

Ekonomia

Modele wielopoziomowe

Wykorzystanie danych regionalnych
w badaniach mikroekonomicznych
i socjologicznych

Wojciech Grabowski



Modele wielopoziomowe

Wykorzystanie danych regionalnych
w badaniach mikroekonomicznych
i socjologicznych



WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

[Kup książkę](#)

Ekonomia

Modele wielopoziomowe

Wykorzystanie danych regionalnych
w badaniach mikroekonomicznych
i socjologicznych

Wojciech Grabowski

Wojciech Grabowski – Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny
Katedra Modeli i Prognoz Ekonometrycznych, 90-214 Łódź, ul. Rewolucji 1905 r. 37/39

RECENZENT

Jerzy Marzec

REDAKTOR INICJUJĄCY

Monika Borowczyk

REDAKCJA

Monika Poradecka

SKŁAD I ŁAMANIE

Mateusz Poradecki

KOREKTA TECHNICZNA

Leonora Gralka

PROJEKT OKŁADKI

Katarzyna Turkowska

Zdjęcie wykorzystane na okładce: © Depositphotos.com/leungchopan

Wydrukowano z gotowych materiałów dostarczonych do Wydawnictwa UŁ

© Copyright by Wojciech Grabowski, Łódź 2019

© Copyright for this edition by Uniwersytet Łódzki, Łódź 2019

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego

Wydanie I. W.08951.18.0.M

Ark. druk. 16,375

ISBN 978-83-8142-418-9

e-ISBN 978-83-8142-419-6

Wydawnictwo Uniwersytetu Łódzkiego

90-131 Łódź, ul. Lindleya 8

www.wydawnictwo.uni.lodz.pl

e-mail: ksiegarnia@uni.lodz.pl

tel. (42) 665 58 63

[Kup książkę](#)

Spis treści

Wstęp	9
Notacja wykorzystywana w monografii	13
1. Podstawowe modele wykorzystujące dane indywidualne	33
1.1. Wprowadzenie	33
1.2. Modele dla ciągłej zmiennej zależnej	34
1.2.1. Klasyczny model regresji liniowej	34
1.2.2. Heteroskedastyczność składnika losowego. Metody estymacji parametrów w przypadku niestajej wariancji	35
1.2.3. Metoda regresji kwantylowej	37
1.2.4. Odporna estymacja parametrów modelu regresji. Estymator M . Estymator S . Estymator MM	38
1.3. Model dwumianowy (dychotomiczny)	42
1.4. Model wielomianowy (polichotomiczny) kategorii uporządkowanych	51
1.5. Model wielomianowy kategorii nieuporządkowanych	52
1.6. Model regresji rankingowej	57
1.7. Problem selekcji próby w modelach ekonometrycznych. Model Heckmana	58
1.8. Model licznikowy	60
1.9. Dwurównaniowy model probitowy	65
1.10. Wielorównaniowy model probitowy	67
1.11. Endogeniczny model probitowy	69
1.12. Podsumowanie	74
2. Dane regionalne wykorzystywane w badaniach ekonomicznych	75
2.1. Podział administracyjny, statystyczny i historyczny Polski	75
2.2. Historyczno-kulturowe różnicowanie terytorium obecnej Rzeczypospolitej Polskiej	78
2.3. Źródła danych, które mogą być wykorzystywane w analizach regionalnych dla Polski	84
2.3.1. Bank Danych Lokalnych	85
2.3.2. Regional Innovation Scoreboard jako źródło informacji o poziomie innowacyjności regionów	91
2.3.3. Inne źródła danych regionalnych wykorzystywane w badaniach empirycznych	95
3. Liniowe modele wielopoziomowe	99
3.1. Wprowadzenie	99
3.2. Zmienne regionalne i sekcyjne w modelach ekonometrycznych opartych na danych indywidualnych	100
3.3. Podstawowy wielopoziomowy model regresji. Estymacja parametrów i predykcja efektów losowych	102
3.4. Efekty krzyżowe w liniowych modelach wielopoziomowych	109

6 Spis treści

3.5. Wykorzystanie liniowego modelu wielopoziomowego uwzględniającego zmienne regionalne do badania czynników wpływających na wynagrodzenia w Polsce	111
3.5.1. Przegląd literatury z zakresu czynników wpływających na wynagrodzenia	111
3.5.2. Konceptcje SBTC i RBTC i ich wykorzystanie do analizy czynników wpływających na różnice między wynagrodzeniami przedstawicieli określonych zawodów	114
3.5.3. Dane dotyczące poziomów wynagrodzeń uzyskiwanych przez pracowników w polskich przedsiębiorstwach. Podział zawodów ze względu na umiejętności posiadane przez pracowników	117
3.5.4. Specyfikacja modelu ekonometrycznego wykorzystywanego do analizy czynników wpływających na wynagrodzenia w Polsce	120
4. Uogólnione liniowe modele wielopoziomowe	141
4.1. Postać uogólnionego liniowego modelu wielopoziomowego	141
4.2. Funkcja wiarygodności w uogólnionym liniowym modelu wielopoziomowym	142
4.3. Estymacja parametrów uogólnionych liniowych modeli wielopoziomowych za pomocą metod aproksymacyjnych	149
4.4. Estymacja parametrów uogólnionych liniowych modeli wielopoziomowych za pomocą metod symulacyjnych	165
4.5. Problem selekcji próby w modelach wielopoziomowych. Estymacja parametrów wielorównaniowych modeli probitowych z efektami losowymi	170
4.6. Wykorzystanie wielopoziomowego modelu zmiennych dyskretnych do analizy zależności między wykorzystywaniem technologii informacyjnych i komunikacyjnych, innowacyjnością a produktywnością	176
4.6.1. Przegląd literatury z zakresu czynników wpływających na innowacyjność firm	176
4.6.2. Model CDM rozszerzony o wykorzystanie TliK oraz uwzględniający czynniki regionalne	179
4.6.2.1. Dane oraz próba badawcza	179
4.6.2.2. Specyfikacja modelu ekonometrycznego	189
4.6.2.3. Wyniki estymacji i interpretacja	193
5. Wielopoziomowy polichotomiczny model logitowy. Wielopoziomowy model regresji rankingowej	215
5.1. Wprowadzenie	215
5.2. Wielopoziomowy model wielomianowy logitowy. Wielopoziomowy model regresji rankingowej	216
5.3. Wykorzystanie wielopoziomowego polichotomicznego nieuporządkowanego modelu logitowego do analizy czynników wpływających na sposób reakcji wobec zaistnienia problemu prawnego	221
5.3.1. Czynniki wpływające na sposób reakcji wobec wystąpienia problemu prawnego – przegląd literatury	221
5.3.2. Estymacja parametrów wielopoziomowego, nieuporządkowanego, polichotomicznego modelu logitowego na podstawie danych pochodzących z badania dla Polski przeprowadzonego przez Instytut Spraw Publicznych w Warszawie	223
Zakończenie	241
Bibliografia	243

	Spis treści	7
Abstract		253
Spis rysunków		257
Spis tabel		259
Od Redakcji		261

Wstęp

W badaniach ekonomicznych i społecznych coraz więcej uwagi poświęca się analizie zależności występujących na poziomie indywidualnym. Powszechne są badania wykorzystujące między innymi dane dotyczące decyzji podejmowanych w przedsiębiorstwach (np. serie badań pt. *Community Innovation Survey* przeprowadzane przez urzędy statystyczne krajów Unii Europejskiej), aktywności ekonomicznej ludności (np. serie badań aktywności ekonomicznej ludności przeprowadzanych w różnych krajach Unii Europejskiej) czy też wynagrodzeń uzyskiwanych przez pracowników (badania struktury wynagrodzeń). Oprócz wymienionych wyżej oraz innych badań cyklicznych przeprowadzane są badania jednorazowe, w których jednostkami są firmy, pracownicy, gospodarstwa domowe, respondenci itp. Niektóre dane pochodzące z tych badań są publicznie dostępne, inne zaś mogą zostać zakupione lub pozyskane przez instytucje naukowo-badawcze. Zwiększa się zatem pole do zastosowań metod mikroekonometrycznych.

W badaniach mikroekonomicznych i społecznych często ignorowana jest rola kontekstu. Przyjmuje się założenia, że zależności występują między jednostkami, a lokalizacja gospodarstwa domowego czy firmy nie ma wpływu na proces podejmowania decyzji. Ewentualne różnice między zachowaniami respondentów mieszkających w innych regionach czy też różnice w procesie decyzyjnym firm znajdujących się w różnych sekcjach PKD traktuje się jako ustalone. Oznacza to zatem, że oszacowania parametrów przy odpowiednich zmiennych zero-jedynkowych odzwierciedlają te różnice. Istnieją jednak metody badania zależności między cechami przedsiębiorstw czy gospodarstw domowych umożliwiające analizę losowych różnic między jednostkami należącymi do innych sekcji czy regionów. Metody te umożliwiają również analizę różniącego się w poszczególnych grupach (którymi mogą być odpowiednie sekcje lub regiony) wpływu określonych cech firm czy gospodarstw domowych na podejmowane przez nie decyzje. Analizowane metody wykorzystuje się podczas estymacji parametrów modeli wielopoziomowych.

Niniejsza monografia poświęcona jest aspektom teoretycznym oraz badaniom empirycznym wykorzystującym modele wielopoziomowe. Prezentowane są zastosowania omawianych modeli w badaniach mikroekonomicznych (dotyczących wynagrodzeń pracowników oraz postaw innowacyjnych przedsiębiorstw wykorzystujących technologie informatyczne i komunikacyjne) oraz socjologicznych (związanych z zagadnieniami z zakresu socjologii prawa). Idea tych badań polega

na dodatkowym uwzględnieniu czynników kontekstowych (zwłaszcza związanych z przynależnością firmy, pracownika czy respondenta do regionu) w modelach wykorzystujących dane indywidualne. Jednocześnie analizowana jest rola poszczególnych czynników w wyjaśnieniu zmienności zmiennej zależnej oraz wskazywane są różnice między wynikami uzyskanymi dla modelu pełnego a tymi otrzymanymi w przypadku nieuwzględnienia zmiennych kontekstowych, czyli związanych z lokalizacją jednostki.

W badaniach empirycznych omawianych w niniejszej monografii wykorzystywane są przede wszystkim dane indywidualne uzyskane podczas realizacji grantów Narodowego Centrum Nauki (zakończonych i będących w trakcie realizacji), w których uczestniczył autor. Dane dotyczące wynagrodzeń pracowników pochodzą z baz związanych z badaniami struktury wynagrodzeń, zakupionymi podczas realizacji grantu pt. „Polaryzacja polskiego rynku pracy w kontekście zmiany technologicznej” o numerze 2016/23/B/HS4/00334. Dane związane z działalnością innowacyjną przedsiębiorstw zostały uzyskane na podstawie badania ankietowego przeprowadzonego w 2015 roku podczas realizacji grantu pt. „Wpływ technologii informacyjnych i telekomunikacyjnych na produktywność – analiza mikro- i makroekonomiczna” o numerze 2013/11/B/HS4/00661. Dane dotyczące faktu zaistnienia problemu prawnego oraz sposobu reakcji na niego pochodzą z badania przeprowadzonego w 2012 roku przez Instytut Spraw Publicznych w Warszawie pt. „Korzystający i niekorzystający z poradnictwa prawnego i obywatelskiego”. Dane te zostały zakupione podczas realizacji grantu Narodowego Centrum Nauki pt. „Nieodpłatna pomoc prawna w Polsce z perspektywy ekonomicznej analizy prawa. Stan obecny i rekomendowany” o numerze 2012/07/B/HS4/02994.

W przypadku każdego z tych trzech badań oprócz informacji indywidualnych wykorzystywane są także dane kontekstowe (np. związane z lokalizacją, przynależnością firmy do sekcji PKD czy też przynależnością pracownika do grupy zawodowej). Należy jednak podkreślić, że prezentowane wyniki są komplementarne względem rezultatów uzyskanych w innych pracach autora wykorzystujących te właśnie dane (por. m.in. Arendt, Grabowski, 2017; 2018; Florczak, Grabowski, 2017; 2018a; 2018b; 2018c; Szczygielski, Grabowski, Woodward, 2017; Szczygielski, Grabowski, Pamukcu, Tandogan, 2017). Różnica polega na dodatkowym uwzględnieniu czynników kontekstowych w modelach mikroekonometrycznych.

Niniejsza monografia składa się z pięciu rozdziałów. W rozdziale pierwszym prezentowane są metody wykorzystywane do analizy zależności na poziomie indywidualnym, ale bez uwzględniania zmiennych kontekstowych. W rozdziale drugim omawiane są różnice w poziomie rozwoju ekonomiczno-społecznego między polskimi regionami historycznymi i administracyjnymi. Jednocześnie prezentowane są bazy danych regionalnych, z których część wykorzystywana jest

w badaniach empirycznych uwzględnionych w monografii. Rozdział trzeci zawiera opis metody estymacji parametrów i predykcji efektów losowych w modelach wielopoziomowych z ciągłą zmienną zależną. Oprócz tego prezentowane są rezultaty badania empirycznego poświęconego determinantom zróżnicowania wynagrodzeń oraz testowaniu hipotezy o występowaniu polaryzacji na polskim rynku pracy. W rozdziale czwartym prezentowane są uogólnione liniowe modele wielopoziomowe. Szczegółowo omawiane są metody estymacji parametrów tych modeli. Jednocześnie prezentowane są wyniki badania empirycznego mającego na celu identyfikację indywidualnych i regionalnych czynników kształtujących decyzje innowacyjne przedsiębiorstw. W rozdziale piątym rozważane są modele wielopoziomowe dla przypadku wielomianowej nieuporządkowanej oraz rankingowej zmiennej zależnej. Dodatkowo prezentowane są wyniki badania empirycznego mającego na celu identyfikację czynników wpływających na prawdopodobieństwo doświadczenia problemu prawnego oraz sposobu reakcji na niego.

Autor pragnie podziękować współautorom wcześniejszych prac, z których pochodziły inspiracje do przeprowadzenia badań empirycznych. Dzięki współpracy autor uzyskał niezbędną wiedzę, która pomogła w specyfikacji odpowiednich modeli ekonometrycznych. Autor składa podziękowania Łukaszowi Arendtowi, Karolowi Korczakowi, Krzysztofowi Szczypielskiemu, Sinanowi Tandoganowi, Teomanowi Pamukcu, Richardowi Woodwardowi. Wyniki niektórych badań były prezentowane podczas zebrań w ramach seminarium naukowego pt. „Modelowanie gospodarki narodowej”. Autor pragnie podziękować uczestnikom tych zebrań, w tym przede wszystkim Aleksandrowi Welfe, Michałowi Majsterkowi, Robertowi Kelmowi, Annie Staszewskiej-Bystrovej, Piotrowi Kębłowskiemu, Piotrowi Karpowi, Emilii Gosińskiej, Katarzynie Leszkiewicz-Kędzior, Aleksandrze Majchrowskiej, Sylwii Roszkowskiej, Iwonie Świeczewskiej, Jakubowi Boratyńskiemu za wnikliwe i szczegółowe uwagi, które często przyczyniały się do poprawy jakości uzyskiwanych rezultatów. Szczególne podziękowania autor składa Ewie Stawasz-Grabowskiej oraz Michałowi Majsterkowi za cierpliwość w lekturze całości monografii. Ewentualne niedociągnięcia i błędy należy zaliczyć na konto autora.

Notacja wykorzystywana w monografii

Ze względu na dużą liczbę wzorów oraz innych symboli pojawiających się w niniejszej monografii użyteczne wydaje się przedstawienie indeksu wzorów i symboli. Czytelnik analizujący wzory i przekształcenia znajdujące się w kolejnych rozdziałach może odwoływać się do niego w celu upewnienia się, jak odczytywać określone oznaczenia.

Indeksowanie

$i = 1, \dots, I$ – jednostki.

$j = 1, \dots, J$ – grupy (klastry) w ogólnym modelu wielopoziomowym.

$qq = 1, \dots, QQ$ – grupy dla efektów losowych.

$w = 1, \dots, W$ – województwa (w modelu ogólnym).

$s = 1, \dots, S$ – sekcje (w modelu ogólnym).

$j1 = 1, \dots, J1$ – gminy w przykładowym modelu wielopoziomowym rozważanym w podrozdziale 3.3.

$j2 = 1, \dots, J2$ – powiaty w przykładowym modelu wielopoziomowym rozważanym w podrozdziale 3.3.

$j3 = 1, \dots, J3$ – województwa w przykładowym modelu wielopoziomowym rozważanym w podrozdziale 3.3.

$jj1 = 1, \dots, JJ1$ – grupy PKD w przykładowym modelu wielopoziomowym z efektami krzyżowymi rozważanym w podrozdziale 3.4.

$jj2 = 1, \dots, JJ2$ – działy PKD w przykładowym modelu wielopoziomowym z efektami krzyżowymi rozważanym w podrozdziale 3.4.

$jj3 = 1, \dots, JJ3$ – sekcje PKD w przykładowym modelu wielopoziomowym z efektami krzyżowymi rozważanym w podrozdziale 3.4.

$ss = 2, \dots, SS$ – poziomy zagnieżdżenia.

$l = 1, \dots, L$ – wybory dokonywane przez jednostkę w uporządkowanym modelu polichotomicznym.

$r = 1, \dots, RA$ – rankingi.

$k = 1, \dots, K$ – zmienne egzogeniczne.

$k = 1, \dots, K1$ – zmienne egzogeniczne wpływające na wynik jednostki na poziomie województw (podrozdział 3.3).

$k = 1, \dots, K2$ – zmienne egzogeniczne wpływające na wynik jednostki na poziomie powiatów (podrozdział 3.3).

$k = 1, \dots, K3$ – zmienne egzogeniczne wpływające na wynik jednostki na poziomie gmin (podrozdział 3.3).

$k = 1, \dots, K4$ – zmienne egzogeniczne wpływające na wynik jednostki na poziomie indywidualnym (podrozdział 3.3).

$k = 1, \dots, \tilde{K}$ – zmienne egzogeniczne dostępne na poziomie indywidualnym, definiowane we wprowadzeniu do modelu wielopoziomowego (podrozdział 3.3).

$k = \tilde{K} + 1, \dots, \tilde{K} + \hat{K}$ – zmienne egzogeniczne obserwowalne na poziomie grupowym, definiowane we wprowadzeniu do modelu wielopoziomowego (podrozdział 3.3).

$k = 1, \dots, \tilde{K}$ – zmienne egzogeniczne, których oddziaływanie na zmienną zależną różni się między grupami.

$m = 1, \dots, M$ – równania w standardowym i wielopoziomowym, wielorównaniowym modelu probitowym.

$n = 1, \dots$ – iteracje w przypadkach stosowania metod symulacyjnych.

$h = 1, \dots, H$ – replikacje w metodzie bootstrap oraz MCMC.

$p = 1, \dots, P$ – progi w modelu polichotomicznym uporządkowanym.

$d = 1, \dots, D$ – strukturyzacje w modelu wielopoziomowym z efektami krzyżowymi.

$g = 1, \dots, G$ – grupy podczas omawiania testu Hosmera-Lemeshowa.

$t = 1, \dots, T$ – indeks czasu.

$b = 1, \dots, B$ – warianty do uszeregowania w modelu regresji rankingowej.

Notacja dla zmiennej zależnej

y – zmienna zależna.

$\mathbf{y} = [y_1 \quad \dots \quad y_I]^T$ – wektor obserwacji na zmiennej objaśnianej.

$\mathbf{y}_j = [y_{j1} \quad \dots \quad y_{jI_j}]^T$ – wektor obserwacji dla j -tej grupy (j -tego klastra).

$\mathbf{y} = [\mathbf{y}_1^T \quad \dots \quad \mathbf{y}_J^T]^T$ – wektor wszystkich obserwacji na zmiennej objaśnianej, składający się z wektorów obserwacji dla poszczególnych klastrów.

y^* – zmienna nieobserwowalna, związana ze zmienną dwumianową lub uporządkowaną.

$y_{(m)}$ – m -ta zmienna zależna w wielorównaniowym modelu probitowym.

$\tilde{\mathbf{y}}$ – wektor endogenicznych regresorów w endogenicznym modelu probitowym.

Notacja dla zmiennych niezależnych

x – zmienna niezależna.

$\mathbf{X}$ – macierz obserwacji na wszystkich zmiennych objaśniających.

x_i – i -ty wiersz macierzy $\mathbf{X}$, odpowiadający wektorowi wartości na zmiennych objaśniających dla i -tej jednostki.

$\mathbf{x}_i^l$ – wektor obserwacji na zmiennych objaśniających w równaniu związanym z l -tym wyborem w modelu polichotomicznym nieuporządkowanym.

- $\tilde{\mathbf{x}}_i$ – wektor wszystkich zmiennych objaśniających w endogenicznym modelu probitowym.
- $\tilde{\mathbf{x}}_{(1)i}$ – wektor zmiennych egzogenicznych wpływających bezpośrednio na zmienną wynikową w endogenicznym modelu probitowym.
- $\tilde{\mathbf{x}}_{(2)i}$ – wektor zmiennych instrumentalnych w endogenicznym modelu probitowym.
- $\mathbf{x}^{(ss)}$ – wektor obserwacji na zmiennych objaśniających na s -tym poziomie zagnieżdżenia (w modelu wielopoziomowym).
- $\mathbf{X}_{[1]}$ – macierz obserwacji na zmiennych objaśniających obserwowanych na poziomie indywidualnym.
- $\mathbf{X}_{[2]}$ – macierz obserwacji na zmiennych objaśniających dostępnych na poziomie grupowym.
- $\mathbf{X}_{[3]}$ – podmacierz macierzy $\mathbf{X}_{[1]}$ zawierająca zmienne obserwowalne na poziomie indywidualnym, których wpływ na regresanta losowo różni się między grupami (np. regionami).
- $\mathbf{X}_{[3]}^{(j)}$ – podmacierz macierzy $\mathbf{X}_{[3]}$ zawierająca wektory zerowe dla jednostek nie należących do j -tej grupy oraz wektory odpowiadające wektorom macierzy $\mathbf{X}_{[3]}$ dla jednostek należących do j -tej grupy.
- $\mathbf{X}_j$ – macierz obserwacji na zmiennych objaśniających dla konkretnej j -tej grupy (j -tego klastra).
- $\mathbf{X} = \begin{bmatrix} \mathbf{X}_1^T & \dots & \mathbf{X}_j^T \end{bmatrix}^T$ – macierz obserwacji na zmiennych objaśniających zawierająca podmacierze składające się z macierzy obserwacji dla poszczególnych grup (klastrów).
- $\mathbf{x}_{(m)i}$ – wektor obserwacji dla i -tej jednostki na zmiennych występujących w m -tym równaniu w wielorównaniowych modelach probitowych.
- $\tilde{\mathbf{x}}_i$ – wektor wszystkich zmiennych egzogenicznych w endogenicznym modelu probitowym.
- $\tilde{\mathbf{x}}_{(1)i}$ – wektor zmiennych egzogenicznych niebędących instrumentami w endogenicznym modelu probitowym.
- $\tilde{\mathbf{x}}_{(2)i}$ – wektor instrumentów w endogenicznym modelu probitowym.
- $\mathbf{w}_i$ – wektor regresorów wpływających na selekcję w modelu Heckmana.
- $\mathbf{zz}_i$ – wektor regresorów wpływających na to, czy zmienna licznikowa przyjmuje wartość 0 w modelu licznikowym z podwyższoną liczbą „zer”.
- $\tilde{\mathbf{x}}_i^{\text{woj},w}$ – wektor obserwacji na zmiennych objaśniających dla i -tej jednostki, pod warunkiem, że należy ona do w -tego województwa. W przeciwnym przypadku wektor ten składa się z elementów zerowych.
- $\tilde{\mathbf{x}}_i^{\text{sek},s}$ – wektor obserwacji na zmiennych objaśniających dla i -tej jednostki, pod warunkiem, że należy ona do s -tej sekcji. W przeciwnym wypadku wektor ten zawiera tylko elementy zerowe.

woj – wektor zmiennych binarnych związanych z przynależnością jednostki do województwa.

sek – wektor zmiennych binarnych związanych z przynależnością jednostki do sekcji.

xw_i – wektor obserwacji na zmiennych objaśniających wyjaśniających selekcję w modelu Heckmana.

ww_i – wektor obserwacji na zmiennych różniących się ze względu na województwa.

wv_i – wektor obserwacji na zmiennych różniących się ze względu na sekcje.

w3 – zmienne wpływające na wartość kategorii wynikowej na poziomie województw (przykład w podrozdziale 3.3).

w2 – zmienne wpływające na wartość kategorii wynikowej na poziomie powiatów (przykład w podrozdziale 3.3).

w1 – zmienne wpływające na wartość kategorii wynikowej na poziomie gmin (przykład w podrozdziale 3.3).

w0 – zmienne wpływające na wartość kategorii wynikowej na poziomie indywidualnym (przykład w podrozdziale 3.3).

Notacja dla pozostałych ważnych zmiennych

UZ_i^l – użyteczność *i*-tej jednostki z wyboru *l*-tego wariantu w modelu polichotomicznym nieuporządkowanym.

VZ_i^l – część deterministyczna użyteczności *i*-tej jednostki z wyboru *l*-tego wariantu w modelu polichotomicznym nieuporządkowanym.

ID_i – zmienna binarna, która w modelu z podwyższoną liczbą „zer” informuje, czy zmienna licznikowa jest równa 0, czy też przyjmuje wartość dodatnią.

K_i – zmienna przekształcająca zmienną binarną w inną zmienną dwuwartościową, przyjmującą wartości -1 (gdy przekształcana zmienna binarna wynosi 0) oraz 1 (dla zmiennej binarnej równej 1).

p_i^l – prawdopodobieństwo wyboru *l*-tego wariantu przez *i*-tą jednostkę.

p_{ij}(*p*) – prawdopodobieństwo, że obserwowalna zmienna zależna w uporządkowanym modelu polichotomicznym przyjmie wartość *p* dla *i*-tej jednostki z *j*-tego klastra.

d_i^l – zmienna binarna przyjmująca wartość 1, jeśli *i*-ta jednostka wybrała *l*-ty wariant i 0 w przeciwnym przypadku.

δ_i^{ll'} – zmienna binarna przyjmująca wartość 1, jeśli dla *i*-tej jednostki wariant *l*-ty jest preferowany w stosunku do wariantu *l'*.

Notacja dla elementów związanych z efektami losowymi w modelach wielopoziomowych

Z – macierz przy efektach losowych. Składa się ona głównie ze zmiennych zero-jedynkowych definiujących przynależność określonych jednostek do poszczególnych klastrów (grup), a także z tych zmiennych wchodzących w skład macierzy **X**, których oddziaływanie na zmienną zależną różni się między grupami (klastrami).

$\mathbf{z}_i^{ZP}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do posiadania technologii informacyjnych i komunikacyjnych w procesach biznesowych związanych z zarządzaniem przedsiębiorstwem.

$\mathbf{z}_i^{ERP}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do posiadania technologii informacyjnych i komunikacyjnych w procesach biznesowych związanych z zarządzaniem zasobami przedsiębiorstwa.

$\mathbf{z}_i^{CAD}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do posiadania technologii informacyjnych i komunikacyjnych w procesach biznesowych związanych ze wsparciem dla projektowania i wytwarzania CAD/CAM.

$\mathbf{z}_i^{SM}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do posiadania technologii informacyjnych i komunikacyjnych w procesach biznesowych związanych ze sterowaniem maszynami lub linią produkcyjną.

$\mathbf{z}_i^{INW}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do inwestowania w rozwój technologii informacyjnych i telekomunikacyjnych.

$\mathbf{z}_i^{BR}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do posiadania własnego wydziału B+R.

$\mathbf{z}_i^{PROD}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do wprowadzania innowacji produktowych.

$\mathbf{z}_i^{PROC}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do wprowadzania innowacji procesowych lub organizacyjnych.

$\mathbf{z}_i^{MARK}$ – wektor przy efektach losowych w równaniu wyjaśniającym skłonność firm do wprowadzania innowacji marketingowych.

$\mathbf{z}_i^{PR}$ – wektor przy efektach losowych w równaniu wyjaśniającym prawdopodobieństwo zaistnienia problemu prawnego.

Notacja dla składników losowych i efektów losowych

$\boldsymbol{\varepsilon}$ – wektor składników losowych.

$\boldsymbol{\varepsilon}_{(m)}$ – wektor składników losowych dla m -tego równania w modelach zawierających więcej niż jedno równanie.

ε_i^l – składnik losowy w równaniu związanym z l -tym wyborem w modelu polichotomicznym nieuporządkowanym.

$\varepsilon_{z_{ij}}^l$ – składnik losowy związany z i -tą jednostką należącą do j -tej grupy oraz l -tym wyborem w wielopoziomowym nieuporządkowanym modelu polichotomicznym.

$\hat{\boldsymbol{\varepsilon}}$ – reszty.

$\tilde{\boldsymbol{\varepsilon}}$ – składniki losowe (po ortogonalizacji) w wielorównaniowym modelu probitowym.

$\mathbf{u}$ – wektor wszystkich efektów losowych.

$\mathbf{u}_{[1]}$ – podwektor wektora efektów losowych związany z losowym wpływem poszczególnych zmiennych egzogenicznych na zmienną zależną.

$\mathbf{u}_{[2]}$ – podwektor wektora efektów losowych związany z losowym wyrazem wolnym.

$\mathbf{u}_j$ – efekty losowe dla poszczególnych klastrów.

$\mathbf{u}^{(ss)}$ – efekty losowe na ss-tym poziomie zagnieżdżenia.

$\mathbf{u} = [\mathbf{u}_1^T \quad \dots \quad \mathbf{u}_J^T]^T$ – wektor efektów losowych.

$\mathbf{u}_{(h)}$ – wektor efektów losowych dla h -tej replikacji podczas wykorzystywania metody MCMC w celu estymacji parametrów uogólnionego liniowego modelu wielopoziomowego.

$\mathbf{u}_{\{j\}}$ – wektor zawierający efekty losowe w przypadku j -tej grupy oraz elementy „zerowe” dla pozostałych grup.

$\mathbf{u}^{ZP}$ – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie zarządzania produkcją.

ε^{ZP} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie zarządzania produkcją.

$\mathbf{u}^{ERP}$ – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie zarządzania zasobami przedsiębiorstwa.

ε^{ERP} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie zarządzania zasobami przedsiębiorstwa.

$\mathbf{u}^{CAD}$ – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie wsparcia dla projektowania i wytwarzania CAD/CAM.

ε^{CAD} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie wsparcia dla projektowania i wytwarzania CAD/CAM.

$\mathbf{u}^{SM}$ – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie sterowania maszynami lub linią produkcyjną.

ε^{SM} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wykorzystywania technologii informacyjnych i komunikacyjnych w zakresie sterowania maszynami lub linią produkcyjną.

$\mathbf{u}^{INW}$ – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do inwestowania w rozwój technologii informatycznych i komunikacyjnych.

ε^{INW} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do inwestowania w rozwój technologii informatycznych i komunikacyjnych.

- u^{BR} – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do posiadania wewnętrznego wydziału B+R.
- ε^{BR} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do posiadania wewnętrznego wydziału B+R.
- u^{PROD} – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wprowadzania innowacji produktowych.
- ε^{PROD} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wprowadzania innowacji produktowych.
- u^{PROC} – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wprowadzania innowacji procesowych lub organizacyjnych.
- ε^{PROC} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wprowadzania innowacji procesowych lub organizacyjnych.
- u^{MARK} – efekty losowe w równaniu wyjaśniającym skłonność przedsiębiorstw do wprowadzania innowacji marketingowych.
- ε^{MARK} – składnik losowy w równaniu wyjaśniającym skłonność przedsiębiorstw do wprowadzania innowacji marketingowych.
- u^{PR} – wektor efektów losowych w modelu wyjaśniającym prawdopodobieństwo zaistnienia problemu prawnego.
- ε^{PR} – składnik losowy w modelu wyjaśniającym prawdopodobieństwo zaistnienia problemu prawnego.

Główne parametry i estymatory

- β – parametr ilustrujący wpływ zmiennej objaśniającej na zmienną zależną w większości przypadków.
- β – wektor parametrów.
- $\beta_{(m)}$ – wektor parametrów związany z m -tym równaniem w modelach wielorównaniowych.
- $\beta_{[1]}$ – wektor parametrów przy zmiennych obserwowalnych na poziomie jednostek.
- $\beta_{[2]}$ – wektor parametrów przy zmiennych dostępnych na poziomie grupowym.
- $\beta_{[3]}^{(j)}$ – wektor parametrów przy zmiennych wchodzących w skład macierzy $X_{[3]}^{(j)}$.
- $\beta_{[q]}$ – wektor parametrów dla kwantyla rzędu q w metodzie regresji kwantylowej.
- $\hat{\beta}_{OLS}$ – estymator uzyskany klasyczną metodą najmniejszych kwadratów.
- $\hat{\beta}_{GLS}$ – estymator uzyskany uogólnioną metodą najmniejszych kwadratów.
- $\hat{\beta}_{FGLS}$ – estymator uzyskany uogólnioną metodą najmniejszych kwadratów z estymacją.
- $\hat{\beta}_{ML}$ – estymator uzyskany metodą największej wiarygodności.

$\hat{\beta}_Q$ – estymator regresji kwantylowej.

$\hat{\beta}_M$ – estymator M .

$\hat{\beta}_S$ – estymator S .

$\hat{\beta}_{MM}$ – estymator MM .

$\hat{\beta}_{BL}$ – estymator parametrów uogólnionego liniowego modelu wielopoziomowego, uzyskany w wyniku zastosowania korekty Breslowa i Lina.

$\hat{\beta}_{\{un\}}$ – estymator nieskorygowany parametrów uogólnionego liniowego modelu wielopoziomowego.

β^l – wektor parametrów związanych z l -tym wyborem w modelu polichotomicznym nieuporządkowanym.

$\tilde{\beta}^{woj,w}$ – wektor parametrów związanych z w -tym województwem

$\tilde{\beta}^{sek,s}$ – wektor parametrów związanych z s -tą sekcją.

$\beta^{\{n\}}$ – wektor parametrów w n -tej iteracji.

β^{ZP} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TliK w procesach biznesowych związanych z zarządzaniem produkcją.

β^{ERP} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TliK w procesach biznesowych związanych z zarządzaniem zasobami przedsiębiorstwa.

β^{CAD} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TliK w procesach biznesowych związanych ze wsparciem dla projektowania i wytwarzania CAD/CAM.

β^{SM} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TliK w procesach biznesowych związanych ze sterowaniem maszynami lub linią produkcyjną.

β^{INW} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do inwestowania w rozwój technologii informacyjnych i komunikacyjnych.

β^{BR} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do posiadania własnego wydziału B+R.

β^{PROD} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wprowadzania innowacji produktowych.

β^{PROC} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wprowadzania innowacji procesowych lub organizacyjnych.

β^{MARK} – wektor parametrów przy zmiennych objaśniających w równaniu wyjaśniającym skłonność do wprowadzania innowacji marketingowych.

β^{PR} – wektor parametrów (stałych) w równaniu wyjaśniającym prawdopodobieństwo zaistnienia problemu prawnego.

$\tilde{b}(\Theta)$ – wartość oczekiwana oszacowania uzyskanego metodą quasi-największej wiarygodności w uogólnionym liniowym modelu wielopoziomowym, gdy Θ jest wektorem prawdziwych parametrów.

$\Sigma = \sigma^2 W$ – macierz kowariancji między składnikami losowymi dla różnych obserwacji.

$\tilde{\Sigma}$ – macierz kowariancji między składnikami losowymi z różnych równań w standardowym i wielopoziomowym, wielorównaniowym modelu probitowym.

$\Sigma_{\varepsilon 2 \varepsilon 2}$ – macierz kowariancji między składnikami losowymi wchodzącymi w skład wektora $\varepsilon_{(2)i}$ (endogeniczny model probitowy).

$\Sigma_{\varepsilon 2 \varepsilon 1}$ – wektor składający się z kowariancji między składnikami losowymi wchodzącymi w skład wektora $\varepsilon_{(2)i}$ a składnikiem losowym $\varepsilon_{(1)i}$ (endogeniczny model probitowy).

$\sigma_{\varepsilon 1}^2$ – wariancja składnika losowego $\varepsilon_{(1)i}$ w endogenicznym modelu probitowym.

Ω – macierz kowariancji między efektami losowymi.

$\Omega^{(ss)}$ – macierz kowariancji między efektami losowymi na s -tym poziomie zagnieźdżenia.

$\tilde{Q}^{(ss)}$ – dekompozycja Choleskiego macierzy $\Omega^{(s)}$.

θ – wektor zawierający unikatowe elementy macierzy Ω .

Θ – wektor zawierający wszystkie parametry do estymacji w uogólnionym liniowym modelu wielopoziomowym.

V – macierz kowariancji między obserwacjami na zmiennej zależnej.

ρ – współczynnik korelacji między składnikami losowymi z dwóch równań w dwurównaniowym modelu probitowym.

$\rho_{mm'}$ – współczynnik korelacji w wielorównaniowym modelu probitowym między składnikami losowymi z równań m' oraz m .

τ_p – parametr progowy w modelu polichotomicznym uporządkowanym.

τ – wektor składający się z parametrów progowych w modelu polichotomicznym uporządkowanym.

$\kappa\kappa$ – parametr progowy związany z funkcją straty (1) w estymacji odpornej.

cc – parametr progowy związany z funkcją straty (2) w estymacji odpornej.

ccc – punkt progowy wyznaczany podczas mierzenia jakości dopasowania w modelu dwumianowym.

$\ddot{M}(s)$ – liczba efektów losowych na poziomie s .

μ_q – wartość kwantyla rzędu q .

- λ – parametr intensywności w modelu Poissona.
- α – parametr związany z nadwyżką wariancji ponad wartość oczekiwaną w modelu ujemnym dwumianowym.
- ϱ – wektor parametrów ilustrujących wpływ przynależności do województw na wartość zmiennej wynikowej (podrozdział 3.2).
- $\ddot{\alpha}$ – wektor parametrów ilustrujących wpływ przynależności do sekcji na wartość zmiennej wynikowej (podrozdział 3.2).
- ς – wektor parametrów mierzących wpływ endogenicznych regresorów na wartość zmiennej wynikowej w endogenicznym modelu probitowym.
- π – wektor parametrów odzwierciedlających wpływ zmiennych egzogenicznych na endogeniczne regresory w endogenicznym modelu probitowym.
- J – macierz przekształcająca wektor obserwacji na wszystkich zmiennych egzogenicznych w wektor niezawierający instrumentów w endogenicznym modelu probitowym.
- MA – macierz zawierająca „zera” i „jedynki” w modelu wielopoziomowym. Przypisanie wartości 1 lub 0 zależy od tego, czy jednostka należy do danej grupy, czy nie.
- W_0 – macierz wag wykorzystywana podczas aproksymacji funkcji wiarygodności za pomocą propozycji Longforda w modelu wielopoziomowym.
- $\tilde{\Theta}$ – estymator dla wektora wszystkich parametrów uzyskany w wyniku maksymalizacji funkcji quasi-największej wiarygodności.
- $\hat{\Theta}_{BC,KUK}^{\{n\}}$ – wektor oszacowań wszystkich parametrów modelu wielopoziomowego w n -tej iteracji, wykorzystywany do symulacji bootstrap zgodnie z propozycją Kuka.
- $\hat{\Theta}_{BC,RM}^{\{n\}}$ – wektor oszacowań wszystkich parametrów modelu wielopoziomowego w n -tej iteracji, wykorzystywany podczas stosowania metody stochastycznej aproksymacji Robbinsa-Monro.
- $\tilde{\Theta}_{\langle h \rangle}^*$ – wektor oszacowań parametrów modelu wielopoziomowego dla h -tej replikacji w modelu wielopoziomowym podczas wykorzystania metod bootstrapowych.
- $\bar{\Theta}^*$ – średnia z oszacowań dla wszystkich replikacji podczas wykorzystania bootstrapowych metod korekty obciążenia.
- ζ – wektor odpowiadający kryterium zbieżności.
- $\ddot{\psi}$ – wektor parametrów mierzących wpływ zmiennych wchodzących w skład wektora ww na wartość zmiennej wynikowej.
- $\ddot{\omega}$ – wektor parametrów mierzących wpływ zmiennych wchodzących w skład wektora vv na wartość zmiennej wynikowej.

γ – wektor parametrów przy zmiennych w równaniu selekcji w modelu Heckmana.

$\dot{\gamma}$ – wektor parametrów przy zmiennych wpływających na fakt przyjmowania przez zmienną zależną „zerowej” wartości w modelu licznikowym z podwyższoną liczbą „zer”.

$\ddot{\pi}$ – parametry mierzące wpływ zmiennych w_3 na wartość zmiennej wynikowej.

$\ddot{\gamma}$ – parametry mierzące wpływ zmiennych w_2 na wartość zmiennej wynikowej.

$\ddot{\beta}$ – parametry mierzące wpływ zmiennych w_1 na wartość zmiennej wynikowej.

$\ddot{\alpha}$ – parametry mierzące wpływ zmiennych w_0 na wartość zmiennej wynikowej.

$\widehat{p_g}$ – średnie prawdopodobieństwo, że zmienna zależna przyjmuje wartość 1 w g -tej grupie (dla testu Hosmera-Lemeshowa).

H – macierz zawierająca macierze π oraz J w endogenicznym modelu probitowym.

$MO = \begin{bmatrix} y^T & u^T \end{bmatrix}^T$ – macierz obserwacji na wszystkich zmiennych w modelu wielopoziomowym (obserwowalnych i nieobserwowalnych).

$\tilde{\mu}_q$ – wartość oczekiwana dla q -tego efektu losowego.

$\tilde{\tau}_q^2$ – wariancja dla q -tego efektu losowego.

uz_{ij}^l – użyteczność i -tej jednostki należącej do j -tego klastra związana z l -tym wyborem w wielopoziomowym, nieuporządkowanym modelu polichotomicznym.

fz_{ij}^l – deterministyczny składnik losowy, reprezentujący obserwowaną heterogeniczność wariantów do wyboru jednostek oraz klastrów w wielopoziomowym, nieuporządkowanym modelu polichotomicznym.

δz_{ij}^l – zmienna sztuczna, reprezentująca nieobserwowalną heterogeniczność w wielopoziomowym, nieuporządkowanym modelu polichotomicznym.

Główne funkcje

$E(\cdot|\cdot)$ – warunkowa wartość oczekiwana.

$F_{\xi}(\cdot)$ – dystrybuanta zmiennej losowej ξ .

$Q_I(\cdot)$ – funkcja celu w metodzie regresji kwantylowej.

$\rho\rho(\cdot)$ – funkcja straty w podrozdziale poświęconym odpornym metodom estymacji parametrów.

$\Phi(\cdot)$ – dystrybuanta standardowego rozkładu normalnego.

$\phi(\cdot)$ – funkcja gęstości standardowego rozkładu normalnego.

$\Phi_M(\cdot)$ – dystrybuanta M -wymiarowego rozkładu normalnego.

$\Lambda(\cdot)$ – dystrybuanta rozkładu logistycznego.

$H(\cdot)$ – dystrybuanta komplementarnego rozkładu log-log.

$I\{\cdot\}$ – funkcja wskaźnikowa przyjmująca wartość 1, gdy warunek zdefiniowany w klamrowym nawiasie jest spełniony.

$\ddot{I}_p(y_{ij})$ – funkcja przyjmująca wartość 1, jeśli zmienna uporządkowana dla i -tej obserwacji należącej do j -tego klastra przyjmuje wartość p oraz 0 w przeciwnym przypadku.

$qL(\Theta)$ – funkcja quasi-największej wiarygodności.

$qL_{S1}(\Theta)$ – funkcja quasi-największej wiarygodności wykorzystująca aproksymację Solomona-Coxa typu pierwszego.

$qL_{S2}(\Theta)$ – funkcja quasi-największej wiarygodności wykorzystująca aproksymację Solomona-Coxa typu drugiego.

$$\widetilde{\ln L_j(\Theta)}_{\langle 0 \rangle}^{[k]} = \frac{\partial^k \ln L_j(\Theta)}{\partial u_j^k} \text{ dla } u_j^k = 0.$$

$\bar{r}$ – funkcja wyrażająca ogólną postać funkcji quasi-największej wiarygodności.

$\pi_f(\cdot)$ – ogólna funkcja nieliniowego modelu wielopoziomowego.

H - L – statystyka testu Hosmera-Lemeshowa.

IMR – odwrócony iloraz Millsa.

$h(\cdot, \cdot)$ – łączna funkcja gęstości zmiennych y oraz $\tilde{y}$ w endogenicznym modelu probitowym.

$f_{ep}(\cdot)$ – funkcja gęstości rozkładu warunkowego y względem $\tilde{y}$ w endogenicznym modelu probitowym.

$g_{ep}(\cdot)$ – brzegowa funkcja gęstości rozkładu $\tilde{y}$ w endogenicznym modelu probitowym.

$g_l(\cdot)$ – funkcja łączącą w uogólnionych liniowych modelach wielopoziomowych.

$g_{uu}(\cdot)$ – funkcja gęstości wektora losowego.

$h_{uu}(\cdot)$ – rozkład próbkowy efektów losowych wykorzystywany podczas stosowania algorytmu Metropolisa-Hastingsa.

$f_{wj}(\cdot | \cdot)$ – funkcja gęstości warunkowego rozkładu zmiennej wynikowej względem efektów losowych w modelach wielopoziomowych dla j -tej grupy.

$f_w(\cdot | \cdot)$ – funkcja gęstości warunkowego rozkładu wektora wartości zmiennej wynikowej względem efektów losowych w modelach wielopoziomowych.

$f_{w_i}(\cdot | \cdot)$ – funkcja gęstości rozkładu warunkowego zmiennej wynikowej względem efektów losowych i parametrów dla i -tej jednostki.

$h_w(\cdot)$ – funkcja zależna od wszystkich parametrów i efektów losowych (wykorzystywana przy definiowaniu wkładu j -tej grupy do funkcji wiarygodności).

$\ln L$ – ogólny zapis dla logarytmu funkcji wiarygodności.

$\ln L_I^g$ – funkcja wiarygodności maksymalizowana w pierwszym kroku dla endogenicznego modelu probitowego.

$\ln L_I^f$ – funkcja wiarygodności maksymalizowana w drugim kroku dla endogenicznego modelu probitowego.

$L_j(\cdot)$ – wkład j -tej grupy do funkcji wiarygodności.

$L_{ij}(\cdot)$ – wkład i -tej jednostki należącej do j -tej grupy do funkcji wiarygodności.

$g_u(\cdot)$ – funkcja gęstości dla pojedynczego efektu losowego.

$\mu_i^u = E(y_i | \mathbf{u})$

$d_i(y_i, \mu_i^u)$ – funkcja zależna od zmiennej wynikowej i warunkowej wartości oczekiwanej zmiennej wynikowej względem efektów losowych.

$\mathbf{\ddot{K}} = (\boldsymbol{\beta}, \varphi)$.

$\tilde{L}_{ij} = \{l_{ij}^1, \dots, l_{ij}^{L_{ij}}\}$ – zbiór wszystkich możliwych wyborów dla i -tej jednostki należącej do j -tego klastra w wielopoziomym nieuporządkowanym modelu polichotomicznym.

Mierniki jakości dopasowania oraz mierniki wpływu zmiennych egzogenicznych na prawdopodobieństwo, że zmienna zależna przyjmuje określone wartości

R_{binary}^2 – pseudo R -kwadrat.

R_{BL}^2 – współczynnik determinacji Ben-Akiva i Lermana oraz Kaya i Little.

R_{Ef}^2 – współczynnik determinacji R^2 -Efrona.

R_{VZ}^2 – współczynnik determinacji Vealla i Zimmermana.

R_{MZ}^2 – współczynnik determinacji Zavoiny i McKelveya.

$R^2 - McFadden$ – współczynnik determinacji McFaddena.

$R^2 - Nagelkerke$ – współczynnik determinacji Nagelkerke'a.

CP – procent poprawnych predykcji.

$SENSITIVITY$ – czułość.

$SPECIFICITY$ – specyficzność.

PPV – pozytywna wartość predyktywna.

NPV – negatywna wartość predyktywna.

MPE_{ik}^l – wpływ marginalnej zmiany wartości k -tej zmiennej egzogenicznej na prawdopodobieństwo, że i -ta jednostka wybierze l -ty wariant.

$\hat{MPE}_k^l$ – efekt krańcowy dla średniej.

$AM\hat{PE}_k^l$ – średni efekt krańcowy.

Zmienne wykorzystywane w badaniach empirycznych

Rozdział 3

WYN_i – wynagrodzenie.

SZK_i – poziom wykształcenia mierzony liczbą lat nauki.

XZ_i – doświadczenie zawodowe mierzone liczbą przepracowanych lat.

$D1_i$ – zmienna zero-jedynkowa związana z wykształceniem podstawowym.

$D2_i$ – zmienna zero-jedynkowa związana z wykształceniem średnim.

$D3_i$ – zmienna zero-jedynkowa związana z wykształceniem wyższym.

ZK_i – zmienna kontrolna w równaniu wyjaśniającym wynagrodzenia (podrozdział 3.5).

$plac_nom_t$ – nominalne płace przeciętne (w modelu wyjaśniającym wynagrodzenia na poziomie makro).

cen_t – indeks cen konsumpcyjnych (CPI) (w modelu wyjaśniającym wynagrodzenia na poziomie makro).

wyd_prac_t – wydajność pracy (w modelu wyjaśniającym wynagrodzenia na poziomie makro).

$bezr_t$ – stopa bezrobocia (w modelu wyjaśniającym wynagrodzenia na poziomie makro).

WYN_NOM_{it} – wynagrodzenie nominalne.

WYN_REL_{it} – wynagrodzenie relatywne, będące ilorazem wynagrodzenia nominalnego i mediany wynagrodzeń.

$WYZSZE_i$ – zmienna binarna przyjmująca wartość 1 dla pracownika z wyższym wykształceniem.

$SREDNIE_ZAWODOWE_i$ – zmienna binarna przyjmująca wartość 1 dla pracownika z wykształceniem średnim technicznym lub policealnym.

$ZASADNICZE_ZAWODOWE_i$ – zmienna binarna przyjmująca wartość 1 dla pracownika z wykształceniem zasadniczym zawodowym.

$PODSTAWOWE_i$ – zmienna binarna przyjmująca wartość 1 dla pracownika z wykształceniem podstawowym.

$DOSW_FIRMA_i$ – liczba przepracowanych lat przez pracownika w firmie, w której obecnie pracuje.

$DOSW_OGOL_i$ – liczba pełnych lat przepracowanych przez pracownika.

$ROZMIAR_10_49_i$ – zmienna binarna przyjmująca wartość 1, jeśli firma (zatrudniająca i -tego pracownika) zatrudnia co najmniej 10 i mniej niż 50 osób.

$ROZMIAR_50_249_i$ – zmienna binarna przyjmująca wartość 1, jeśli firma (zatrudniająca i -tego pracownika) zatrudnia co najmniej 50 i mniej niż 250 osób.

$ROZMIAR_250_499_i$ – zmienna binarna przyjmująca wartość 1, jeśli firma (zatrudniająca i -tego pracownika) zatrudnia co najmniej 250 i mniej niż 500 osób.

$ROZMIAR_CON500_i$ – zmienna binarna przyjmująca wartość 1, jeśli firma (zatrudniająca i -tego pracownika) zatrudnia co najmniej 500 osób.

$KOBIETA_i$ – zmienna binarna przyjmująca wartość 1 w przypadku kobiet oraz 0 dla mężczyzn.

$SEKTOR_PRYWATNY_i$ – zmienna binarna przyjmująca wartość 1 w przypadku pracownika firmy z sektora prywatnego.

$NIEOKRESLONY_i$ – zmienna binarna przyjmująca wartość 1 w przypadku, gdy pracownik jest zatrudniony na czas nieokreślony.

$BEZR_{it}^w$ – stopa bezrobocia w okresie t w województwie w , będącym siedzibą firmy, w której pracuje i -ty pracownik.

$BEZR_t^k$ – stopa bezrobocia w Polsce w okresie t .

$\widehat{BEZR}_{it}^w$ – relatywna stopa bezrobocia w okresie t w województwie w , będącym siedzibą firmy, w której pracuje i -ty pracownik. Jest ona ilorazem stopy bezrobocia dla województwa do stopy bezrobocia dla całego kraju.

WYD_{it}^w – poziom wydajności pracy w okresie t w województwie w , będącym siedzibą firmy, w której pracuje i -ty pracownik.

$\widehat{WYD}_{it}^w$ – relatywny poziom wydajności pracy w okresie t w województwie w , będącym siedzibą firmy, w której pracuje i -ty pracownik.

Rozdział 4

ICT_KS – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w księgowości.

ICT_ZZL – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych z zarządzaniem zasobami ludzkimi.

ICT_ZZ – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych z zarządzaniem zaopatrzeniem.

ICT_ZP – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych z zarządzaniem produkcją.

ICT_CRM – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych z zarządzaniem sprzedażą i kontaktem z klientami.

ICT_ERP – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych z zarządzaniem zasobami przedsiębiorstwa.

ICT_CAD – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych ze wsparciem dla projektowania i wytwarzania CAD/CAM.

ICT_SM – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych ze sterowaniem maszynami lub linią produkcyjną.

ICT_ZPAB – zmienna binarna przyjmująca wartość 1 dla firm wykorzystujących TIiK w procesach biznesowych związanych z zarządzaniem pracami administracyjno-biuroowymi.

INNOW_PROD – zmienna binarna przyjmująca wartość 1 dla firm, które w ciągu ostatnich 24 miesięcy wprowadziły innowację produktową.

INNOW_PROC – zmienna binarna przyjmująca wartość 1 dla firm, które w ciągu ostatnich 24 miesięcy wprowadziły innowację procesową lub organizacyjną.

INNOW_MARKT – zmienna binarna przyjmująca wartość 1 dla firm, które w ciągu ostatnich 24 miesięcy wprowadziły innowację marketingową.

INWESTYCJE_ICT – zmienna binarna przyjmująca wartość 1 dla firm inwestujących w technologie informatyczne i komunikacyjne.

BR – zmienna binarna przyjmująca wartość 1 dla firm posiadających własny dział B+R.

ROZMIAR – logarytm z liczby osób zatrudnionych w firmie.

WWKK – zmienna binarna przyjmująca wartość 1 w przypadku firm, w których większość kadry kierowniczej posiada wyższe wykształcenie.

WWP – zmienna binarna przyjmująca wartość 1 w przypadku firm, w których większość szeregowych pracowników posiada wyższe wykształcenie.

MSWKK – zmienna binarna przyjmująca wartość 1 w przypadku firm stosujących motywacyjny system wynagradzania kadry kierowniczej.

MSWP – zmienna binarna przyjmująca wartość 1 w przypadku firm stosujących motywacyjny system wynagradzania pracowników.

Zasieg_KZ – zmienna binarna przyjmująca wartość 1 w przypadku firm o ogólnokrajowym lub zagranicznym zasięgu oddziaływania.

Ocena_okresowa – zmienna binarna przyjmująca wartość 1 w przypadku firm prowadzących okresową ocenę kompetencji pracowników pod kątem ich przydatności do potrzeb firmy.

SZKOLENIA_TIiK – zmienna binarna przyjmująca wartość 1 w przypadku firm organizujących dodatkowe szkolenia dla pracowników w związku z wdrażaniem TIiK.

BRANZA_PRZEM – zmienna binarna przyjmująca wartość 1 w przypadku firm z branży przemysłowej.

BRANZA_BUD – zmienna binarna przyjmująca wartość 1 w przypadku firm z branży budownictwo.

BRANZA_PHU – zmienna binarna przyjmująca wartość 1 w przypadku firm z branży produkcyjno-handlowo-usługowej.

ZES_ROB – zmienna binarna przyjmująca wartość 1 dla firm, w których tworzone są zespoły robocze.

DZIEL_INF – zmienna binarna przyjmująca wartość 1 w przypadku firm, w których istnieje zwyczaj dzielenia się informacjami istotnymi dla funkcjonowania firmy z pracownikami.

NOW_UM_INF – zmienna binarna przyjmująca wartość 1 dla firm, w których wszyscy nowo przyjmowani pracownicy mają wysokie umiejętności informatyczne.

ORG – zmienna ilustrująca gotowość firmy do przeprowadzenia zmiany organizacyjnej.

WZROSTOWA – zmienna binarna przyjmująca wartość 1 dla przedsiębiorstw, które odpowiedziały, że zarówno w 2014, jak i 2013 roku przychód w firmie był większy niż w poprzednim roku.

RIS1 – udział mieszkańców (w województwie, w którym zlokalizowana jest dana firma) z wyższym wykształceniem w populacji wszystkich osób w wieku 25–64 lat.

RIS2 – wydatki na badania i rozwój w sektorze publicznym w relacji do PKB.

RIS3 – wydatki na badania i rozwój w sektorze przedsiębiorstw w relacji do PKB.

RIS4 – wydatki na innowacje firm małych i średnich (niezwiązane z wydatkami na badania i rozwój) w relacji do PKB.

RIS5 – odsetek małych i średnich przedsiębiorstw wprowadzających innowacje wewnętrzne.

RIS6 – odsetek innowacyjnych małych i średnich przedsiębiorstw współdziałających z innymi.

RIS7 – wartość zgłoszeń patentów do Europejskiego Urzędu Patentowego w relacji do PKB.

RIS8 – odsetek małych i średnich przedsiębiorstw wprowadzających innowacje produktowe lub procesowe.

RIS9 – odsetek małych i średnich przedsiębiorstw wprowadzających innowacje marketingowe.

RIS10 – odsetek zatrudnionych w przemysłach wysokiej i średniowysokiej technologii oraz usługach opartych na wiedzy.

RIS11 – relacja sprzedaży produktów stanowiących innowacje nowe dla firmy lub nowe dla rynku do całkowitych obrotów.

$\mathbf{x}_i^{ZP}$ – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TIiK w procesach biznesowych związanych z zarządzaniem produkcją.

$\mathbf{x}_i^{ERP}$ – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TIiK w procesach biznesowych związanych z zarządzaniem zasobami przedsiębiorstwa.

$\mathbf{x}_i^{CAD}$ – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TIiK w procesach biznesowych związanych ze wsparciem dla projektowania i wytwarzania CAD/CAM.

- x_i^{SM} – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wykorzystywania TIiK w procesach biznesowych związanych ze sterowaniem maszynami lub linią produkcyjną.
- x_i^{INW} – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do inwestowania w rozwój technologii informacyjnych i komunikacyjnych.
- x_i^{BR} – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do posiadania własnego wydziału B+R.
- x_i^{PROD} – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wprowadzania innowacji produktowych.
- x_i^{PROC} – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wprowadzania innowacji procesowych lub organizacyjnych.
- x_i^{MARK} – wektor zmiennych objaśniających w równaniu wyjaśniającym skłonność do wprowadzania innowacji marketingowych.
- $x_i^{PRODUKT}$ – wektor zmiennych objaśniających w równaniu wyjaśniającym produktywność.

Rozdział 5

- PR* – zmienna dychotomiczna przyjmująca wartość 1 w przypadku osoby doświadczającej problemu prawnego oraz 0 w przeciwnym przypadku.
- LBUL* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa budowlanego.
- LCIV* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa cywilnego.
- LROA* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa drogowego.
- LPEN* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa karnego.
- LCON* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa konsumenckiego.
- LDAM* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy odszkodowań.
- LFIN* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy problemów finansowych.
- LFAM* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa rodzinnego.
- LJOB* – zmienna binarna przyjmująca wartość 1, jeśli problem prawny doświadczany przez respondenta dotyczy prawa pracy.
- LVAL* – zmienna binarna przyjmująca wartość 1, jeśli respondent uznał, że problem prawny, którego doświadczył, jest ważny.

FEM – zmienna binarna przyjmująca wartość 1 w przypadku kobiet oraz 0 dla mężczyzn.

AGE – wiek respondenta.

EDU1 – zmienna binarna przyjmująca wartość 1 w przypadku respondenta z wykształceniem ponadpodstawowym.

EDU2 – zmienna binarna przyjmująca wartość 1 w przypadku respondenta z wykształceniem ponadśrednim.

SCZO – zmienna binarna przyjmująca wartość 1, jeśli respondent jest żonatym mężczyzną lub zamężną kobietą.

SCRO – zmienna binarna przyjmująca wartość 1, jeśli respondent jest osobą rozwiedzioną.

SCWD – zmienna binarna przyjmująca wartość 1, jeśli respondent jest wdowcem lub wdową.

NFAM – liczba osób w gospodarstwie domowym.

DOCH – dochód na osobę w gospodarstwie domowym.

RESD – zmienna binarna przyjmująca wartość 1, jeśli respondent mieszka w mieście powyżej 20 000 mieszkańców.

SLAB – zmienna binarna przyjmująca wartość 1 w przypadku respondentów niezatrudnionych lub pracujących dorywczo.

PAWR – zmienna ilustrująca poziom świadomości prawnej u osoby ankietowanej.

PCPL – zmienna binarna przyjmująca wartość 1 w przypadku respondentów twierdzących, że zawsze należy przestrzegać prawa.

PUSE – zmienna związana z postawą wobec stosowania prawa.

PTRU – zmienna ilustrująca poziom zaufania do palestry.

PAVA – zmienna binarna ilustrująca subiektywną ocenę dostępności usług prawnych.

APAR – zmienna binarna przyjmująca wartość 1, jeśli respondent przynależy do organizacji społecznych.

AACT – zmienna binarna przyjmująca wartość 1 w przypadku osób, które pozytywnie odpowiedziały na pytanie dotyczące działalności społecznej.

$\mathbf{x}_i^{PR}$ – wektor zmiennych objaśniających w równaniu wyjaśniającym prawdopodobieństwo doświadczenia problemu prawnego przez respondenta.

Notacja dla zbieżności

$_d$

→ – zbieżność według rozkładu.

$_p$

→ – zbieżność według prawdopodobieństwa.

Inne funkcje, parametry i macierze

m(end) – liczba endogenicznych regresorów w endogenicznym modelu probitowym.