

Rozdział 1. Wprowadzenie do modelowania wielopoziomowego

1.1. Geneza i rozwój metod analizy wariancji

Początków modelowania wielopoziomowego należy upatrywać w rozwoju metod związanych z analizą struktury wariancji – w szczególności z metodami ANOVA. Modele wielopoziomowe zyskały popularność w naukach humanistycznych i społecznych, w szczególności w tych sferach badań, w których wykorzystywane są dane indywidualne. Jak zauważa J.W. Twisk [2010, s. 17], w ciągu niespełna 10 lat (1995–2004), w samej tylko medycynie, liczba prac naukowych, w których zastosowano analizę wielopoziomową, wzrosła z 22 do ponad 170. O tym, iż modele wielopoziomowe uznawane są za interesujący sposób analizy danych, może świadczyć również to, że znalazły się na liście 50 zagadnień uznanych przez Radę Badań Ekonomicznych i Społecznych¹ za istotne z punktu widzenia nauk społecznych. Obecnie, mimo że modelowanie wielopoziomowe wciąż łączy się z badaniami edukacyjnymi, socjologicznymi czy też związanymi z ochroną zdrowia, modele te coraz chętniej wykorzystywane są przez badaczy zajmujących się analizą danych zlokalizowanych geograficznie.

Szczegółowy opis rozwoju metod modelowania wielopoziomowego, obejmujący okres od połowy XIX wieku do końca lat 60. XX wieku, można znaleźć w monografii pt. *Variance components* autorstwa S.R. Searle, G. Casella i C.E. McCullocha [1992, s. 19–43]. Rozwój metody od lat 70. XX wieku omówiony został syntetycznie przez J. de Leeuw, E. Meijera i H. Goldsteina [2008, s. 1–5] w *Handbook of Multilevel Analysis*. I.G.G. Kreft i J. de Leeuw [1998, s. 15–17] w *Introducing Multilevel Modeling* przedstawili rozwój metod modelowania wielopoziomowego z punktu widzenia upowszechniania kolejnych technik analizy. Ze względu na pokrewieństwo modeli wielopoziomowych i panelowych, znaczna część informacji dotyczących rozwoju modelowania wielopoziomowego jest zbieżna z historią modelowania panelowego – szczególnie odwołania do początków analizy wariancji (por. np. Nerlove [2002] – rozdział pt. *The history of panel data econometrics, 1861–1997*).

¹ Rada Badań Ekonomicznych i Społecznych (*Economic and Social Research Council – ESRC*) jest jedną z najważniejszych w Wielkiej Brytanii instytucji przyznających środki finansowe na prowadzenie badań ekonomicznych i społecznych. Istnieje od 1965 r.

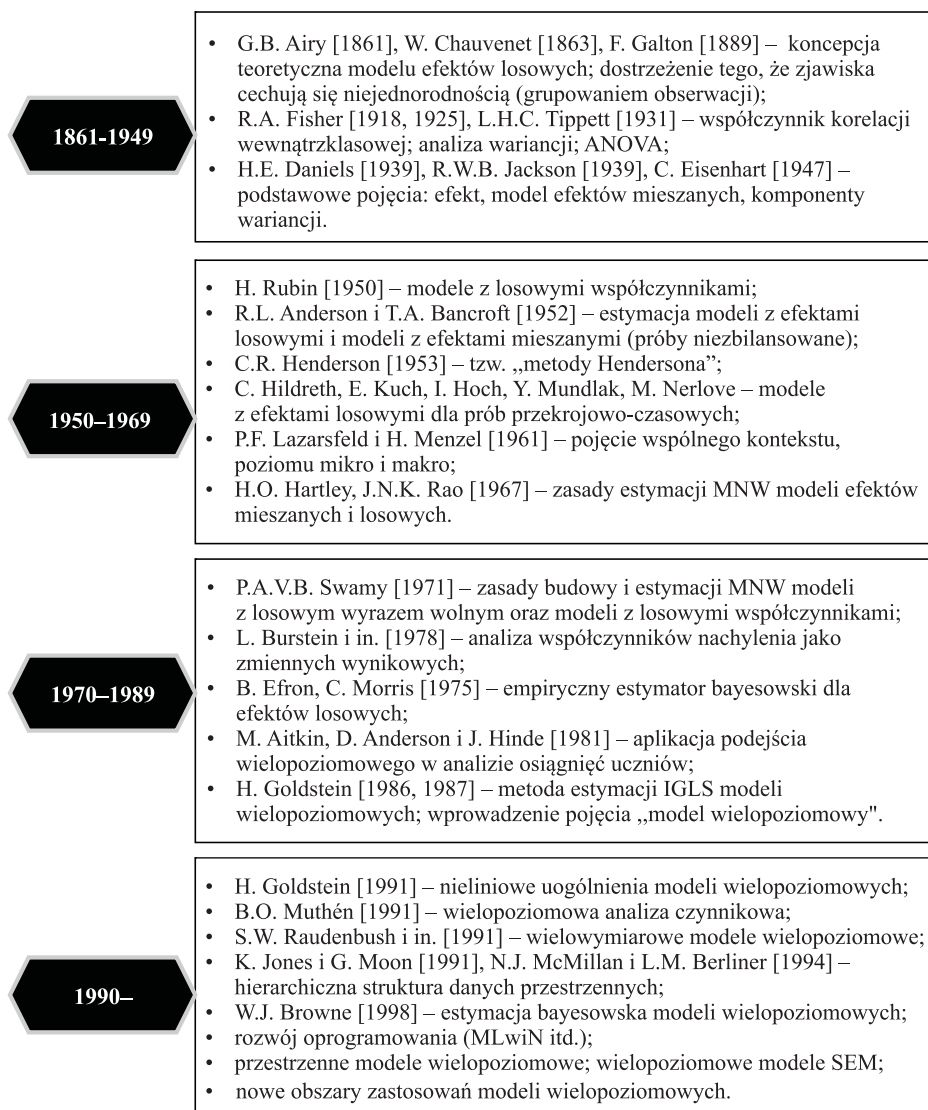
1.1.1. Okres do 1949 roku

Historia modelowania wielopoziomowego sięga połowy XIX wieku, kiedy to pojawiły się pierwsze rozważania dotyczące niejednorodności zjawisk. Opis nawiązujący do budowy modelu efektów losowych (choć termin ten nie został użyty), pojawił się po raz pierwszy w pracy G.B. Airy'ego z 1861 r. [1875], dotyczącej obserwacji astronomicznych. Analizując pracę Airy'ego, Scheffé [za: Searle i in., 1992, s. 23–25] zauważył, że rozważany problem (analiza tego samego obiektu astronomicznego przez j nocy, w ciągu których dokonywano i obserwacji) znajduje odzwierciedlenie w zapisie $y_{ij} = \alpha_0 + \mu_j + \varepsilon_{ij}$ (gdzie μ_j określa efekt związany z dniem wykonywania obserwacji), który jest równoważny podstawowemu modelowi struktury wariancji. Co ciekawe, mimo iż w literaturze przyjęło się, iż to Airy pierwszy nawiązał do modelu efektów losowych, jego praca rozpoczyna się słowami: „*No novelty, I believe, of fundamenal character, will be found in these pages*”. Brak powołań na inne prace uniemożliwia jednak weryfikację możliwości nawiązania do koncepcji efektów losowych we wcześniejszych badaniach.

Kolejnym badaczem, który nawiązał w swojej pracy do modeli efektów losowych był W. Chauvenet. W roku 1863 opublikował on dwutomowe dzieło o astronomii, w którym przedstawił sposób dekompozycji wariancji procesu jako $\text{var}(\bar{y}) = (\sigma_\mu^2 + \sigma_\varepsilon^2/J)/N$. Choć w jego pracy nie ma matematycznego zapisu formuły modelu (tak samo jak u Airy'ego), stosowany przez obu koncept myślowy wiąże się nierozdzielnie z analizą wariancji. Koniec XIX wieku wiąże się również z pracami F. Galtona, w których pojawiły się załączki modelu struktury wariancji [Nerlove, 2002, s. 12].

Jako pioniera w dziedzinie analizy wariancji wskazuje się najczęściej R.A. Fishera. Jego zasługi wiążą się z zaproponowaniem w 1918 r. jednoczynnikowej analizy ANOVA oraz użyciem po raz pierwszy terminów: „wariancja”, „analiza wariancji” i „współczynnik korelacji wewnątrzklasowej”. Koncepcja analizy, opracowana przez Fishera [1925], ukierunkowana była na dekompozycję wariancji przez wprowadzenie efektów stałych lub losowych. Fisher nie dokonał jednak dokładnego rozróżnienia i zdefiniowania obu rodzajów efektów. Co więcej, chociaż wskazał możliwości rozwinięcia proponowanej analizy dla prób niezbilansowanych (o różnej liczbie obserwacji w grupach) oraz rozważał analizę dwuczynnikową, nie określił sposobu estymacji tych modeli.

Z pierwszym okresem rozwoju modelowania wielopoziomowego związany był również L.H.C. Tippett [1931]. W wydanej w 1931 r. monografii poświęconej metodom statystycznym opisał on procedurę estymacji modelu struktury wariancji dla prób zbilansowanych oraz sposób estymacji dwuczynnikowego (z klasyfikacją krzyżową) modelu efektów losowych. Wraz z wyznaczeniem metody estymacji parametrów modeli dekompozycji wariancji Tippett zainicjował rozważania na temat optymalnej wielkości i struktury próby w modelach struktury wariancji.



Rysunek 1.1. Rozwój metod modelowania wielopoziomowego

Źródło: opracowanie własne.

Lata 30. i 40. XX wieku zaowocowały klaryfikacją pojęć, szczególnie istotnych z punktu widzenia modelowania wielopoziomowego. J. Neyman wraz z K. Iwaszkiewicz i S. Kołodziejczykiem [1935] wprowadzili termin „komponent błędu” (*error component*), H.E. Daniels [1939] zaś wprowadził sformułowanie: „komponenty wariancji” (*components of variance*). W tym samym roku R.W.B. Jackson [1939] użył słowa „efekt” w odniesieniu do składowych dekompozycji wariancji oraz przyjął założenia dotyczące normalności rozkładu efektów losowych. W roku 1947 C. Eisenhart [1947] użył po raz pierwszy pojęć: „model efek-

tów mieszanych” i „mieszana analiza wariancji” (*mixed analysis of variance*), jak również precyzyjnie określił różnice między efektami stałymi i losowymi w modelu.

1.1.2. Lata 50. i 60. XX wieku

Okres od 1950 do 1969 r. wiązał się z intensywnym rozwojem metod modelowania wielopoziomowego, określanych wciąż jako metody analizy wariancji. Searle i in. [1992, s. 33] nazwał te lata czasem wielkich postępów w dziedzinie badań nad strukturą wariancji. Kluczowe znaczenie miało, w tym okresie, zaproponowanie procedury estymacji bazującej na metodzie największej wiarygodności. W roku 1952 ukazała się książka pt. *Statistical Theory In Research* autorstwa R.L. Andersona i T.A. Bancrofta [1952]. Autorzy przedstawili w niej procedurę estymacji modeli z efektami losowymi oraz modeli z efektami mieszanymi (stałymi i losowymi) dla prób zbilansowanych. Rozpatrzyli oni również przypadek klasyfikacji zagnieżdżonej dla prób niezbilansowanych, przy czym nie rozwiązali oni szeregu problemów związanych z estymacją tego typu modeli. Mimo to książka ta została uznana [Searle i in., 1992, s. 33] za kamień milowy w rozważaniach dotyczących modeli struktury wariancji.

Osobą najczęściej kojarzoną z omawianym okresem jest C.R. Henderson. W roku 1953 w czasopiśmie „*Biometrics*” ukazał się artykuł pt. *Estimation of variance and covariance components*, w którym Henderson zaproponował trzy metody wyznaczania ocen komponentów wariacyjnej dla prób niezbilansowanych, określanych obecnie jako „metody Hendersona” [1953]. W metodzie I (tzw. metodzie analizy wariancji) zaproponował wykorzystanie dla prób niezbilansowanych takiej samej procedury dekompozycji wariancji (bazującej na fakcie, iż wariancja jest funkcją sum kwadratów odchyłeń od średniej) jak w przypadku zbiorów zbilansowanych. Metoda II (tzw. metoda korekty błędu w modelach mieszanych) wiązała się z wprowadzeniem korekty ze względu na obecność efektów stałych i odnosiła się do procedury dekompozycji wariancji w modelach mieszanych dla prób niezbilansowanych. Głównym ograniczeniem metody II był brak możliwości wprowadzania interakcji efektów stałych i losowych. Metoda III (tzw. metoda dopasowania stałych) opierała się zaś na obliczeniu odchyłeń od średniej na podstawie wartości teoretycznych z modelu i submodeli².

Zainicjowane przez Hendersona prace nad wyznaczeniem estymatorów BLUE (*Best Linear Unbiased Estimation*) dla efektów stałych i predykcji BLUP (*Best Linear Unbiased Prediction*) dla efektów losowych były kontynuowane m.in. przez Searlego [1958], Searlego i Hendersona [1961], Lowa [1964], Blischkego [1966]. Dotyczyły one przede wszystkim poszukiwania rozwiązań, w zakresie metod analizy wariancji, dla bardziej złożonych struktur (analiza dwuczynnikowa z klasyfikacją krzyżową lub zagnieżdżeniem, analiza trzyczynnikowa itd.).

² Szczegółowe omówienie metod Hendersona można znaleźć m.in. w artykule Djordjević i Lepojević [2003].

Równoległe z pracami nad różnymi formami analizy ANOVA prowadzone były badania w zakresie metod wprowadzania efektów grupowych w modelach szacowanych na podstawie prób uwzględniających wymiar czasowy, tj. danych czasowo-przekrojowych. Rolę efektów grupowych i czasowych analizowali m.in.: C. Hildreth, E. Kuch, I. Hoch, Y. Mundlak i M. Nerlove [szerzej: Dańska-Borsiak, 2011, s. 23]. Nurt ten, ukierunkowany na wyznaczenie efektów grupowych dla obiektów i/lub jednostek w czasie, przekształcił się później w odrębny obszar modelowania. Jego wyznacznikiem stało się wykorzystywanie danych panelowych, adekwatnym zaś sposobem modelowania – modele panelowe.

Porównując prace poświęcone rozwojowi metod analizy wariancji (które leżą u podstaw modelowania wielopoziomowego) z tymi, w których prezentowana jest geneza modelowania panelowego, można zauważyć, iż począwszy od lat 50. XX wieku nastąpiło stopniowe rozdzielenie obu kierunków analizy niejednorodności. W ekonomii kolejne lata wiązały się z koncentracją na rozwoju modeli panelowych, podczas gdy w pozostałych naukach społecznych i humanistycznych preferowanym podejściem stała się analiza wariancji za pomocą modeli wielopoziomowych.

Z końcem lat 60. ubiegłego wieku zaczęto rozważać szerzej możliwości estymacji modeli efektów losowych metodą największej wiarygodności (MNW). Mając na względzie wcześniejsze prace [Crump, 1951; Herbach, 1959] poświęcone temu zagadnieniu, H.O. Hartley i J.N.K. Rao [1967] sformułowali zasady estymacji MNW modeli efektów mieszanych i losowych, szacowanych na podstawie prób zbilansowanych i niezbilansowanych. Jak wskazują Kreft i de Leeuw [1998, s. 15] lata 70. były istotne dla metod analizy wariancji ze względu na rozwój technik obliczeniowych. Z jednej strony, powstało w tym czasie szereg prac pogłębiających procedury estymacji MNW modeli struktury wariancji [m.in. Saerle, 1970; Harville, 1977]. Z drugiej strony, w okresie tym podjęto szersze badania nad modelami losowych współczynników, których konstrukcję zaproponował H. Rubin [1950].

1.1.3. Lata 70. i 80. XX wieku

P.A.V.B. Swamy [1971] w monografii pt. *Statistical Inference in Random Coefficient Regression Models* szczegółowo omówił zasady budowy i estymacji MNW modeli z losowym wyrazem wolnym oraz modeli z losowymi współczynnikami, przedstawiając dodatkowo możliwość ich aplikacji w obszarze analiz ekonomicznych (badania na podstawie danych panelowych dotyczące inwestycji i konsumpcji). Co ważne, przedstawił również model z losowymi współczynnikami, zakładający, iż składnik losowy jest skorelowany przestrzennie. Badania w tym zakresie prowadził również B. Rosenberg [1973], a E. Spjøtvoll [1977] dokonał przeglądu metod budowy modeli z losowymi współczynnikami.

Postęp, jaki dokonał się w latach 70. ubiegłego wieku w obszarze metod budowy i estymacji modeli z efektami losowymi, miał dwojaki charakter. Z jed-

nej strony, L. Burstein i in. [1978] zaproponowali metodę analizy, stanowiącą alternatywę dla techniki dekompozycji wariancji Cronbacha, nazwaną analizą współczynników nachylenia jako zmiennych wynikowych (*slopes-as-outcomes analysis*). Proponowaną metodą estymacji takiego modelu była dwustopniowa MNK. Z drugiej strony, w pracy B. Efrona i C. Morrisa [1975] pojawiła się propozycja stosowania empirycznego estymatora bayesowskiego do wyznaczenia wartości średnich grupowych. Idea ta została później wykorzystana w predykcji efektów losowych w modelach wielopoziomowych [m.in. Dempster i in., 1977].

Pod koniec lat 70. dostrzeżono, że efekt grupowania obserwacji może wynikać ze wspólnego kontekstu, który można próbować wyjaśniać. Rozważania te prowadzono w odniesieniu do badań edukacyjnych, które stały się w tym czasie głównym obszarem aplikacyjności modeli z efektami losowymi. Ówcześni badacze odwoływali się przy tym do osiągnięć P.F. Lazarsfelda i H. Menzela [1961], którzy wprowadzili takie pojęcia jak: wspólny kontekst, poziom mikro i makro – będące kluczowymi z punktu widzenia modelowania wielopoziomowego.

W roku 1981 M. Aitkin, D. Anderson i J. Hinde opublikowali artykuł pt. *Statistical Modelling of Data on Teaching Styles*, w którym zastosowali model efektów mieszanych dla prób niezbilansowanych w celu określenia, na ile różnice w ocenach uzyskiwanych przez uczniów są związane z grupowaniem według klas i kompetencjami nauczycieli. Takie rozumienie grupowania było tożsame z wprowadzeniem poziomu indywidualnego (uczniowie) i poziomu II (klasa, prowadzona przez danego nauczyciela). Koncepcja grupowania jednostek wprowadzona przez Aitkina i in. [1981] zyskała popularność i jest do dziś przytaczana w większości podręczników modelowania wielopoziomowego. Mimo iż model stosowany w omawianym badaniu można uznać za wielopoziomowy, pojęcie to nie padło w artykule wprost.

Analiza wielopoziomowa w postaci, jaka jest stosowana obecnie, wykształciła się jako rezultat połączenia wiedzy płynącej z prac poświęconych metodom analizy wariancji z ideą analizy kontekstowej (*contextual analysis*) [Snijders, Bosker, 2012, s. 2–3]. Rola przynależności do danej grupy (wspólnego kontekstu) akcentowana była już wcześniej m.in. przez Robinsona [1950] – pojęcie błędu ekologicznego, Davisa i in. [1961] – rozróżnienie między regresją międzygrupową i wewnątrzgrupową oraz u Bursteina i in. [1978] – modele współczynników nachylenia jako zmiennych wynikowych. Mimo to, aż do lat 80. nie podjęto prób połączenia obu koncepcji. Analiza kontekstowa bazowała zatem, do tego czasu, na estymacji MNK tradycyjnych modeli liniowych, metody analizy wariancji zaś wykluczały możliwość wyjaśnienia przyczyn zróżnicowania za pomocą zmiennych kontekstowych.

Dopiero w latach 80. XX wieku zaczęto łączyć ze sobą fakt, iż skala zróżnicowania na poszczególnych poziomach agregacji (modelowana jako wariancja efektów losowych) może podlegać wyjaśnieniu przez adekwatne dla danego poziomu analizy predyktory. Wspomnianą grupę modeli zaczęto określać wtedy

modelami wielopoziomowymi [Goldstein, 1986]. W szczególności H. Goldstein [1986] dokonał klasyfikacji nazewnictwa, wskazując m.in. na to, iż dane mogą mieć strukturę hierarchiczną. Zaproponował ponadto procedurę estymacji (rozważanych modeli) iteracyjną uogólnioną metodą najmniejszych kwadratów (IGLS) oraz odróżnił model jednopoziomowy (*single model*) od wielopoziomowego (*multilevel model*). Ostatecznie, w roku 1987 H. Goldstein wydał monografię pt. *Multilevel Models in Educational and Social Research*, będącą pierwszą książką w pełni poświęconą modelowaniu wielopoziomowemu.

1.1.4. Okres od lat 90. do chwili obecnej

Z początkiem lat 90. nastąpił gwałtowny wzrost zainteresowania modelowaniem wielopoziomowym, który zaowocował wydaniem pierwszych kluczowych pozycji książkowych przez: Raudenbusha i Bryka w 1991 r. (*Hierarchical Linear Models: Applications and Data Analysis Methods*), Goldsteina w 1995 r. (*Multilevel Statistical Models*), Krefta i de Leeuwa w 1998 r. (*Introducing Multilevel Modeling*).

Pod względem metod budowy nastąpiło zainteresowanie m.in. nieliniowymi uogólnieniami modelu wielopoziomowego [Goldstein, 1991], wielopoziomową analizą czynnikową [Muthén, 1991] oraz wielowymiarowymi modelami wielopoziomowymi [Raudenbush i in., 1991]. Modele wielopoziomowe zaczęto dostosowywać na szerszą skalę do analizy struktur danych z przyporządkowaniem do wielu grup i krzyżowym [Goldstein, 1994] oraz danych z powtarzającymi pomiarami [DiPrete, Grusky, 1990; Goldstein i in., 1994]. Z punktu widzenia rozwoju technik estymacji kluczowego znaczenia nabrały zaś metody bayesowskie [Browne, 1998].

Początkowo głównym obszarem zastosowań modeli wielopoziomowych były analizy rynku edukacyjnego [np. Lee, Bryk, 1989; Chen, Vazsonyi, 2013]. Stopniowo obszar aplikacyjności uległ poszerzeniu o analizy związane z ochroną zdrowia [np. Duncan i in., 1996; Bostan i in., 2015] i procesami społecznymi [np. Fernandez i Kulik, 1981; Manfredo i in., 2009]. Do końca lat 90. to właśnie te trzy obszary badań zdominowały ogół analiz wielopoziomowych.

Na początku lat 90. XX wieku dostrzeżono, że bazując na danych zlokalizowanych przestrzennie, można sformułować strukturę hierarchiczną, odzwierciedlającą kolejne agregaty przestrzenne, a w efekcie, że modele wielopoziomowe mogą być naturalnym kierunkiem analizy danych przestrzennych. W tym okresie nieliczni autorzy wykazali w swoich pracach możliwość wykorzystania modeli wielopoziomowych w analizie danych o przestrzennej strukturze hierarchicznej.

Jednymi z nich byli Jones i Moon [1991], którzy zaproponowali, aby w modelu indywidualnego stanu zdrowia uwzględnić: cechy danej osoby (np. wiek, płeć, czynniki ryzyka, takie jak palenie), cechy środowiskowe (m.in. poziom zanieczyszczenia środowiska), interakcje między cechami indywidualnymi i środowiskowymi oraz efekty losowe indywidualne i przestrzenne. Tym samym dostrzegli oni, że zróżnicowanie stanu zdrowia jednostek może wynikać z czynni-

ków, które są trudno mierzalne bądź których w ogóle nie można skwantyfikować. Analogicznie, trudno zakładać, że stan zdrowia jest jednakowy w każdym punkcie przestrzeni geograficznej, a skoro tak, to zestaw czynników wyjaśniających zróżnicowanie terytorialne może nie być wyczerpujący. Włączenie do modelu zróżnicowania indywidualnego i terytorialnego umożliwiło zastosowanie modelu wielopoziomowego.

Hierarchię przestrzenną zaproponowali także McMillan i Berliner [1994]. Analizując wielkość plonów kukurydzy w Stanach Zjednoczonych, wykorzystali dane pochodzące z 3852 farm. Ogół gospodarstw został przez nich podzielony na trzy rodzaje według wielkości oraz na 88 grup według lokalizacji w hrabstwach. W celu określenia zróżnicowania plonów w poszczególnych hrabstwach McMillan i Berliner zastosowali model z dekompozycją wariancji składnika losowego, w którym poziom hrabstw uzyskano, wprowadzając efekty losowe. Ważnym elementem rozważanego modelu było wykorzystanie technik modelowania, bazujących na polach losowych Markowa, w specyfikacji macierzy kowariancji efektów specyficznych dla hrabstw.

Możliwość wykorzystania technik modelowania wielopoziomowego w analizach danych o strukturze przestrzennej wskazali ponadto Kreft i de Leeuw [1998, s. 7]. Zauważyli oni, że rozwój geograficznych systemów informacji przestrzennej oraz coraz większa dostępność danych mikro, zlokalizowanych geograficznie, sprzyjać będą rozwojowi podejścia wielopoziomowego w analizach przestrzennych. Twierdzenie to potwierdziły późniejsze, coraz liczniejsze badania, w których zaczęto zwracać uwagę na hierarchiczną strukturę danych przestrzennych, np.: Langford i in. [1999], Pierewan i Tampubolon [2014], Lacombe i McIntyre [2016].

Współczesne kierunki rozwoju metod modelowania wielopoziomowego wiążą się z coraz powszechniejszym stosowaniem metod bayesowskich, poszukiwaniem nowych obszarów badań, w których można stosować modelowanie wielopoziomowe (np. modele hedonicznej wyceny rynku nieruchomości, dobrostanu psychospołecznego), budową wielopoziomowych modeli równań strukturalnych oraz takich, które umożliwiają jednoczesną estymację efektów interakcji przestrzennych i/lub społecznych (przestrzennych modeli wielopoziomowych). Ostatnia kategoria modeli została omówiona szerzej w rozdziale 3.

1.1.5. Rozwój oprogramowania

Wzrost zainteresowania wspomnianą klasą modeli doprowadził do zainicjowania prac nad oprogramowaniem umożliwiającym estymację modeli wielopoziomowych. W roku 1988 w Wielkiej Brytanii H. Goldstein wraz z J. Rasbashem i zespołem, w ramach *The Multilevel Models Project*, przygotowali program ML2 umożliwiający szacowanie parametrów modeli dwupoziomowych. Program ten był zintegrowany z oprogramowaniem statystycznym NANOSTAT autorstwa M.J.R. Healy'ego. W ciągu kolejnych kilku lat ML2 został rozwinięty do postaci ML3 (1990 r.) – estymacja modeli trzypoziomowych i ML_n (1995 r.) –

szacowanie parametrów modeli o dowolnej liczbie poziomów. Oprogramowanie to było pierwowzorem dla programu MLwiN, który pojawił się w 1998 r. i był już samodzielnym pakietem. W przygotowanie kolejnych wersji programu zaangażowany był przede wszystkim J. Rasbash. Rozbudową programu MLwiN w kierunku estymacji dodatkowych modeli wielopoziomowych (z klasyfikacją krzyżową, przynależnością do wielu grup i przestrzennych wielopoziomowych) kierował zaś W. Browne. Ostatecznie, kolejne aktualizacje programu MLwiN, wymagające uelastycznienia procedur m.in. dla dużych baz danych, przygotował C. Charlton.

Równoległe z ośrodkiem w Wielkiej Brytanii, pracę nad oprogramowaniem do modelowania wielopoziomowego rozpoczęto w Stanach Zjednoczonych. W połowie lat 80. XX wieku N. Longford opracował program VARCL, który wykorzystał do obliczeń w swoich pracach [m.in. Aitkin, Longford, 1986], Mason i in. [1988] zaś zaproponowali program GENMOD. VARCL składał się z dwóch subprogramów: VARL3 i VARL9. Pierwszy z nich umożliwiał estymację modeli wielopoziomowych z maksymalnie trzema poziomami (wprowadzanymi w postaci efektów losowych), drugi zaś z maksymalnie dziewięcioma (jako efekty stałe). Program HLM (*Hierarchical Linear and Nonlinear Modeling*), specjalizujący się w estymacji modeli wielopoziomowych, zaproponowali A.S. Bryk (z uniwersytetu w Chicago), S.W. Raudenbush (uniwersytet w Michigan) i R. Congdon (Uniwersytet Harvarda).

Początki programu HLM są zbliżone do MLwiN. Pierwsze wersje oprogramowania umożliwiały szacowanie wyłącznie parametrów modeli dwupoziomowych, kolejne zaś aktualizacje wiązały się z wprowadzeniem modeli o większej liczbie poziomów agregacji. O ile VARCL był programem niekomercyjnym i przeznaczonym na użytek autora, o tyle MLwiN oraz HLM są programami komercyjnymi. W obecnej postaci oba umożliwiają modelowanie ciągłych i dyskretnych zmiennych wynikowych oraz wprowadzanie dowolnej liczby poziomów analizy. Analogicznie do MLwiN program HLM daje możliwość analizy bardziej zaawansowanych struktur (z klasyfikacją krzyżową i przynależnością do wielu grup).

Od połowy lat 90. ubiegłego wieku nastąpił wzrost liczby programów ukierunkowanych wyłącznie na analizę wielopoziomową oraz dodatków i rozszerzeń umożliwiających estymację takich modeli. W latach 1994–1996 D. Hedeker i R.D. Gibbons z Uniwersytetu Illinois w Chicago (wraz z D. Pattersonem jako programistą) stworzyli serię programów do modelowania wielopoziomowego opatrzonych wspólną nazwą MIXFOO. W jego skład wchodziły min.: MIXREG – liniowe modele wielopoziomowe (w tym modele z autokorelacją składnika losowego), MIXOR i MIXNO – modele wielopoziomowe dla zmiennej wynikowej nieciągłej, MIXGSUR – wielopoziomowa analiza przeżycia [de Leeuw, Kreft, 2011]. W roku 2008 na bazie programów wchodzących w skład MIXFOO Hedeker i Gibbons stworzyli komercyjny program SuperMix.

W roku 2009 z inicjatywy J. Rasbasha, w Centrum Modelowania Wielopoziomowego na uniwersytecie w Bristolu, W. Browne z zespołem opracowali program Stat-JR, który w przeciwieństwie do MLwiN zawiera rozbudowany moduł do analiz statystycznych. Umożliwia on również szacowanie parametrów modeli wielopoziomowych. Spośród innych pakietów opracowanych w Centrum Modelowania Wielopoziomowego można wskazać: pakiet REALCOM (zintegrowany z Matlabem i MLwiN) i MLPowSim oraz biblioteki komend *runmlwin* i *runmixregls* (integracja MLwiN ze Stata), R2MLwiN (integracja MLwiN z R Cran).

Aktualnie większość pakietów statystycznych ma oprogramowane moduły do analiz wielopoziomowych. W Stata procedury umożliwiające modelowanie wielopoziomowe są oprogramowane pod nazwą „*Multilevel mixed-effects models*”. Ponadto dla bardziej zaawansowanych struktur danych (np. klasyfikacja krzyżowa i do wielu grup) dostępny jest zestaw procedur z biblioteki GLLAMM (*General Linear Latent and Mixed Models*), przygotowanej przez S. Rabe-Hesketh, A. Picklesa i A. Skrondala w 1999 r. [szerzej: Rabe-Hesketh i Skrondal, 2008]. W programie SPSS, począwszy od wersji 11, możliwe jest przeprowadzenie liniowej analizy wielopoziomowej (dla ciągłej zmiennej wynikowej). W programie R Cran najpopularniejszą biblioteką, zawierającą funkcje umożliwiające estymację modeli wielopoziomowych, jest *lme4*. Oprócz tego, istnieje szereg bibliotek z funkcjami estymującymi bardziej zaawansowane modele wielopoziomowe (m.in. HSAR dla hierarchicznych modeli autoregresji przestrzennej).

1.1.6. Wiodące ośrodki badawcze

Obecnie modele wielopoziomowe są stosowane powszechnie w naukach społecznych i humanistycznych (w szczególności socjologii, pedagogice, psychologii, kryminologii, ochronie zdrowia). Coraz większe zainteresowanie budzą również w tych obszarach badań, w których wykorzystywane są dane zlokalizowane geograficznie. Spośród wielu ośrodków akademickich, w których badania mają na względzie rozwój metod budowy, estymacji i aplikacyjności modeli wielopoziomowych, można wskazać kilka ośrodków, których pozycja jest niewątpliwie dominująca ze względu na zakres prowadzonych prac badawczych oraz potencjał zespołów. Ośrodki te określić można „szkołami” modelowania wielopoziomowego.

„Szkoła” brytyjska (bristolska)

Koncentruje badaczy zajmujących się rozwojem metod modelowania wielopoziomowego i ich aplikacją głównie w takich dziedzinach jak: demografia, nauki edukacyjne, ochrona zdrowia, geografia. Jej założycielem jest H. Goldstein – kierownik „*The Multilevel Models Project*” na uniwersytecie w Bristolu, a następnie dyrektor Centrum Modelowania Wielopoziomowego (obecnie dyrektorem jest W. Browne), będącego jednym z nielicznych na świecie i wiodących jednostek ba-

dawczych ukierunkowanych na modelowanie wielopoziomowe. Centrum skupia badaczy pracujących w Szkole Nauk Edukacyjnych, Szkole Weterynarii i Szkole Nauk Geograficznych Uniwersytetu w Bristolu. Działania jednostki koncentrują się na obszarze: rozwoju oprogramowania (MLwiN, Stat-JR), kształcenia kadr (organizacja kursów i seminariów z zakresu modelowania wielopoziomowego, prowadzenie studiów doktoranckich), upowszechniania podejścia wielopoziomowego (prowadzenie mailowej grupy dyskusyjnej, internetowego katalogu prac, w których wykorzystano podejście wielopoziomowe) oraz prowadzenia badań nad budową i aplikacją wspomnianej klasy modeli.

„Szkola” holenderska

Dwa podstawowe ośrodki w Holandii zajmujące się modelowaniem wielopoziomowym znajdują się na uniwersytetach w Groningen i Utrechcie. Z pierwszym z nich związani są R.J. Bosker i T.A.B. Snijders, z drugim zaś J. Hox. Badania Snijdersa ukierunkowane są na połączenie analizy sieci społecznych z modelowaniem wielopoziomowym. Te dwa nurty zajmują kluczową pozycję w jego pracach, które można wpisać w obszar badań socjologicznych. Dla porównania, badania Boskera koncentrują się w obszarze modelowania wielopoziomowego rynku edukacyjnego. Efektem prac Snijdersa i Boskera jest podręcznik *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling* [2012], będący jednym z najpopularniejszych podręczników do modelowania wielopoziomowego. W obszarze nauk społecznych swoje badania prowadzi również Hox. W przeciwieństwie do ośrodka w Wielkiej Brytanii, „szkola” holenderska nie wykształciła odrębnego ośrodka, który zrzeszałby wyżej wymienionych badaczy.

„Szkola” amerykańska

Do najpopularniejszych badaczy zajmujących się modelowaniem wielopoziomowym w Stanach Zjednoczonych należą obecnie, m.in.: S. Rabe-Hesketh (uniwersytet w Kalifornii, Berkley), D. Hedeker i R.D. Gibbons (uniwersytet w Chicago), A. Gelman (Uniwersytet Columbia, Nowy York), S.W. Raudenbush (uniwersytet w Chicago). Poszczególne badacze specjalizują się w różniących się od siebie dyscyplinach (Raudenbush jest socjologiem, podczas gdy Hedeker biostatystykiem), ale łączy ich to, że w swoich badaniach wykorzystują powszechnie podejście wielopoziomowe. Dodatkowo część ich prac poświęcona jest rozwojowi metod modelowania i budowie oprogramowania (np. Hedeker i Rabe-Hesketh). Analogicznie jak w przypadku Holandii, nie ma w Stanach Zjednoczonych jednego ośrodka, który koncentrowałby przedstawicieli wspomnianego nurtu modelowania.

Poza wyżej wymienionymi można wskazać szereg ośrodków – odznaczających się znacznie krótszą tradycją badań w wyżej wymienionym zakresie – w których prowadzone są prace dotyczące budowy i możliwości wykorzystania modeli wielopoziomowych. Szereg uczelni prowadzi również kursy dla studentów i doktorantów poświęcone zasadom modelowania wielopoziomowego.

1.2. Nazewnictwo i wybrane pojęcia związane z modelami wielopoziomowymi

Szereg wątpliwości, szczególnie wśród osób, które do tej pory nie miały styczności z modelowaniem wielopoziomowym, może budzić terminologia. Dotyczy ona nazw poszczególnych składowych modelu (np. zmienne kontekstowe, interakcje krzyżowe) odnosi się do samego pojęcia modelu wielopoziomowego i możliwości zastępowania go innymi nazwami, np. model hierarchiczny lub model z efektami losowymi [Twisk, 2010, s. 17].

1.2.1. Podstawowe zagadnienia związane z terminologią

Termin „model wielopoziomowy” (*multilevel model* – MLM) przez większość badaczy jest traktowany jako synonim nazw: **hierarchiczny model liniowy** (*hierarchical linear model* – HLM), **liniowy model efektów mieszanych** (*linear mixed model* – LMM), **model efektów losowych** (*random effects model* – REM), **model zagnieżdżony** (*nested model*), **model struktury wariancji** lub inaczej **komponentów wariancyjnych** (*variance components model*). Różnice w stosowanym nazewnictwie wynikają z obszaru badań oraz sposobu postrzegania składowych modelu [Garson, 2013, s. 4]. W statystyce najczęściej przyjmowaną nazwą jest model struktury wariancji, podczas gdy biostatystycy używają terminu liniowy model efektów mieszanych. W ekonomii preferowanym terminem jest model efektów losowych, w socjologii zaś model wielopoziomowy. W praktyce wszystkie te nazwy odnoszą się jednak do tego samej konstrukcji myślowej.

Jak już wspomniano, przyjmowane nazewnictwo dość często odzwierciedla tę cechę modelu, która jest najbardziej eksponowana przez badacza. Na przykład dla ekonomistów cechą tą będzie modelowanie tzw. efektów specyficznych dla obiektów (przejawiających się w postaci grupowania wartości zmiennej wynikowej) jako efektów losowych, a nie stałych. Zdecydowanie mniej uwagi przywiązuje się w takich badaniach do możliwości strukturyzowania danych do postaci wielopoziomowej, bądź hierarchicznej. Nieeksponowany jest w nich również wpływ zmiennych kontekstowych, czy też możliwość wprowadzania interakcji między poziomami. Wprowadzenie efektów losowych traktowane jest w tych badaniach raczej jako sposób „kontrolowania” efektu grupowania obserwacji niż jako punkt wyjścia analizy skali i źródeł heterogeniczności.

Dla porównania, w badaniach edukacyjnych używa się powszechnie terminu model wielopoziomowy. Nazwa ta akcentuje nie sam fakt modelowania różnic między grupami, przez wprowadzenie efektów stałych bądź losowych, ale możliwość symultanicznej analizy procesu na wielu poziomach agregacji. Wynika to z faktu, iż w badaniach tych szczególną rolę odgrywa identyfikacja czynników indywidualnych i grupowego kontekstu. Samo wprowadzenie zróżnicowania międzygrupowego jako wariancji efektów losowych nie jest niczym niezwykłym, a rozważania dotyczące wyboru między efektami stałymi a losowymi ustępują

miejsca analizie wpływu zmiennych kontekstowych i interakcji. Zastosowanie terminu model wielopoziomowy wskazuje również na dość powszechną możliwość grupowania obserwacji jednocześnie według kilku równorzędnych kryteriów (np. podział uczniów ze względu na szkołę i miejsce zamieszkania).

Z perspektywy analizy danych przestrzennych wydaje się uzasadnione używanie terminu model hierarchiczny lub wielopoziomowy. Pierwszy z wymienionych podkreśla jedną z podstawowych cech danych przestrzennych, jaką jest możliwość hierarchicznego porządkowania jednostek w przestrzeni geograficznej. Drugie określenie odzwierciedla zaś to, że porządkowanie jednostek terytorialnych nie musi odbywać się wyłącznie według szczebli agregacji przestrzennej, które tworzą układ hierarchiczny. Możliwe jest łączenie klasyfikacji wynikającej z przynależności terytorialnej z innymi sposobami grupowania, niepowiązanych z lokalizacją w przestrzeni geograficznej. Używanie terminu „model wielopoziomowy”, dzięki jego ogólności, wydaje się zatem najlepiej odzwierciedlać rozważaną klasę modeli ekonometrii.

Mając na uwadze różnice w stosowanej terminologii, Garson [2013, s. 7–12] dokonał rozróżnienia stosowanego nazewnictwa, wydzielając typy modeli (w niniejszej monografii nazwanych modelami wielopoziomowymi) różniące się od siebie obecnością określonych składowych. Zgodnie z jego podziałem najbardziej ogólną kategorię modeli stanowią **modele efektów mieszanych**. Model taki zawiera:

- zmienne objaśniające na poziomie indywidualnym x_{ij} (opcjonalnie również na poziomie grupowym x_j , których wpływ na zmienną wynikową jest stały (β),
- zmienne objaśniające na poziomie indywidualnym, których wpływ na zmienną wynikową różni się w poszczególnych $j = 1, \dots, J$ grupach, co jest uwzględniane przez wprowadzenie losowych nachyleń współczynników (β_j),
- dekompozycję wyrazu wolnego, która jest uwzględniana w postaci efektów losowych (α_j).

Przykładowy model efektów mieszanych można zatem wyrazić wzorem:

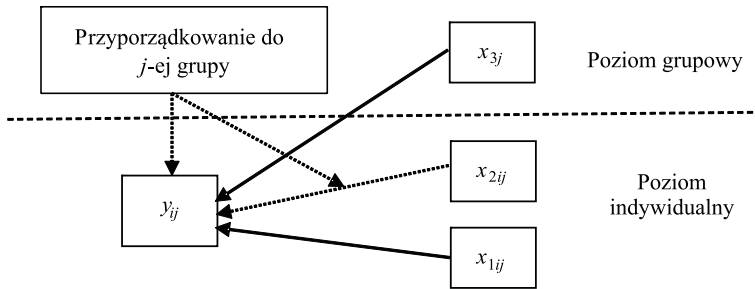
$$y_{ij} = \alpha_0 + \alpha_j + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3j} + \varepsilon_{ij}$$

lub graficznie za pomocą rysunku 1.2.

Pod względem budowy model efektów mieszanych stanowi połączenie modelu efektów losowych i modelu efektów stałych. **Model efektów losowych** cechuje się występowaniem dekompozycji wyrazu wolnego (za pomocą efektów losowych) i losowymi nachyleniami współczynników stojących przy zmiennych objaśniających dla poziomu indywidualnego. W przypadku **modelu efektów stałych** wszystkie zmienne objaśniające mają stałą wartość parametru β , zróżnicowanie międzygrupowe zaś wprowadzane jest jako ciąg zmiennych binarnych, z których każda określa przynależność obserwacji do danej grupy.

Modele efektów losowych można podzielić ze względu na typ struktury wariancji na dwie grupy:

- **modele z losowym wyrazem wolnym** (*random intercept model*),

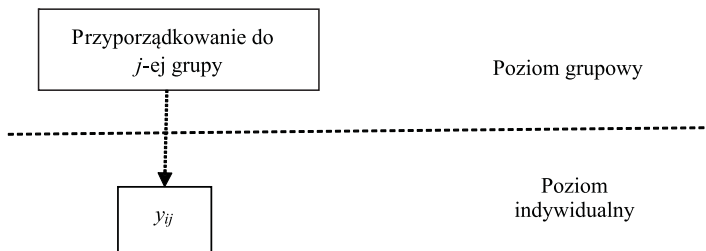


Rysunek 1.2. Powiązania w modelu efektów mieszanych

Źródło: opracowanie własne.

- **modele z losowymi współczynnikami nachylenia** (*random slope model*) lub **modele losowych współczynników** (*random coefficients model*).

Modele z losowym wyrazem wolnym podlegają podziałowi ze względu na obecność zmiennych objaśniających na poziomie indywidualnym i zmiennych objaśniających grupowych. Jeśli w modelu, oprócz dekompozycji wyrazu wolnego, nie ma żadnego predyktora, to model taki odpowiada **jednoczynnikowej analizie ANOVA z efektami losowymi** (*one-way ANOVA with random effects*). Nazywany jest on również modelem pustym (*null model*) lub bezwarunkowym. Celem zastosowania takiego modelu jest wstępna ocena skali zróżnicowania zmiennej wynikowej na poszczególnych poziomach analizy. Model ten nie daje możliwości wyjaśnienia przyczyn niejednorodności (rys. 1.3).

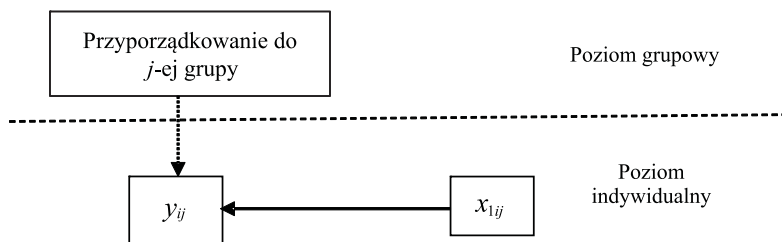


Rysunek 1.3. Powiązania w modelu pustym

Źródło: opracowanie własne na podstawie [Garson, 2013, s. 7–12].

Kolejnym modelem należącym do grupy modeli z losowym wyrazem wolnym jest **jednoczynnikowa ANCOVA z efektami losowymi** (*one-way ANCOVA with random effects model*). W odróżnieniu od poprzedniego, oprócz dekompozycji wyrazu wolnego, modele te zawierają predyktory umożliwiające wyjaśnienie przyczyn zróżnicowania wartości zmiennej wynikowej na poziomie indywidualnym (rys. 1.4). Nie dopuszcza się jednak różnicowania wartości współczynników nachylenia dla tych zmiennych, jak również wprowadzania zmiennych kontekstowych. Model ten może być interesującym rozwiązaniem w przypadku, gdy mamy do czynienia z grupowaniem obserwacji (przy czym liczba możliwych do wy-

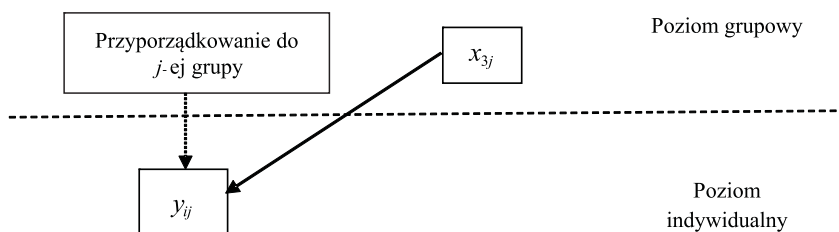
dzielenia grup jest znacząca), ale przedmiotem analizy nie jest próba wyjaśnienia przyczyn różnic międzygrupowych. Dla odpowiednio małego J model ten można zastąpić analizą ANCOVA z efektami stałymi.



Rysunek 1.4. Powiązania w jednoczynnikowej ANCOVA z efektami losowymi

Źródło: opracowanie własne na podstawie [Garson, 2013, s. 7–12].

Wyjaśnienie przyczyn zróżnicowania grupowego umożliwia **model regresji z losowym wyrazem wolnym** (*random intercept regression model*), określany również mianem **modelu średnich jako wynikowych** (*means-as-outcomes regression model*). Jego istotą jest określenie skali heterogeniczności (za pomocą wariancji efektów losowych), przy jednoczesnej próbie wyjaśnienia źródeł zróżnicowania, przez wprowadzenie zmiennych objaśniających zagregowanych do poziomu grupowego (*group composition variable*). Model ten pozbawiony jest predyktorów dla poziomu indywidualnego, a te zmienne objaśniające, które są dostępne na poziomie I, podlegają agregacji do poziomu II i tylko w takiej formie mogą pojawić się w modelu (rys. 1.5). Model w tej postaci może być wykorzystany wtedy, gdy uwaga badacza skoncentrowana jest wyłącznie na analizie relacji makro–mikro.

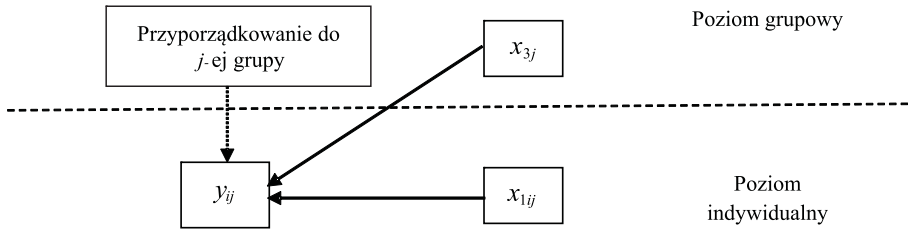


Rysunek 1.5. Powiązania w modelu średnich wynikowych

Źródło: opracowanie własne na podstawie [Garson, 2013, s. 7–12].

Jako model z losowym wyrazem wolnym traktowany jest również taki model, który oprócz dekompozycji wyrazu wolnego, uzyskiwanej przez wprowadzenie efektów losowych dla J grup, zawiera predyktory dla poziomu indywidualnego i grupowego. Model taki określany jest mianem **modelu ANCOVA średnich jako wynikowych** (*means-as-outcomes ANCOVA model*). W odróżnieniu od modelu ANOVA średnich wynikowych umożliwia on jednoczesną analizę po-

wiązań makro–mikro (wpływ zmiennych grupowych na wynikową) oraz relacji mikro–mikro (wpływ zmiennych objaśniających na poziomie I). Model ten można zatem traktować jako połączenie modelu ANOVA średnich wynikowych z jednozmienną ANCOVA z efektami losowymi.



Rysunek 1.6. Powiązania w modelu ANCOVA średnich wynikowych

Źródło: opracowanie własne na podstawie [Garson, 2013, s. 7–12].

Drugą grupę modeli efektów losowych stanowią **modele z losowymi współczynnikami nachylenia**. W modelach tych przyjmuje się, że wpływ zmiennej objaśniającej na zmienną wynikową nie jest jednakowy w poszczególnych grupach. Międzygrupowe zróżnicowanie siły wpływu danego predyktora uwzględniane jest przez wprowadzenie losowych wartości współczynników nachylenia. Modele te nie podlegają dalszej klasyfikacji i podziałowi na podgrupy, tak jak w przypadku modeli z losowym wyrazem wolnym.

Dla modeli z losowymi współczynnikami nachylenia można zauważyć pewną nieścisłość dotyczącą nazewnictwa. Mimo że nazwa modelu sugeruje, iż mamy do czynienia wyłącznie ze różnicowaniem siły wpływu predyktora, w literaturze [m.in. Snijders i Bosker, 2012, s. 74–75] przyjmuje się, że taki model oprócz różnicowania wartości parametru β odznacza się jednoczesną dekompozycją wyrazu wolnego. Modele zawierające wyłącznie losowe współczynniki nachylenia nie są spotykane w badaniach empirycznych.

Analogiczne wymagania, tj. dekompozycja wyrazu wolnego i obecność losowych współczynników nachylenia, stanowią podstawy wydzielenia klasy modeli określanej mianem **modeli losowych współczynników**. Innymi słowy, obie nazwy odnoszą się do tej samej grupy modeli efektów losowych i mogą być stosowane zamiennie. Zgodnie z ujęciem proponowanym przez Garsona [2013, s. 7–12] do modeli losowych współczynników można zaliczyć dwie kategorie modeli, różniące się od siebie obecnością zmiennych grupowych.

Pierwszą grupę tworzą **modele regresji losowych współczynników** (*random coefficient regression models*). W modelach tych przyjmuje się, że przynależność do określonej grupy skutkuje zróżnicowaniem wartości średniej zmiennej wynikowej (losowy wyraz wolny) oraz modyfikacją siły wpływu zmiennych objaśniających na zmienną wynikową (losowe współczynniki nachylenia). Model ten pozbawiony jest jednak zmiennych grupowych, które umożliwiłyby wyjaśnienie przyczyn zróżnicowania (rys. 1.7).