

Adam Sagan

Analiza inwariancji pomiaru w badaniach przekrojowych

Problemy Zarządzania, Finansów i Marketingu 33, 187-197

2014

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

*ADAM SAGAN*¹

Uniwersytet Ekonomiczny w Krakowie

ANALIZA INWARIANCJI POMIARU W BADANIACH PRZEKROJOWYCH

Streszczenie

Artykuł jest poświęcony zagadnieniom oceny inwariancji pomiaru w badaniach marketingowych w przypadku liczebnego zróżnicowania populacji. Ocena ta została przeprowadzona na przykładzie skali strategii podejmowania decyzji w gospodarstwie domowym z wykorzystaniem dwóch podejść. Dla dużej liczby grup zastosowano procedurę wyrównywania na podstawie funkcji upraszczającej, a dla małej liczby przekrojów – wielogrupową confirmacyjną analizę czynnikową dostępną w programie Mplus 7.11.

Słowa kluczowe: inwariancja pomiaru, porządkowanie, wielogrupowa confirmacyjna analiza czynnikowa

Wprowadzenie

Inwariancja pomiaru polega najczęściej na zastosowaniu wielogrupowej analizy czynnikowej w celu oceny stopnia niezmienniczości (stabilności) modelu czynnikowego (struktury ładunków czynnikowych, wyrazów wolnych, warian-

¹ sagana@uek.krakow.pl

cji błędów, kowariancji między czynnikami) w przekroju analizowanych grup. Wyróżnia się cztery podstawowe poziomy inwariancji skali:

1. Konfiguracywna ekwiwalencja. Zakłada jednorodny charakter ładunków czynnikowych (istotny statystycznie i różny od zera) oraz brak ograniczeń parametrów modelu (uwolnienie parametrów w przekroju porównywalnych grup).
2. Słaba ekwiwalencja (metryczna). Przyjmuje założenie o równości ładunków czynnikowych w przekroju grup (kultur), co oznacza, że wariancje i kowariancje czynników mogą być porównywalne. Jeżeli kultury różnią się ze względu na ładunki czynnikowe, to jednostkowa zmiana zmiennej ukrytej (wartości czynnikowych) nie odpowiada takiej samej zmianie w wartości wskaźników w przekroju grup. Zróznicowana zmiana w wartościach zmiennych obserwowalnych wskazuje, że czynniki nieuwzględnione w modelu pomiarowym tłumią lub wzmacniają wartości zmiennych w sposób specyficzny dla danej subpopulacji i efekt ten wzrasta (lub maleje) wraz ze zmianą wartości czynnika. Wskaźniki mierzą więc inne konstrukty w poszczególnych grupach. Ekwiwalencja metryczna zakłada stałość relacji między wskaźnikiem a czynnikiem (zmienną ukrytą).
3. Silna ekwiwalencja (skalarna). Oznacza równość ładunków czynnikowych i wyrazów wolnych w przekroju grup. Obserwowalne różnice w wartościach średnich ocen dla wskaźników są wynikiem różnic w ocenach zmiennych ukrytych. Jeżeli grupy różnią się wartościami wyrazów wolnych, to wartości zmiennej ukrytej w jednej z grup są wyższe ze względu na wpływ nieznanymi czynników nieuwzględnianymi w modelu (niezależnie od różnic w wartościach czynnikowych). Brak ekwiwalencji skalarnej jest efektem wpływu systematycznych czynników kulturowych (np. nastawienia lub stylu odpowiedzi: ARS/DARS. ERS).
4. Ścisła ekwiwalencja (błędów pomiaru). Nakłada restrykcje na równość ładunków czynnikowych, wyrazów wolnych i wariancji reszt w przekroju grup. Wariancja resztowa jest kombinacją wariancji swoistej i wariancji wynikającej z błędu losowego. Różnice w wariancjach resztowych w przekroju kultur wskazują na specyficzne zewnętrzne czynniki, które tłumią lub wzmacniają indywidualne różnice w obserwacjach dla poszczególnych kultur. Nieekwiwalencja błędów oznacza wpływ losowych czynników kulturowych.

W celu identyfikacji stopnia inwariancji można wyróżnić dwie strategie postępowania. W pierwszej (z dołu go góry) analizę rozpoczyna się od modelu nieograniczonego (konfiguracywnego) i w kolejnych krokach ustala poszczególne parametry jako równe w przekroju grup. W podejściu drugim (z góry do dołu) postępowanie jest odwrotne: rozpoczyna się od modelu o ścisłej inwariancji, a następnie uwalnia poszczególne parametry dla kolejnych estymowanych modeli.

Kluczowy problem w ocenie inwariancji stanowi liczba porównywanych grup. W sytuacji, gdy analizujemy inwariancję w porównawczych badaniach międzynarodowych, ocena stabilności parametrów modelu obejmuje porównania niewielkiej liczby grup (krajów), najczęściej ograniczonej do dwóch–trzech przekrojów. W tego typu analizach wielogrupowa analiza czynnikowa jest najczęściej przyjmowaną strategią oceny inwariancji i pozwala na pełną jej ocenę na podstawie analizy różnic w wartościach statystyk c^2 między modelem ograniczonym (przy założeniu równości parametrów w przekrojach grup) a modelem nieograniczonym (przy uwolnieniu założenia równości parametrów).

Problem jest bardziej skomplikowany, gdy porównujemy inwariancję dla dużej liczby grup (np. większej od dziesięciu). W takim przypadku stosowanie klasycznej wielogrupowej analizy czynnikowej jest utrudnione i niezbyt praktyczne, ponieważ rośnie liczba porównań międzygrupowych (dla dziesięciu grup wynosi ona $(10 \times 9)/2 = 45$ porównań). Stosuje się wówczas dwa podejścia: a) analizę wielogrupową z wykorzystaniem procedury porządkowania (*alignment*), która jest dostępna w wersji 7.1 programu Mplus, oraz b) analizę wielopoziomową z efektami losowymi dla pozycji skali (*random item parameters*).

Charakterystyka skali pomiarowej

Ocena inwariancji pomiaru w przekrojach przestrzennych i geograficznych dotyczy zastosowań skali służącej do pomiaru postaw wobec podejmowania decyzji w gospodarstwie domowym. Najogólniej można w tym obszarze wyodrębnić dwie podstawowe strategie. Pierwsza zakłada „altruistyczną” strategię podejmowania decyzji, maksymalizującą dobro wspólnego gospodarstwa domowego, w której tzw. głowa rodziny podejmuje decyzje w imieniu całej rodziny. Ten unitarny model (model wspólnych preferencji) charakteryzuje się brakiem wewnętrznej rywalizacji o zasoby poszczególnych członków gospodarstwa².

² H. Alderman i in., *Unitary versus Collective Models of the Household: Is it Time to Shift the Burden of Proof?* „The World Bank Research Observer” 1995, nr 1, s. 1–19.

Druga strategia polega na wewnętrznym procesie negocjacyjno-rywalizacyjnym w gospodarstwie domowym, swoistej „walce o zasoby” wynikającej z nierównego poziomu kapitału społecznego członków rodziny, wartości osobowych lub posiadania odrębnych zasobów finansowych przez członków rodziny³.

W realizowanych przez Katedrę Analizy Rynku i Badań Marketingowych ogólnopolskich badaniach ankietowych została opracowana skala altruizmu w procesie podejmowania decyzji przez członków gospodarstwa domowego⁴. Skala ta obejmuje sześć pozycji skali Likerta, która identyfikuje postawy członków gospodarstwa domowego wobec strategii podejmowania decyzji na jednowymiarowym continuum altruizm – rywalizacja⁵. Analiza struktury czynnikowej skali wskazuje na występowanie dwóch czynników głównych wskazujących na dominujący typ postaw konsumentów.

Struktura ładunków czynnikowych i korelacji między czynnikami (rotacja Geomin) została przedstawiona w tabeli 1.

Na podstawie ładunków czynnikowych można stwierdzić, że pierwszy czynnik opisuje głównie pozycje dotyczące strategii altruistycznej podejmowania decyzji (za wyjątkiem pozycji P26). Natomiast najbardziej charakterystyczna dla strategii rywalizacyjnej jest pozycja P22. Biorąc pod uwagę strukturę czynnikową skali, został zbudowany dwupoziomowy model pomiarowy uwzględniający

³ M. Manser, M. Brown, *Marriage and Household Decision-Making*, „International Economic Review” 1989, nr 21, s. 31–44.

⁴ Badania zostały zrealizowane przez Centrum Badań i Ekspertyz Uniwersytetu Ekonomicznego w Katowicach w ramach Grantu NCN Katedry Analizy Rynku i Badań Marketingowych UEK nr 211/01/B/HS4/04812 w latach 2011–2013. Badania zostały przeprowadzone na kwotowej próbie 1100 respondentów w 420 gospodarstwach domowych. Próba miała charakter zagnieżdżony. Wywiady były przeprowadzane wśród członków gospodarstwa domowego.

⁵ Pozycje skali mierzące continuum strategii podejmowania decyzji (altruizm – rywalizacja) są następujące:

^{P21} Staram się tak podzielić dochody/budżet/oszczędności mojego gospodarstwa domowego, aby wszyscy członkowie rodziny, w miarę możliwości, mieli wystarczające środki na zaspokojenie indywidualnych potrzeb.

^{P22} Zawsze dążę do wydzielenia osobnych środków z dochodów wspólnych rodziny na zaspokojenie moich indywidualnych potrzeb.

^{P23} Uważam, że rodzina powinna ograniczać wydatki na indywidualne potrzeby poszczególnych osób na rzecz zaspokojenia wspólnych potrzeb

^{P24} Dobro wspólne całej rodziny jest ważniejsze niż zaspokajanie zachcianek, dążeń i przyjemności każdego z osobna.

^{P25} Lepiej realizować się dzięki dobrom i usługom służącym całemu gospodarstwu domowemu niż indywidualnym potrzebom każdego z osobna.

^{P26} Radość życia czerpię w większym stopniu z dóbr i usług, które służą wszystkim członkom rodziny, niż tych kupowanych na moje osobiste potrzeby.

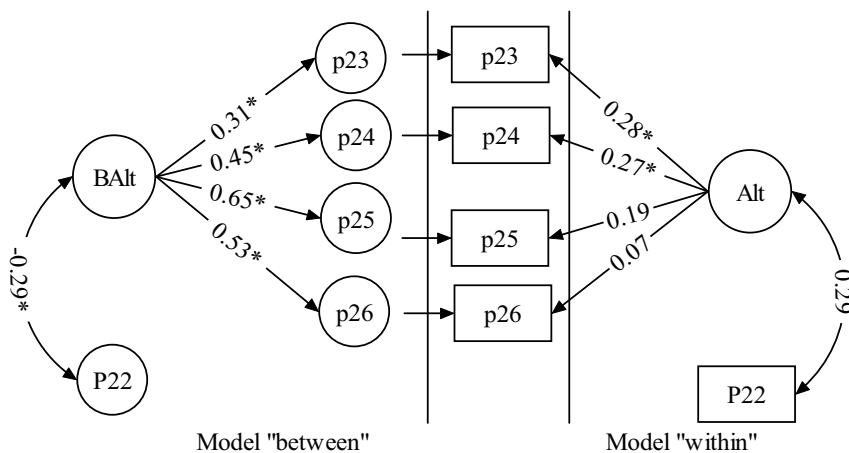
zarówno poziom indywidualny analizy (członków rodziny – model *within*), jak również poziom gospodarstw domowych (poziom grupowy – model *between*). Struktura modelu dwupoziomowego została przedstawiona na rysunku 1.

Tabela 1

Struktura ładunków czynnikowych skali postaw wobec strategii podejmowania decyzji w gospodarstwie domowym

Zmienna	Czynnik 1	Czynnik 2
P21	0.019	0.126
P22	0.044	-0.381*
P23	0.390*	0.030
P24	0.654*	-0.021
P25	0.562*	0.167
P26	0.003	0.747*
Dopasowanie modelu	Chi-kwadrat = 4.726 (4), poziom p = 0,31 RMSEA = 0.00 CFI = 0.99	

Źródło: opracowanie własne.



Rys. 1. Dwupoziomowa konfirmacyjna analiza czynnikowa skali postaw wobec strategii podejmowania decyzji

Źródło: opracowanie własne na podstawie programu Mplus 7.1.

Dopasowanie modelu znajduje się na akceptowalnym poziomie. Wartość statystyki χ^2 wynosi 59,27 dla dziesięciu stopni swobody ($p=0.00$). Wartość średniokwadratowego pierwiastka błędu RMSEA równa się 0,06, a przyrostowe wskaźniki CFI i TLI wynoszą 0,90 i 0,79. Standaryzowany średniokwadratowy pierwiastek reszt (SRMR) dla modelu z poziomu I i II wynosi odpowiednio 0,05 i 0,18. Należy zwrócić uwagę na fakt, że na poziomie gospodarstwa domowego ładunki czynnikowe wskazują na klarowne występowania wymiaru altruizmu w strategii podejmowania decyzji (wskaźnik P26 ma już istotnie dodatni ładunek czynnikowy).

Ocena inwariancji pomiaru w zróżnicowanej populacji

W ocenie skali ważnym zagadnieniem jest analiza inwariancji pomiarowej skali. Prezentowana skala została zastosowana na próbie ogólnopolskiej i stabilność ocen respondentów może zależeć od różnic w przekrojach geograficzno-kulturowych. Znaczenie analizy inwariancji skal jest, jak już podkreślono, szczególnie ważne w badaniach międzykulturowych, lecz w przestrzennych analizach porównawczych rola oceny stabilności modelu pomiarowego również stanowi istotny czynnik umożliwiający diagnozę obciążenia skali w badaniach ogólnokrajowych.

W przypadku znacznej liczby przekrojów analizy (wiele krajów, województw, regionów) w ocenie inwariancji pomiaru można zastosować procedurę wyrównywania (*alignment*), pozwalającą na ocenę ekwiwalencji pomiaru w przekroju wielu jednostek (np. województw). Twórcami metody są T. Asparouhov i B. Muthen⁶. Jej celem jest minimalizacja zakresu braku inwariancji metrycznej na podstawie oszacowanych wartości średnich dla czynników i wariancji ładunków czynnikowych. Identyfikacji parametrów modelu czynnikowego dokonuje się przy zastosowaniu ograniczeń maksymalizujących inwariantną strukturę czynnikową i tzw. funkcję upraszczającą (*simplicity function*). Jej idea jest bliska koncepcji rotacji czynnikowej maksymalizującej zróżnicowanie ładunków czynnikowych w przekrojach czynników (uzyskujemy wówczas czynniki o bardzo niskich lub bardzo wysokich ładunkach zamiast czynników o ładunkach przeciętnych utrudniających ich interpretację). W tym przypadku efektem maksymalizacji staje się uzyskanie wielu silnie inwariantnych parametrów i jednocześnie niewielkiej liczby silnie nieinwariantnych parametrów w przekroju grup

⁶ B. Muthen, T. Asparouhov, *New Methods for the Study of Measurement Invariance with Many Groups*, www.statmodel.com, (27.10.2013).

(w porównaniu do „nierotowanego” układu odniesienia, w którym występuje wiele parametrów o umiarkowanym poziomie braku inwariancji).

Procedura wyrównywania składa się z dwóch etapów. W pierwszym szacowane są modele konfiguracywne (z wolnymi ładunkami i wyrazami wolnymi, średnimi czynnikowymi ustalonymi jako zero i wariancjami ładunków ustalonymi na poziomie 1,0). W drugim następuje optymalizacja minimalizująca całkowitą inwariancję modelu w przekroju grup na podstawie funkcji upraszczającej (dla każdej pary grup j_1, j_2 wyrazu wolnego n_{pj} i ładunku czynnikowego l_{pj}). Funkcja upraszczająca ma postać⁷:

$$F = \sum_P \sum_{j_1 < j_2} w_{j_1, j_2} f(\lambda_{pj_1} - \lambda_{pj_2}) + \sum_P \sum_{j_1 < j_2} w_{j_1, j_2} f(v_{pj_1} - v_{pj_2}) \quad (1)$$

Oceny jakości minimalizacji dokonuje się na podstawie współczynnika determinacji R^2 , wskazującego, jaka część zmienności parametrów modelu konfiguracyjnego jest tłumaczona za pomocą zmienności średnich czynnikowych i wariancji ładunków. Wysoka wartość R^2 wskazuje na silną inwariancję pomiaru.

W tabeli 2 przedstawiono wyniki oceny inwariancji pomiaru skali altruizmu w przekroju jedenastu badanych województw. Kody województw znajdujące się w nawiasach wskazują na brak inwariancji pomiaru dla tych jednostek administracyjnych.

Z tabeli 2 wynika, że jedynie w przypadku województwa podlaskiego (6) nie jest spełniona inwariancja pomiaru dla wskaźnika P24 („Dobro wspólne całej rodziny jest ważniejsze niż zaspokajanie zachcianek, dążeń i przyjemności każdego z osobna”). Wyrazy wolne w modelu pomiarowym dla tych województw istotnie różnią się od pozostałych. Ładunki czynnikowe spełniają zasadę inwariancji metrycznej, co pozwala na akceptację modelu pomiaru w badaniach ogólnopolskich i porównywanie wyników w przekrojach województw. Można więc wysnuć wniosek, że dla skali altruizmu jest spełniona ekwiwalencja metryczna, a w przypadku województwa podlaskiego nie jest spełniona ekwiwalencja skalarna (wśród mieszkańców tego województwa występuje efekt nastawienia lub stylu odpowiedzi). Ocena wkładu dopasowania modeli jest zaprezentowana w tabeli 3.

⁷ T. Asparouhov, B. Muthen, *Multiple-group Factor Analysis Alignment*, www.statmodel.com, (28.10.2013).

Tabela 2

Ocena inwariancji pomiaru skali strategii altruizmu w podejmowaniu decyzji

Zmienne	Kody województw
Wyrazy wolne:	
P23	1 2 3 4 5 6 7 8 9 10 11
P24	1 2 3 4 5 (6) 7 8 9 10 11
P25	1 2 3 4 5 6 7 8 9 10 11
P26	1 2 3 4 5 6 7 8 9 10 11
Ładunki czynnikowe:	
P23	1 2 3 4 5 6 7 8 9 10 11
P24	1 2 3 4 5 6 7 8 9 10 11
P25	1 2 3 4 5 6 7 8 9 10 11
P26	1 2 3 4 5 6 7 8 9 10 11

Kody województw:

1. dolnośląskie
2. kujawsko-pomorskie
3. lubelskie
4. małopolskie
5. mazowieckie
6. podlaskie
7. pomorskie
8. śląskie
9. warmińsko-mazurskie
10. wielkopolskie
11. zachodniopomorskie

Źródło: opracowanie własne na podstawie programu Mplus 7.1.

Tabela 3

Wkłady dopasowania modelu w procedurze porządkowania

Pozycja skali	Wyrazy wolne		Ładunki czynnikowe	
	Wkłady do funkcji dopasowania	R ²	Wkłady do funkcji dopasowania	R ²
P23	-36.526	0,36	-24.563	0,11
P24	-30.026	0,21	-24.089	0,02
P25	-21.440	0,41	-31.804	0,49
P26	-18.432	0,23	-30.792	0,11

Źródło: opracowanie własne na podstawie obliczeń w programie Mplus 7.1.

Wartości wkładów przedstawione w tabeli 3 wskazują, że pozycje P26 i P25 mają najslabszy wkład w wartość funkcji upraszczającej w odniesieniu do wyrazów wolnych (inwariancji skalarnej), a pozycje P23 i P24 – najslabszy wkład w wartość tej funkcji dla ładunków czynnikowych (inwariancji metrycznej). Zróznicowanie wkładów dla ładunków (w porównaniu do wyra-

zów wolnych) jest zdecydowanie słabsze. Oznacza to, że najsilniejszy wkład w brak inwariancji pomiaru w przekroju województw mają pozycje P23 i P24 (dla wyrazów wolnych) oraz P25 (dla ładunków czynnikowych). Współczynniki R^2 wskazują, jaka część zmienności wyrazów wolnych i ładunków jest wyjaśniana za pomocą zmienności średnich i wariancji czynnika w przekroju województw.

Ocena inwariancji pomiaru w układach wielogrupowych

W przypadku, gdy liczba przekrojów analizy nie jest zbyt duża (dwie–cztery grupy), wówczas stosowane jest najczęściej klasyczne podejście w ocenie inwariancji pomiaru. Polega ono na porównaniu modeli z nałożonymi ograniczeniami na parametry. Testowanie inwariancji konfiguracywnej polega na oszacowaniu modelu z wolnymi parametrami i ocenie ładunków różnych od zera. W teście inwariancji metrycznej zakłada się równość ładunków czynnikowych w przekroju grup. Ocena inwariancji skalarnej dodatkowo zakłada równość wyrazów wolnych.

W analizie skali strategii podejmowania decyzji oceny inwariancji pomiaru dokonano w przekroju członków gospodarstw domowych (ojca, matki i dziecka). Jest to szczególnie istotne dla jakości pomiaru z uwzględnieniem interakcji zachodzących między członkami gospodarstwa. W tabeli 4 przedstawiono wyniki porównania inwariancji konfiguracywnej, metrycznej i skalarnej skali w przekrojach osób w gospodarstwie domowym. Pierwszy typ inwariancji wskazuje na podobieństwo wzorów ładunków czynnikowych w przekroju porównywanych grup. Drugi odnosi się do równości ładunków w grupach (możliwe jest porównywanie wariancji czynników między grupami). Trzeci typ inwariancji związany jest z równością ładunków i wyrazów wolnych w modelu pomiarowym (możliwe jest porównywanie średnich dla czynników w grupach). Model w analizach porównawczych powinien spełniać co najmniej warunek ekwiwalencji metrycznej.

Tabela 4

Ocena ekwiwalencji pomiaru w przekroju osób w gospodarstwie domowym

Model z ekwiwalencją	Liczba parametrów	Chi-kwadrat, stopnie swobody i poziom p
konfiguratywną	36	13.32, 6, 0.04
metryczną	30	17.98, 12, 0.12
skalarną	24	32.25, 18, 0.02
Porównanie modeli		
	Chi-kwadrat	Stopnie swobody i poziom p
konfiguralny względem metrycznego	4.66	6, 0.59
metryczny względem skalarnego	14.26	6, 0.03
konfiguralny względem skalarnego	18.92	12, 0.09

Źródło: opracowanie własne.

Wszystkie modele (oprócz ostatniego) charakteryzują się akceptowalnymi statystykami dopasowania. Model z założoną ekwiwalencją metryczną jest nieistotny statystycznie, co pozwala na przyjęcie założenia metrycznej ekwiwalencji pomiaru i porównywania postaw członków gospodarstwa domowego. Model bardziej ograniczony (z ekwiwalencją skalarną) jest istotnie gorszy od modelu z ekwiwalencją metryczną. Należy więc przyjąć występowanie ekwiwalencji metrycznej w przekroju członków gospodarstwa domowego. Pozwala to na dokonywanie porównań w przekroju ról zakupowych ojca, matki i dziecka.

Podsumowanie

Wykorzystywanie skal w badaniach porównawczych (międzyregionalnych, międzynarodowych) wymaga spełnienia założenia inwariancji pomiaru. Zasada inwariancji pozwala na porównywanie wartości średnich i wariancji zmiennych ukrytych w przekroju porównywanych jednostek. Niespełnienie tego założenia i przeprowadzanie porównań bez uprzedniego sprawdzenia ekwiwalencji pomiaru może prowadzić do otrzymywania obciążonych wartości skal i błędów w interpretacji wielkości różnic między średnimi dla porównywanych przekrojów. Metoda analizy inwariancji zależy od liczebności grup.

W przypadku niewielkiej liczebności porównywanych grup ocena jest dokonywana na podstawie istotności różnic między miarami dopasowania modelu nieograniczonego, modelu z założoną inwariancją metryczną (z równymi ładunkami czynnikowymi w grupach) i z inwariancją skalarną (z równymi ładunkami

i wyrazami wolnymi). Dla większej liczby grup porównania tego typu są niepraktyczne i złożone. Metodę oceny inwariancji stanowi procedura wyrównywania, umożliwiającą porównywanie na podstawie funkcji upraszczającej.

W analizowanej skali podejmowania strategii altruistycznej w alokacji zasobów skala jest inwariantna zarówno w przekroju województw, jak i członków gospodarstw domowych. Może stanowić podstawę porównań strategii podejmowania decyzji w danym przekroju geograficzno-społecznym.

Bibliografia

- Alderman H. i in. *Unitary versus Collective Models of the Household: Is it Time to Shift the Burden of Proof?* „The World Bank Research Observer” 1995, nr 1.
- Asparouhov T., Muthen B., *Multiple-group factor analysis alignment*, www.statmodel.com (28.10.2013).
- Manser M., Brown M, *Marriage and Household Decision-Making*, „International Economic Review” 1989, nr 21.
- Muthen B., Asparouhov T., *New Methods for the Study of Measurement Invariance with Many Groups*, www.statmodel.com (28.10.2013).

ANALYSIS OF MEASUREMENT INVARIANCE IN CROSS-SECTIONAL RESEARCH

Summary

The article is devoted to the issues of assessment of measurement invariance in marketing research for different number of groups. The invariance check has been carried out on the example of the scale of decision-making strategies in the household using two approaches. For a large number of groups, a procedure of alignment based on simplicity function are used. For a small number of groups, the multi-group confirmatory factor analysis has been used for invariance testing.

Keywords: measurement invariance, alignment method, confirmatory factor analysis

Translated by Adam Sagan