

**Waldemar FLORCZAK<sup>1</sup>**  
**Wojciech GRABOWSKI<sup>2</sup>**

## Co wpływa na wielkość popytu na porady prawne? Analiza logitowa z wykorzystaniem metody klasycznego uśredniania międzymodelowego<sup>3, 4</sup>

### 1. WPROWADZENIE

Powszechny dostęp do pomocy prawnej jest warunkiem podmiotowości obywateli i niezbędną przesłanką równości szans wszystkich członków społeczeństwa. Jednakże dopiero ustawa z dnia 5 sierpnia 2015 r. o nieodpłatnej pomocy prawnej oraz edukacji prawnej, wdrożona z dniem 1 stycznia 2016 r., zapewniła najuboższym w Polsce instytucjonalny dostęp do profesjonalnych porad prawnych. Tym samym kraj nasz dołączył do szerokiego grona państw demokratycznych, w których pomoc taka świadczona jest nieodpłatnie.

Akt uchwalenia ustawy poprzedziły szerokie badania prawnych i społeczno-ekonomicznych uwarunkowań funkcjonowania nieformalnego systemu pomocy prawno-obywatelskiej, obowiązującego – i wciąż funkcjonującego równolegle do obecnego systemu formalnego – przed wdrożeniem reformy. W badaniach tych, realizowanych w ramach projektu pt. „Opracowanie kompleksowych i trwałych mechanizmów wsparcia dla poradnictwa prawnego i obywatelskiego w Polsce” – współfinansowanego przez Unię Europejską w ramach Europejskiego Funduszu Społecznego i kierowanych przez Instytut Spraw Publicznych w Warszawie – uczestniczyli liczni interesariusze reformy poradnictwa prawnego (ISP, Departament Pożytku Publicznego (DPP) w Ministerstwie Pracy i Polityki Społecznej, Fundacja Instytut Prawa i Społeczeństwa (INPRIS), Związek Biur Porad Obywatelskich (ZBPO) oraz Fundacja Uniwersyteckich Poradni Prawnych (FUPP)).

---

<sup>1</sup> Uniwersytet Jagielloński, Wydział Zarządzania i Komunikacji Społecznej, Instytut Ekonomii, Finansów i Zarządzania, ul. Prof. S. Łojasiewicza 4, 30–348 Kraków, Polska.

<sup>2</sup> Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, Katedra Modeli i Prognoz Ekonometrycznych, Rewolucji 1905 r. 41, 90-214 Łódź, Polska, autor prowadzący korespondencję – e-mail: emfwog@uni.lodz.pl.

<sup>3</sup> Opracowanie powstało w ramach realizacji grantu NCN nr 2012/07/B/HS4/02994.

<sup>4</sup> Autorzy artykułu serdecznie dziękują anonimowym Recenzentom za cenne uwagi i sugestie dotyczące pierwotnej wersji artykułu. Ich wprowadzenie przyczyniło się do poprawy jakości tekstu.

Jednym z wymiernych efektów projektu było przeprowadzenie w 2012 roku ogólnopolskiego badania ankietowego beneficjentów i usługodawców nieformalnego systemu poradnictwa prawnego, którego wyniki omówiono m.in. w dwóch obszernych opracowaniach (Burdziej, Dudkiewicz, 2013; Peisert i inni, 2013). Nie zastosowano jednak metod wielowymiarowej analizy statystyczno-ekonometrycznej, ograniczając się jedynie do narzędzi statystyki opisowej. Tym samym ilościowe wnioski płynące z tych analiz pomijają szereg kluczowych założeń wnioskowania statystycznego, przede wszystkim z powodu nie uwzględnienia jednoczesności wpływu różnorodnych czynników na wielkość popytu na usługi poradnicze.

Nie jest to słabość jedynie wyżej wymienionych badań. Przegląd literatury przedmiotu pokazuje (Winczorek, 2015), że większość z istniejących badań empirycznych, stawiających sobie za cel odpowiedź na pytanie sformułowane w tytule niniejszego opracowania, bazuje na nieskomplikowanych metodach analitycznych, wykorzystujących metody statystyki opisowej (np. Legal Service Corporation, 1994; Currie, 2006; Murayama, 2007), lub proste narzędzia statystyki matematycznej, ograniczone do przestrzeni dwuwymiarowej (np. Gramatikov, 2008; Kritzer 2009). Nie podejmując się odpowiedzi na pytanie o przyczyny takiego stanu rzeczy, skonstatować należy jednak, że metodyka taka nie jest poprawna. Fakt bowiem, że na gruncie rozważań logiczno-teoretycznych wskazuje się na bardzo liczne, potencjalnie istotne czynniki mogące wpływać na prawdopodobieństwo wystąpienia problemu prawnego, zaś w praktyce stosuje się metody jednowymiarowe, nadszarpuje wiarygodność rezultatów (i wniosków) uzyskiwanych przy ich użyciu. Dlatego metody takie należałoby uznać za wskazane jedynie w kontekście badania pomocniczego, nie zaś głównego.

Wystąpienie problemu prawnego stanowi jeden z niezliczonych przejawów funkcjonowania jednostki w społeczeństwie. Aby lepiej zrozumieć ideę badania przeprowadzonego w niniejszym artykule, należy przytoczyć definicję pojęcia „problem prawny”. Respondenci byli pytani czy w okresie ostatnich 5 lat znaleźli się w sytuacji wymagającej pomocy prawnika, osoby z wykształceniem prawniczym lub instytucji/organizacji udzielającej porad prawnych. Na podstawie odpowiedzi respondentów definiowana była zmienna zależna, przyjmująca wartość 1 dla tych respondentów, którzy odpowiedzieli „Tak” na zadane pytanie oraz 0, jeśli odpowiedzieli „Nie”.

Celem artykułu jest identyfikacja i kwantyfikacja wpływu wszystkich grup czynników, wskazywanych w literaturze tematu jako ważnych, na prawdopodobieństwo wystąpienia problemu prawnego. Cel ten zrealizowano przy użyciu kilku metod/procedur ekonometrycznych. W kroku pierwszym wyspecyfikowano i oszacowano parametry dychotomicznego modelu logitowego, obejmującego 26 regresorów. Uzyskano wyniki zgodne z ustaleniami merytorycznymi, jednakże znaczna część zmiennych objaśniających okazała się statystycznie nieistotna. Dlatego też

w kolejnym kroku zastosowano strategię modelowania od ogółu do szczegółu, w efekcie czego otrzymano wariant, w którym wszystkie regresory są statystycznie istotne. I wreszcie w celu zbadania odporności otrzymanych rezultatów, zastosowano klasyczną metodę uśredniania międzymodelowego.

Do obliczeń empirycznych wykorzystano dane mikroekonomiczne pochodzące, ze wspomnianego wcześniej, ogólnopolskiego, reprezentatywnego badania ankietowego beneficjentów i usługodawców, przeprowadzonego przez Instytut Spraw Publicznych w Warszawie w 2012 roku (Burdziej, Dudkiewicz, 2013; Peisert i inni, 2013). Rezultaty analizy pokazują, że jedynie nieliczne czynniki uznać należy za faktycznie zwiększające lub zmniejszające prawdopodobieństwo wystąpienia problemu prawnego. Są to: płeć, wiek, stan cywilny, aktywność społeczna i postawa wobec przestrzegania prawa.

Struktura pracy jest następująca. W kolejnym punkcie omówiono kwestie metodyczne. Sekcja 3-cia zawiera dyskusję nad specyfikacją dychotomicznego modelu logitowego w nawiązaniu do ustaleń teoretycznych i wyników wcześniejszych badań empirycznych. W punkcie 4-tym przytoczono rezultaty badania autorskiego i dokonano ich interpretacji. Artykuł zamykają uwagi końcowe.

## 2. METODYKA

### 2.1. Model logitowy

Gdy zmienna zależna jest zero-jedynkowa, wyjaśnianie zależności pomiędzy tą zmienną a innymi kategoriami ekonomiczno-społecznymi odbywa się po dokonaniu estymacji parametrów modelu dychotomicznego:

$$y_i^* = \alpha_0 + \alpha_1 x_{1i} + \dots + \alpha_K x_{Ki} + \varepsilon_i, \quad i = 1, \dots, N, \quad \varepsilon_i \sim F, \quad (1.a)$$

$$y_i = I\{y_i^* > 0\}, \quad (1.b)$$

gdzie  $i$  indeksuje obserwacje,  $I\{\cdot\}$  oznacza zmienną wskaźnikową, przyjmującą wartość 1, gdy warunek zdefiniowany w „klamrowym nawiasie” jest spełniony oraz 0 w przeciwnym przypadku, natomiast  $x_{1i}, \dots, x_{Ki}$  są zmiennymi objaśniającymi,  $\varepsilon_i$  jest nieobserwowalnym składnikiem losowym, natomiast  $F$  oznacza dystrybucję jego rozkładu.  $\alpha_0, \alpha_1, \dots, \alpha_K$  są parametrami podlegającymi estymacji.  $y_i$  oznacza obserwowalną zmienną dychotomiczną, natomiast  $y_i^*$  jest zmienną nieobserwowalną związaną ze zmienną dychotomiczną (por. np. Marzec, 2008; Gruszczyński, 2010). Najczęściej przyjmuje się, że składnik losowy pochodzi z rozkładu normalnego lub logistycznego. Wówczas mamy do czynienia odpowiednio z modelem probitowym lub logitowym. W przypadku niesymetrycznego rozkładu skład-

nika losowego, najczęściej przyjmuje się, że jest to komplementarny rozkład log-log. Z rozkładem tym związany jest model komplementarny log-log.

W badaniach ekonomicznych, kluczowym zagadnieniem jest interpretacja siły wpływu zmiennych objaśniających na zmienną objaśnianą. W modelu dwumianowym oszacowania parametrów nie mają takiej interpretacji jak w klasycznym modelu regresji liniowej. Niemniej jednak po dokonaniu estymacji parametrów modelu logitowego możliwa jest interpretacja funkcji oszacowań w kontekście wpływu jednostkowej zmiany wartości zmiennej objaśniającej na iloraz szans. Jeśli w wyniku estymacji parametrów modelu logitowego uzyskane zostanie oszacowanie parametru wynoszące  $\hat{\alpha}_1$ , należy je interpretować następująco:

- jednostkowy wzrost wartości zmiennej  $x_1$  powoduje przy innych czynnikach niezmiennych zmianę ilorazu szans o  $(\exp(\hat{\alpha}_1) - 1) \cdot 100\%$ .

W badaniach opartych na modelach dychotomicznych, często pojawiają się zmienne binarne w charakterze regresorów. Załóżmy, że zmienna  $x_2$  jest zmienną binarną. Jeśli oszacowanie parametru przy tej zmiennej wynosi  $\hat{\alpha}_2$ , należy je interpretować następująco:

- iloraz szans różni się o  $(\exp(\hat{\alpha}_2) - 1) \cdot 100\%$  w przypadku gdy zmienna  $x_2$  przyjmuje wartość 1 w porównaniu z sytuacją, gdy jej wartość wynosi 0.

W badaniach opartych na modelach dychotomicznych, mogą pojawić się także zmienne nominalne o więcej niż 2 kategoriach (np. rozmiar firmy – mikro, mała, średnia, duża). Wówczas jednym ze sposobów rozwiązania problemu jest stworzenie zmiennych zero-jedynkowych dla wszystkich wariantów oprócz wariantu bazowego (por. Świadek, Wojciech, 2015). Niech zmienna  $x_{3l}$  oznacza zmienną zero-jedynkową przyjmującą wartość 1, gdy zmienna nominalna  $x_3$  przyjmuje wartość  $l$  ( $l=2, 3, \dots, L$ ), natomiast wariant, w którym przyjmuje ona wartość 1 jest wariantem bazowym. Wówczas oszacowanie parametru przy zmiennej  $x_{3l}$  można interpretować następująco:

- iloraz szans różni się o  $(\exp(\hat{\alpha}_{3l}) - 1) \cdot 100\%$  w przypadku gdy zmienna  $x_3$  przyjmuje wartość  $l$  w porównaniu z sytuacją, gdy jej wartość wynosi 1.

Dla modeli dwumianowych jakość dopasowania można mierzyć poprzez porównanie teoretycznych wartości zmiennej zależnej z wartościami empirycznymi. Funkcja wykorzystywana do obliczania teoretycznych wartości zmiennej zależnej przyjmuje następującą postać:

$$\hat{y}_i = I\{\hat{P}(y_i = 1|\mathbf{x}_i) > c\}, \quad (2)$$

gdzie  $\hat{P}$  oznacza oszacowanie prawdopodobieństwa, że zmienna zależna przyjmuje wartość 1.  $\hat{y}_i$  jest wyznaczoną ze wzoru (2) prognozą wewnątrzpróbkową zmiennej  $y_i$ . Po wyznaczeniu wewnątrzpróbkowych prognoz dla zmiennej  $y_i$ , definiowana jest tabela trafności predykcji (patrz tabela 1).

Tabela 1. TRAFNOŚĆ PREDYKCJI

Empiryczne Teoretyczne	$y = 0$	$y = 1$
$\hat{y} = 0$ .....	$N_{00}$	$N_{01}$
$\hat{y} = 1$ .....	$N_{10}$	$N_{11}$

Źródło: opracowanie własne.

$N_{00}$  jest liczbą obserwacji, dla których zarówno obserwowana jak i prognozowana wartość zmiennej zależnej wynosi 0.  $N_{01}$  jest liczbą obserwacji, dla których zmienna zależna przyjmuje wartość 1, jednak jej wewnątrzpróbkowa prognoza wynosi 0, zaś  $N_{10}$  jest liczbą takich obserwacji, dla których prognozowana wartość wynosi 1 a rzeczywista 0. Liczba obserwacji, dla których zarówno empiryczna jak i prognozowana wartość zmiennej zależnej jest równa 1 wynosi  $N_{11}$ . Oznacza to zatem, że procent poprawnych predykcji obliczany jest za pomocą następującej formuły:

$$CP = \frac{N_{00} + N_{11}}{N_{00} + N_{11} + N_{01} + N_{10}} \tag{3}$$

Podczas generowania tabeli trafności predykcji, ważnym zagadnieniem jest wybór parametru  $c$ . W przypadku, gdy udział „jedynek” w próbie nie jest zbyt odległy od 0,5, to wówczas przyjmuje się, że  $c = 0,5$ . Jeśli jednak próba ma charakter niezbilansowany (udział „zer” lub „jedynek” jest bliski 100%) to wówczas punkt progowy  $c$  powinien być równy udziałowi „jedynek” w próbie. Oprócz wartości zmiennej  $CP$  na podstawie wartości z tabeli 1 obliczane są także następujące wielkości:

- czułość (ang. *sensitivity*):

$$SENSITIVITY = \frac{\sum_{i=1}^N I\{\hat{y}_i = 1 | \mathbf{x}_i\} \cdot I\{y_i = 1\}}{\sum_{i=1}^N I\{y_i = 1\}}, \tag{4}$$

- specyficzność:

$$SPECIFICITY = \frac{\sum_{i=1}^N I\{\hat{y}_i = 0 | \mathbf{x}_i\} \cdot I\{y_i = 0\}}{\sum_{i=1}^N I\{y_i = 0\}}, \tag{5}$$

- udział empirycznych „jedynek”, jeśli prognozowana jest wartość 1:

$$PPV = \frac{\sum_{i=1}^N I\{\hat{y}_i = 1 | \mathbf{x}_i\} \cdot I\{y_i = 1\}}{\sum_{i=1}^N I\{\hat{y}_i = 1 | \mathbf{x}_i\}}, \tag{6}$$

– udział empirycznych „zer”, jeśli prognozowana wartość wynosi 0:

$$NPV = \frac{\sum_{i=1}^N I\{\hat{P}(y_i = 1|\mathbf{x}_i) \leq c\} \cdot I\{y_i = 0\}}{\sum_{i=1}^N I\{\hat{P}(y_i = 1|\mathbf{x}_i) \leq c\}}, \quad (7)$$

Analizowane powyżej mierniki zależą od poziomu  $c$ . Nietrudno zauważyć, że wraz ze wzrostem parametru  $c$ , wzrasta wartość miernika (5) i obniża się wartość miernika (4). Udział „jedynek” w próbie jest tzw. optymalną wartością graniczną Cramera (por. Gruszczynski, 2010). Możliwe jest ustalenie  $c$  na poziomie minimalizującym prawdopodobieństwo popełnienia błędu predykcji. A zatem optymalna wartość parametru wyznaczana jest na podstawie wzoru:

$$C_{OTP} = \arg \min_c \left( \frac{\sum_{i=1}^N I\{\hat{P}(y_i = 1|\mathbf{x}_i) > c\} \cdot I\{y_i = 0\}}{\sum_{i=1}^N I\{y_i = 0\}} + \frac{\sum_{i=1}^N I\{\hat{P}(y_i = 1|\mathbf{x}_i) \leq c\} \cdot I\{y_i = 1\}}{\sum_{i=1}^N I\{y_i = 1\}} \right). \quad (8)$$

Krzywa ROC (ang. *Receiver Operating Characteristic*) wykorzystywana jest w celu analizy zależności pomiędzy czułością a specyficznością dla różnych wartości współczynnika  $c$ . Ilustruje ona wszystkie kombinacje obu rodzajów błędów dla różnych wartości progowych. Pokazuje zależność między *SENSITIVITY* oraz *1-SPECIFICITY*. Pole pod krzywą ROC wykorzystywane jest także do oceny jakości dopasowania modelu do danych. Przyjmuje ono wartości z przedziału  $(0,5; 1)$ . Jeśli powierzchnia pod krzywą ROC jest minimalna, to wówczas model nie ma żadnej mocy predykcyjnej. Drugi krańcowy przypadek (pole=1) oznacza, że model idealnie prognozuje wartości empiryczne. Koncepcja krzywej ROC wykorzystywana była m.in. w analizie ubóstwa miejskich i wiejskich gospodarstw domowych (por. Sączewska-Piotrowska, 2016).

Porównanie średniego prawdopodobieństwa, że zmienna zależna przyjmuje wartość 1 czyli wielkości  $\frac{\sum_{i=1}^N \hat{P}(y_i = 1|\mathbf{x}_i)}{N}$  z udziałem „jedynek” w próbie czyli wielkością  $\bar{y}$  nie dostarcza odpowiednich informacji na temat jakości dopasowania modelu do danych. W przypadku modelu logitowego zawierającego wyraz wolny, te dwie wielkości są sobie równe (por. m.in. Lechner, 1991). Dlatego też w celu analizy jakości dopasowania modelu do danych często wykorzystywany jest test specyfikacji zaproponowany przez Hosmera i Lemeshowa (1980). Ocenia on jakość dopasowania poprzez porównanie udziału „jedynek” w próbie z prawdopodobieństwem, że zmienna zależna przyjmuje wartość 1 w podgrupach.

Ponieważ zgodność pomiędzy teoretycznym a empirycznym rozkładem składnika losowego jest warunkiem koniecznym zgodności estymatora największej wiarygodności, po dokonaniu estymacji parametrów modelu logitowego lub probitowego, uzasadnione jest testowanie symetryczności rozkładu składnika

losowego. W tym celu Stukel (1988) zaproponował testowanie symetryczności rozkładu składnika losowego za pomocą testu mnożnika Lagrange'a. Sposób przeprowadzenia tego testu jest następujący:

- 1) Dokonuje się estymacji parametrów modelu logitowego/probitowego.
- 2) Następnie wyznaczane są wartości teoretyczne dla zmiennej nieobserwowalnej oraz ich kwadraty:  $(\mathbf{x}_i^T \hat{\boldsymbol{\alpha}}^{ML})^2$ .
- 3) Następnie dokonuje się estymacji parametrów następującego modelu logitowego/probitowego:

$$y_i^* = \mathbf{x}_i^T \boldsymbol{\alpha} + \lambda (\mathbf{x}_i^T \hat{\boldsymbol{\alpha}}^{ML})^2 + \varepsilon_i, \quad \varepsilon_i \sim F, \quad (9.a)$$

$$y_i = I\{y_i^* > 0\}. \quad (9.b)$$

- 4) Za pomocą testu Walda, weryfikowana jest następująca hipoteza:

$$H_0: \lambda = 0, \quad (10.a)$$

$$H_1: \lambda \neq 0. \quad (10.b)$$

W celu weryfikacji hipotezy (10) wykorzystywana jest następująca statystyka testu Walda:

$$WALD = \frac{(\hat{\lambda}^{ML})^2}{\hat{\sigma}_{\lambda}^2}, \quad (11)$$

gdzie  $\hat{\sigma}_{\lambda}^2$  jest oszacowaniem wariancji estymatora parametru  $\lambda$  z modelu (9.a)–(9.b).

Uzyskana wartość statystyki (11) porównywana jest z wartością krytyczną dla rozkładu chi-kwadrat z jednym stopniem swobody. Obszar odrzuceń ma charakter jednostronny. Hipoteza zerowa wskazuje na symetryczność rozkładu składnika losowego.

Podczas testowania za pomocą testu Stukela (1988) należy uwzględnić fakt, że omówiona powyżej procedura jest nieformalna, gdyż nie testuje *par excellence* względem siebie modelu z asymetrią i bez asymetrii. Formalne testowanie symetryczności przy wykorzystaniu technik Bayesowskich zostało zaproponowane przez m.in. Osiewalskiego, Marca (2004). Wykorzystując techniki klasyczne, formalne testowanie symetryczności może zostać wykonane za pomocą statystyki zaproponowanej przez Lechnera (1991). Polega ono na wykorzystaniu faktu, że rozkład logistyczny jest szczególnym przypadkiem rozkładu Burra 2-ego typu z parametrem  $J$ . Jeśli w modelu (1.a)–(1.b) składnik losowy  $\varepsilon_t$  pochodzi z rozkładu Burra drugiego typu z parametrem  $J$ , to wówczas dystrybuenta jego rozkładu przyjmuje postać:

$$F(\varepsilon_t) = \frac{1}{(1 + \exp(-\varepsilon_t))^{-J}}, \quad J \geq 0. \quad (12)$$

Model dwumianowy ze składnikiem losowym pochodzącym z rozkładu Burra drugiego typu sprowadza się do modelu logitowego, jeśli  $J = 1$ . Jeśli  $J < 1$  mamy do czynienia z asymetrią lewostronną natomiast w przypadku gdy  $J > 1$  pojawia się problem asymetrii prawostronnej. Dlatego też testowanie symetryczności polega na testowaniu prawdziwości restrykcji  $J = 1$ . Lechner (1991) oraz Thomas (1993) proponują oszacowanie parametrów modelu ze składnikiem losowym pochodzącym z rozkładu Burra drugiego typu oraz/lub składnikiem losowym pochodzącym z rozkładu logistycznego metodą największej wiarygodności a następnie wykorzystanie jednej z trzech klasycznych statystyk (mnożnika Lagrange'a, ilorazu wiarygodności, Walda). Niech  $\hat{\alpha}^{LOGIT}$  oraz  $\hat{\alpha}^{BURR}$  będą estymatorami największej wiarygodności parametrów odpowiednio przy założeniu, że składnik losowy pochodzi z rozkładu logistycznego oraz rozkładu Burra drugiego typu natomiast  $\hat{j}^{BURR}$  jest estymatorem największej wiarygodności parametru  $J$  przy założeniu, że składnik losowy pochodzi z rozkładu Burra drugiego typu. Wówczas statystyka testu ilorazu wiarygodności przyjmuje postać:

$$LR = -2 \left[ \begin{array}{l} \sum_{i=1}^N I\{y_i = 1\} \log \left( 1 - \left( 1 + \exp(\mathbf{x}_i^T \hat{\alpha}^{LOGIT}) \right)^{-1} \right) + \\ \sum_{i=1}^N I\{y_i = 0\} \log \left( \left( 1 + \exp(\mathbf{x}_i^T \hat{\alpha}^{LOGIT}) \right)^{-1} \right) + \\ - \sum_{i=1}^N I\{y_i = 1\} \log \left( 1 - \left( 1 + \exp(\mathbf{x}_i^T \hat{\alpha}^{BURR}) \right)^{-j^{BURR}} \right) - \\ - \sum_{i=1}^N I\{y_i = 0\} \log \left( \left( 1 + \exp(\mathbf{x}_i^T \hat{\alpha}^{BURR}) \right)^{-j^{BURR}} \right) \end{array} \right]. \quad (13)$$

Przy prawdziwości hipotezy zerowej, statystyka (13) ma rozkład chi-kwadrat z jednym stopniem swobody. Jeśli odrzucona zostanie hipoteza o symetryczności rozkładu należy bazować na oszacowaniach parametrów modelu z rozkładem Burra drugiego typu dla składnika losowego lub oszacować parametry komplementarnego modelu log-log. W przypadku braku symetryczności rozkładu składnika losowego, można też oszacować parametry modelu dychotomicznego ze skośnym rozkładem t-Studenta (por. Osiewalski, Marzec, 2004).

W modelu dwumianowym zakłada się, że dla wszystkich obserwacji rozkład składnika losowego jest taki sam. W związku z tym przyjmuje się założenie o stałości wariancji. Alternatywnym modelem, w którym przyjmuje się założenie o heteroskedastyczności składnika losowego jest heteroskedastyczny model probitowy:



$$P(y_i = 1 | \mathbf{x}_i) = \Phi\left(\frac{\mathbf{x}_i^T \boldsymbol{\alpha}}{\sigma_i}\right), \quad (14)$$

gdzie:  $\sigma_i^2 = \exp(\mathbf{z}_i^T \boldsymbol{\delta})$ ,  $\mathbf{z}_i$  jest wektorem zmiennych egzogenicznych niezawierającym stałej.

W celu sprawdzenia, czy mamy do czynienia z heteroskedastycznością składnika losowego, weryfikowana jest następująca hipoteza:

$$H_0: \boldsymbol{\delta} = \mathbf{0}, \quad (15.a)$$

$$H_1: \boldsymbol{\delta} \neq \mathbf{0}. \quad (15.b)$$

Prawdziwość hipotezy (15) weryfikowana jest za pomocą statystyki testu Walda.

## 2.2. Klasyczne uśrednianie w modelu dwumianowym

Jeśli zbiór potencjalnych zmiennych objaśniających jest duży, trudno jest zdecydować się na optymalny wariant modelu. Zastosowanie klasycznej strategii „od ogółu do szczegółu” może doprowadzić do takiej sytuacji, że istotna i ważna zmienna zostanie pominięta ze względu na wysoki graniczny poziom istotności przy tej zmiennej wynikający z wysokiej wartości współczynnika korelacji pomiędzy tą zmienną a innym regresorem. Oprócz tego zastosowanie uśredniania klasycznego (lub alternatywnej strategii czyli uśredniania bayesowskiego) umożliwia porównanie poszczególnych regresorów pod kątem ich wkładu do wartości informacyjnej modeli zawierających te zmienne. Jeśli liczba potencjalnych zmiennych objaśniających wynosi  $M$ , to wówczas liczba możliwych podzbiorów wynosi  $P = 2^M$ , włączając w to zbiór pusty. Niech  $p = 1, 2, \dots, P$  numeruje modele zawierające różne warianty zestawu zmiennych objaśniających. Dla każdego modelu obliczana jest wartość kryterium informacyjnego Akaike'a (AIC) (por. Akaike, 1974):

$$AIC_p = -2 \ln L_{U,p} + 2K_p, \quad (16)$$

gdzie  $K_p$  jest liczbą parametrów do estymacji w  $p$ -tym modelu. Spośród wielu modeli, preferowane są warianty o najmniejszych wartościach kryterium (16).

Możliwa jest jednak estymacja parametrów modelu dwumianowego zgodnie z zaproponowaną przez Burnhama, Anderssona (2002, 2004) metodą uśredniania modeli. Niech  $\hat{\boldsymbol{\alpha}}_p$  będzie  $M+1$ -wymiarowym wektorem oszacowań parametrów modelu dwumianowego w przypadku, gdy wybierany jest  $p$ -ty model. Oznacza to zatem, że jeśli  $m$ -ta zmienna nie występuje w  $p$ -tym wariantcie, to wówczas odpo-

wiedni element wektora  $\hat{\alpha}_p$  wynosi 0. Ponieważ najlepszy model posiada najniższą wartość kryterium informacyjnego Akaike, dla każdego z  $P$  modeli obliczana jest tzw. odległość od wzorca czyli różnica pomiędzy wartością kryterium informacyjnego dla tego modelu oraz minimalną wartością kryterium:

$$\Delta_p = AIC_p - AIC_{\min}. \quad (17)$$

Następnie obliczane są tzw. wagi Akaike dla każdego modelu w następujący sposób:

$$w_p = \frac{\exp\left(-\frac{\Delta_p}{2}\right)}{\sum_{s=1}^P \exp\left(-\frac{\Delta_s}{2}\right)}. \quad (18)$$

Jak pokazali Burnham, Andersson (2002, 2004), wielkości typu  $\frac{w_p}{w_r}$  są równoważne ilorazom funkcji wiarygodności dla modeli  $p$  oraz  $r$ . Dlatego też wielkość (18) należy interpretować jako prawdopodobieństwo, że model  $p$  jest właściwy. A zatem estymator największej wiarygodności po zastosowaniu klasycznego warunkowego<sup>5</sup> uśredniania międzymodelowego dla  $m$ -tego elementu wektora  $\alpha_p$  jest następujący:

$$\hat{\alpha}_{UM,m} = \frac{\sum_{p=1}^P w_p \hat{\alpha}_{p,m}}{\sum_{p=1}^P w_p J_m^p}, \quad (19)$$

gdzie  $\hat{\alpha}_{p,m}$  jest oszacowaniem parametru przy  $m$ -tej zmiennej, jeśli  $p$ -ty model został wybrany, natomiast  $J_{p,m}$  jest indykátorem określającym czy  $m$ -ta zmienna występuje jako objaśniająca w  $p$ -tym wariancie.

Wnioskowanie o istotności wpływu poszczególnych zmiennych objaśniających na zmienną binarną wymaga obliczenia oszacowań wariancji estymatorów parametrów. Oszacowanie wariancji dla  $m$ -tego elementu wektora  $\alpha_p$  obliczane jest następująco:

$$\text{var}(\hat{\alpha}_{UM,m}) = \left[ \sum_{p=1}^P J_m^p \left[ \text{var}(\hat{\alpha}_{p,m} | g_p) + (\hat{\alpha}_{p,m} - \hat{\alpha}_{UM,m}) \right]^{1/2} \right]^2, \quad (20)$$

gdzie  $(\hat{\alpha}_{p,m} | g_p)$  jest oszacowaniem wariancji estymatora parametru przy  $m$ -tej zmiennej w przypadku gdy wybierany jest  $p$ -ty model.

<sup>5</sup> Uśrednianie warunkowe polega na tym, że przy wyliczaniu oszacowania uwzględniane są tylko te warianty, w których dana zmienna występuje. W przypadku uśredniania Bayesowskiego uwzględniane są wszystkie warianty, czyli także takie, w których wartość parametru wynosi 0.

W przypadku klasycznego uśredniania międzymodelowego, oprócz wnioskowania o istotności zmiennych na podstawie oszacowań parametrów i średnich błędów szacunku, możliwe jest wnioskowanie o ważności zmiennej objaśniającej w kształtowaniu się danej zmiennej binarnej. W tym celu dla każdej zmiennej obliczana jest jej ważność, którą oblicza się następująco:

$$waz_m = \sum_{p=1}^P J_m^p w_p. \quad (21)$$

Miernik (21) przyjmuje zatem wartości z przedziału  $<0,1>$ . Im wyższa jest wartość tego miernika dla danej zmiennej tym większa jest jej moc w wyjaśnieniu kształtowania się zmiennej zależnej. Wynika to z faktu, że wagi dla modeli niezawierających tej zmiennej muszą być niskie, natomiast wagi dla modeli uwzględniających daną zmienną są wysokie.

### 3. SPECYFIKACJA DYCHOTOMICZNEGO MODELU LOGITOWEGO OBJAŚNIAJĄCEGO PRAWDOPODOBIENSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO

Wykorzystywana w badaniu empirycznym zmienna zależna przyjmuje wartość 1 dla tych respondentów, którzy mieli problem prawny w okresie 5 lat poprzedzających badanie. Respondenci Ci byli pytani, czy w analizowanym okresie znaleźli się w sytuacji wymagającej pomocy prawnika, osoby z wykształceniem prawniczym lub instytucji/organizacji udzielającej porad prawnych. Prezentowane wyniki badania empirycznego oparte są na reprezentatywnej próbie 572 respondentów. Tabela 2 przedstawia rozkład częstości dla dychotomicznej zmiennej zależnej. Ponad 70% respondentów nie spotkało się z problemem prawnym w okresie ostatnich 5 lat, natomiast mniej niż 30% respondentów odpowiedziało twierdząco na zadane w ankiecie pytanie.

Tabela 2. ROZKŁAD CZĘSTOŚCI DLA ZMIENNEJ DYCHOTOMICZNEJ

Wartość	Częstość
0 .....	0,717
1 .....	0,283

Źródło: obliczenia własne z wykorzystaniem danych z ogólnopolskiego badania ankietowego ISP.

Przegląd literatury teoretycznej i empirycznej (np. Murayama, 2007; Kritzer, 2008; Pleasence i inni, 2011; Winczorek, 2015) wskazuje na następujące grupy czynników ryzyka wystąpienia problemu prawnego:

- indywidualne cechy społeczno-ekonomiczne i demograficzne,
- indywidualne postawy wobec przestrzegania i stosowania prawa,
- indywidualną aktywność społeczną,
- środowisko społeczno-ekonomiczne i bariery dostępu do poradnictwa prawnego.

Tabela 3. LISTA CZYNNIKÓW, KTÓRE MOGA WPLÝWAĆ NA PRAWDOPODOBIEŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO

Czynniki	Symbol zmiennej/ jednostka miary	Uzasadnienie teoretyczne i empiryczne z obszaru nauk społecznych (socjologicznych, prawnych, ekonomicznych)	Oczekiwany znak efektu krańcowego, kategoria referencyjna, uwagi
<b>Cechy społeczno-demograficzne</b>			
Płeć	FEM (0-1)	Kobiety częściej doświadczają problemów prawnych, co wiązać się może z ich gorszą pozycją społeczno-ekonomiczną w większości społeczeństw (Moorhead i inni, 2004).	(+) Kategoria referencyjna: mężczyźni.
Wiek	AGE	Występowanie problemów prawnych jest zależne od wieku respondentów, przy czym zależność ta nie ma charakteru liniowego. Wynika to szeroko rozumianego cyklu życia jednostki (Pleasence i inni, 2004; Currie, 2006; Coumarelos i inni, 2006; Curran, 2007).	(+)
Wiek do kwadratu	AGE2		(-)
Wykształcenie ponadpodstawowe: podstawowe nieukończone	EDUP(0-1)	<i>Ceteris paribus</i> lepsze wykształcenie powinno skutkować obszerniejszą wiedzą ogólną i tym samym bardziej aktywną postawą wobec zaistniałego problemu prawnego (Legal Service Corporation, 1994; Murayama, 2007; Feisert i inni, 2013).	(+) Kategoria referencyjna: wykształcenie podstawowe lub niższe. Należy oczekiwać dodatnich oszacowań parametrów.
Wykształcenie ponadśrednie	EDUS (0-1)		
Stan cywilny (kawaler/panna)	SCKA (0-1)	Stan cywilny w dużym stopniu wpływa na inne cechy osobnicze, które mogą oddziaływać na percepcję rzeczywistości i warunkują przyjmowanie określonych postaw i zachowań, które w konsekwencji mogą sprzyjać lub oddalać prawdopodobieństwo wystąpienia problemu prawnego (Legal Service Corporation, 1994).	(?) Kategoria referencyjna: „inne”. Trudno <i>a priori</i> przesądzać o kierunku, zależności, jej rzędzie i relacji względem pozostałych parametrów określających stan cywilny.
Stan cywilny (żonaty/mężatka)	SCZO (0-1)		
Stan cywilny (rozwódnik/rozwódka)	SCRO (0-1)		
Stan cywilny (wdowiec/wdówka)	SCWD (0-1)		
<b>Cechy ekonomiczno-społeczne</b>			
Logarytm dochodu osobę w gospodarstwie	lnDOCH (zł)	Wyższy dochód zmniejsza bariery dostępu do usług prawnych, zwłaszcza komercyjnych. Osoby z relatywnie niskim i relatywnie wysokim dochodem częściej doświadczają problemów prawnych, chociaż z innych przyczyn (Legal Service Corporation, 1994).	(-)
Logarytm dochodu na osobę w gospodarstwie do kwadratu	lnDOCH2		(+)
Miejsce zamieszkania: miasto powyżej 20 tys. mieszkańców	RESCITY (0-1)	Relacje społeczne, warunki materialno-bytowe, dostępność do usług prawnych itd. są bardzo zróżnicowane względem miejsca zamieszkania. Sprawdza to, że indywidualna podatność na problemy prawne może silnie zależeć od miejsca zamieszkania (Pleasence i inni, 2004).	(?) Zmienna referencyjna: wieś lub miasto poniżej 20 tys. mieszkańców.

Tabela 3. LISTA CZYNNIKÓW, KTÓRE MOGĄ WPLYWAĆ NA PRAWDOPODOBIEŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO (cd.)

Czynniki	Symbol zmiennej/ jednostka miary	Uzasadnienie teoretyczne i empiryczne z obszaru nauk społecznych (socjologicznych, prawnych, ekonomicznych)	Oczekiwany znak efektu krańcowego, kategoria referencyjna, uwagi
<b>Cechy ekonomiczno-społeczne (cd.)</b>			
Sytuacja zawodowa: osoba pracująca	WORKA (0-1)	Sytuacja zawodowa wpływa na wysokość zarobków, relacje w miejscu pracy, więzi społeczne, samoocenę, itp. Nieintencjonalny brak pracy lub jej niedomiar prowadzić może do negatywnych konsekwencji psychologicznych, ekonomicznych i społecznych, co zmniejsza podmiotowość osób mało- lub nieaktywnych zawodowo, a w konsekwencji może oddziaływać na podatność na problemy prawne (Pleasence i inni, 2004; Murayama, 2007; Peisert i inni, 2013).	Zmienna referencyjna: (?) osoba niepracująca. Należy spodziewać się regularnej reakcji większościowej, mniejszościowej lub parabolicznej pomiędzy poszczególnymi parametrami stojącymi przy zmiennych określających sytuację zawodową respondenta.
Sytuacja zawodowa (2): osoba pracująca w jednym lub więcej miejscach pracy	WORKB (0-1)		Zmienna referencyjna: (?) Nie pracuję zawodowo lub Nie mam stałej pracy, pracuję dorywczo w różnych miejscach. Należy spodziewać się regularnej reakcji większościowej lub mniejszościowej pomiędzy poszczególnymi parametrami stojącymi przy zmiennych określających sytuację zawodową respondenta.
Subiektywnie postrzegany dobrobyt	SUB/ (0-1)	Subiektywnie postrzegany poziom zamożności może być istotnym czynnikiem wpływającym na percepcję rzeczywistości, w tym możliwości adekwatnej identyfikacji wystąpienia problemu prawnego oraz możliwości jego rozwiązania (Legal Service Corporation, 1994; Moorhead i inni, 2004; Coumarelos i inni, 2006).	Zmienna referencyjna: (?) pozostate odpowiedzi Wartość „1” — udzielenie odpowiedzi Żyjemy dobrze/bardzo dobrze na następujące pytanie: Które z wymienionych określeń najlepiej charakteryzuje sytuację P. gospodarstwa domowego.

Tabela 3. LISTA CZYNNIKÓW, KTÓRE MOGĄ WPŁYWAĆ NA PRAWDOPODOBIEŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO (cd.)

Czynniki	Symbol zmiennej/ jednostka miary	Uzasadnienie teoretyczne i empiryczne z obszaru nauk społecznych (sociojogicznych, prawnych, ekonomicznych)	Oczekiwany znak efektu krańcowego, kategoria referencyjna, uwagi
<b>Cechy ekonomiczno-społeczne (dok.)</b>			
Stan zdrowia własnego lub członka rodziny	HTH (0-1)	<p>Afirmatywy wybór przynajmniej jednej z opcji odpowiedzi na następujące pytanie: <i>Czy w ciągu ostatnich dwunastu miesięcy:</i></p> <p>a) <i>cierpiał/a P. z powodu długotrwałej choroby,</i>  b) <i>cierpiał/a P. z powodu niepełnosprawności,</i>  c) <i>był/a P. pacjentem/ką szpitala,</i>  d) <i>był/a P. przez miesiąc lub dłużej niezdolny/a do wykonywania swojej pracy zawodowej lub codziennych obowiązków w związku z chorobą lub obrażeniami ciała?</i>  (Curran, 2007; Coumarelos i inni, 2006).</p>	<p>(+)</p> <p>Zmienna referencyjna: odpowiedź <i>Nie</i>.</p>
<b>Świadomość prawna i postawa wobec przestrzegania i stosowania prawa</b>			
Świadomość prawna	PAWR liczba całkowita z przedziału 0-15	<p>Wraz ze wzrostem świadomości prawnej oczekiwać należy adekwatnej oceny zaistniałej sytuacji i pełniejszego uswiadomienia niekorzystnych następstw pozostawienia problemu nierozwiązanym (Pleasence i inni, 2003; Barendrecht; de Langen, 2008; Pleasence i inni, 2011; Winczorek, 2015).</p>	<p>(+)</p> <p>Zerowa znajomość prawa: PAWR = 0; doskonała znajomość prawa: PAWR = 15 (komentarz w tekście zasadniczym).</p>
Postawa wobec przestrzegania prawa: <i>powinno się zawsze przestrzegać prawa nawet jeśli jest ono niesłuszne</i>	PCPL (0-1)	<p>Postawa wobec przestrzegania prawa rzutuje na sposób rozwiązywania problemów prawnych (Kritzer i inni, 1991; Kurczewski, Fuszara, 2004; Pleasence i inni, 2004).</p>	<p>(-)</p> <p>Kategoria referencyjna: <i>Do przepisów prawa, które uważam za niesłuszne nie należy się wcale stosować lub Jeśli napotyka się przepisy prawa naszym zdaniem niesłuszne, należy się do nich stosować tylko na pozór.</i> Komentarz w tekście zasadniczym.</p>
Postawa wobec stosowania prawa	PUSE (liczba całkowita z przedziału 0-3)	<p>Postawa wobec stosowania prawa jest ważką i trwałą cechą osobniczą rzutującą na sposób rozwiązywania problemów prawnych (Kurczewski, Fuszara, 2004).</p>	<p>(+)</p> <p>Postawa ugodowa: PUSE=0; postawa „skargliwa”: PUSE=3 (komentarz w tekście zasadniczym).</p>

Tabela 3. LISTA CZYNNIKÓW, KTÓRE MOGĄ WPŁYWAĆ NA PRAWDOPODOBIENIŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO (dok.)

Czynniki	Symbol zmiennej/ jednostka miary	Uzasadnienie teoretyczne i empiryczne z obszaru nauk społecznych (socjologicznych, prawnych, ekonomicznych)	Oczekiwany znak efektu krańcowego, kategoria referencyjna, uwagi
<b>Postawa społeczna, kapitał społeczny, aktywność społeczna i intelektualna</b>			
Kapitał społeczny/zaufaanie do obcych osób	SCAP (0-1)	Afirmatywna odpowiedź na pytanie, cyt. <i>Ogólnie rzecz biorąc, czy uważa P., że można ufać większości ludzi, czy też sądzi P., że w postępowaniu z ludźmi ostrożności nigdy za wiele.</i>	(?) Kategoria referencyjna: odpowiedź <i>Nie.</i>
Przynależność do organizacji społecznych	APAR (0-1)	Aktywność społeczna świadczy o operatywności, inicjatywności upodmiotowieniu i uspołecznieniu osób aktywnych, co powinno skutkować innymi sposobami reakcji na zaistnienie problemu prawnego niż osób mniej społecznie aktywnych (por. Murayama, 2007; Peisert i inni, 2013).	(?) Nie należący do żadnej organizacji: APAR=0; Członek przynajmniej jednej organizacji: APAR=1.
Działalność społeczna	AACT (0-1)		(?) Nie: AACT=0 Tak: AACT=1
Aktywność intelektualna	INTA (0-1)	Następująca odpowiedź na pytanie, cyt. <i>Czy i jak w ciągu tygodnia czyta P. prasę codzienną lub czasopisma, albo poszukuje informacji w internecie czyta prasę codzienną lub czasopisma: „Tak, więcej niż trzy godziny w tygodniu”</i> (Peisert i inni 2013).	(?)
Dostęp do Internetu	INTB (0-1)	Afirmatywna odpowiedź na pytanie, cyt. <i>Czy ma P. dostęp do internetu w domu?</i> (Peisert i inni, 2013).	(?)
Korzystanie z Internetu	INTC (0-1)	Następująca odpowiedź na pytanie, cyt. <i>Jak często zagląda P. do serwisów i portali internetowych: „Raz w tygodniu lub częściej”</i> (Peisert i inni, 2013).	(?)

Źródło: opracowanie własne.

Pytania kwestionariuszowe, zawarte we wspomnianej we wprowadzeniu ogólnopolskiej ankiecie Instytutu Spraw Publicznych w Warszawie, objęły wszystkie wyżej wymienione uwarunkowania, co dało możliwość konstrukcji modelu logitowego objaśniającego prawdopodobieństwo wystąpienia problemu prawnego. W tabeli 3 przedstawiono listę potencjalnie istotnych czynników oddziałujących na to prawdopodobieństwo, które uwzględniono w ankiecie ISP, i które zakwalifikowano do jednej z wymienionych w poprzedzającym akapicie grup tematycznych.

Na podstawie przesłanek logicznych i dotychczasowych ustaleń teoretycznych, każdemu regresorowi przypisano oczekiwany znak parametru („+” lub „-”). Dla większości zmiennych w tabeli 3 znaki te są z góry ustalone, ale w przypadku niektórych regresorów nie było to możliwe. Taką sytuację oznaczono symbolem (?). Należy przy tym zaznaczyć, że zarówno znaki parametrów, jak i ich interpretacja jest warunkowa zarówno względem sposobu kodowania danej zmiennej objaśniającej<sup>6</sup>, jak i kategorii referencyjnej w dychotomicznym modelu logitowym<sup>7</sup>.

Wystąpienie problemu prawnego – tak jak inne wymiary życia – w dużym stopniu zależy od cech osobniczych respondenta i jego otoczenia społecznego. Dlatego we wszystkich badaniach empirycznych cytowanych w niniejszym artykule spotkać można zmienne objaśniające, które uwzględniają takie uwarunkowania. Tabela 3 zawiera listę zmiennych omawianego typu, która obejmuje czynniki społeczno-ekonomiczno-demograficzne oraz środowiskowe (miejsce zamieszkania). Większość z tych zmiennych przedstawiona jest w postaci zmiennych 0–1, co – po podaniu w odpowiednich kolumnach tabeli 3 informacji o kategorii referencyjnej (opcja, dla której dana zmienna przyjmuje wartość „0”) – umożliwia zrozumienie ich konstrukcji oraz oczekiwanego znaku parametru (tam gdzie takie ustalenia w odniesieniu do kierunku zależności pomiędzy regresorem a regresantem były *a priori* możliwe). Każdej zmiennej towarzyszy przy tym zwięzła wzmianka uzasadniająca jej obecność w modelu<sup>8</sup>.

Pewne zmienne z omawianych grup regresorów mają charakter polichotomiczny (stan cywilny, wykształcenie), w którym to przypadku uwagi sformułowane w poprzedzającym akapicie pozostają w mocy po uprzednim usta-

---

<sup>6</sup> Np. w przypadku zmiennej dychotomicznej, takiej jak „płeć”, wartość „0” przypisano do mężczyzn, zaś wartość „1” do kobiet, dlatego oczekiwany znak parametru stojącego przy tej zmiennej jest dodatni (kobiety częściej doświadczają problemów prawnych). Oczywiście jednoczesna zmiana sposobu kodowania („0” – kobiety, „1” – mężczyźni) skutkowałaby apriorycznym wskazaniem ujemnej *ceteris paribus* zależności pomiędzy taką zmienną a regresantem.

<sup>7</sup> Zmienną referencyjną w modelu jest brak problemu prawnego.

<sup>8</sup> Założyć można przy tym, że ze względu na renomę instytucji odpowiedzialnej za opracowanie i przeprowadzenie badań ankietowych (ISP w Warszawie) wszystkie pytania zawarte w ankiecie służą określonej celowi analitycznemu, zwłaszcza iż w dużym stopniu treści kwestionariusza czerpią z wcześniej ustalonych wzorców (np. Murayama, 2007; Gramatikov, 2008).



leniu kategorii referencyjnej. Z kolei kilka zmiennych objaśniających przyjmuje wartości z przedziałów całkowitoliczbowych (świadomość prawna, postawa wobec stosowania prawa), zaś jedna zmienna objaśniająca – dochód na osobę w gospodarstwie – wyrażona jest w jednostkach pieniężnych (w zł na osobę).

Odrębnego komentarza wymaga grupa czynników aproksymujących świadomość prawną i postawę wobec prawa. Do konstrukcji zmiennej mierzącej stopień świadomości/wiedzy prawnej danego respondenta, PAWR, wykorzystano sumę poprawnych odpowiedzi na wieloczkłonowe pytanie dotyczące zakresu znajomości prawa przez respondenta. Wartość tej zmiennej waha się w przedziale od 0 do 15. Oczekiwać należy przy tym, że wraz ze wzrostem znajomości prawa dana osoba będzie w stanie częściej zidentyfikować pojawiający się problem życiowy jako problem prawny. Zestaw pytań, na które respondenci udzielali odpowiedzi „zgadzam się” lub „nie zgadzam się” był następujący (na końcu pytania przytoczono poprawną na nie odpowiedź<sup>9</sup>):

- a) *Każdy ma prawo kopiować i robić odpisy z akt w postępowaniu wszczętym przez urząd przeciwko niemu (P),*
- b) *Każdy ma prawo przeczytać protokoły, uchwały z posiedzeń rady swojej gminy (P),*
- c) *Każdy ma prawo zwołać zgromadzenie publiczne bez pisemnego zezwolenia urzędu gminy (F),*
- d) *W postępowaniach przed sądem każdy ma prawo do uzyskania pomocy adwokata z urzędu, opłacanego z pieniędzy publicznych (F),*
- e) *Każdy może wejść na posiedzenia rady swojej gminy (P),*
- f) *Każdy ubezpieczony w NFZ może zmienić swojego lekarza rodzinnego (P),*
- g) *Bez względu na wiek dziecka lekarz ma prawo poinformować rodziców o stanie zdrowia ich dziecka, nawet jeśli dziecko sobie tego nie życzy (F),*
- h) *Osoba bezdomna ma prawo do bezpłatnych świadczeń zdrowotnych (F),*
- i) *Przewóz karetką do Zakładu Opieki Zdrowotnej i z powrotem osoby starszej lub niepełnosprawnej jest bezpłatny (F),*
- j) *Uczniowie liceum mają prawo wypowiadać się na temat programu swoich lekcji (P),*
- k) *Urząd Gminy musi konsultować z mieszkańcami lokalizację spalarni śmieci (P),*
- l) *Urząd Gminy może zamknąć szkołę podstawową, nawet jeśli nie zgadzają się na to rodzice dzieci, które do niej uczęszczają (P),*
- m) *Sąd Najwyższy może w Polsce wydawać ustawy (F),*
- n) *Prawo zabrania świadkowi składania niezgodnych z prawdą zeznań (P),*
- o) *Najważniejszym źródłem prawa w Polsce są rozporządzenia (F),*

---

<sup>9</sup> Autorzy pragną podziękować dr M. Araszkieviczowi za opracowanie poprawnych odpowiedzi na pytanie w ankiecie ISP, cyt. *Odczytam P. teraz kilka stwierdzeń, proszę powiedzieć, które z nich są P. zdaniem prawdziwe, a które nieprawdziwe?*

Postawy wobec przestrzegania i stosowania prawa są kolejnymi indywidualnymi cechami osobniczymi, które - w świetle literatury tematu (np. Kritzer i inni, 1991) – mogą wpływać na umiejętność poprawnej identyfikacji problemu życiowego jako problemu prawnego, i tym samym oddziaływać na podatność na występowanie problemów prawnych. W badaniu niniejszym pierwszy z wymienionych obszarów jest reprezentowany wyrażanymi przez respondenta poglądami na temat przestrzegania prawa w postaci jednej z czterech odpowiedzi na następujące pytanie: *Na temat praktycznego przestrzegania prawa w życiu są różne poglądy. Prosimy wybrać z przytoczonych zdań to, które bardziej P. odpowiada:*

1. *Powinno się zawsze przestrzegać prawa, nawet jeśli jest ono niesłuszne,*
2. *Jeśli napotyka się przepisy prawa naszym zdaniem niesłuszne, należy się do nich stosować tylko na pozór, a praktycznie starać się je omijać,*
3. *Do przepisów, które uważamy za niesłuszne nie należy się w ogóle stosować,*
4. *(nie czytać) nie wiem / trudno powiedzieć,*

przy czym zmienna objaśniana przyjmuje wartość „1” dla odpowiedzi pierwszej, zaś wartość „0” dla pozostałych opcji (zmienna PCPL w tabeli 3).

W odniesieniu do poglądów dotyczących stosowania prawa, wymiar ten jest reprezentowany jest zmienną PUSE, przyjmującą wartości całkowite z przedziału  $<0,3>$ . Odpowiednia zmienna objaśniająca powstaje w wyniku zsumowania afirmatywnych odpowiedzi nr 1 na pytanie ankietowe, cyt. *W sytuacji konfliktu, sporu z innymi, ludzie często podejmują różne działania. Co P. zdaniem jest lepsze w takiej sytuacji?* Respondent zaznaczał jedną z odpowiedzi dla każdego z trzech zbiorów:

Zbiór 1:

1. *Zaspokojenie w pełni żądań jednej ze stron, choćby druga była z tego niezadowolona...*
2. *Doprowadzenie do obopólnej zgody na zasadzie, że każdy rezygnuje po trosze ze swych roszczeń,*
3. *(nie czytać) nie wiem / trudno powiedzieć.*

Zbiór 2:

1. *Rozstrzygnięcie sporu ściśle według przepisów prawa, choćby druga strona była z tego niezadowolona,*
2. *Rozstrzygnięcie sporu ku zadowoleniu obu stron, choćby nie było to ściśle według przepisu prawa,*
3. *(nie czytać) nie wiem / trudno powiedzieć.*

Zbiór 3:

1. Rozwiązanie sporu przez instytucję, np. sąd, która ma władzę i może narzucić swoje rozwiązanie,
2. Rozwiązanie sporu przez postronnych ludzi, którzy mogą jedynie doradzić zważnionym stronom, jak postępować,
3. (nie czytać) nie wiem / trudno powiedzieć.

Jeśli respondent wskazał pierwszą odpowiedź w poszczególnych zbiorach, przypisywana była wartość 1, natomiast w przeciwnym przypadku przypisywana była wartość 0. Dlatego też im wyższa jest wartość zmiennej *PUSE* tym bardziej pryncypialna i skargliwa postawa wobec stosowania prawa reprezentowana jest przez tę osobę. Wydaje się, że im wyższa jest wartość analizowanej zmiennej, tym wyższe jest prawdopodobieństwo zaistnienia problemu prawnego.

Postać analityczna pozostałych zmiennych objaśniających wyszczególnionych w tabeli 3 jest – jak się wydaje – na tyle jednoznaczna, że zrezygnowano z ich dokładniejszego opisu w tekście zasadniczym. Warto jednak zaznaczyć, że ze względu na brak wielowymiarowych badań empirycznych w temacie będącym przedmiotem niniejszego artykułu, nie można uzyskanych w nim wyników porównać z teoretycznym, czy empirycznym wzorcem. Dodać trzeba również, że hipotezy dotyczące związków pomiędzy indywidualnymi regresorami a regresantem, lakonicznie wyrażone oczekiwanymi znakami parametrów stojących przy danej zmiennej objaśniającej – w tym również hipotezy niekonkluzywne w odniesieniu do wpływu danego czynnika ryzyka na wystąpienie problemu prawnego – sformułowane są na gruncie socjologii i ekonomii behawioralnej, nie zaś matematycznej. Nie istnieją bowiem sformalizowane modele socjologii czy ekonomii matematycznej, które umożliwiłyby bardziej sformalizowane podejście do kwestii modelowania prawdopodobieństwa wystąpienia problemu prawnego. Dlatego kluczowa rola przypada metodom eksploracyjnym i indukcyjnym, do których zaliczyć należy niniejsze badanie.

Postać ogólna rozważanego modelu jest następująca (symbole zmiennych i oczekiwane znaki parametrów podano w tabeli 3):

$$\begin{aligned}
 y_i^* = & \alpha_0 + \alpha_1 FEM_i + \alpha_2 AGE_i + \alpha_3 AGE2_i + \alpha_4 EDUP_i + \alpha_5 EDUS_i + \\
 & + \alpha_6 SCKA_i + \alpha_7 SCZO_i + \alpha_8 SCRO_i + \alpha_9 SCWD_i + \alpha_{10} \ln DOCH_i + \\
 & + \alpha_{11} \ln DOCH2_i + \alpha_{12} RESCITY_i + \alpha_{13} WORKA_i + \alpha_{14} WORKB_i + \\
 & + \alpha_{15} SUBI_i + \alpha_{16} HTH_i + \alpha_{17} PAWR_i + \alpha_{18} PCPL_i + \alpha_{19} PUSE_i + \\
 & + \alpha_{20} SCAP_i + \alpha_{21} APAR_i + \alpha_{22} AACT_i + \alpha_{23} INTA_i + \alpha_{24} INTB_i + \\
 & + \alpha_{25} INTC_i + \varepsilon_i,
 \end{aligned} \tag{22}$$

$$y_i = I\{y_i^* > 0\},$$

$$\varepsilon_i \sim F, \quad F(a) = \frac{\exp(a)}{1 + \exp(a)}$$

gdzie

$y_i$  – zmienna dychotomiczna przyjmująca wartość 1 jeśli  $i$ -ty respondent doświadczył problemu prawnego w ostatnich 5 latach oraz 0 w przeciwnym przypadku,  $i=1, \dots, N$ ,

$y_i^*$  – zmienna nieobserwowalna, którą można interpretować jako skłonność  $i$ -tego respondenta do doświadczania problemu prawnego,  $i=1, \dots, N$ ,

$\alpha_j$  – parametr stojący przy  $j$ -tym regresorze,  $j=0, \dots, 25$ .

Omówieniu wyników empirycznych badania, dla którego punktem wyjścia jest równanie, poświęcony jest kolejny punkt artykułu.

#### 4. WYNIKI I ICH INTERPRETACJA

W tabeli 4 przedstawiono wyniki oszacowań parametrów strukturalnych równania (22). Znaki oszacowań parametrów są w większości zgodne z oczekiwaniami. Jednakże liczne ze zmiennych objaśniających równania ogólnego charakteryzują się statystyczną nieistotnością. Trudno zatem uzasadnić ich obecność w tym równaniu, zwłaszcza że fakt ten nie pozostaje bez wpływu na oszacowania parametrów strukturalnych pozostałych regresorów. Dlatego w kolejnym kroku zastosowano strategię modelowania od ogółu do szczegółu, przyjmując za kryterium obecności w modelu szczegółowym poziom istotności 20%<sup>10</sup>. Wyniki zawiera tabela 5.

Tabela 4. REZULTATY OSZACOWANIA PARAMETRÓW MODELU ZAGREGOWANEGO OBJAŚNIAJĄCEGO PRAWDOPODOBIENSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO – WARIANT OGÓLNY

Zmienna	Oszacowanie (błąd stand.)	GPI <sup>a</sup>	Zmienna	Oszacowanie (błąd stand.)	GPI
Wyraz wolny	7,02 (5,99)	0,241	<i>WORKA</i>	-0,46 (0,47)	0,325
<i>FEM</i>	0,40 (0,21)	0,060	<i>WORKB</i>	-0,50 (0,47)	0,289
<i>AGE</i>	0,08 (0,04)	0,059	<i>SUBI</i>	0,10 (0,30)	0,743
<i>AGE2</i>	-0,001 (0,0005)	0,059	<i>HTH</i>	0,30 (0,24)	0,205

<sup>10</sup> Przyjęcie tak wysokiego poziomu istotności wynika z eksploracyjnego charakteru pracy. Zgodnie ze stanem wiedzy autorów jest to bowiem pierwsze tak obszerne badanie, zaś liczne hipotezy sformułowane w tabeli 4 mają charakter pilotażowy i nie zakładają jednoznacznie określonego kierunku relacji (znaku parametru) pomiędzy regresorem a regresantem.

Tabela 4. REZULTATY OSZACOWANIA PARAMETRÓW MODELU ZAGREGOWANEGO OBJASNIAJĄCEGO PRAWDOPODOBIEŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO – WARIANT OGÓLNY (dok.)

Zmienna	Oszacowanie (błąd stand.)	GPI <sup>a</sup>	Zmienna	Oszacowanie (błąd stand.)	GPI
<i>EDUP</i>	0,92 (0,45)	0,042	<i>PAWR</i>	-0,02 (0,05)	0,625
<i>EDUS</i>	1,48 (0,52)	0,005	<i>PCPL</i>	-0,27 (0,22)	0,216
<i>SCKA</i>	-0,13 (0,75)	0,864	<i>PUSE</i>	0,45 (0,10)	0,000
<i>SCZO</i>	0,46 (0,70)	0,513	<i>SCAP</i>	0,07 (0,23)	0,770
<i>SCRO</i>	1,56 (0,77)	0,044	<i>APAR</i>	0,09 (0,46)	0,842
<i>SCWD</i>	1,55 (0,79)	0,049	<i>AACT</i>	0,47 (0,28)	0,095
<i>lnDOCH</i>	-3,07 (1,75)	0,080	<i>INTA</i>	-0,18 (0,24)	0,447
<i>lnDOCH2</i>	0,20 (0,13)	0,130	<i>INTB</i>	-0,20 (0,26)	0,431
<i>RESCITY</i>	-0,33 (0,22)	0,131	<i>INTC</i>	0,76 (0,32)	0,017
R <sup>2</sup> McFaddena		0,138	Testowanie symetryczności rozkładu za pomocą statystyki testu ilorazu wiarygodności		Statystyka=1,13 GPI=0,29
Udział poprawnych predykcji		65,7%	Testowanie jakości dopasowania za pomocą testu Hosmera-Lemeshowa		Statystyka=5,6 GPI=0,69
Wrażliwość		69,8%	Test Stukela symetryczności rozkładu składnika losowego		Statystyka=0,02 GPI=0,89
Specyficzność		64,2%	Test homoskedastyczności składnika losowego		Statystyka=1,4 GPI=0,50
Pole pod krzywą ROC		0,75	Liczba obserwacji		572

a We wszystkich tabelach GPI oznacza graniczny poziom istotności.

Źródło: obliczenia własne z wykorzystaniem danych z ogólnopolskiego badania ankietowego ISP.

Liczba regresorów zmalała z wyjściowych 26-ciu w wariancie ogólnym do 13-stu w wariancie szczegółowym. Znaki parametrów przy wszystkich zmiennych objaśniających zgodne są przesłankami teoretycznymi, tam gdzie ustalenia takie były możliwe (tabela 3). Regresory reprezentują wszystkie grupy czynników wyszczególnione w tabeli 3. Za wyjątkiem dochodu i jego kwadratu oraz zamiennej *ACCT* (aktywność społeczna), wszystkie pozostałe regresory (nie licząc wyrazu wolnego) są istotne statystycznie na poziomie niższym od 10%. Miary dobroci modelu oraz wszystkie testy diagnostyczne przemawiają za uznaniem tego modelu za w pełni poprawny.

Tabela 5. REZULTATY OSZACOWANIA PARAMETRÓW MODELU ZAGREGOWANEGO OBJASNIAJĄCEGO PRAWDOPODOBIEŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO – WARIANT SZCZEGÓŁOWY

Zmienna	Oszacowanie (Błąd stand.)	GPI	Zmienna	Oszacowanie (Błąd stand.)	GPI
Wyraz wolny	4,42 (5,76)	0,442	<i>SCWD</i>	1,65 (0,48)	0,001
<i>FEM</i>	0,36 (0,21)	0,085	<i>lnDOCH</i>	-2,59 (1,73)	0,134

Tabela 5. REZULTATY OSZACOWANIA PARAMETRÓW MODELU ZAGREGOWANEGO OBJAŚNIAJĄCEGO PRAWDOPODOBIEŃSTWO WYSTĄPIENIA PROBLEMU PRAWNEGO – WARIANT SZCZEGÓŁOWY (dok.)

Zmienna	Oszacowanie (Błąd stand.)	GPI	Zmienna	Oszacowanie (Błąd stand.)	GPI
AGE	0,08 (0,04)	0,052	lnDOCH2	0,17 (0,13)	0,199
AGE2	-0,0008 (0,0004)	0,057	PUSE	0,41 (0,10)	0,000
EDUP	0,88 (0,45)	0,050	AACT	0,41 (0,25)	0,106
EDUS	1,53 (0,51)	0,003	INTC	0,57 (0,27)	0,034
SCZO	0,55 (0,31)	0,081	Test Walda zerowych restrykcji na parametry przy zmiennych nieuwzględnionych Statystyka=7,44; GPI=0,83		
SCRO	1,71 (0,46)	0,000			
R <sup>2</sup> McFaddena		0,127	Testowanie symetryczności rozkładu za pomocą statystyki testu ilorazu wiarygodności		Statystyka=0,95 GPI=0,33
Udział poprawnych predykcji		65,8%	Testowanie jakości dopasowania za pomocą testu Hosmera–Lemeshowa		Statystyka=5,1 GPI=0,74
Wrażliwość		70,4%	Test Stukela symetryczności rozkładu składnika losowego		Statystyka=0,01 GPI=0,91
Specyficzność		63,9%	Test homoskedastyczności składnika losowego		Statystyka=4,8 GPI=0,19
Pole pod krzywą ROC		0,74	Liczba obserwacji		572

Źródło: obliczenia własne z wykorzystaniem danych z ogólnopolskiego badania ankietowego ISP.

Jeśli model powyższy uznać za wiążący, to wówczas okazuje się, że wśród czynników ryzyka największe znacznie mają czynniki demograficzne (wiek i płeć) oraz te określające stan cywilny i wykształcenie. Są to zatem uwarunkowania bądź pozawolitionalne bądź trudne do zmiany w krótkim okresie. Drugą grupę stanowią czynniki związane z postawami społecznymi (AACT, INTC) oraz postawami wobec prawa (PUSE). Uwarunkowania ściśle ekonomiczne reprezentowane są jedynie przez wielkość dochodu na osobę w gospodarstwie i są to jednocześnie zmienne o najniższej istotności statystycznej.

Zastosowanie strategii modelowania od ogółu do szczegółu prowadzi do wyboru jednego tylko modelu – o ustalonej liczbie regresorów – który z punktu widzenia określonych kryteriów jest modelem najlepszym. Jest to zatem sytuacja, w której „zwycięzca bierze wszystko”, zaś wszystkie inne modele objaśniające to samo zjawisko – ale o innej niż „zwycięzca” liczbie regresorów – są *implicit*e dyskredytowane, nawet w przypadku, gdy ze względu na przyjęte kryterium selekcji modele takie ustępują modelowi najlepszemu jedynie nieznacznie (szerzej w: Florczak, 2012). Ich odrzucenie jest równoznaczne z pominięciem tej wiedzy o analizowanym zjawisku, która zawarta jest w modelach zdyskredytowanych, a tym samym zaprzepaszczone zostaje możliwość wzbogacenia analizy o aspek-

ty nieobecne w modelu uznanym za najlepszy. W celu zapobieżenia tej sytuacji zastosować można metodę klasycznego uśredniania międzymodelowego (ang. *frequentist model averaging*), co uczyniono w dalszej części badania<sup>11</sup>.

W tabeli 6 przedstawiono wyniki estymacji parametrów modelu warunkowego uśredniania klasycznego dla 16-stu zmiennych objaśniających (nie licząc wyrazu wolnego) o największej statystycznej istotności, które ustalono przy użyciu strategii modelowania od ogółu do szczegółu<sup>12</sup>. Prezentowane są oszacowania parametrów uzyskane za pomocą wzoru (19), oszacowania błędów standardowych uzyskane za pomocą wzoru (20) oraz mierniki ważności poszczególnych zmiennych, uzyskane za pomocą wzoru (21). W tabeli 7 wymienione są regresory w czterech najlepszych modelach charakteryzujących się najniższymi wartościami kryterium informacyjnego Akaike. Rezultaty badania pokazują, że jedynie nieliczne zmienne – wskazywane w literaturze przedmiotu oraz uproszczonych analizach statystycznych jako potencjalnie ważne – rzeczywiście uznać można za czynniki ryzyka w kontekście występowania problemu prawnego. Przy przyjęciu bardzo tolerancyjnego poziomu istotności i zastosowaniu strategii modelowania od ogółu do szczegółu liczba takich czynników wynosi kilkanaście, zaś zastosowanie bardziej restrykcyjnego – i w kontekście badań eksploracyjnych również bardziej wskazanego – podejścia analitycznego, jakim jest metoda klasycznego uśredniania międzymodelowego, liczba takich czynników okazuje się jeszcze niższa, wynosząc w przypadku uśredniania warunkowego jedynie 7 (dla poziomu istotności równego 10%).

Tabela 6. REZULTATY OSZACOWANIA PARAMETRÓW MODELU LOGITOWEGO PRZY UŻYCIU WARUNKOWEGO KLASYCZNEGO UŚREDNIANIA MIĘDZYMODELOWEGO

Zmienna	Oszacowanie (błąd stand.)	GPI	Ważność	Zmienna	Oszacowanie (błąd stand.)	GPI	Ważność
Wyraz wolny	-0,38 (5,05)	0,940	–	SCWD	1,59 (0,51)	0,002	0,99
FEM	0,38 (0,24)	0,113	0,64	lnDOCH	-1,36 (1,45)	0,348	0,64
AGE	0,09 (0,06)	0,134	0,64	lnDOCH2	0,07 (0,10)	0,484	0,55
AGE2	-0,0009 (0,0007)	0,199	0,65	RESCITY	-0,23 (0,19)	0,226	0,39
EDUP	0,93 (0,50)	0,063	0,79	HTH	0,26 (0,18)	0,149	0,40
EDUS	1,47 (0,52)	0,005	0,98	PCPL	-0,22 (0,14)	0,116	0,38
SCZO	0,68 (0,38)	0,074	0,76	PUSE	0,42 (0,10)	0,000	1,00
SCRO	1,71 (0,52)	0,001	1,00	AACT	0,43 (0,28)	0,125	0,59
				INTC	0,59 (0,27)	0,029	0,80

Źródło: obliczenia własne z wykorzystaniem danych z ogólnopolskiego badania ankietowego ISP.

<sup>11</sup> Wydaje się, że metody uśredniania międzymodelowego są szczególnie wskazane w przypadku badań eksploracyjnych, gdzie trudno o teoretyczne wsparcie w postaci modeli formalnych, i gdzie – m.in. właśnie za sprawą wcześniej wzmiankowanej przyczyny – liczba potencjalnych regresorów jest bardzo duża.

<sup>12</sup> Oznaczało to ustalenie poziomu istotności równego 30%.

Tabela 7. ZMIENNE OBJAŚNIAJĄCE WYSTĘPUJĄCE W CZTERECH NAJLEPSZYCH MODELACH

Model	1	2	3	4
Zmienne	INTC AATC FEM lnDOCH PUSE SCRO SCWD EDUP EDUS SCZO	INTC AATC FEM lnDOCH PUSE SCRO SCWD AGE AGE2 EDUP EDUS SCZO	INTC AATC lnDOCH PUSE SCRO SCWD EDUP EDUS SCZO	INTC AATC FEM lnDOCH lnDOCH2 PUSE SCRO SCWD EDUP EDUS SCZO

Źródło: obliczenia własne z wykorzystaniem danych z ogólnopolskiego badania ankietowego ISP.

Na podstawie uzyskanych wyników można wnioskować, że ze zbioru czynników ryzyka wykluczyć należy te z nich, które w tabeli 3 przypisano do grupy uwarunkowań społeczno-ekonomicznych (dochód, miejsce zamieszkania, stan zdrowia). Wykluczenie takie może nieco zaskakiwać, biorąc pod uwagę wnioski płynące z wcześniejszych – ale dodajmy jednowymiarowych – badań (Legal Service Corporation, 1994; Murayama, 2007; Gramatikov, 2008).

Kluczowymi – jak wynika z analizy ważności zmiennych (tabele 6 i 7) – czynnikami zwiększającymi ryzyko doświadczenia problemu prawnego jest indywidualna postawa wobec prawa, aproksymowana zmienną *PUSE*, oraz stan cywilny wskazujący na „życiowe przejścia” i samotność (rozwodnicy (*SCRO*), osoby owdowiałe (*SCWD*)). Okazuje się zatem, że subiektywny stosunek wobec stosowania prawa jest trwałym czynnikiem wpływającym na indywidualną skłonność do interpretowania pojawiających się zdarzeń życiowych w kategoriach problemów prawnych, co sprawia, iż osoby o pryncypialnym stosunku wobec prawa częściej doświadczają takich problemów (Kurczewski, Fuszara, 2004). W grupie osób charakteryzujących się wyższą wartością „indeksu skargliwości” (zmienna *PUSE*) wyższa jest szansa zaistnienia problemu prawnego w porównaniu z osobami charakteryzującymi się niskimi wartościami tego indeksu.

Osoby rozwiedzione i owdowiałe wyraźnie częściej niż osoby stanu wolnego lub żyjące w konkubinacie doświadczają problemów prawnych (ilorazy szans wystąpienia problemu u takich osób wynoszą odpowiednio 5,53 ( $\exp(1,71)$ ) i 4,90 ( $\exp(1,59)$ )). Bezspornie ważnymi okolicznościami wpływającymi na prawdopodobieństwo zaistnienia problemu prawnego są zmienne zastane, takie jak płeć i wiek. Kobiety częściej niż mężczyźni doświadczają problemów prawnych (iloraz równy 1,46), zaś wraz z wiekiem maleje podatność na jego doświadczanie. Zgodnie z wnioskami płynącymi z wcześniejszych analiz wzrost wykształcenia implikuje *ceteris paribus* wyższą podatność na wystąpienie problemów prawnych (ilorazy szans dla zmiennych *EDUP* i *EDUS* wynoszą odpowiednio 2,53 i 4,35).



Ostatnią grupę stanowią zmienne opisujące aktywność społeczną. Okazuje się, że – w świetle wyników niniejszego badania – czynny udział w życiu społecznym uznać należy za stymulantę pojawienia się problemu prawnego (ilorazy szans dla zmiennych *AACT* i *INTC* wynoszą odpowiednio 1,54 i 1,80). Trudno jednak wskazać na behawioralny mechanizm takiego związku. Być może pasem transmisyjnym jest możliwość szerszego korzystania z nieformalnej pomocy innych osób w procesie identyfikowania doświadczanego problemu życiowego jako prawnego, albo działanie w imieniu innych, czyli przedstawianie problemów innych osób, jako własnych.

Zastosowanie adekwatnych narzędzi analizy wielowymiarowej ograniczyło liczebność potencjalnych czynników ryzyka zaistnienia problemu prawnego do zaledwie kilku. Obok konkluzji szczegółowych, omówionych powyżej, wskazać można na dwa meta-wnioski płynące z badania. Po pierwsze, identyfikacja indywidualnych czynników ryzyka i kwantyfikacja ich wpływu na prawdopodobieństwo wystąpienia problemu prawnego może być szczególnie użyteczna w przypadku prób ustanowienia systemu ubezpieczeń od takiego zdarzenia losowego. Po drugie, zidentyfikowane i skwantyfikowane czynniki ryzyka są raczej mało użyteczne w kontekście rozważań nad alokacją środków publicznych na funkcjonowanie obecnego systemu nieodpłatnej pomocy prawnej, głównie ze względu na brak odpowiednich danych na poziomie mezo (powiatów). Dlatego wydaje się, że alokacja tych środków jedynie według kryterium liczebności populacji jest rozwiązaniem właściwym (Florczak, 2015, 2017).

## 5. UWAGI KOŃCOWE

Przedstawione w artykule badanie jest pierwszą tak szeroką próbą zastosowania narzędzi wielowymiarowej analizy statystyczno-ekonometrycznej do badania zjawisk standardowo przypisywanych do dziedziny nauk prawnych, i które do tej pory analizowane były za pomocą jednowymiarowych narzędzi statystyki opisowej. W celu zrealizowania zamierzonych celów badawczych zastosowano sekwencję metod analitycznych nieczęsto stosowaną w praktyce modelowania (model logitowy + strategia modelowania od ogółu do szczegółu + uśrednianie klasyczne).

Rezultaty i wnioski płynące z badania dają asumpt do dalszych analiz, w tym np. nad sposobami reakcji osób dotkniętych problemem prawnym, czy indywidualnymi efektami pomocy prawnej. Wszystkie takie badania ograniczone są jednak do wymiaru mikroekonomicznego, co wynika z rodzaju użytych w nich danych. Warto jednak wskazać na możliwość jakościowego poszerzenia kontekstu poznawczego tego typu badań poprzez ich połączenie z danymi mezo, czy nawet makroekonomicznymi, co umożliwiają metody modelowania hierarchicznego. Stanowić to będzie kolejny cel naukowy autorów niniejszego artykułu.

## LITERATURA

- Akaike H., (1974), A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, 19, 716–723.
- Barendrecht J. M., de Langen M., (2008), Legal Empowerment of the Poor: Innovating Access to Justice, w: *The State of Access: Success and Failure of Democracies to Create Equal Opportunities*, Brookings Institution Press, Washington, 250–271.
- Burdziej S., Dudkiewicz M., (2013), *Korzystający i niekorzystający z poradnictwa prawnego i obywatelskiego*, Instytut Spraw Publicznych, ISP, Warszawa.
- Burnham K. P., Anderson D. R., (2002), *Model Selection and Multimodel Inference. A Practical Information – Theoretic Approach*, Second Edition, Springer-Verlag, New York.
- Burnham K. P., Anderson D. R., (2004), Multimodel Inference. Understanding AIC and BIC in Model Selection, *Sociological Methods and Research*, 33, 261–304.
- Cameron A. C., Trivedi P. K., (2009), *Microeconometrics. Methods and Applications*, Cambridge University Press, Cambridge.
- Coumarelos C., Wei Z., Zhou A. Z., (2006), *Justice Made to Measure. NSW Legal Survey in Disadvantaged Areas*, Law and Justice Foundation, South Wales.
- Curran L., (2007), *Ensuring Justice and Enhancing Human Rights: A Report on Improving Legal Aid Service Delivery to Reach Vulnerable and Disadvantaged People*, La Trobe University and the Victoria Law Foundation, La Trobe.
- Currie A., (2006), A National Survey of the Civil Justice Problems of Low- and Moderate-Income Canadians: Incidence and Patterns, *International Journal of Legal Profession*, 13 (3), 217–242.
- Florczak W., (2012), O możliwości zintegrowanej weryfikacji empirycznej alternatywnych teorii na przykładzie teorii przestępczości, *Ekonomista*, 6/212, 735–764.
- Florczak W., (2015), Ile może kosztować reforma systemu pomocy prawno-obywatelskiej w Polsce, *Studia Prawno-Ekonomiczne*, 95, 189–206.
- Florczak W., (2017), Szacunki kosztów systemu pomocy prawno-obywatelskiej w Polsce, *Wiadomości Statystyczne*, 1, 7–25.
- Gramatikov M., (2008), Multiple Justiciable Problems in Bulgaria, *Tilburg University Legal Studies Working Paper*, 16/2008.
- Gruszczyński M., (2010), Modele zmiennych jakościowych dwumianowych, w: *Mikroekonometria. Modele i metody analizy danych indywidualnych*, Oficyna Wolters Kluwer Business, Warszawa, 53–101.
- Hosmer D. W., Lemeshow S., (1980), Goodness-of-fit Tests for the Multiple Logistic Regression Model, *Communications in Statistics: Theory and Methods*, 9, 1043–1069.
- Hunter R., De Simone T., (2009), Women, Legal Aid and Social Inclusion, *Australian Journal of Social Issues*, 44 (4), 379–398.
- Kritzer H. M., (2008), To Lawyer or Not to Lawyer: Is that the Question?, *Journal of Empirical Legal Studies*, 5, 875–906.
- Kritzer H. M., (2009), Examining the Real Demand for Legal Services, *Fordham Urban Law Journal*, 37, 255–272.
- Kritzer H. M., Bogart W. A., Vidmar N., (1991), *Context, Context, Context: A Cross-problem, Cross-Cultural Comparison of Compensation Seeking Behaviour*, Paper Prepared for Law and Society, Amsterdam.
- Kurczewski J., Fuszara M., (red.), (2004), *Polskie spory i sądy*, Ośrodek Badań Społecznych ISNS. Uniwersytet Warszawski, Warszawa.
- Lechner M., (1991), Testing Logit Models in Practice, *Empirical Economics*, 16, 177–198.
- Legal Service Corporation, (1994), *Legal Needs and Civil Justice, A survey of Americans. Major Finding, from Comprehensive Legal Needs Study*, American Bar Association, Chicago, Illinois.

- Marzec J., (2008), *Bayesowskie modele zmiennych jakościowych i ograniczonych w badaniach niespłacalności kredytów*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.
- McFadden D., (1974), The Measurement of Urban Travel Demand, *Journal of Public Economics*, 3, 303–328.
- Moorhead R., Douglas G., Doughty J., Sefton M., (2004), The Advice Needs of Lone Parents, *Family Law*, 34, 667–673.
- Murayama M., (2007), Experiences of Problems and Disputing Behaviour in Japan, *Meiji Law Journal*, 14, 1–59.
- Osiewlaski J., Marzec M., (2004), Uogólnienie dychotomicznego modelu probitowego z wykorzystaniem skośnego rozkładu Studenta, *Przegląd Statystyczny*, 51, 13–24.
- Pleasence P., Balmer N., Reimers S., (2011), What Really Drives Advice Seeking Behaviour? Looking Beyond the Subject of Legal Disputes, *Onati Socio-Legal Series*, 1, 1–21.
- Pleasence P., Buck A., Balmer N., O'Grady A., Genn H., Smith H., (2004), Causes of Action: Civil Law and Social Justice. The Final Report of the First LSRC Survey of Justiciable Problems, *Legal Service Commission Working Paper*.
- Peisert A., Schimanek T., Waszak M., Winiarska A., (red.), (2013), *Poradnictwo Prawne i Obywatelskie w Polsce, Stan obecny i wizje przyszłości*, Instytut Spraw Publicznych, ISP, Warszawa.
- Sączewska-Piotrowska A., (2016), Zastosowanie krzywych ROC w analizie ubóstwa miejskich i wiejskich gospodarstw domowych, *Przegląd Statystyczny*, 63, 211–232.
- Stukel T. A., (1988), Generalized Logistic Models, *Journal of American Statistical Association*, 83, 426–431.
- Świadek A., Wojciech M., (2015), Analiza skłonności przedsiębiorstw z wiodących regionów Polski do inwestowania w nowe środki trwałe z wykorzystaniem modelu logitowego uwzględniającego interakcje zmiennych jakościowych, *Przegląd Statystyczny*, 62, 117–138.
- Thomas J. M., (1993), On Testing the Logitsic Assumption in Binary Dependent Variable Models, *Empirical Economics*, 18, 381–392.
- Winczorek J., (2015), *Przegląd literatury na temat dostępności i korzystania z pomocy prawnej*, INPRIS, Instytut Prawa i Społeczeństwa, Warszawa.

## **CO WPŁYWA NA WIELKOŚĆ POPYTU NA PORADY PRAWNE? ANALIZA LOGITOWA Z WYKORZYSTANIEM METODY KLASYCZNEGO UŚREDNIANIA MIĘDZYMODELOWEGO**

### **Streszczenie**

*Celem artykułu jest identyfikacja – na podstawie przeglądu literatury – i kwantyfikacja, przy użyciu modelu logitowego, wpływu czynników ryzyka na prawdopodobieństwo wystąpienia problemu prawnego w Polsce. Zbiór regresorów objął czynniki obiektywne – takie jak dochód, wiek, wykształcenie, stan cywilny, miejsce zamieszkania, sytuację zawodową – oraz subiektywne: postawę wobec prawa, świadomość prawną, kapitał ludzki oraz aktywność społeczną. Z badania, przeprowadzonego na reprezentatywnej próbie dorosłych Polaków, wynika, że czynniki ryzyka znajdują się poza obszarem oddziaływania instytucjonalnego. Dlatego alokacja środków publicznych na funkcjonowanie nieodpłatnego systemu pomocy prawnej w oparciu o kryterium liczebności populacji*

wydaje się właściwa. Ze względu na brak analogicznych analiz oraz eksploracyjny charakter badania, dla ostatecznego ustalenia zbioru statystycznie istotnych regresorów zdecydowano się na zastosowanie metody logitowego uśredniania klasycznego. Jej użycie nie zmieniło znacząco wniosków płynących z analizy klasycznej.

**Słowa kluczowe:** problemy prawne, popyt na usługi prawne, model logitowy, klasyczna metoda uśredniania międzypodelowego

## WHAT INFLUENCES DEMAND FOR LEGAL AID? LOGIT ANALYSIS WITH USE OF CLASSICAL MODEL AVERAGING

### Abstract

*The aim of this paper to identify, by means of relevant literature survey, and then to quantify – using representative panel of individual questionnaire data and a logit model – the impact of relevant factors affecting demand for legal aid in Poland. The set of explanatory variables contains objective factors – such as income, age, education, gender, marital status, place of residence, occupational status – as well as subjective ones, such as personal attitude towards law, knowledge of law, social capital or social activity. It follows from the results obtained on a representative sample of adult Poles that the number of factors influencing the occurrence of a legal problem is large, the factors themselves being beyond the scope of direct and intentional impact of the state. Thereby, it seems reasonable to allocate the funds devoted to the funding and functioning of the reformed legal aid system in Poland on the basis of the population size criterion. In view of the relative scarcity of quantitative research into the issues raised in the paper frequentist model averaging method has been also used to confirm/reject the conclusions draw on the basis of the logit model. However, this has not altered the afore-mentioned conclusions in any significant way.*

**Keywords:** logit model, legal problems, demand for legal services, classical model averaging