

Adam Sagan

Uniwersytet Ekonomiczny w Krakowie

PREZENTACJA GRAFICZNA MODELI RÓWNAŃ STRUKTURALNYCH


1. Wstęp


Analizy wykorzystujące zmienne ukryte są jednymi z najbardziej popularnych metod analizy i prezentacji danych w przyczynowych interpretacjach relacji między zmiennymi w nieeksperymentalnych podejściach badawczych w naukach społecznych. Szczególną popularność zdobyły w badaniach marketingowych. W tym obszarze zastosowań znaczenie ma zarówno dążenie do trafnego, ale jednocześnie uproszczonego wyjaśnienia badanych zjawisk, jak i odpowiednia graficzna prezentacja uzyskanych wyników. Jak podają B.J. Babin, J.F. Hair Jr i J.S. Boles [2008], artykuły w czasopiśmie marketingowych bez zastosowanych podejść z rodziny modelowania strukturalnego ze zmiennymi ukrytymi (SEM) mają większą szansę odrzucenia przez recenzentów już we wstępnej recenzji (autorzy podają, że 70% odrzuceń dotyczy prac z zastosowanym podejściem SEM w porównaniu z innymi metodami, dla których występuje aż 84% odrzuceń), a zakwalifikowane wstępnie artykuły z SEM mają wyższą średnią ocenę (ocena 3,5) niż te wykorzystujące inne metody analizy (ocena 2,7).


Szczególnie istotne znaczenie w modelach ze zmiennymi ukrytymi ma wizualizacja graficzna wyników. W metodach eksploracyjnych, takich jak analiza głównych składowych, analiza głównych współrzędnych, analiza korespondencji i analiza czynnikowa, wizualizacja jest nierozdzielnie związana z procesem analizy i stanowi podstawę interpretacji wyników uzyskanych w procesie rzutowania w zredukowanej przestrzeni o mniejszej liczbie wymiarów. W metodach confirmacyjnych, takich jak confirmacyjna analiza czynnikowa i ogólne modele strukturalne, graficzna prezentacja wyników oparta jest wyłącznie na przyjętej przez badaczy konwencji wynikającej najczęściej z wykorzystywanego oprogramowania komputerowego. Obecnie większość programów komputerowych służących do budowy modeli strukturalnych (jak np. LISREL, EQS, AMOS, Mx) zawiera odpowiednie graficzne narzędzia do podstawowej prezentacji wyników modelu. W dziedzinie tej panuje umiarkowany


consensus między badaczami dotyczący tego, w jaki sposób prezentować poszczególne kategorie występujące w modelu.


Poniżej przedstawione są podstawowe sposoby prezentacji elementów występujących w modelu strukturalnym stanowiące swoisty „język” komunikowania wyników modelowania.


 – zmienna obserwowalna (*manifest variable*) będąca wskaźnikiem zmiennej ukrytej w modelu pomiarowym confirmacyjnej analizy czynnikowej lub egzogeniczna zmienna obserwowalna (zmienna kowariancyjna).


 – zmienna ukryta (*latent variable*) odzwierciedlająca mierzony konstrukt teoretyczny, zmienna nieobserwowalna reprezentująca ukrytą klasę obserwacji (przypadków) lub zmienna ukryta odzwierciedlająca zmienność wyrazów wolnych lub współczynników kierunkowych w modelach krzywych rozwojowych (*latent growth curve*).


 – sztuczna zmienna ukryta fantomowa (*phantom variable*) pozwalająca na wyrażanie relacji między zmiennymi stanowiącymi kombinacje współczynników ścieżkowych (np. współczynnik ścieżkowy stanowiący sumę dwóch współczynników dla innych ścieżek) [Loehlin 2004].


 – ukryta zmienna resztowa (*residual, disturbance*) w modelach pomiarowych lub ukryta zmienna resztowa reprezentująca zakłócenia (część niewyjaśnionej zmienności zależnej zmiennej ukrytej).


 – stała (*constants*) lub średnia w modelu.

 – zmienna definicyjna (*definitional variable*) w modelach wielopoziomowych ustalająca parametry modelu na poziomie indywidualnym.

 – nieobserwowalna heterogeniczność (*unobserved heterogeneity*) w modelach wielopoziomowych odzwierciedlająca losowe wyrazy wolne lub losowe współczynniki kierunkowe w przekroju zespołów losowania w doborze grupowym lub analizie.

 – ścieżka (*path*) określająca wpływ obserwowalnych lub ukrytych zmiennych niezależnych na obserwowalne lub ukryte zmienne zależne.

 – kowariancja (*covariance*) lub korelacja między zmiennymi.

 – relacje między zmiennymi zachodzące w przekrojach klas ukrytych (*broken path*).

Przedstawione symbole stanowią umowny i wygodny sposób prezentacji wyników analizy i znajdują zastosowanie w odpowiednich pakietach statystycznych służących do modelowania strukturalnego.

2. Rozwiązania graficzne w programach komputerowych

Pierwszym pakietem statystycznym służącym do modelowania strukturalnego był program LISREL, który został zaprezentowany na konferencji „Structural Equation Models in the Social Sciences” w 1970 r. przez K. Joreskoga. W 1971 r. pojawiła się wersja LISREL II w postaci – jak to wspomina jego pierwszy doktorant

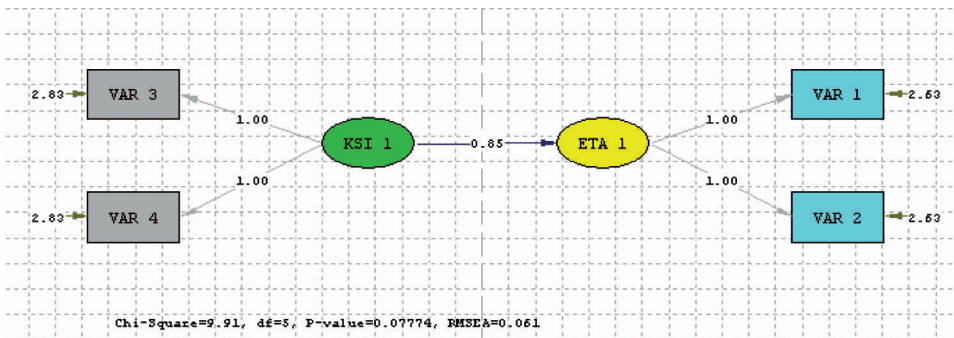
D. Sorbom – „pudełka kart perforowanych” zapisanych językiem FORTRAN [Sorbom 2001, s. 6]. Pierwsza komercyjna wersja z podręcznikiem użytkownika LISREL III została wprowadzona na rynek w 1974 r. Kolejne wersje to LISREL IV (1978) uwzględniająca słowa kluczowe, dane w postaci swobodnej, dynamiczną alokację pamięci, LISREL V (1981) mający automatyczne punkty startowe, ocenę efektów całkowitych i metodę ważonych i uogólnionych najmniejszych kwadratów, LISREL VI (1984) zawierający wykres parametrów modelu, indeksy modyfikacyjne i automatyczną modyfikację modeli, LISREL VII (1988) obejmujący moduł PRELIS, ważne najmniejsze kwadraty i rozwiązanie kompletnie standaryzowane. Przełomem w rozwoju programu okazała się wersja LISREL VIII (1993) wyposażona w SIMPLIS, diagramy ścieżkowe i nieliniowe ograniczenia modelu. Kolejne odmiany tej wersji umożliwiają analizę wielopoziomową, krzywe rozwojowe, wieloraką imputację danych, modelowanie sprzężeń zwrotnych i zmiennych kategoryalnych. Z historii rozwoju LISREL wynika, że graficzne formy wizualizacji wyników wkomponowane w programy statystyczne zaczęły się rozwijać dopiero od połowy lat 90.

Współcześnie do najbardziej popularnych programów wizualizacji modeli strukturalnych należą programy AMOS, LISREL, EQS i Mx. Są to najpowszechniejsze komercyjne programy SEM umożliwiające zarówno budowę modelu, jak i zintegrowaną z analizą graficzną prezentację wyników modelowania.

a/ LISREL

Zaletą programu opracowanego przez K. Joreskoga i D. Sorboma jest możliwość zarówno graficznej budowy, jak i automatycznej prezentacji standaryzowanych i niestandaryzowanych modeli strukturalnych. Reprezentacja zmiennych (w postaci owali i prostokątów) jest w pełni skalowalna, co pozwala na automatyczne dostosowanie rozmiaru narzędzi graficznych do wielkości tekstu w nazwach zmiennych. Dodatkowo program umożliwia różnicowanie funkcji zmiennych w modelu, odróżniając kolorami ukryte i obserwowalne zmienne niezależne od ukrytych i obserwowalnych zmiennych niezależnych. Na rysunku 1 niezależna zmienna ukryta jest oznaczona owalem w kolorze zielonym, a zmienna zależna – owalem w kolorze żółtym. Wskaźniki niezależnej zmiennej ukrytej są przedstawione w postaci szarych prostokątów, a wskaźniki zależnej zmiennej ukrytej – prostokątów turkusowych. Wariancje błędów pomiaru znajdują się przy strzałkach dochodzących do zmiennych obserwowalnych, wartości ładunków czynnikowych leżą na strzałkach wiążących relacje między zmiennymi ukrytymi a wskaźnikami, parametr ścieżkowy mieści się na strzałce wskazującej na relację między zmiennymi ukrytymi.

Domyślnie wydruk graficzny jest uzupełniony o podstawowe statystyki dopasowania modelu. Wadą programu jest ograniczenie nazw zmiennych jedynie do 8 znaków.

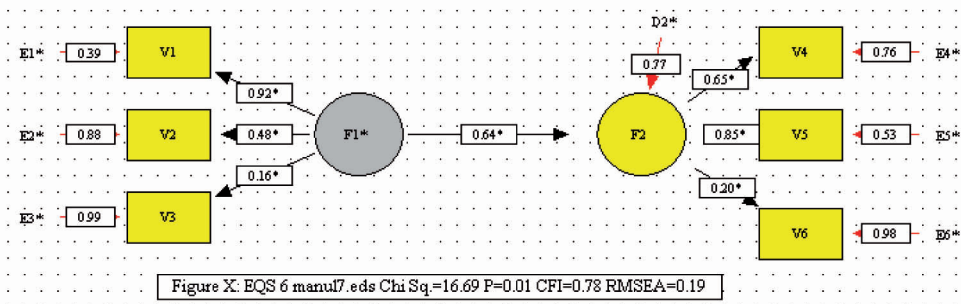


Rys. 1. Graficzna prezentacja modelu strukturalnego w programie LISREL

Źródło: opracowanie własne.

b/EQS

Program stworzony przez P. Bentlera pozwala na graficzną budowę i prezentację wyników modelowania. Dodatkowo program zawiera kreator graficznej budowy modeli ścieżkowych, pomiarowych (konfirmacyjnej analizy czynnikowej) i modeli krzywych rozwojowych. W zaletach programu należy wyróżnić automatyczne skalowanie grafiki, różnicowanie kolorami funkcji zmiennych i ścieżek oraz możliwość pełnego opisu zmiennych obserwowalnych i ukrytych. Oprócz statystyk podstawowych dopasowania modelu wyświetla domyślnie także wskaźnik porównawczy CFI.



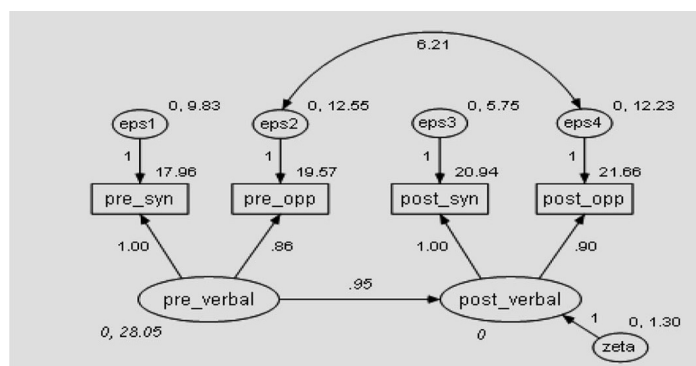
Rys. 2. Graficzna prezentacja modelu strukturalnego w programie EQS

Źródło: opracowanie własne.

Rysunek 2 prezentuje wydruk modelu strukturalnego w programie EQS. Forma graficzna jest bardzo podobna do prezentacji w pakiecie LISREL. Wartości oszacowania parametrów (ładunki czynnikowe, parametry ścieżkowe i wariancje błędów są podane w kwadratach na odpowiednich strzałkach.

c/AMOS

Program opracowany przez J. Arbuckle'a i zintegrowany z pakietem SPSS daje możliwość budowy i prezentacji graficznej modeli strukturalnych. Rozbudowane menu graficzne pozwala na dodawanie całych układów modeli pomiarowych i ich rotację, umożliwia pełny opis zmiennych. Do istotnych wad programu można zaliczyć brak automatycznego skalowania grafiki i różnicowania funkcji zmiennych i typów ścieżek. Domyślnie prezentowane są podstawowe statystyki dopasowania modelu.



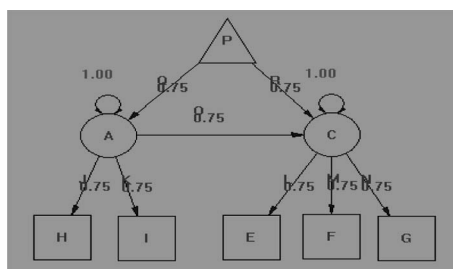
Rys. 3. Graficzna prezentacja modelu strukturalnego w programie AMOS

Źródło: opracowanie własne.

Na rysunku 3 zaprezentowany jest typowy wynik graficznej analizy w programie AMOS. Wszystkie oszacowania parametrów są podane obok symboli graficznych. Zmienne ukryte reprezentujące błędy pomiaru są również przedstawione w formie graficznej (owale).

d/Mx

Mx jest programem niekomercyjnym stworzonym przez M.C. Neala, pozwalającym na graficzną budowę i interpretację modelu. Ma on zastosowanie przede wszyst-



Rys. 4. Graficzna prezentacja modelu strukturalnego w programie Mx

Źródło: opracowanie własne.

kim w psychologii i genetyce. W programie można wprowadzać zmienne obserwowalne, ukryte oraz stałe do modelu, a także dokonywać ich pełnego opisu. Nie ma automatycznego skalowania grafiki oraz różnicowania funkcji zmiennych.

3. Wizualizacja złożonych modeli strukturalnych

Przedstawione formy graficzne prezentacji modeli w popularnych programach pozwalają na budowę i interpretację podstawowych i względnie prostych jednopoziomowych modeli obejmujących niewielką liczbę wskaźników i kilka zmiennych ukrytych budowanych w latach 80. Silny rozwój tego podejścia w latach 90. był związany z budową modeli z coraz to większą liczbą zmiennych ukrytych zarówno pierwszego, jak i drugiego rzędu, wieloma ścieżkami i korelacjami między zmiennymi ukrytymi i błędami ich pomiaru utrudniającymi interpretację modelu.

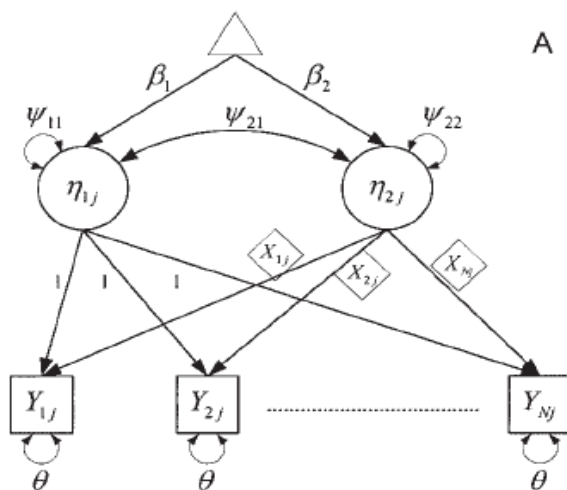
Równoległe w tym okresie rozwijano złożone modele obejmujące nie tylko dużą liczbę czynników i ich wskaźników, ale także integrujące wiele podejść dotychczas nieobecnych w modelach strukturalnych. Należą do nich modele wielopoziomowe (*multilevel SEM*), modele z ukrytymi trajektoriami i krzywymi rozwojowymi (*latent growth curves*), modele mieszanek integrujące klasyczne modelowanie strukturalne z modelami klas ukrytych (*mixture factors*) oraz modele różnicowe obejmujące zjawiska dynamiczne, samosprężenia zmiennych i różnice między zmiennymi ukrytymi (*latent difference models*). Modele te są niezwykle trudne w budowie przy wykorzystaniu tradycyjnych graficznych narzędzi budowy modelu.

A/Modele wielopoziomowe

Pierwszym problemem w prezentacji modeli strukturalnych okazał się problem ujęcia poziomów analizy w hierarchicznych modelach ze zmiennymi ukrytymi reprezentującymi losowość wyrazów wolnych i współczynników nachylenia w przekroju grup lub zespołów losowania. Modele wielopoziomowe mają zastosowanie w analizie danych charakteryzujących się „zagnieżdżeniem”. Przykłady tego typu analiz to np. ocena osiągnięć uczniów w szkołach w zależności od charakteru szkoły i kompetencji nauczycieli (uczniowie są „zagnieżdżeni” w klasach uczonych przez określonych nauczycieli, a nauczyciele są „zagnieżdżeni” w określonych szkołach) czy ocena satysfakcji pacjentów „zagnieżdżonych” w podstawowych ośrodkach opieki zdrowotnej.

W prezentacji tych modeli można wyróżnić dwie tradycje związane z twórcami i popularyzatorami tych podejść. Pierwsza jest reprezentowana przez twórców programów Mx i RAMONA, a druga przez B. Muthena, który zaproponował inne przedstawienie wielopoziomowych modeli estymowanych w programie Mplus. W tradycji Mx wykorzystywane są tzw. zmienne definicyjne.

Rysunek 5 przedstawia graficzny model wielopoziomowy równań strukturalnych z losowymi współczynnikami nachyleń (*random slopes*), w którym Y_{ij} stano-

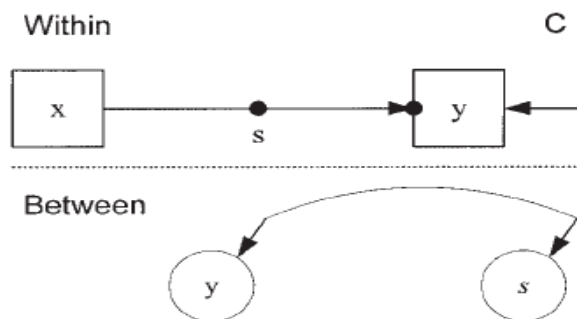


Rys. 5. Prezentacja modelu wielopoziomowego w programie Mx

Źródło: [Duncan i in. 2009].

wią indywidualne obserwacje dla zmiennych obserwowalnych Y respondenta i należącego do grupy j . Zmienna ukryta η_{1j} określa losowy wyraz wolny, a zmienna ukryta η_{2j} przedstawia międzygrupową zmienność współczynników kierunkowych określających wpływ zmiennej niezależnej X na zmienną Y . Zmienna X_{ij} jest traktowana jako „zmienna definicyjna” (\diamond) pozwalająca na ustalenie parametrów modelu (ładunków czynnikowych $\eta_{2j}-Y_{ij}$) na poziomie indywidualnym.

Alternatywny sposób prezentacji modeli wielopoziomowych został przyjęty przez B. Muthena, który zaproponował oddzielną prezentację modelu na poziomie indywidualnym i grupowym przedstawioną na rys. 6.



Rys. 6. Prezentacja modelu wielopoziomowego w programie Mplus

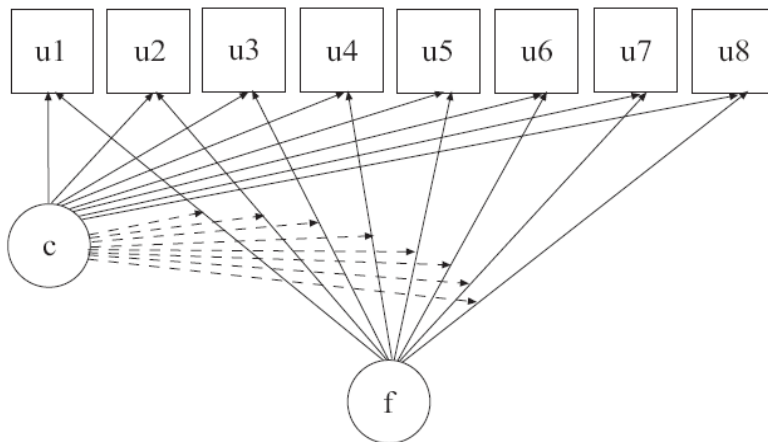
Źródło: [Duncan i in. 2009].

Na poziomie indywidualnym (*within*) zarówno wyrazy wolne zmiennej zależnej (kropka przy prostokącie reprezentującym obserwowalną zmienną zależną), jak i współczynniki kierunkowe (kropka na strzałce reprezentującej zależność regresyjną między X a Y) charakteryzują się zmiennością w przekroju grup. Na poziomie grupowym (*Between*) zmienne losowe wyrażające wyrazy wolne (γ) i nachylenia (s) są traktowane jako zmienne ukryte (w owalach).

B/Modele mieszane czynnikowe

Drugi etap w rozwoju modeli złożonych stanowiły modele mieszane integrujące klasyczne podejście konfirmacyjnej analizy czynnikowej z analizą klas ukrytych. Szczególne zastosowanie tych modeli jest w segmentacji rynku dokonywanej na podstawie kryteriów psychograficznych lub związanych z poszukiwanymi przez konsumenta korzyściami w produkcie.

Na rysunku 7 zaprezentowany jest mieszany model konfirmacyjnej analizy czynnikowej. Jest on uogólnieniem modelu klas ukrytych, w którym kategoriałne wskaźniki (u) danej klasy ukrytej są zależne od ciągłej zmiennej ukrytej (f). Odzwierciedla ona zmienność prawdopodobieństw zgody z pozycjami w ramach danej klasy. Przerwane strzałki wskazują, że ładunki czynnikowe zmieniają się w przekroju klas ukrytych (c), co wskazuje na swoiste modele pomiarowe dla danego kryterium segmentacji w obrębie każdego segmentu.

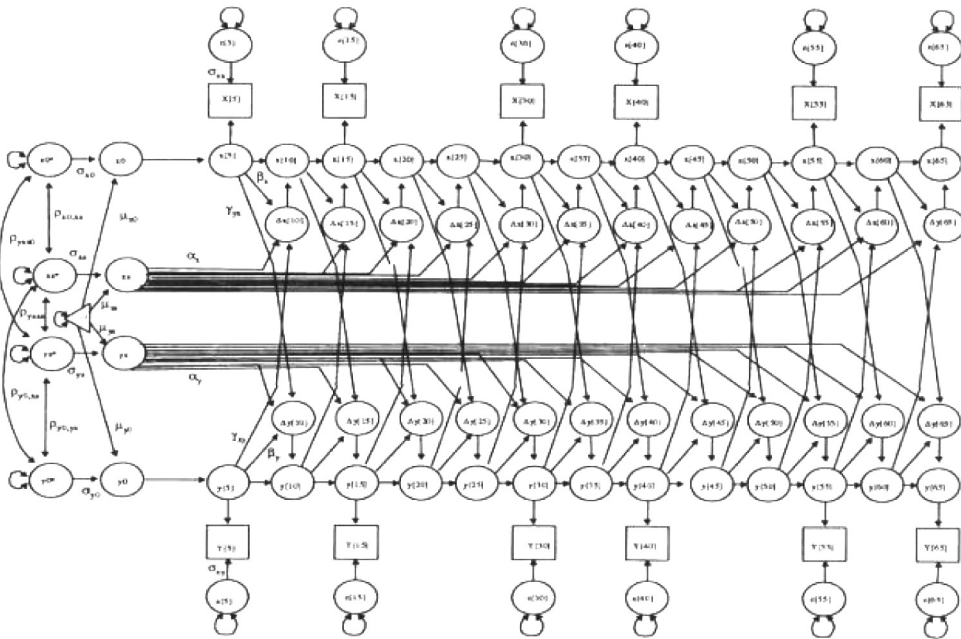


Rys. 7. Graficzna prezentacja mieszanego modelu czynnikowego

Źródło: [Muthen, Muthen 2007, s. 176].

C/Modele mieszane krzywych rozwojowych

Uwzględnienie zmian wartości zmiennych obserwowalnych w czasie pozwoliło na wykorzystanie modeli strukturalnych w analizie danych wzdłużnych. Zmienności wartości początkowych i współczynników kierunkowych linii trendu (liniowych lub kwadratowych) są przedstawiane jako zmienne ukryte.



Rys. 9. Model różnicowy krzywych rozwojowych

Źródło: [Mc Ardle, Hamagami 2004, s. 320].

miczne relacje wzajemne między dwiema krzywymi rozwojowymi z efektami zmiany i samosprzężeń. Występują tutaj dwie zmienne obserwowalne (x) i (y) oraz ich dualne zmiany reprezentowane jako zmienne ukryte (Δx) i (Δy).

4. Podsumowanie

Prezentacja graficzna modeli strukturalnych staje się coraz bardziej złożona. Wynika to nie tylko z dużej liczby zmiennych obserwowalnych i ukrytych, ale przede wszystkim ze zróżnicowania typów zmiennych obserwowalnych i ukrytych w analizie (kategorialne, metryczne, uwzględniające ukrytą losową heterogeniczność, fantomowe) oraz rodzajów zależności między nimi (kowariancyjne, ścieżkowe, interakcyjne). Powoduje to dużą trudność w odbiorze wyników modelowania nie tylko przez odbiorców niezwiązanych ze środowiskiem SEM (np. klientów firm badawczych, decydentów, pracowników działów marketingu), ale nawet przez specjalistów z tej dziedziny. Problem czytelności graficznych form prezentacji wyników jest jedną z najważniejszych barier w szerszym upowszechnianiu modelowania strukturalnego, szczególnie w komercyjnych badaniach marketingowych.

Literatura

- Babin B.J., Hair Jr J.F., Boles J.S., *Publishing Research in Marketing Journals using Structural Equation Modeling*, „Journal of Marketing Theory and Practice”, Fall 2008.
- Duncan T.E., Duncan S., Strycker C., Okut L.A., Li F., *Growth Mixture Modeling of Adolescent Alcohol Use Data*, <http://oregonstate.edu/~acock/growth-curves/mixture1-30-02.pdf>, 1.06.2009.
- Loehlin J.C., *Latent Variable Models*, Lawrence Erlbaum, Mahwah, N.J. 2004.
- Mc Ardle J., Hamagami F., *Longitudinal Structural Equation Modeling Methods for Dynamic Change Hypothesis*, [w:] *Recent Developments on Structural Equation Models*, red. K. van Monfort, J. Oud, A. Satorra, Kluwer Academic Pub., 2004.
- Muthen B., Muthen L., *Mplus User's Guide*, Muten&Muthen, Los Angeles 2007.
- Sorbom D., *Karl Joreskog and LISREL: a Personal Story*, [w:] *Structural Equation Modeling: Present and Future*, R. Cudeck, S. du Toit, D. Sorbom (red.), SSI Scientific Software International, Lincolnwood 2001.

GRAPHICAL VISUALIZATION OF STRUCTURAL EQUATION MODELS

Summary

The development of structural equation models is inherently combined with commonly accepted graphical methods of its presentation. The aim of the paper is to outline the evolution of graphical tools of SEM models visualization and the issues concerning the presentation of contemporary complex models. Additionally, the graphical tools used in popular SEM models are presented.