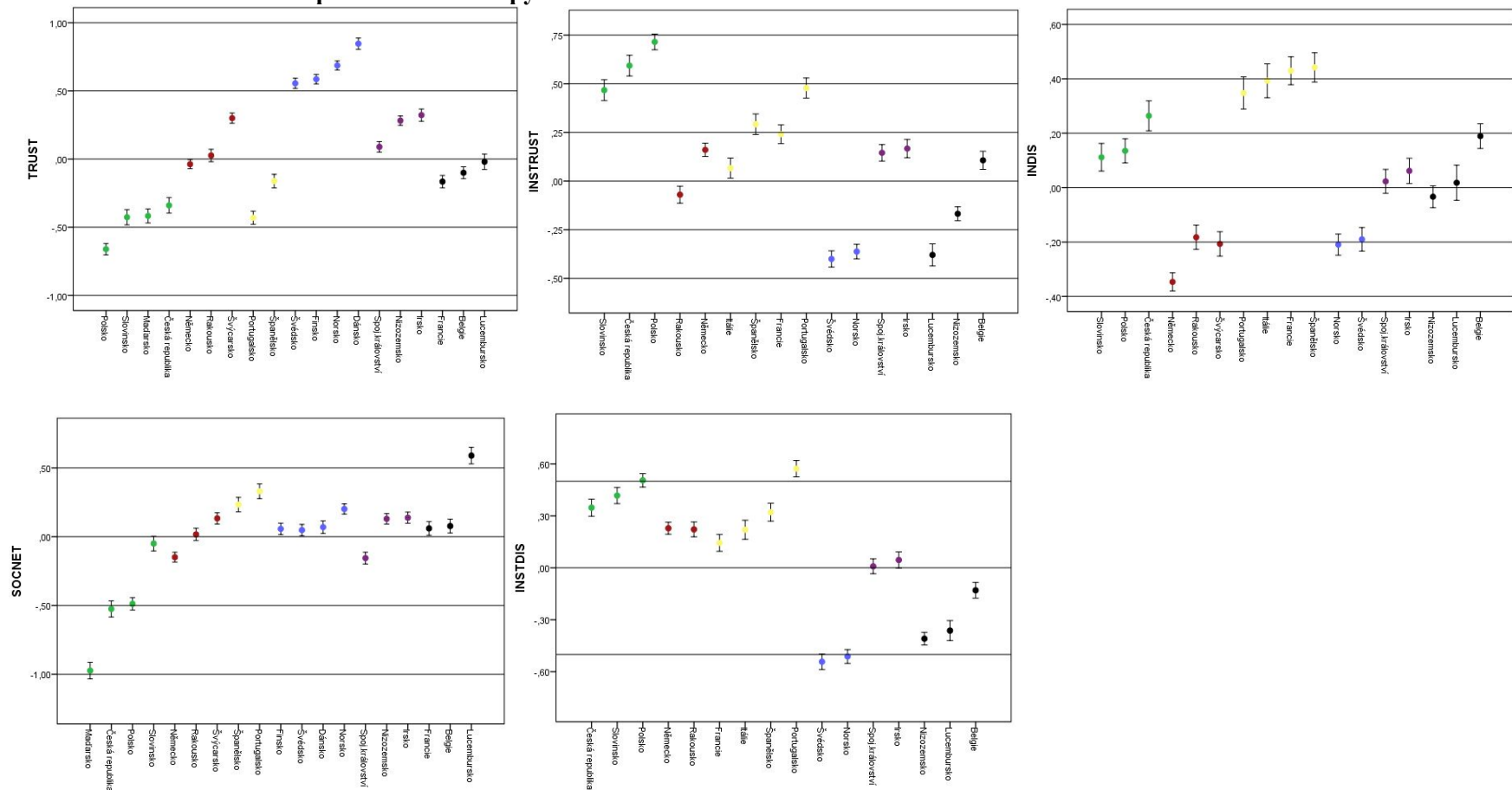
```
Path Diagram
```

```
Lisrel output: SC MI PC
```

```
End of Problem
```

Příloha 4: grafické srovnání latentních průměrů sociálního kapitálu a politického odcizení

Graf 2 Srovnání průměrů faktorů sociálního kapitálu, důvěry v instituce a politického odcizení a jejich 95 % intervalů spolehlivosti mezi kulturními skupinami zemí Evropy



Zdroj: ESS 2002; vlastní výpočty

Poznámka: Barevně jsou rozlišeny kulturní skupiny zemí střední, jižní, východní, západní a severní Evropy.

Příloha 5: separátní analýzy základního modelu škál důvěry v šesti vlnách šetření ESS

Tabulka 86 Separátní analýza základního modelu škál důvěry v první vlně šetření (2002) - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko	34,547	8	4,318	0,995	0,986	0,993	0,040	(0,027;0,054)	0,883
Belgie	26,625	8	3,328	0,995	0,987	0,994	0,037	(0,022;0,053)	0,913
Česká republika	7,037	8	0,880	0,998	0,995	1,000	0,000	(0,000;0,031)	0,999
Německo	51,691	8	6,461	0,994	0,984	0,990	0,044	(0,033;0,056)	0,786
Dánsko	22,313	8	2,789	0,995	0,987	0,994	0,035	(0,018;0,053)	0,912
Španělsko	31,151	8	3,894	0,993	0,982	0,992	0,044	(0,028;0,060)	0,714
Finsko	52,174	8	6,522	0,992	0,978	0,989	0,053	(0,040;0,067)	0,330
Francie	20,737	8	2,592	0,995	0,988	0,994	0,033	(0,016;0,051)	0,940
Spojené království	31,672	8	3,959	0,995	0,987	0,994	0,039	(0,025;0,053)	0,900
Řecko	35,134	8	4,392	0,995	0,987	0,994	0,037	(0,025;0,051)	0,942
Maďarsko	16,300	8	2,037	0,996	0,991	0,997	0,026	(0,006;0,044)	0,988
Švýcarsko	55,989	8	6,999	0,990	0,974	0,985	0,056	(0,043;0,070)	0,214
Irsko	12,944	8	1,618	0,998	0,994	0,998	0,018	(0,000;0,035)	1,000
Izrael	33,471	8	4,184	0,995	0,988	0,992	0,037	(0,025;0,051)	0,942
Itálie	15,978	8	1,997	0,995	0,988	0,995	0,030	(0,005;0,051)	0,943
Lucembursko	41,487	8	5,186	0,989	0,971	0,983	0,059	(0,042;0,077)	0,185
Nizozemí	83,052	8	10,382	0,988	0,969	0,984	0,064	(0,052;0,077)	0,027
Norsko	64,514	8	8,064	0,990	0,973	0,981	0,059	(0,046;0,073)	0,116
Polsko	9,548	8	1,194	0,998	0,996	0,999	0,010	(0,000;0,030)	1,000
Portugalsko	34,561	8	4,320	0,992	0,978	0,988	0,050	(0,034;0,068)	0,461
Švédsko	38,997	8	4,875	0,993	0,983	0,991	0,045	(0,032;0,060)	0,684
Slovinsko	17,44	8	2,180	0,996	0,989	0,996	0,029	(0,010;0,048)	0,969

Zdroj: ESS 2002; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: $GFI \geq 0,95$, $AGFI \geq 0,95$, $CFI \geq 0,95$ (příp. $CFI \geq 0,92$), $RMSEA \leq 0,05$ (příp. $RMSEA \leq 0,08$), $PCLOSE \geq 0,5$; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Tabulka 87 Separátní analýza základního modelu škál důvěry ve druhé vlně šetření (2004) - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko	33,065	8	4,133	0,995	0,987	0,994	0,039	(0,026;0,053)	0,904
Belgie	28,395	8	3,549	0,995	0,986	0,993	0,038	(0,024;0,054)	0,879
Česká republika	38,364	8	4,795	0,995	0,988	0,996	0,037	(0,026;0,049)	0,960
Německo	52,288	8	6,536	0,994	0,984	0,992	0,045	(0,034;0,057)	0,745
Dánsko	46,212	8	5,777	0,989	0,972	0,985	0,058	(0,043;0,075)	0,184
Španělsko	8,935	8	1,117	0,998	0,995	1,000	0,009	(0,000;0,032)	1,000
Finsko	76,018	8	9,502	0,988	0,968	0,980	0,066	(0,053;0,080)	0,025
Francie	17,873	8	2,234	0,997	0,991	0,996	0,026	(0,009;0,043)	0,992
Spojené království	75,379	8	9,422	0,986	0,964	0,981	0,068	(0,055;0,083)	0,015
Řecko	28,418	8	3,552	0,996	0,990	0,996	0,033	(0,020;0,047)	0,981
Maďarsko	23,990	8	2,999	0,994	0,985	0,994	0,038	(0,021;0,056)	0,856
Švýcarsko	34,010	8	4,251	0,994	0,985	0,991	0,041	(0,027;0,055)	0,846
Irsko	26,879	8	3,360	0,996	0,989	0,995	0,033	(0,020;0,047)	0,975
Izrael									
Itálie	36,399	8	4,550	0,992	0,979	0,989	0,049	(0,033;0,066)	0,510
Lucembursko	14,197	8	1,775	0,997	0,991	0,997	0,024	(0,000;0,044)	0,988
Nizozemí	54,799	8	6,850	0,990	0,975	0,988	0,057	(0,043;0,071)	0,196
Norsko	58,110	8	7,264	0,989	0,972	0,981	0,060	(0,046;0,075)	0,113
Polsko	10,170	8	1,271	0,998	0,994	0,999	0,013	(0,000;0,034)	0,999
Portugalsko	68,722	8	8,590	0,988	0,969	0,975	0,063	(0,050;0,077)	0,052
Švédsko	69,179	8	8,647	0,988	0,970	0,984	0,064	(0,051;0,079)	0,042
Slovinsko	6,952	8	0,869	0,998	0,995	1,000	0,000	(0,000;0,029)	1,000

Zdroj: ESS 2004; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Poznámka 2: Šedá pole značí absenci země ve výzkumu.

Tabulka 88 Separátní analýza základního modelu škál důvěry ve třetí vlně šetření (2006) - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko	63,960	8	7,995	0,990	0,975	0,987	0,056	(0,044;0,070)	0,188
Belgie	40,927	8	5,116	0,993	0,981	0,989	0,048	(0,034;0,063)	0,543
Česká republika									
Německo	56,982	8	7,123	0,993	0,983	0,991	0,047	(0,036;0,059)	0,655
Dánsko	30,160	8	3,770	0,993	0,982	0,991	0,044	(0,028;0,061)	0,700
Španělsko	28,085	8	3,511	0,995	0,986	0,991	0,038	(0,023;0,054)	0,888
Finsko	95,958	8	11,995	0,984	0,957	0,976	0,077	(0,064;0,091)	0,001
Francie	17,023	8	2,128	0,997	0,992	0,996	0,024	(0,007;0,040)	0,997
Spojené království	39,067	8	4,883	0,994	0,985	0,992	0,041	(0,029;0,055)	0,853
Řecko									
Maďarsko	43,189	8	5,399	0,990	0,974	0,988	0,057	(0,041;0,074)	0,232
Švýcarsko	49,944	8	6,243	0,990	0,974	0,984	0,056	(0,042;0,071)	0,236
Irsko	15,529	8	1,941	0,997	0,992	0,997	0,024	(0,001;0,042)	0,994
Izrael									
Itálie									
Lucembursko									
Nizozemí	48,220	8	6,028	0,991	0,978	0,988	0,052	(0,039;0,067)	0,360
Norsko	52,009	8	6,501	0,990	0,974	0,985	0,056	(0,042;0,071)	0,216
Polsko	19,232	8	2,404	0,996	0,989	0,995	0,030	(0,013;0,047)	0,975
Portugalsko	37,537	8	4,692	0,994	0,984	0,992	0,043	(0,030;0,058)	0,761
Švédsko	48,261	8	6,033	0,991	0,978	0,988	0,053	(0,039;0,067)	0,356
Slovinsko	26,042	8	3,255	0,994	0,983	0,992	0,041	(0,024;0,059)	0,776

Zdroj: ESS 2006; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: $GFI \geq 0,95$, $AGFI \geq 0,95$, $CFI \geq 0,95$ (příp. $CFI \geq 0,92$), $RMSEA \leq 0,05$ (příp. $RMSEA \leq 0,08$), $PCLOSE \geq 0,5$; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Poznámka 2: Šedá pole značí absenci země ve výzkumu.

Tabulka 89 Separátní analýza základního modelu škál důvěry ve čtvrté vlně šetření (2008) - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko	81,206	8	10,151	0,988	0,968	0,989	0,066	(0,053;0,079)	0,020
Belgie	27,917	8	3,490	0,995	0,986	0,993	0,038	(0,023;0,054)	0,890
Česká republika	49,819	8	6,227	0,992	0,978	0,991	0,052	(0,039;0,066)	0,386
Německo	55,469	8	6,934	0,993	0,982	0,989	0,047	(0,036;0,059)	0,624
Dánsko	60,517	8	7,565	0,988	0,968	0,981	0,065	(0,050;0,080)	0,049
Španělsko	27,197	8	3,400	0,996	0,990	0,995	0,032	(0,020;0,046)	0,984
Finsko	65,468	8	8,183	0,990	0,974	0,987	0,058	(0,045;0,071)	0,149
Francie	70,119	8	8,765	0,989	0,970	0,979	0,062	(0,049;0,076)	0,062
Spojené království	32,129	8	4,016	0,995	0,988	0,995	0,036	(0,024;0,050)	0,950
Řecko	31,047	8	3,881	0,995	0,987	0,996	0,038	(0,024;0,052)	0,918
Maďarsko	24,983	8	3,123	0,994	0,985	0,994	0,038	(0,022;0,056)	0,857
Švýcarsko	46,962	8	5,870	0,990	0,975	0,987	0,054	(0,040;0,070)	0,289
Irsko	28,891	8	3,611	0,994	0,985	0,992	0,039	(0,024;0,055)	0,869
Izrael	51,631	8	6,454	0,993	0,981	0,987	0,049	(0,037;0,062)	0,512
Itálie									
Lucembursko									
Nizozemí	25,834	8	3,229	0,995	0,987	0,995	0,036	(0,021;0,052)	0,927
Norsko	25,391	8	3,174	0,995	0,986	0,994	0,038	(0,022;0,055)	0,873
Polsko	27,983	8	3,498	0,994	0,984	0,991	0,041	(0,025;0,058)	0,787
Portugalsko	64,999	8	8,125	0,990	0,974	0,984	0,058	(0,045;0,071)	0,144
Švédsko	48,388	8	6,049	0,991	0,977	0,988	0,054	(0,040;0,069)	0,318
Slovinsko	12,915	8	1,614	0,996	0,991	0,998	0,023	(0,000;0,044)	0,983

Zdroj: ESS 2008; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Poznámka 2: Šedá pole značí absenci země ve výzkumu.

Tabulka 90 Separátní analýza základního modelu škál důvěry v páté vlně šetření (2010) - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko									
Belgie	46,123	8	5,765	0,991	0,976	0,986	0,053	(0,039;0,069)	0,328
Česká republika	57,133	8	7,142	0,992	0,979	0,993	0,052	(0,039;0,064)	0,395
Německo	80,805	8	10,101	0,991	0,976	0,986	0,056	(0,045;0,067)	0,170
Dánsko	49,021	8	6,128	0,989	0,972	0,984	0,058	(0,043;0,074)	0,177
Španělsko	7,794	8	0,974	0,999	0,996	1,000	0,000	(0,000;0,027)	1,000
Finsko	114,624	8	14,328	0,981	0,949	0,973	0,085	(0,072;0,099)	0,000
Francie	50,740	8	6,343	0,990	0,974	0,982	0,056	(0,042;0,071)	0,226
Spojené království	62,002	8	7,750	0,991	0,977	0,989	0,055	(0,043;0,068)	0,248
Řecko	55,315	8	6,914	0,993	0,982	0,992	0,047	(0,036;0,059)	0,628
Maďarsko	34,790	8	4,349	0,992	0,979	0,992	0,048	(0,032;0,065)	0,552
Švýcarsko	44,803	8	5,600	0,990	0,973	0,984	0,058	(0,042;0,075)	0,201
Irsko	96,981	8	12,123	0,987	0,965	0,982	0,068	(0,056;0,081)	0,006
Izrael	19,774	8	2,472	0,997	0,992	0,997	0,027	(0,012;0,042)	0,996
Itálie									
Lucembursko									
Nizozemí	53,574	8	6,697	0,990	0,974	0,987	0,057	(0,043;0,071)	0,204
Norsko	57,225	8	7,153	0,988	0,969	0,982	0,064	(0,049;0,080)	0,065
Polsko	26,251	8	3,281	0,995	0,986	0,994	0,038	(0,022;0,054)	0,878
Portugalsko	22,507	8	2,813	0,996	0,990	0,996	0,030	(0,016;0,046)	0,985
Švédsko	53,033	8	6,629	0,988	0,968	0,982	0,063	(0,047;0,079)	0,083
Slovinsko	10,663	8	1,333	0,997	0,993	0,999	0,016	(0,000;0,038)	0,996

Zdroj: ESS 2010; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: $GFI \geq 0,95$, $AGFI \geq 0,95$, $CFI \geq 0,95$ (příp. $CFI \geq 0,92$), $RMSEA \leq 0,05$ (příp. $RMSEA \leq 0,08$), $PCLOSE \geq 0,5$; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Poznámka 2: Šedá pole značí absenci země ve výzkumu.

Tabulka 91 Separátní analýza základního modelu škál důvěry v šesté vlně šetření (2012) - statistiky vhodnosti modelu

země	χ^2	stupně volnosti df	χ^2/df	GFI	AGFI	CFI	RMSEA	RMSEA 90 % C. I.	PCLOSE
Rakousko									
Belgie	72,366	8	9,046	0,987	,967	0,983	0,066	(0,053;0,080)	0,026
Česká republika	39,319	8	4,915	0,993	0,982	0,995	0,046	(0,032;0,060)	0,669
Německo	80,092	8	10,012	0,991	0,976	0,985	0,056	(0,045;0,067)	0,173
Dánsko	96,317	8	12,040	0,980	0,948	0,973	0,083	(0,069;0,098)	0,000
Španělsko	28,472	8	3,559	0,995	0,986	0,993	0,038	(0,023;0,053)	0,896
Finsko	86,302	8	10,788	0,987	0,966	0,983	0,067	(0,055;0,080)	0,012
Francie									
Spojené království	29,810	8	3,726	0,995	0,988	0,995	0,036	(0,023;0,050)	0,951
Řecko		8							
Maďarsko									
Švýcarsko	68,212	8	8,527	0,984	0,958	0,973	0,074	(0,058;0,091)	0,006
Irsko	34,594	8	4,324	0,995	0,988	0,995	0,037	(0,025;0,050)	0,955
Izrael	18,100	8	2,263	0,997	0,993	0,997	0,023	(0,009;0,038)	0,999
Itálie									
Lucembursko		8							
Nizozemí	69,417	8	8,677	0,987	0,967	0,985	0,065	(0,052;0,080)	0,033
Norsko	70,054	8	8,757	0,985	0,961	0,979	0,070	(0,055;0,085)	0,013
Polsko	27,981	8	3,498	0,995	0,986	0,993	0,038	(0,023;0,054)	0,895
Portugalsko	12,127	8	1,516	0,998	0,995	0,999	0,016	(0,000;0,033)	1,000
Švédsko	22,496	8	2,812	0,996	0,989	0,996	0,032	(0,017;0,048)	0,968
Slovinsko	16,239	8	2,030	0,995	0,988	0,997	0,029	(0,006;0,050)	0,950

Zdroj: ESS 2012; vlastní výpočty – každý řádek reprezentuje samostatnou analýzu

Poznámka: Všechny χ^2 testy jsou statisticky významné (p value < 0,05). Model reprezentuje data dobře v případě, že: GFI \geq 0,95, AGFI \geq 0,95, CFI \geq 0,95 (příp. CFI \geq 0,92), RMSEA \leq 0,05 (příp. RMSEA \leq 0,08), PCLOSE \geq 0,5; 90 % C. I. je 90 % interval spolehlivosti indexu RMSEA.

Poznámka 2: Šedá pole značí absenci země ve výzkumu, k modře označeným zemím nebylo ještě možné v lednu 2014 získat nové datové soubory za šestou vlnu šetření ESS (2012).

SUMMARY

This methodological study focuses on the international comparability of two of the most frequent robust attitude scales: namely social capital and political disaffection scales as they are measured in the European Social Survey first data set. The first goal of this study is to evaluate the levels of cross-country and cross-time attitude scales comparability. The key analytic part of the study is based on structural modelling which is the most popular technique used in this context today. It can provide a statistical test for the measurement invariance of factors over groups or rounds of survey. Furthermore, it enables researches to decide which type of statistics may be compared, given different levels of invariance. The second, rather theoretical goal is to introduce the concept of equivalence in the context of cross-national surveys to the Czech sociology. This issue hasn't been analysed and published in a systematic way in the Czech social sciences yet, and it is still overshadowed by the social survey quality issues and by the concept of measurement error in cross-national surveys.

The text is divided into several parts which focus successively on the theoretical background, methodological issues and empirical findings. The first chapter introduces the development of the cross-national surveys in the 20th century. The second chapter summarizes the concept of equivalence in the cross-national surveys, introduces basic terms, suggests brief categorization of the basic types of equivalence and outlines the portfolio of statistical techniques which are typically used for testing the level of scales comparability. The third chapter summarizes theoretical background of social capital and political disaffection concepts and shows the ways of their measurement as attitude scales in the European Social Survey. In this context, there are all systematic measurement errors which can jeopardize scales comparability pointed out. These theoretical and methodological parts of the study are used for switching to the second analytic part where the level of attitude scales comparability is actually analysed.

Therefore the fourth chapter describes the basic techniques, which verify the consistency and reliability of social capital and political disaffection attitude scales, and brings an overview of first outcomes of scales comparability level across 22 countries. The fifth chapter introduces structural modelling which is a more technically advanced method for the evaluation of scales equivalence. The way and the analysis procedure are described there in more detail. In the sixth chapter, the basic model of attitude scales is illustrated, incl. the description of all relations between the selected measurement indicators and latent variables. Then, it is the subject of further testing in comparison analyses. The seventh, the eighth and the ninth chapters represent separate multi-group analyses of selected attitude scales comparability complemented by rich empiric material on the basis of which detailed evaluation of the level of scales and item equivalence are carried out. The tenth chapter deals with the explanation of the character of time dependent multi-group analysis based on interpersonal general trust and institutional trust attitude scales, and it intentionally complements the outcomes of previous analyses. The last chapter summarizes the most important outcomes proceeding from all comparability analyses and universal recommendations for substantive comparative work options suitable for dealing with these attitude scales are suggested.

The results of the data analysis are the following. The attitude scales were verified as configural and at least partially metric equivalent in all participating countries. It is obvious that the social capital and political disaffection concepts are understood and interpreted in the same way in all countries. The scalar equivalence of these scales hasn't been verified in these European countries. Therefore it is probable that extent of the measurement scales of items and individual answer categories of these scales aren't perceived by respondents identically. At the same time, respondents use these measurement scales in different way. Therefore, even

though respondents have the same opinion about trust in people or institutions, their score at measurement scale doesn't have to be the same. Because plenty of culturally and linguistically various countries were chosen to this analyses, obviously there is certain kind of the measurement error which caused the deviation in the intercept parameter of the item. High sensitivity of the scalar equivalence to the combination of slight cultural, linguistic and historic differences between countries and to the gentle changes of public opinion in time can cause this problem.

The implications of the given conclusions into practise is then of such a nature that in all the participating countries in ESS (2002) there has been no problem in comparing the relations between these two attitude scales, or relations of these attitude scales toward other metric equivalent attitude scales, or, perhaps the relations between the scales attitudes and other metric equivalent variable out of the data analysis set. In this process, we can use correlation, or regression analyses, or structural modelling. This type of meritory analyses is applied at cross-national level more frequently. The complications arises when we want to compare latent means or item means across countries because outcomes show that it is not possible due to absence of the level of scalar equivalence. It would be an invalid comparison which is not allowed. Only in case of social capital factors the comparison of averages is possible among ten European countries. However, a detailed multi-group analysis has revealed that latent means can be mutually compared in small homogenous groups of countries. Mostly, these countries are homogenous from the point of view of geographical location, culture, common history and even linguistic origins. Apparently, it is because the subtle cultural and linguistic deviations make the comparison of variable means impossible.

Not surprisingly, the multi-group analyses of scalar equivalence show that items which measure the latent concept of social capital are much more problematic as regards cross-country comparability than items which measure latent concept of political disaffection or institutional trust. These results correspond to results of other empirical research projects and methodology studies. Structural modelling has revealed, as the most problematic the following three items in the social capital attitude scale (HELPFUL, VOLUNORG, a LEGAL) and two items in the political disaffection attitude scale (OPINION and VOTE). These items probably deviate due to methodological systematic error but these outcomes should be confirmed in the subsequent cognitive testing or by a more appropriate statistical technique such as hierarchical structural modelling or item response theory.

General conclusion of this study is that it is not possible to rely exclusively on high-quality work of the international research team which plans and prepares the cross-national survey. The fact that any attitude scales are thoroughly formulated, pretested and analysed because of their validity and reliability and that every item is expertly translated into various languages and cognitive tested doesn't mean that these scales are fully comparable across all countries in the data set. It is always necessary to verify attitude scales comparability over and over in every new round of data collection.