PRACE NAUKOWE UNIWERSYTETU EKONOMICZNEGO WE WROCŁAWIU
RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS nr 508 • 2018

Taksonomia 31
Klasyfikacja i analiza danych – teoria i zastosowania

ISSN 1899-3192
e-ISSN 2392-0041

Joanna Małgorzata Landmesser
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: joanna_landmesser@sggw.pl

DEKOMPOZYCJA NIERÓWNOŚCI PŁACOWYCH W POLSCE ZA POMOCĄ METOD BAZUJĄCYCH NA KONSTRUKCJI ROZKŁADÓW WARUNKOWYCH

DECOMPOSITION OF INCOME INEQUALITIES IN POLAND USING METHODS BASED ON THE CONSTRUCTION OF CONDITIONAL DISTRIBUTIONS

DOI: 10.15611/pn.2018.508.12
JEL Classification: J31, D31

Streszczenie: W artykule rozważono różne metody estymacji warunkowej funkcji dystrybuanty dla rozkładu dochodów za pomocą takich modeli regresji, jak dwumianowe modele logitowe, wielomianowe modele probitowe oraz modele hazardu. Koncepcje te pozwoliły na modelowanie prawdopodobieństwa tego, że dochody osoby nie przekroczą pewnego poziomu. Oszacowane dystrybuanty wraz z dystrybuantą kontrfaktyczną wykorzystano przy realizacji celu pracy, jakim była dekompozycja nierówności płacowych wśród kobiet i mężczyzn wzdłuż rozkładu dochodów. Całkowite różnice zdekomponowano na część wyjaśnioną i niewyjaśnioną. Wysoki udział dodatniej części niewyjaśnionej świadczy o tym, że rynek wyżej „wycenia" charakterystyki mężczyzn niż kobiet (efekt dyskryminacji), ujemne zaś wartości luki wyjaśnionej wskazują na redukcję nierówności dzięki „lepszym" charakterystykom kobiet. Wykorzystano dane empiryczne z badania EU-SILC dla Polski w 2014 roku.

Słowa kluczowe: luka płacowa, dekompozycja nierówności, rozkład warunkowy.

Summary: In the paper, we consider various estimation methods of conditional distribution functions for income distribution using regression models such as binomial logit, multinomial probit and hazard models. These concepts allow to model the probability that a person's income will not exceed a certain level. Estimated distribution functions and counterfactual distributions were used to achieve the goal of decomposition of income inequalities among men and women along the whole income distribution. The total differences were decomposed into the explained and the unexplained part. The high share of the positive unexplained part indicates a higher "valuation" by the market of men's characteristics than of women's (the effect of discrimination), while the negative values of the explained component show a reduction of inequality due to "better" characteristics of women. Empirical data from the EU-SILC study for Poland in 2014 were used.

Keywords: pay gap, decomposition of inequalities, conditional distribution.

1. Wstęp

W powszechnym rozumieniu fakt, że przeciętne miesięczne wynagrodzenie brutto dla mężczyzn jest wyższe niż dla kobiet (na przykład w 2014 roku dla mężczyzn wynosiło ono 4481,75 zł, a dla kobiet 3717,57 zł [GUS 2015]) interpretowany jest jako przejaw dyskryminacji. Nowoczesne techniki mikroekonometryczne pozwalają na przeprowadzenie wnikliwej analizy w tym zakresie – analizy, która wykracza poza zwykłe porównywanie średnich wartości płac. Celem artykułu jest przeprowadzenie dekompozycji nierówności wzdłuż całego rozkładu dochodów dla kobiet i mężczyzn w Polsce. Jego realizacja wymagała oszacowania warunkowych dystrybuant dla rozkładu dochodów, co uzyskano za pomocą alternatywnych modeli regresyjnych. Następnie otrzymane różnymi metodami wyniki dekompozycji porównano ze sobą.

Techniki dekompozycji nierówności dochodowych dla licznych kwantyli rozkładu w światowej literaturze ekonomicznej cieszą się dużą popularnością. Przykłady zastosowań w tym zakresie regresji kwantylowej stanowią prace: [Albrecht, Bjorklund, Vroman 2003] dla Szwecji, [Albrecht, van Vuuren, Vroman 2009] dla Holandii [De la Rica, Dolado, Llorens 2005] czy dla Hiszpanii. Inną technikę, zdecentrowaną funkcję wpływu (RIF-regresję), zastosowali [Khanna, Goel, Morissette 2016] dla danych pochodzących z Indii. Powodzeniem cieszą się również metoda imputacji reszt czy metoda oparta na współczynniku ważącym.

W Polsce techniki dekompozycji nierówności wzdłuż rozkładu dochodów są wciąż mało popularne. Znacznie większą popularnością cieszą się metody Oaxaca-Blindera lub Nopo, służące dekompozycji różnic dla średnich poziomów dochodów. Nieliczne przykłady zastosowań regresji kwantylowej stanowią prace: [Newell, Socha 2005; Magda, Szydłowski 2008; Rokicka, Ruzik 2010]. Konstrukcję rozkładu kontrfaktycznego symulacyjną metodą Machado-Mata (bazującą na równaniach regresji kwantylowej) przeprowadziła autorka artykułu w swoich wcześniejszych pracach ([Landmesser, Karpio, Łukasiewicz 2015; Landmesser 2016; Karpio i in. 2016]). Z kolei inni autorzy ([Matysiak, Baranowska, Słoczyński 2010] oraz [Słoczyński 2012]) budowali rozkłady kontrfaktyczne metodą imputacji reszt Juhna, Murphy'ego i Price'a.

Konstrukcja rozkładu kontrfaktycznego nie zawsze jest łatwym zadaniem. Spośród zaproponowanych w tym celu w literaturze anglojęzycznej technik regresyjnych wyróżniają się te, które polegają na estymacji rozkładu dochodów z pomocą dwumianowego modelu logitowego [Chernozhukov, Fernandez-Val, Melly 2013], uporządkowanego modelu probitowego [Fortin, Lemieux 1998] oraz modelu hazardu [Donald, Green, Paarsch 2000]. Zastosowane w niniejszej pracy metody stanowią modyfikacje powyższych i stanowią alternatywę dla podejść wykorzystywanych w Polsce do tej pory.

2. Metody analizy wykorzystane w badaniu

Niech y – zmienna objaśniana (np. dochody), X – wektor zmiennych objaśniających (np. wiek, wykształcenie), β – parametry modelu. Wtedy, o ile oszacujemy w grupach mężczyzn i kobiet równania regresji $y_g = X_g\beta_g + v_g$, $g = M, K$, możliwa jest dekompozycja [Oaxaca 1973; Blinder 1973]:

$$\hat{\Delta}^{\mu} = \overline{X}_M\hat{\beta}_M - \overline{X}_K\hat{\beta}_K = \underbrace{\overline{X}_M(\hat{\beta}_M - \hat{\beta}_K)}_{\hat{\Delta}^{\mu}_{\text{efekt niewyjaśniony}}} + \underbrace{(\overline{X}_M - \overline{X}_K)\hat{\beta}_K}_{\hat{\Delta}^{\mu}_{\text{efekt wyjaśniony}}}.$$

Część niewyjaśniona mierzy różnicę w wartościach y związaną z różnymi parametrami regresji i odzwierciedla dokonaną przez rynek odmienną „wycenę" dla identycznych cech posiadanych przez osoby z różnych grup (tzw. efekt dyskryminacji). Część wyjaśniona mierzy różnicę wynikającą z odmiennych charakterystyk osób.

Niech dystrybuanta zmiennej Y w grupie g wyraża się wzorem $F_{Y_g}(y) = \int F_{Y_g|X_g}(y|X)\cdot dF_{X_g}(X)$, $g = M, K$. Aby przeanalizować różnice względem całego rozkładu badanej cechy, skonstruujemy rozkład kontrfaktyczny stanowiący mieszankę warunkowego rozkładu zmiennej zależnej oraz rozkładu zmiennych objaśniających (rozkład dochodów, który obowiązywałby dla osób z grupy K, gdyby miałyby one charakterystyki zgodne z rozkładem cech w grupie M):

$$F_{Y_K^C}(y) = \int F_{Y_K|X_K}(y|X)\cdot dF_{X_M}(X).$$

O ile znane są rozkłady warunkowe dla Y, możliwa jest dekompozycja:

$$F_{Y_M}(y) - F_{Y_K}(y) = \underbrace{[F_{Y_M}(y) - F_{Y_K^C}(y)]}_{\hat{\Delta}_{\text{efekt niewyjaśniony}}} + \underbrace{[F_{Y_K^C}(y) - F_{Y_K}(y)]}_{\hat{\Delta}_{\text{efekt wyjaśniony}}}.$$

2.1. Konstrukcja rozkładu za pomocą dwumianowych modeli logitowych

Warunkowy rozkład dochodów można oszacować za pomocą dwumianowych modeli logitowych, estymowanych dla różnych wartości y:

$$\Pr[Y \le y|X] = F(y|X) = \Lambda(X\beta(y)) = 1/(1+\exp(-X\beta(y))).$$

Oceny parametrów $\beta(y)$ uzyskiwane są poprzez konstrukcję 0-1 zmiennej $\mathrm{I}\{Y_i \le y\}$, identyfikującej, czy Y_i jest poniżej poziomu y, a następnie estymację modelu regresji logistycznej dla $\mathrm{I}\{Y_i \le y\}$ względem zmiennych X_i (por. podobne podejście [Chernozhukov, Fernandez-Val, Melly 2013]). Rozpatrując podział

zakresu wartości zmiennej Y taki, że $0 = c_0 < c_1 < ... < c_P = \infty$, dla każdego progu $y := c_k$, $k = 1, \ldots, P-1$ szacowany jest odrębny model logitowy.

Rozkłady warunkowe i kontrfaktyczny otrzymywane są na drodze uśredniania:

$$\hat{F}_{Y_g}(y) = \frac{1}{n_g} \sum_{i \in g} \Lambda(X_i \hat{\beta}_g(y)),\ g = K, M, \qquad \hat{F}_{Y_K^C}(y) = \frac{1}{n_M} \sum_{i \in M} \Lambda(X_i \hat{\beta}_K(y)).$$

Takie uśrednianie odzwierciedli wyrównany w przedziałach, aczkolwiek zróżnicowany w skali całej osi Y, wpływ cech X na dochód osób.

2.2. Konstrukcja rozkładu za pomocą wielomianowych uporządkowanych modeli probitowych

Rozkład $F(y \mid X)$ można również oszacować za pomocą wielomianowego uporządkowanego modelu probitowego (por. [Fortin, Lemieux 1998]). Wtedy dla $Y^* = X\beta + \varepsilon, \varepsilon \sim \mathrm{N}(0,1)$:

$$F(y|X) = \Phi(-X\beta + \Lambda(y)), \qquad Y = \Lambda^{-1}(Y^*).$$

Rozpatrując podział zakresu wartości zmiennej Y na przedziały (c_{j-1}, c_j), $j = 1, ..., P$, warunkowe prawdopodobieństwa przynależności do przedziałów wyznacza się jako:

$$\Pr\left[c_{j-1} \le Y \le c_j \middle| X\right] = \Phi(-X\beta + \Lambda(c_j)) - \Phi(-X\beta + \Lambda(c_{j-1})),$$

gdzie $\Lambda(c_j)$ są szacowanymi progami modelu. Dla każdego punktu $y := c_j$:

$$\hat{F}_{Y_g}(c_j) = \frac{1}{n_g} \sum_{i \in g} \Phi(-X_i \hat{\beta}_g + \hat{\Lambda}_g(c_j)),\ g = K, M,$$

$$\hat{F}_{Y_K^C}(c_j) = \frac{1}{n_M} \sum_{i \in M} \Phi(-X_i \hat{\beta}_K + \hat{\Lambda}_K(c_j)).$$

2.3. Konstrukcja rozkładu za pomocą modelu hazardu

Rozkład wynagrodzeń można oszacować za pomocą modelu hazardu, ponieważ model taki służy do opisu rozkładu nieujemnych zmiennych losowych (czasów trwania). W przeprowadzonej analizie rolę zmiennej T przyjmuje nieujemna zmienna Y dotycząca poziomu dochodów (por. [Donald, Green, Paarsch 2000]). Funkcja przeżycia $S(y) = \Pr[Y \ge y]$ wyraża wtedy prawdopodobieństwo tego, że dochody osoby będą wyższe lub przynajmniej równe pewnemu poziomowi. Funkcja hazardu $h(y) = f(y)/S(y)$ może być oszacowana za pomocą następującego modelu hazardu

wykładniczego o różnym hazardzie bazowym w przedziałach oraz stałych parametrach β: $h(y \mid X) = h_{0k}(y)\exp(X\beta)$ dla $y \in (c_{k-1}, c_k)$, $k = 1,...,P$. Wtedy

$$S(y|X) = \exp\left[-\sum_{j=1}^{k-1}(c_j - c_{j-1})h_{0j}(y)\exp(X\beta) - (y - c_{k-1})h_{0k}(y)\exp(X\beta) \right] \text{ dla } y \in (c_{k-1}, c_k).$$

Rozkłady warunkowe i kontrfaktyczny otrzymywane są na drodze uśredniania po odpowiednich przedziałach, z wykorzystaniem zależności $F(y \mid X) = 1 - S(y \mid X)$.

3. Dane wykorzystane w badaniu

Próba empiryczna, na której oparto badanie, pochodziła z Europejskiego Badania Dochodów i Warunków Życia EU-SILC dla Polski w roku 2014 i składała się z 5177 obserwacji dla mężczyzn oraz 4727 dla kobiet (tab. 1).

Tabela 1. Charakterystyki zmiennych wykorzystanych w badaniu

Zmienna		Mężczyźni	Kobiety	Zmienna		Mężczyźni	Kobiety
dochod, średnia		7165,94	5900,21	*wykszt*	= 1	4,91%	3,89%
latapracy, średnia		20,09	18,46		= 2	1,45%	0,55%
stancywilny	= 1	72%	70%		= 3	68,57%	47,32%
niepelny	= 1	4,31%	10,09%		= 4	2,55%	7,91%
kierownik	= 1	18,68%	15,74%		= 5	22,52%	40,32%

Źródło: opracowanie własne na podstawie danych EU-SIlC dla Polski z 2014 roku.

Każda osoba została scharakteryzowana za pomocą następujących cech: *stancywilny* (1 dla pozostających w związku małżeńskim, zero w przeciwnym razie), *wykszt* (1 dla wykształcenia podstawowego, … , 5 dla wyższego), *latapracy* (staż pracy w latach), *niepelny* (1 dla zatrudnionych w niepełnym wymiarze czasu pracy, zero w przeciwnym razie), *kierownik* (1 dla zatrudnionych na stanowiskach kierowniczych, zero w przeciwnym razie). Porównywano ze sobą roczne dochody netto kobiet i mężczyzn wyrażone w euro (zmienna Y o symbolu *dochod*). Charakterystyki zmiennych wykorzystanych w badaniu podano w tab. 1.

4. Wyniki

W pierwszym kroku przeprowadzono dekompozycję nierówności dochodowych Oaxaca-Blindera dla średnich poziomów dochodów. Oszacowana surowa luka dochodowa wyniosła 1265,73 euro. Część niewyjaśniona luki wyniosła 1560,19 euro, co stanowiło 123,26% całej luki, a wyjaśniona osiągnęła poziom –294,46 (–23,26% luki). Przeprowadzona dekompozycja szczegółowa wskazała istotnie redukujący

nierówności wpływ czynnika *wyksz* (efekt wyższego przeciętnie poziomu wykształcenia kobiet) oraz powiększające lukę skutki zatrudniania kobiet w niepełnym wymiarze czasu pracy i na gorzej opłacanych stanowiskach niekierowniczych.

W dalszej kolejności dla zmiennej *dochod* oszacowano po 9 dwumianowych modeli logitowych dla mężczyzn oraz dla kobiet, w których jako progi c_k, $k = 1,\ldots,P-1$, przyjęto kolejne decyle rozkładu dochodów w połączonej próbie (wyniki estymacji przedstawia tab. 2). Ujemne oceny parametrów przy zmiennych charakterystykach osób oznaczają, że wraz ze wzrostem wartości takich zmiennych dochodzi do spadku prawdopodobieństwa dochodu niższego od rozważanego progu.

Tabela 2. Oszacowania dwumianowych modeli logitowych

Mężczyźni	Przedział								
	1	2	3	4	5	6	7	8	9
stancywilny	-0,45***	-0,41***	-0,58***	-0,54***	-0,59***	-0,69***	-0,76***	-0,70***	-0,82***
wyksz	-0,41***	-0,46***	-0,52***	-0,62***	-0,65***	-0,69***	-0,71***	-0,75***	-0,78***
latapracy	-0,04***	-0,03***	-0,03***	-0,03***	-0,02***	-0,02***	-0,02***	-0,02***	-0,02***
niepelny	2,96***	2,55***	2,48***	2,42***	2,19***	2,04***	1,82***	1,67***	1,44***
kierownik	-1,33***	-1,18***	-1,18***	-1,05***	-1,05***	-1,05***	-1,17***	-1,23***	-1,30***
cons	-0,34	0,73***	1,58***	2,34***	2,95***	3,61***	4,35***	5,13***	6,24***
lnL	-1148,10	-1987,98	-2478,42	-2846,10	-3039,12	-3008,51	-2754,67	-2305,39	-1580,47
Kobiety	Przedział								
	1	2	3	4	5	6	7	8	9
stancywilny	-0,33***	-0,20**	-0,21***	-0,22***	-0,22***	-0,20**	-0,16*	-0,16	0,12
wyksz	-0,43***	-0,61***	-0,81***	-1,00***	-1,14***	-1,14***	-1,20***	-1,19***	-1,16***
latapracy	-0,07***	-0,06***	-0,05***	-0,06***	-0,07***	-0,07***	-0,06***	-0,06***	-0,05***
niepelny	2,66***	2,38***	2,16***	1,98***	1,85***	1,70***	1,63***	1,75***	1,43***
kierownik	-0,92***	-0,82***	-0,81***	-0,77***	-0,81***	-0,90***	-0,99***	-1,00***	-1,20***
cons	0,44**	1,98***	3,39***	4,83***	6,05***	6,65***	7,44***	8,10***	8,76***
lnL	-1318,37	-2039,41	-2419,85	-2493,12	-2414,70	-2286,47	-2008,49	-1597,68	-953,07

***,**,* – istotność odpowiednio na poziomie 1, 5 i 10%.

Źródło: obliczenia własne.

Następnie oszacowano dwa pozostałe modele: probitowy dla zmiennej wielomianowej uporządkowanej (tab. 3) oraz model hazardu (tab. 4). Dla tego pierwszego oceny parametrów nie pozwalają jednoznacznie określić znaku efektów marginalnych dla poszczególnych kategorii zmiennej *Y*. W przypadku modelu hazardu dodatnie oceny parametrów przy zmiennych charakterystykach osób oznaczają, że wraz ze wzrostem wartości takich zmiennych dochodzi do wzrostu ryzyka spadku zarobków poniżej poziomu *y*. Na przykład: im wyższy poziom wykształcenia, tym mniejsze narażenie na utratę otrzymywanego wynagrodzenia.

Tabela 3. Oszacowania wielomianowych modeli probitowych

Zmienna	Mężczyźni	Kobiety	Zmienna	Mężczyźni	Kobiety
stancywilny	0,354***	0,111***	/cut3	1,064***	2,056***
wyksz	0,371***	0,498***	/cut4	1,389***	2,414***
latapracy	0,015***	0,033***	/cut5	1,716***	2,755***
niepelny	–1,405***	–1,320***	/cut6	2,007***	3,078***
kierownik	0,707***	0,552***	/cut7	2,351***	3,415***
/cut1	0,177***	1,041***	/cut8	2,732***	3,836***
/cut2	0,712***	1,617***	/cut9	3,245***	4,436***
liczba obs.	5177	4727	lnL	-10 954,42	-9669,68

*** – istotność na poziomie 1%.

Źródło: obliczenia własne.

Tabela 4. Oszacowania modeli hazardu wykładniczego

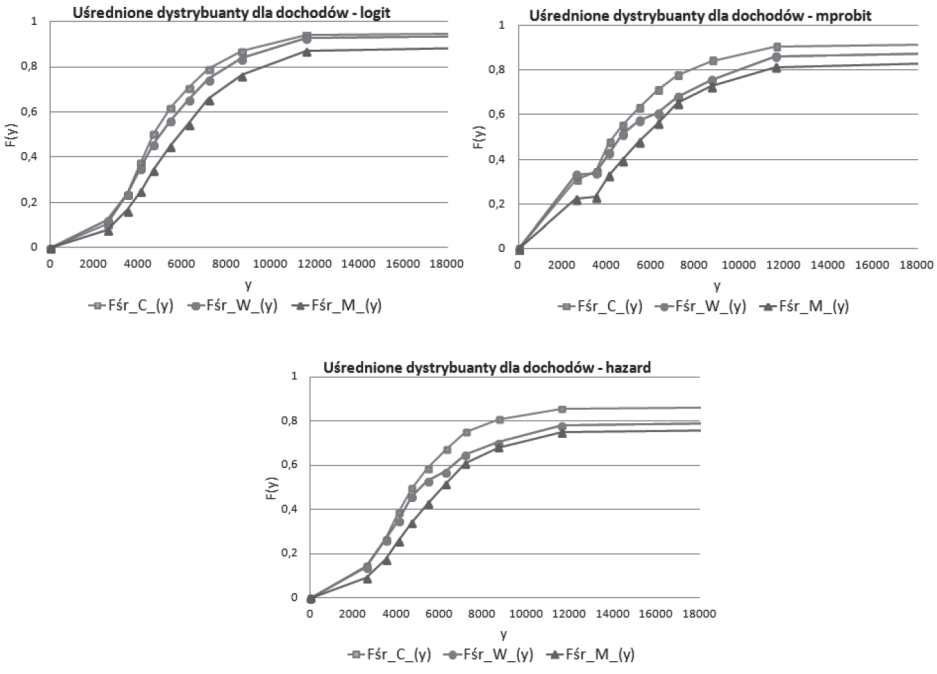
Zmienna	Mężczyźni		Kobiety		Zmienna	Mężczyźni		Kobiety	
	b	exp(*b*)	*b*	exp(*b*)		*b*	exp(*b*)	*b*	exp(*b*)
tp1	-8,833***	0,0001	-7,849***	0,0004	*tp9*	-6,424***	0,0016	-5,378***	0,0046
tp2	-7,507***	0,0005	-6,629***	0,0013	*tp10*	-6,577***	0,0014	-5,501***	0,0041
tp3	-6,948***	0,0010	-5,889***	0,0028	stancywilny	-0,343***	0,7094	-0,072**	0,9303
tp4	-6,702***	0,0012	-5,752***	0,0032	*wyksz*	-0,340***	0,7117	-0,456***	0,6335
tp5	-6,610***	0,0013	-5,715***	0,0033	*latapracy*	-0,012***	0,9878	-0,026***	0,9740
tp6	-6,560***	0,0014	-5,608***	0,0037	*niepelny*	1,153***	3,1692	0,927***	2,5274
tp7	-6,225***	0,0020	-5,388***	0,0046	*kierownik*	-0,592***	0,5533	-0,528***	0,5901
tp8	-6,387***	0,0017	-5,409***	0,0045					
licz. obs.	5177		4727		lnL	-4222,57		-3484,62	

***,**,* – istotność odpowiednio na poziomie 1, 5, 10%.

Źródło: obliczenia własne.

Uzyskane na podstawie powyższych modeli regresyjnych oszacowania dystrybuant dla dochodów przedstawia rys. 1. Położenie dystrybuant dla kobiet powyżej dystrybuant dla mężczyzn świadczy o tym, że dla kobiet zachodzi większe prawdopodobieństwo nieprzekroczenia poziomu płac y niż dla mężczyzn, czyli że kobiety zarabiają mniej od mężczyzn. Ponadto oddzielnie sporządzone wykresy względem płci wykazały, że dystrybuanty wynikające z modelu probitowego są położone zawsze wyżej niż dystrybuanty z modelu hazardu.

Na koniec przeprowadzono dekompozycję różnic w skonstruowanych rozkładach dochodów kobiet i mężczyzn. Różnice dekomponowano dla poszczególnych kwantyli rozkładu (odległości poziome między funkcjami rozkładu płac kobiet



Rys. 1. Uzyskane różnymi metodami wykresy dystrybuant dochodów
(Fśr_C – uśredniona dystrybuanta kontrfaktyczna, Fśr_W – uśredniona dystrybuanta dla kobiet, Fśr_M – uśredniona dystrybuanta dla mężczyzn)

Źródło: opracowanie własne.

i mężczyzn). Wartości poszczególnych kwantyli wyznaczano jako $\hat{Q}_{g,\tau} = \hat{F}^{-1}_{Y_g}(\tau)$, dokonując liniowych interpolacji punktowych uśrednionych dystrybuant.

Wyniki wskazują na istnienie dodatnich, aczkolwiek niemonotonicznych, różnic pomiędzy dochodami mężczyzn i kobiet (tab. 5). Modele logitowy i hazardu sugerują ich wyższą wartość dla kwantyla rzędu 0,1 (wśród najuboższych), a modele probitowy i hazardu – dla kwantyla 0,5. Wszystkie trzy modele zgodnie wykazują wyższe nierówności na prawym krańcu rozkładu dochodów. Szczególnie duża luka ma miejsce dla kwantyla rzędu 0,9 (czyli u najbogatszych).

Zaobserwowano pogłębiający się – wraz ze wzrostem dochodów – przyrost niewyjaśnionej części luki dochodowej (czyli rynek wyżej „wycenia" charakterystyki mężczyzn niż kobiet) (tab. 6). Luka wyjaśniona przez charakterystyki indywidualne osób ma niższy udział w luce całkowitej. Niemniej uzyskane w tym zakresie ujemne wartości świadczą o redukcji nierówności płacowych zapewne dzięki „lepszym" charakterystykom kobiet niż mężczyzn. Ta korzystna dla kobiet redukcja pogłębia się w miarę rozpatrywania coraz wyższych grup dochodowych. W modelu logitowym i probitowym, inaczej niż w modelu hazardu, część wyjaśniona jest początko-

wo dodatnia, co sugeruje, że najbiedniejsze kobiety mają „gorsze" charakterystyki od najbiedniejszych mężczyzn.

Tabela 5. Wyniki dekompozycji różnic w dochodach

τ	Całkowita różnica		
	logit	mprobit	hazard
0,1	714,30	375,72	844,36
0,2	527,85	751,45	657,68
0,3	583,83	1 546,21	651,32
0,4	729,28	829,07	854,51
0,5	880,89	1 054,58	1 017,28
0,6	925,88	550,33	490,21
0,7	1 055,20	570,31	915,64
0,8	1 581,96	1 218,47	29 902,30
0,9	16 764,90	13 195,69	40 163,33

Źródło: opracowanie własne.

Tabela 6. Wyniki dekompozycji różnic w dochodach

τ	Część niewyjaśniona			Część wyjaśniona		
	logit	mprobit	hazard	logit	mprobit	hazard
0,1	366,00	317,50	893,25	348,30	58,22	–48,90
0,2	479,19	635,00	673,44	48,66	116,45	–15,75
0,3	633,02	1371,54	724,59	–49,19	174,67	–73,26
0,4	885,78	983,64	1 053,52	–156,50	–154,57	–199,01
0,5	1 200,17	1 385,99	1 422,43	–319,28	–331,41	–405,15
0,6	1 385,46	1 506,83	1 507,74	–459,58	–956,51	–1 017,53
0,7	1 576,47	1 957,65	2 916,01	–521,27	–1 387,33	–2 000,37
0,8	2 317,84	3 417,75	47 438,18	–735,88	–2 199,28	–17 535,88
0,9	17 524,48	34 749,18	76 712,12	–759,58	–21 553,49	–36 548,80

Źródło: opracowanie własne.

5. Podsumowanie

Celem pracy było przeprowadzenie dekompozycji nierówności wzdłuż rozkładu dochodów dla kobiet i mężczyzn w Polsce, posługując się przy tym wybranymi modelami regresji. Oszacowano warunkowe funkcje rozkładu dochodów za pomocą: dwumianowych modeli logitowych, wielomianowych uporządkowanych modeli probitowych, modeli hazardu wykładniczego przedziałami stałego, a następnie wy-

znaczono różnice wzdłuż rozkładów dochodów. W dalszej kolejności różnice te zdekomponowano na część wyjaśnioną i część niewyjaśnioną.

Zastosowane podejścia cechuje łatwa implementacja dzięki wykorzystaniu popularnych modeli mikroekonometrycznych. Silne założenia wynikające z podejmowanych parametryzacji stanowią jednak pewną wadę. Im bardziej elastyczne podejście, tym większe nakłady obliczeniowe (np. konieczność szacowania wielu regresji logitowych), a podział skali dochodów na węższe podprzedziały może powodować niemonotoniczność dystrybuanty kontrfaktycznej.

Wykorzystane w pracy modele są modelami regresyjnymi, więc w toku dalszych planowanych badań możliwe będzie wyciągniecie wniosków na temat wpływu indywidualnych charakterystyk osób, takich jak wykształcenie czy staż pracy, na obserwowane różnice. Możliwa do przeprowadzenia dekompozycja szczegółowa nie jest jednak prosta, a nawet może okazać się *path dependent*.

Literatura

Albrecht J., Bjorklund A., Vroman S., 2003, *Is there a glass ceiling in Sweden?*, Journal of Labor Economics, 21, s. 145-177.

Albrecht J., van Vuuren A., Vroman S., 2009, *Counterfactual Distributions with Sample Selection Adjustments: Econometric Theory and an Application to the Netherlands*, Labour Economics, 16(4), s. 383-396.

Blinder A., 1973, *Wage Discrimination: Reduced Form and Structural Estimates*, Journal of Human Resources, 8(4), s. 436-455.

Chernozhukov V., Fernandez-Val I., Melly B., 2013, *Inference on counterfactual distributions*, Econometrica, 81(6), s. 2205-2268.

De la Rica S., Dolado J., Llorens V., 2005, *Glass Ceiling or Floors? Gender Wage Gaps by Education in Spain*, IZA Discussion Paper, 1483, Bonn.

Donald S.G., Green D.A., Paarsch H.J., 2000, *Differences in Wage Distributions between Canada and the United States: An Application of a Flexible Estimator of Distribution Functions in the Presence of Covariates*, Review of Economic Studies, 67(4), s. 609-633.

Fortin N.M., Lemieux T., 1998, *Rank Regressions, Wage Distributions, and the Gender Gap*, Journal of Human Resources, 33, s. 610-643.

GUS, 2015, *Rocznik Statystyczny Pracy*, Warszawa.

Karpio K., Landmesser J.M., Łukasiewicz P., Orłowski A.J., 2016, *The Quantile Decomposition of Personal Income Distributions in the USA*, Acta Physica Polonica A, 129(5), s. 965-970.

Khanna S., Goel D., Morissette R., 2016, *Decomposition analysis of earnings inequality in rural India: 2004-2012*, IZA Journal of Labor & Development, 5(1), 18.

Landmesser J.M., 2016, *Decomposition of Differences in Income Distributions Using Quantile Regression*, Statistics in Transition – new series, 17(2), s. 331-348.

Landmesser J.M., Karpio K., Łukasiewicz P., 2015, *Decomposition of Differences Between Personal Incomes Distributions in Poland*, Quantitative Methods in Economics, XVI (2), s. 43-52.

Magda I., Szydłowski A., 2008, *Płace w makro i mikroperspektywie*, [w:] Bukowski M. (ed.), *Zatrudnienie w Polsce 2007 – Bezpieczeństwo na elastycznym rynku pracy*, Ministry of Labour and Social Policy, Warszawa.

Matysiak A., Baranowska A., Słoczyński T., 2010, *Kobiety i mężczyźni na rynku pracy*, [w:] Bukowski M. (red.), *Zatrudnienie w Polsce 2008 – Praca w cyklu życia*, Human Resources Development Center, Warszawa.

Newell A., Socha M., 2005, *The Distribution of Wages in Poland*, IZA Discussion Paper, 1485, Bonn.

Oaxaca R., 1973, *Male-Female Wage Differentials in Urban Labor Markets*, International Economic Review, 14(3), s. 693-709.

Rokicka M., Ruzik A., 2010, *The Gender Pay Gap in Informal Employment in Poland*, CASE Network Studies and Analyses, 406/2010.

Słoczyński T., 2012, *Próba wyjaśnienia regionalnego zróżnicowania międzypłciowej luki płacowej w Polsce*, Studia Regionalne i Lokalne, 3(49).