

Adam Sagan

Modele PLS-PM i ich zastosowania w predykcji i wyjaśnianiu zjawisk ekonomicznych

Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania 39/2, 127-138

2015

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*

Uniwersytet Ekonomiczny w Krakowie

MODELE PLS-PM I ICH ZASTOSOWANIA W PREDYKCJI I WYJAŚNIANIU ZJAWISK EKONOMICZNYCH

STRESZCZENIE

W artykule przedstawiono problem zastosowań modeli typu PLS-PM do predykcji i wyjaśniania zjawisk ekonomicznych w kontekście ogólnych modeli strukturalnych SEM. Wskazano na podstawowe różnice między podejściami z punktu widzenia budowy wskaźników, metod estymacji i ocen dopasowania. Szczególna uwaga została poświęcona traktowaniu ich jako modeli predykcyjnych pozwalających (w stosunku do SEM) na uzyskanie wyższej oceny wyjaśnianej wariancji i tym samym wyższą trafność przewidywania zjawisk.

Słowa kluczowe: modele PLS-PM, wskaźniki formatywne, moc predykcyjna modelu

Wstęp

Budowa modeli empirycznych w szeroko rozumianych naukach ekonomicznych zakłada opis, wyjaśnianie lub przewidywanie zdarzeń gospodarczych. W zależności od dojrzałości danej dyscypliny, charakteru założeń teoretycznych, celu badań i typu danych, w budowie modelu może być położony nacisk na jakość opisu zjawisk, po-

* Adres e-mail: sagana@uek.krakow.pl

prawną specyfikację zależności przyczynowo-skutkowych pozwalających na wyjaśnienie badanej rzeczywistości lub trafną ich predykcję. W obszarze nauk ekonomicznych, a w szczególności w zarządzaniu i marketingu, dużą popularnością cieszą się modele strukturalne ze zmiennymi ukrytymi, a w szczególności dwa ich rodzaje: 1) modele równań strukturalnych (SEM) i 2) modele ścieżkowe cząstkowych najmniejszych kwadratów (PLS-PM). Pierwsze z nich częściej są stosowane w badaniach naukowych związanych z wyjaśnianiem zjawisk (funkcja eksplanacyjna), a drugie z nich w badaniach komercyjnych eksponujących silniej funkcję predykcyjną.

Modele eksplanacyjne są budowane w celu poprawnego odzwierciedlenia zależności przyczynowo-skutkowych. W modelach tych dużą rolę odgrywa diagnoza egzogeniczności zmiennych, ocena zależności warunkowych i efektów mediacji oraz kontrola wyodrębnionych relacji przyczynowo-skutkowych. Stąd rola poprawnej specyfikacji modelu i oceny obciążenia (*bias*) wynikającego z niebezpieczeństwa błędnego oszacowania jego parametrów strukturalnych. Modele tego typu powinny charakteryzować się zarówno wysoką mocą eksplanacyjną (niskim błędem specyfikacji), jak i wysoką mocą predykcyjną (niskim błędem przewidywania).

Modele predykcyjne mają na celu poprawne przewidywanie przyszłych (nowych) obserwacji na podstawie modelu istniejącego. Dużą rolę odgrywa w nich wykorzystywanie prób uczących (budowa modelu) i prób testowych (predykcja nowych obserwacji). Jakość predykcji jest najczęściej uzyskiwana na podstawie różnorodnych metod tzw. walidacji krzyżowej (wieloraki podział prób na testowe i uczące). Modele predykcyjne powinny cechować się wysokim poziomem konfirmacji i posiadać wysoką trafność przewidywania (moc predykcyjną), a jednocześnie ich trafność wyjaśniania zjawisk może być mniejsza (niska moc eksplanacyjna). Moc i trafność predykcyjna jest oceniana na podstawie miar zmienności losowej oszacowań parametrów z kolejnych prób losowanych w procedurach walidacyjnych (*sampling variance*). Należy podkreślić, że modele charakteryzujące się dużym błędem specyfikacji (niską mocą eksplanacyjną) mogą posiadać równocześnie dużą moc predykcyjną (*bias – variance trade off*). Jest to sytuacja częsta w marketingu relacji, analitycznym CRM i modelach *data mining*, w których wykorzystywana jest zwykle duża liczba skorelowanych (współliniowych) predyktorów o dużym zakresie błędów pomiarowych¹.

¹ Jest to szczególnie widoczne w instrumentalistycznej koncepcji modeli ekonomicznych (ekonomia pozytywna M. Friedmana). Krytycy tego podejścia (np. przedstawiciele ekonomii neoaustriackiej) podkreślają jałowość modeli predykcyjnych, niepozwalających na zrozumienie natury zjawisk i mechanizmów zależności przyczynowych. Ilustracją kompromisu *bias-variance* są znane modele

Celem artykułu jest ocena porównawcza modeli SEM i PLS-PM z punktu widzenia ich własności predykcyjnych. Jest ona dokonana na podstawie wskaźników determinacji modelu (R^2), redundancji i wariancji oszacowań parametrów (przedziałów ufności). W artykule wskazano na problemy w interpretacji współczynników R^2 jako miary mocy predykcyjnej modelu.

1. Modele PLS-PM a modele SEM

Szczególne miejsce w wyjaśniającej i predykcyjnej funkcji modelu zajmują modele równań strukturalnych. Z jednej strony klasyczne modele równań strukturalnych ze zmiennymi ukrytymi (SEM) kładą szczególny nacisk na aspekt wyjaśniania przyczynowo-skutkowego zjawisk (z tego powodu niekiedy nazywane są błędnie „modelami przyczynowymi”). Z drugiej strony inna wersja tych modeli, związana z tradycją PLS-PM, podkreśla predykcyjny charakter modeli budowanych w tym nurcie. W tabeli 1 przedstawiono podstawowe różnice między podejściami do modelowania.

Tabela 1. Porównanie podejść SEM i PLS-PM

Kryterium	SEM	PLS-PM
1	2	3
Cel modelowania	Eksplanacyjno-falsyfikacyjny	Eksploracyjno-predykcyjny
Założenia metodologiczne	Podejście realistyczne, korespondencyjna teoria prawdy naukowej	Podejście konstrukcjonistyczne, koherencyjna teoria prawdy naukowej
Dane wejściowe	Macierz kowariancji lub korelacji	Dane surowe, centrowane lub standaryzowane
Podstawa wyjaśniania relacji między zmiennymi ukrytymi	Kowariancje	Korelacje kanoniczne
Podejście do estymacji parametrów	Pełna informacja (<i>full information</i>)	Ograniczona informacja (<i>limited information</i>)
Kryterium dopasowania	Minimalizacja reszt między wejściową a odtworzoną macierzą kowariancji	Maksymalizacja wyjaśnionej wariancji zmiennych zależnych

Układu Słonecznego. Model Ptolemeusza cechował się (na owe czasy) dobrą mocą predykcyjną w porównaniu do modelu Kopernika, lecz posiadał o wiele gorszą moc eksplanacyjną (przedstawiał fałszywy obraz rzeczywistości).

1	2	3
Wybrane miary dopasowania modelu	χ^2 , GFI, RMSEA, SRMR, RMSR, CFI, TLI	Brak wskaźników ogólnego dopasowania. GoF, R ² , redundancja, współczynnik Stone-Geissera Q ²
Ocena modelu pomiarowego	Klasyczna teoria testu, teoria reakcji na pozycję	Klasyczna teoria testu, C-OAR-SE
Zmienne ukryte	Szacowane na podstawie parametrów modelu	Liniowa kombinacja wskaźników korygowana iteracyjnie wagami maksymalizującymi wyjaśnianą wariancję zmiennych zależnych
Dominujący typ wskaźników	Refleksywne	Formatywne (i refleksywne)
Charakter parametrów	Nieobciążone przy spełnionych założeniach (normalności rozkładu, liniowości relacji)	Obciążone (spójne dla dużych prób – <i>consistency at large</i>)
Liczebność próby	Względnie duże (powyżej 200 obserwacji)	Względnie niewielkie (powyżej 30 obserwacji)

Źródło: A. Sagan, *Zmienne ukryte w badaniach marketingowych*, UEK Kraków 2014.

Z tabeli 1 wynika, że podejścia te charakteryzują się odrębnymi własnościami statystycznymi, celem zastosowań i charakterem wyników. To, co szczególnie je odróżnia w odniesieniu do kontekstu wyjaśniania i predykcji zjawisk, to: a) sposób definiowania zmiennych ukrytych, b) metoda estymacji parametrów i c) ocena dopasowania modelu. Sposób definiowania zmiennych ukrytych jest kluczowym czynnikiem różnicującym oba podejścia. W metodologii badań zmienna jest zmienną ukrytą, jeżeli równanie w modelu pomiarowym tej zmiennej nie może być tak przekształcone, że zmienna ta jest wyrażona jako funkcja wyłącznie zmiennych obserwowalnych. Można wyróżnić trzy sposoby ich definiowania: 1) jako wariancję wspólną zestawu wskaźników refleksywnych (*common factor model*), 2) jako liniową kombinację wskaźników formatywnych (*composite latent variable*), 3) jako liniową kombinację wskaźników formatywnych z uwzględnieniem ukrytych zakłóceń (*formative latent variable*). W przypadku modeli SEM przyjmowana jest pierwsza definicja, a modelem pomiarowym zmiennych ukrytych jest konfirmacyjna analiza czynnikowa. Modele PLS-PM przyjmują drugie rozumienie zmiennej ukrytej i model pomiarowy stanowi analiza regresji z iteracyjnie szacowanymi wagami. Ostatni typ zmiennych ukrytych jest niezwykle trudny do oszacowania, wiąże się z to z dekompozycją czynnika resztowego ze względu na błąd pomiaru tej zmiennej (*error*) oraz część niewyjaśnionej zmienności przez zestaw zmiennych niezależnych

lub pominiętych w modelu (*disturbance*). Zakłócenia mogą reprezentować nieznaną wymiar danej zmiennej ukrytej, całkowity błąd pomiaru dokonywany za pomocą zestawu wskaźników formatywnych lub trafność wskaźników formatywnych². A. Diamantopoulos podkreśla, że nie należy utożsamiać zakłóceń w modelu z błędami pomiaru i tym samym z rzetelnością wskaźników formatywnych³.

Druga kluczowa różnica między modelami wynika z metody estymacji. W modelach PLS-PM zmienne ukryte są najczęściej związane z metodą estymacji częściowych najmniejszych kwadratów (*partial least squares*) opracowanej odrębnie przez H. Wolda oraz J.-B. Lohmollera i K. Joreskoga. H. Wold jest związany z podejściem iteracyjnym najmniejszych kwadratów (*partial least squares regression*). W drugim podejściu (PCA-OLS) zmienne ukryte są tożsame z pierwszą, najważniejszą składową główną wyodrębnioną na podstawie bloku wskaźników danego konstruktów. Zmienne ukryte są traktowane jako kombinacje liniowe wskaźników z iteracyjnie określaną strukturą wag minimalizujących wariancję resztową dla ukrytych i obserwowalnych zmiennych endogenicznych⁴. Proces estymacji jest dwuetapowy: po obliczeniu wartości zmiennych ukrytych oblicza się współczynniki regresji wielorakiej (w przypadku wskaźników formatywnych) lub serii regresji liniowych prostych (w przypadku wskaźników refleksywnych) między wskaźnikami a oszacowaną zmienną ukrytą. W efekcie każda zmienna ukryta jest określana ze względu na minimalizację wariancji reszt łącznie w modelu pomiarowym i strukturalnym.

W modelach SEM estymacja zmiennych ukrytych jest dokonywana na podstawie macierzy kowariancji i ma charakter estymacji jednoczesnej z wykorzystaniem metody największej wiarygodności.

Różnice w definiowaniu zmiennych ukrytych i metodach estymacji powodują także inne podejście do problemu identyfikacji modelu. W modelach SEM identyfikacja modelu pomiarowego wynika z reguły trzech wskaźników dla nieskorelowanych zmiennych ukrytych (każda zmienna ukryta musi być powiązana z co najmniej trzema jej wskaźnikami).

² A. Diamantopoulos, P. Riefler, K.P. Roth, *Advancing Formative Measurement Models*, „Journal of Business Research” 2008, No. 61 (12), s. 1203–1218.

³ A. Diamantopoulos, *The Error Term in Formative Measurement Models: Interpretations and Modelling Implications*, „Journal of Modelling in Management” 2006, No. 1 (1), s. 7–17.

⁴ M. Tenenhaus, V.E. Vinzi, Y.-M. Chatelin, C. Lauro, *PLS Path Modeling*, „Computational Statistics & Data Analysis” 2005, No. 48, s. 159–205.

W modelach PLS-PM identyfikacja modelu jest dokonywana poprzez włączenie tych zmiennych w szerszy układ analityczny i jest nieodłącznie związana z powiązaniem danej zmiennej ukrytej z innymi czynnikami lub wskaźnikami. Model ten jest identyfikowalny, jeżeli ze zmiennej ukrytej są wyprowadzone co najmniej dwie ścieżki do innych nieskorelowanych zmiennych ukrytych (ze wskaźnikami refleksywnymi), jest ona powiązana z dwoma innymi własnymi wskaźnikami refleksywnymi (tzw. *model multiple indicators – multiple causes*, MIMIC) lub zawiera jedną ścieżkę ze zmienną ukrytą ze wskaźnikami refleksywnymi i jedną z własnym wskaźnikiem refleksywnym⁵.

A. Diamantopoulos i H.M. Winklhofer zauważają, że ten rodzaj wskaźników formatywnych jest dominujący w badaniach społeczno-ekonomicznych i dlatego większość błędów w pomiarze i nietrafności modeli wynika z: 1) braku poprawnej specyfikacji modelu pomiarowego, 2) stosowania wskaźników refleksywnych w sytuacji, gdzie bardziej poprawnym modelem pomiarowym jest model ze wskaźnikami formatywnymi, 3) problemów związanych z efektem maskowania i stosowaniem (błędnie) wskaźników formatywnych, w sytuacji niskiej rzetelności skal lub wskaźników refleksywnych w celu ukrycia problemów z współliniowością zmiennych⁶.

Zmienne ukryte ze wskaźnikami formatywnymi mogą być traktowane w modelu w dwojaki sposób – zarówno jako zmienne egzogeniczne, jak i endogeniczne. O ile ich rola jako predyktorów w modelu nie jest kwestionowana, to występowanie endogenicznych formatywnych zmiennych ukrytych wzbudza pewne wątpliwości. Jest to związane z brakiem dostatecznego wyjaśnienia mechanizmu, poprzez który zestaw egzogenicznych zmiennych ukrytych wyjaśnia formatywne zmienne endogeniczne, których wariancja jest w pełni tłumaczona przez zestaw wskaźników formatywnych i ukryty czynnik zakłóceń⁷. Do momentu, w którym badacz nie określi całkowitej „populacji” wskaźników formatywnych i tym samym nie przekształci for-

⁵ D. Temme, L. Hildebrandt, *Formative Measurement Models in Covariance Structure Analysis. Specification and Identification*, „SFB 649 Discussion Paper”, Humboldt-Universität zu Berlin, 2006-083, s. 1–18.

⁶ A. Diamantopoulos, H.M. Winklhofer, *Index Construction with Formative Indicators: An Alternative to Scale Development*, „Journal of Marketing Research” 2001, No. 38 (2), s. 269–277.

⁷ A. Diamantopoulos, P. Riefler, K.P. Roth, *Advancing Formative...*

matywnej zmiennej ukrytej w składową (bez błędu związanego z zakłóceniami), nie dysponuje wiedzą dotyczącą źródeł wariancji zależnych zmiennych formatywnych⁸.

Ocena dopasowania modeli SEM i PLS-PM opiera się na odrębnych założeniach. W pierwszych występują globalne miary dopasowania modelu do danych związane z jakością odwzorowania populacyjnej macierzy kowariancji przez macierz kowariancji odtworzoną przez parametry modelu. Poza statystyką χ^2 do najczęściej stosowanych miar opisowych należy średniokwadratowy pierwiastek błędu aproksymacji (RMSEA), wskaźnik dobroci dopasowania (GFI) czy dostosowany wskaźnik dobroci dopasowania (AGFI). Występuje także pewna liczba wskaźników przyrostowych porównujących dany model do modelu odniesienia (najczęściej najlepszego), taki jak wskaźnik przyrostowy dopasowania (IFI), wskaźnik porównawczy (CFI) czy wskaźnik Tuckera-Lewisa (TLI). Wskaźniki te testują moc eksplanacyjną modelu SEM, wskazując, jak dobrze zależności występujące w macierzy kowariancji są oddawane przez model.

W przypadku PLS-PM nie istnieją miary ogólnego dopasowania. Ocena modelu jest dokonywana na podstawie miar redundancji, która stanowi iloczyn zasobów zmienności wspólnej (C_j) oraz współczynników determinacji R^2 .

$$R_j = C_j \times R^2(y_j) \quad (1)$$

Budowany na tej podstawie tzw. globalny indeks dopasowania (GoF) jest średnią geometryczną z przeciętnej wyjaśnionej wariancji C_j i średniej z współczynników determinacji R^2 .

$$GoF = \sqrt{\bar{C} \times \bar{R}^2} \quad (2)$$

W świetle powyższych miar dopasowania należy podkreślić, że współczynnik determinacji R^2 jest w istocie miarą pozwalającą na ocenę mocy eksplanacyjnej modelu. Traktowanie go jako miary jakości predykcji jest pewnym nadużyciem (R^2 określa zakres wyjaśnianej, a nie przewidywanej wariancji).

⁸ J.W. Cadogan, A.L. Souchon, D.B. Procter, *The Quality of Market-Oriented Behaviors: Formative Index Construction*, „Journal of Business Research” 2008, No. 61 (12), s. 1263–1277.

2. Predykcja w modelach PLS-PM

Jedną z podstawowych zalet stosowania modeli PLS-PM w badaniach zjawisk ekonomicznych jest stosowanie wskaźników formatywnych i predykcijny charakter modelu. Możliwość wprowadzania wskaźników formatywnych (szczególnie jako zmiennych zależnych) jest obciążona dużym błędem specyfikacji i obciążeniem źródeł wyjaśnianej wariancji związanej z brakiem możliwości jej dekompozycji ze względu na wpływ wskaźników formatywnych, pominiętych zmiennych (zakłóceń) i egzogenicznych zmiennych ukrytych (predyktorów). Również walor lepszej predykcji modelu jest kwestionowany. Po pierwsze, sposób konstruowania wag w procesie szacowania zmiennych ukrytych jest funkcją maksymalizacji wyjaśnianej wariancji, po drugie, kryterium jakości predykcji, jakim jest miara determinacji R^2 , w istocie również wskazuje na moc eksplanacyjną modelu PLS-PM.

W celu porównania własności predykcyjnych obu podejść została dokonana analiza porównawcza modeli satysfakcji konsumentów zbudowanych na próbie 250 hiszpańskich klientów instytucji kredytowych⁹. Zmienne ukryte dotyczące wizerunku (IMAG), oczekiwań (EXPE), postrzeganej jakości (QUAL), korzyści (VAL), satysfakcji z oferty (SAT) i lojalności (LOY) były mierzone za pomocą 10-punktowych skal Likerta. Model SEM (estymacja największej wiarygodności w programie Mplus) został zbudowany na podstawie wskaźników refleksywnych, natomiast w modelu PLS-PM (biblioteka *plspm* programu R) przyjęto formatywny charakter wskaźników (*mode B*). Wagi wewnętrzne zostały oszacowane metodą ścieżkową (*path weighting*). W celu porównania zmienności rozkładu parametrów i wskaźników dopasowania w obu metodach zastosowano 1000-krotną próbę bootstrapową. Struktura i parametry obu modeli w części strukturalnej są przedstawione w tabelach 2 i 3.

⁹ Dane zostały pobrane z repozytorium danych biblioteki *plspm* w programie R.

Tabela 2. Ładunki czynnikowe modeli pomiarowych

Metoda		SEM		PLS-PM	
Zmienne ukryte	Zmienne obserwowalne	Oszacowania	Przedział ufności	Oszacowania	Przedział ufności
LOY	L1	0.51	0.39-0.63	0.93	0.83-0.98
	L2	0.35	0.25-0.46	0.58	0.37-0.73
	L3	0.51	0.38-0.63	0.92	0.84-0.97
	L4	0.34	0.28-0.43	0.50	0.29-0.72
SAT	S1	0.23	0.07-0.36	0.96	0.92-0.98
	S2	0.22	0.07-0.35	0.92	0.86-0.96
	S3	0.50	0.05-0.29	0.74	0.60-0.82
	S4	0.19	0.05-0.30	0.79	0.67-0.88
VAL	W1	0.33	0.15-0.41	0.91	0.86-0.95
	W2	0.29	0.11-0.37	0.77	0.62-0.88
	W3	0.24	0.12-0.31	0.66	0.51-0.77
	W4	0.30	0.12-0.38	0.85	0.77-0.91
QUAL	Q1	0.09	0.05-0.12	0.78	0.63-0.88
	Q2	0.11	0.06-0.14	0.90	0.84-0.94
	Q3	0.09	0.05-0.11	0.74	0.61-0.84
	Q4	0.09	0.05-0.13	0.81	0.70-0.90
	Q5	0.10	0.05-0.15	0.81	0.72-0.88
EXPE	O1	0.51	0.43-0.60	0.71	0.53-0.83
	O2	0.59	0.52-0.70	0.88	0.77-0.93
	O3	0.48	0.40-0.58	0.70	0.54-0.80
	O4	0.51	0.45-0.60	0.78	0.65-0.87
	O5	0.56	0.48-0.67	0.80	0.70-0.87

Źródło: opracowanie własne na podstawie programu Mplus i *plspm*.

Tabela 2 przedstawia wartości ładunków czynnikowych dla obu metod w modelach pomiarowych (zewnętrznych). Uwzględniając inny typ skalowania zmiennych ukrytych w obu metodach, należy zwrócić uwagę na bootstrapowe przedziały ufności dla parametrów. Modele SEM zawierają szersze przedziały ufności w porównaniu do PLS-PM. Pozwala to wnioskować o większej wariancji oszacowań w klasycznych modelach strukturalnych i tym samym mniejszym prawdopodobieństwie popełnienia błędu I rodzaju (odrzućenia hipotezy o zerowej wartości ładunku, w sytuacji gdy hipoteza ta jest prawdziwa). Mniejsza wariancja oszacowań i tym samym większa precyzja estymacji parametrów w PLS-PM potwierdza predykcyjny charakter metody.

Tabela 3. Parametry strukturalne i współczynniki dopasowania

Typ parametru	Zmienne i relacje	SEM		PLS-PM		
		Oszacowania	Przedział ufności	Oszacowania		Przedział ufności
Ścieżkowe	IMAG-> EXPE	0.90*	0.63-1.13	0.60*		0.51-0.70
	IMAG->LOY	0.35*	0.11-0.60	0.23*		0.10-0.39
	IMAG->SAT	0.65*	0.15-2.50	0.20*		0.10-0.32
	EXPE-> QUAL	5.88*	2.09-7.67	0.85*		0.81-0.89
	EXPE->VAL	-3.22	-6.87- 2.00	0.15*		0.02-0.30
	EXPE->SAT	-2.73	-12.37-7.18	0.00		-0.12-0.13
	QUAL->VAL	0.82	-0.09-1.92	0.65*		0.49-0.78
	QUAL->SAT	0.38	-3.04-2.54	0.08		-0.08-0.26
	VAL->SAT	1.47*	0.70- 6.44	0.62*		0.47-0.76
	SAT->LOY	0.28*	0.09-0.46	0.56*		0.42-0.69
R2 i redundancja	EXPE	0.45	-	0.37	0.22	0.26-0.49
	QUAL	0.98	-	0.72	0.48	0.65-0.79
	VAL	0.84	-	0.61	0.39	0.50-0.72
	SAT	0.93	-	0.73	0.54	0.68-0.80
	LOY	0.65	-	0.55	0.31	0.46-0.66
Ogólny wskaźnik dopasowania (GFI/GoF)		0,756		0.609		

Źródło: opracowanie własne na podstawie programu Mplus, Statistica i *plspm*.

Tabela 3 zawiera parametry ścieżkowe dla obu metod oraz wartości współczynników determinacji i ogólną ocenę dopasowania modelu. Podobnie jak w przypadku modelu zewnętrznego, również w części strukturalnej modelu (wewnętrznym) obserwuje się węższe przedziały ufności dla metody PLS-PM, a szersze dla SEM. Istotne parametry w modelu PLS-PM (relacja między korzyściami a oczekiwaniami i postrzeganą jakością) okazały się nieistotne w modelu SEM. Biorąc pod uwagę wartości współczynnika determinacji, wyższe, w porównaniu do modelu PLS-PM, poziomy wyjaśnianej wariancji mają zmienne ukryte dla modeli SEM. Świadczy to o większej mocy eksplanacyjnej SEM w porównaniu do PLS-PM.

Analiza porównawcza z wykorzystaniem metody bootstrap potwierdziła założenia wynikające z efektu kompromisu między wariancją oszacowań parametrów ścieżkowych z prób (*sampling variance*) i obciążeniem parametrów ścieżkowych (*bias*). Zaletą PLS-PM jest mniejsza wariancja oszacowań i tym samym większa precyzja estymacji parametrów oraz wyższe prawdopodobieństwo uzyskania istotnych współczynników ścieżkowych (w tym przejawia się moc predykcyjna mode-

li PLS-PM). Z drugiej strony modele SEM cechują się wyższą mocą eksplanacyjną (mierzoną współczynnikiem determinacji R^2) i dostarczają mniej obciążonych oszacowań¹⁰.

Podsumowanie

W marketingu i naukach o zarządzaniu podejście PLS-PM ma szerokie zastosowania w modelowaniu zmiennych ukrytych. Jego popularność wynika zarówno z predykcyjnego celu analiz, jak i względnie uproszczonej interpretacji parametrów i wskaźników dopasowania, braku konieczności spełnienia założeń normalności rozkładu, wykorzystywania dwuetapowej procedury estymacji mniej wrażliwej na nietypowe rozwiązania oraz możliwości stosowania modelu dla prób o niewielkiej liczebności. Walory tego podejścia, związane z wyższą precyzją oszacowania parametrów, kontrastują z problemami związanymi z fałszywie dodatnimi rozstrzygnięciami w zakresie istotności związków oraz niższą mocą eksplanacyjną, która jest jednym z kluczowych kryteriów akceptacji modelu. Pytanie o wybór podejścia do analizy zmiennych ukrytych w naukach ekonomicznych (PLS-PM czy SEM) pozostaje więc nadal otwarte.

Literatura

- Cadogan J.W., Souchon A.L., Procter D.B., *The Quality of Market-Oriented Behaviors: Formative Index Construction*, „Journal of Business Research” 2008, No. 61 (12), s. 1263–1277.
- Diamantopoulos A., *The Error Term in Formative Measurement Models: Interpretations and Modelling Implications*, „Journal of Modelling in Management” 2006, No. 1 (1), s. 7–17.
- Diamantopoulos A., Riefler P., Roth K.P., *Advancing Formative Measurement Models*, „Journal of Business Research” 2008, No. 61 (12), s. 1203–1218.
- Diamantopoulos A., Winklhofer H.M., *Index Construction with Formative Indicators: An Alternative to Scale Development*, „Journal of Marketing Research” 2001, No. 38 (2), s. 269–277.
- Sagan A., *Zmienne ukryte w badaniach marketingowych*, UEK, Kraków 2014.

¹⁰ Analiza potwierdza, że kryterium mocy predykcyjnej modelu nie jest wysoka wartość współczynnika R^2 , który odzwierciedla zakres wyjaśnianej wariancji, lecz niska wariancja oszacowań parametrów ścieżkowych w replikacjach (*sampling variance*).

Temme D., Hildebrandt L., *Formative Measurement Models in Covariance Structure Analysis. Specification and Identification*, „SFB 649 Discussion Paper”, Humboldt-Universität zu Berlin, 2006-083, s. 1–18.

Tenenhaus M., Vinzi V.E., Chatelin Y.-M., Lauro C., *PLS Path Modeling*, „Computational Statistics & Data Analysis” 2005, No. 48, s. 159–205.

PLS-PM MODEL AND ITS APPLICATION IN EXPLANATION AND PREDICTION OF ECONOMIC PHENOMENA

Abstract

In the paper, the problems of application of PLS-PM/SEM models in explanation and prediction of economic phenomena are outlined. The basic differences between two approaches are explicated with respect to the indicators formation, estimation methods and goodness of fit measures. Special attention is paid to predictive aspect of PLS-PM model that enables to obtaining the higher precision of estimates and higher validity of predicted phenomena.

Translated by Adam Sagan

Keywords: PLS-PM models, formative indicators, predictive power

Kod JEL: C1, M3