Fundusze Europejskie
Pomoc Techniczna

Rzeczpospolita Polska

100 lat GUS

Unia Europejska
Fundusz Spójności

Rozszerzenie Badania Aktywności Ekonomicznej Ludności Moduł II

Marcin Szymkowiak

Urząd Statystyczny w Poznaniu

Warszawa, 23.11.2018 r.

Plan prezentacji

1. Moduł II - główny cel
2. Podejście kalibracyjne w ujęciu teoretycznym
3. Jedno- i dwukrokowa kalibracja
4. Kalibracja wag w BAEL
5. Kalibracja wag w BAEL z uwzględnieniem podregionów
6. Modele klasy SMO
7. Podsumowanie
8. Literatura

Moduł II - główny cel

- Głównym celem modułu drugiego było **zbadanie możliwości i opracowanie metodologii szacowania podstawowych wskaźników z zakresu rynku pracy** (stopa bezrobocia, wskaźnik zatrudnienia, liczba bezrobotnych, liczba pracujących, liczba biernych oraz aktywnych zawodowo) na podregiony NTS 3 z użyciem metod statystyki małych obszarów (SMO) i przy wykorzystaniu dostępnych danych administracyjnych oraz wyników badań reprezentacyjnych.

Moduł II - I etap prac

W pierwszym etapie drugiego modułu stworzono raport pośredni, który obejmował następujące działania:

- kwerendę zastosowań metodologii estymacji pośredniej w obszarze rynku pracy,
- rozpoznanie potencjalnych źródeł zmiennych pomocniczych możliwych do wykorzystania podczas procesu estymacji wybranych charakterystyk rynku pracy,
- wybór estymatorów pośrednich, które zostaną użyte w procesie estymacji,
- przeprowadzenie estymacji wybranych wskaźników z obszaru rynku pracy w ujęciu rocznym w latach 2010–2015 w wybranych przekrojach,
- przeprowadzenie statystycznej oceny jakości uzyskanych wyników estymacji.

Moduł II - II etap prac

W drugim etapie drugiego modułu stworzono raport końcowy, w którym:

- zbadano możliwości estymacji wybranych wskaźników rynku pracy wraz z ich statystyczną oceną, w ujęciu kwartalnym w latach 2010–2015,
- dokonano oceny możliwości włączenia wyników badania do statystyki publicznej w ramach programu badań statystycznych realizowanego przez GUS,
- opracowano wytyczne, wnioski i rekomendacje z zakresu stosowania estymacji pośredniej dla wskaźników związanych z rynkiem pracy na poziomie NTS 3 i dodatkowych przekrojach w kontekście przyszłego cyklicznego wykorzystania rozpatrywanej metodologii.

Moduł II - szacowane parametry

- liczba pracujących, liczba bezrobotnych, liczba biernych zawodowo, liczba aktywnych zawodowo, wskaźnik zatrudnienia oraz stopa bezrobocia ogółem (w ujęciu rocznym i kwartalnym),
- liczba pracujących, liczba bezrobotnych, liczba biernych zawodowo, liczba aktywnych zawodowo, wskaźnik zatrudnienia oraz stopa bezrobocia (w ujęciu rocznym) w domenach:
  - płeć (mężczyźni, kobiety),
  - miejsce zamieszkania (miasto, wieś),
  - grupy wieku (15–24, 25–54, 55–64, 20–64).

Podejście kalibracyjne w ujęciu teoretycznym

- Załóżmy, że z populacji U pobieramy n elementową próbę s zgodnie z określonym planem jej losowania. Zakładamy, że $d_k = \frac{1}{\pi_k}$, $k = 1, \ldots, N$ jest odpowiednią wagą, gdzie π_k oznacza prawdopodobieństwo inkluzji pierwszego rzędu.
- Zdarza się jednak, że wagi d_k wynikające z planu losowania próby nie odtwarzają znanych wartości globalnych w odniesieniu do niektórych kluczowych zmiennych.
- Wartości globalne takich zmiennych są znane zazwyczaj ze spisów powszechnych czy rejestrów administracyjnych.

Podejście kalibracyjne w ujęciu teoretycznym

- Przyjmując, że:

$$\sum_{k \in U} \mathbf{x}_k = \left(\sum_{k \in U} x_{k1}, \ldots, \sum_{k \in U} x_{kJ} \right)^T \tag{1}$$

stanowi wektor wartości globalnych wszystkich zmiennych pomocniczych oznacza to, że co najmniej dla jednej zmiennej $j = 1, \ldots, J$ nie jest spełniony poniższy warunek:

$$\sum_{k \in s} d_k x_{kj} = \sum_{k \in U} x_{kj}, \tag{2}$$

gdzie x_{kj} oznacza wartość j–tej zmiennej pomocniczej dla k–tej jednostki badania oraz $\sum_{k \in U} x_{kj}$ jest wartością globalną tej zmiennej.

Podejście kalibracyjne w ujęciu teoretycznym

- Załóżmy w dalszym ciągu, że $\mathbf{d} = (d_1, \ldots, d_n)^T$ jest wektorem wag wynikającym z planu losowania próby, a $\mathbf{w} = (w_1, \ldots, w_n)^T$ poszukiwanym wektorem końcowych wag kalibracyjnych.
- Wagi w_k są rozwiązaniem zadania optymalizacyjnego:
  - (W1) minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in s} \frac{d_k}{q_k} G\left(\frac{w_k}{d_k}\right) \to min, \tag{3}$$

  - (W2) równania kalibracyjne:

$$\sum_{k \in s} w_k \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k. \tag{4}$$

  - (W3) warunki ograniczające:

$$L \leq \frac{w_k}{d_k} \leq U, \text{ gdzie: } \quad 0 \leq L \leq 1 \leq U, \; k = 1, \ldots, n. \tag{5}$$

Podejście kalibracyjne w ujęciu teoretycznym

- Przy wyborze funkcji $G(\cdot)$ w procesie wyznaczania wag kalibracyjnych w_k istnieje pewna dowolność. Funkcja $G(\cdot)$ powinna jednak spełniać pewne własności matematyczne, do których należy zaliczyć: $G(\cdot)$ jest ściśle wypukła i dwukrotnie różniczkowalna, $G(\cdot) \geq 0$, $G(1) = 0$, $G'(1) = 0$ oraz $G''(1) = 1$.
- Do wyznaczania wag kalibracyjnych w_k szczególnie przydatna jest znajomość funkcji $F(\cdot)$ odwrotnej do pierwszej pochodnej funkcji $G(\cdot)$ tj.:

$$F(\cdot) = G'^{-1}(\cdot), \tag{6}$$

przy czym $F(0) = 1$. Jest to tzw. funkcja kalibracyjna (ang. *calibration function*).

Podejście kalibracyjne w ujęciu teoretycznym

- Ostatecznie wagi kalibracyjne można przedstawić w następującej postaci:

$$w_k = d_k F\left(q_k \mathbf{x}_k^T \boldsymbol{\lambda}\right) = d_k g_k, \tag{7}$$

gdzie $g_k = F\left(q_k \mathbf{x}_k^T \boldsymbol{\lambda}\right)$ to tzw. mnożniki wagowe (kalibracyjne). Do wyznaczenia wag kalibracyjnych (7) w dalszym ciągu niezbędna jest znajomość wektora mnożników Lagrange'a $\boldsymbol{\lambda}$, który można znaleźć rozwiązując odpowiednie równanie kalibracyjne – por. (4):

$$\sum_{k \in s} d_k F\left(q_k \mathbf{x}_k^T \boldsymbol{\lambda}\right) \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k. \tag{8}$$

Wybrane postacie funkcji $G(\cdot)$

L.p.	$G(x)$	$F(u)$	Ograniczenia
1	$\frac{(x-1)^2}{2}, \quad x \in \mathbb{R}$	$1+u, \quad u \in \mathbb{R}$	Brak
2	$x \ln(x) - x + 1, \quad x \in \mathbb{R}^+$ $-x+1, \quad x = 0$	$e^u, \quad u \in \mathbb{R}$	Brak
3	$\left[(x-L)\ln\frac{x-L}{1-L} + (U-x)\ln\frac{U-x}{U-1}\right] A^{-1}, \quad x \in (L, U)$ $\left[(U-L)\ln\frac{U-L}{U-1}\right] A^{-1}, \quad x \leq L$ $\left[(U-L)\ln\frac{U-L}{1-L}\right] A^{-1}, \quad x \geq U$	$\frac{L(U-1) + U(1-L)\, e^{Au}}{(U-1) + (1-L)\, e^{Au}},$ gdzie $u \in \mathbb{R}$ oraz $A = \frac{U-L}{(U-1)(1-L)}$	$0 \leq L < 1 < U$
4	$\frac{1}{2\alpha}\int_1^x \sinh\left[\alpha\left(t - \frac{1}{t}\right)\right] dt, \quad \alpha > 0$	$\frac{\operatorname{arsinh}(2\alpha u) + \sqrt{\operatorname{arsinh}^2(2\alpha u) + 4\alpha^2}}{2\alpha}, \quad u \in \mathbb{R}$	$0 \leq L < 1 < U$
5	$\frac{(x-1)^2}{2}, \quad x \in \mathbb{R}$	$1+u, \quad u \in [L-1, U-1]$ $L, \quad u = L-1$ $U, \quad u = U-1$	$x \in [L, U]$ oraz $0 \leq L \leq 1 \leq U$

Jedno- i dwukrokowa kalibracja

- W dalszym ciągu załóżmy, że r oznacza zbiór respondentów tj. jednostek, które wzięły udział w badaniu. Oznacza to, że $r \subset s$.
- Załóżmy ponadto, że m oznacza liczebność zbioru respondentów.
- Podejście kalibracyjne można w takim przypadku opisać jako:
  - (W1') minimalizacja funkcji odległości:

$$D(\mathbf{d}, \mathbf{w}) = \sum_{k \in r} \frac{d_k}{q_k} G\left(\frac{w_k}{d_k}\right) \rightarrow min, \tag{9}$$

  - (W2') równania kalibracyjne:

$$\sum_{k \in r} w_k \mathbf{x}_k = \sum_{k \in U} \mathbf{x}_k, \tag{10}$$

  - (W3') warunki ograniczające:

$$L \leq \frac{w_k}{d_k} \leq U, \text{ gdzie: } \quad 0 \leq L \leq 1 \leq U, \; k = 1, \ldots, m. \tag{11}$$

Jedno- i dwukrokowa kalibracja

- Oznaczmy przez $\mathbf{x}_k^*$ wektor, dla którego wektor wartości globalnych $\sum_U \mathbf{x}_k^*$ jest znany (na przykład ze spisu bądź rejestru administracyjnego) i dla każdej jednostki $k \in r$ wartości wektora $\mathbf{x}_k^*$ są również znane.
- Załóżmy ponadto, że $\mathbf{x}_i^\circ$ oznacza wektor, w którym dla każdego $k \in s$ wektor wartości $\mathbf{x}_k^\circ$ jest znany, podczas gdy wektor wartości globalnych $\sum_U \mathbf{x}_k^\circ$ pozostaje nieznany. Oznacza to zatem, że dla $k \in r$ wektor wartości $\mathbf{x}_k^\circ$ jest również znany.
- Särndal i Lundström (2005) określają wektor $\mathbf{x}_k^\circ$ mianem „księżycowego" (ang. *moon vector*), a wektor $\mathbf{x}_k^*$ jako „gwiezdny" (ang. *star vector*).

Podejście jednokrokowe

- Wektor zmiennych pomocniczych $\mathbf{x}_k$ można przedstawić w następujący sposób:

$$\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_k^* \\ \mathbf{x}_k^\circ \end{pmatrix}. \tag{12}$$

- Wektor wartości globalnych $\mathbf{X}$ można zapisać jako:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}^* \\ \widehat{\mathbf{X}}^\circ \end{pmatrix}, \tag{13}$$

gdzie $\mathbf{X}^* = \sum_{k\in U} \mathbf{x}_k^*$ i $\widehat{\mathbf{X}}^\circ = \sum_{k\in s} d_k \mathbf{x}_k^\circ$.

- Równanie kalibracyjne (10) przyjmuje w takiej sytuacji postać:

$$\begin{pmatrix} \sum_{k\in r} w_k \mathbf{x}_k^* \\ \sum_{k\in r} w_k \mathbf{x}_k^\circ \end{pmatrix} = \begin{pmatrix} \sum_{k\in U} \mathbf{x}_k^* \\ \sum_{k\in s} d_k \mathbf{x}_k^\circ \end{pmatrix}. \tag{14}$$

Podejście dwukrokowe

- W dwukrokowej procedurze wyznaczania wag kalibracyjnych kalibrację przeprowadza się z poziomu zbioru respondentów r na poziom próby s wykorzystując wartości wektora $\mathbf{x}_i^\circ$ celem uzyskania wag pośrednich.
- Następnie wagi pośrednie wykorzystywane są jako wagi wejściowe w drugim kroku kalibracji z poziomu zbioru respondentów r na poziom populacji U z wykorzystaniem wektorów $\mathbf{x}_k^*$ i $\mathbf{x}_k^\circ$.
- Należy tutaj rozróżnić dwa przypadki określane mianem dwukrokowego podejścia kalibracyjnego typu A i B.

Podejście dwukrokowe typu A

- W pierwszym kroku wykorzystujemy wagi $d_k = \frac{1}{\pi_k}$ jako wagi wejściowe i kalibrujemy je celem uzyskania wag pośrednich w_k°, tak aby spełnione było następujące równanie:

$$\sum_{k \in r} w_k^\circ \mathbf{x}_k^\circ = \sum_s d_k \mathbf{x}_k^\circ. \tag{15}$$

- W drugim kroku wykorzystujemy pośrednie wagi w_k° jako wagi wejściowe w procesie kalibracji i znajdujemy końcowe wagi kalibracyjne w_k kalibrując ze zbioru respondentów r na poziom populacji U tj., tak aby spełnione było równanie kalibracyjne:

$$\sum_{k \in r} w_k \mathbf{x}_k = \mathbf{X}, \tag{16}$$

gdzie $\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_k^* \\ \mathbf{x}_k^\circ \end{pmatrix}$ oraz $\mathbf{X} = \begin{pmatrix} \mathbf{X}^* \\ \widehat{\mathbf{X}}^\circ \end{pmatrix}$.

Podejście dwukrokowe typu B

- W dwukrokowej procedurze kalibracyjnej typu B, pierwszy krok jest dokładnie taki sam jak w dwukrokowej procedurze typu A.
- Również drugi krok w podejściu kalibracyjnym typu B jest bardzo zbliżony do drugiego kroku w podejściu kalibracyjnym typu A z tym wyjątkiem tego, że wektory $\mathbf{x}_k = \begin{pmatrix} \mathbf{x}_k^* \\ \mathbf{x}_k^\circ \end{pmatrix}$ oraz $\mathbf{X} = \begin{pmatrix} \mathbf{X}^* \\ \widehat{\mathbf{X}}^\circ \end{pmatrix}$ są zastępowane przez $\mathbf{x}_k = \mathbf{x}_k^*$ i $\mathbf{X} = \mathbf{X}^*$.
- Wyniki uzyskane w podejściu jednokrokowym, dwukrokowym typu A oraz dwukrokowym typu B prowadzą zazwyczaj do odmiennych zestawów wag kalibracyjnych.

Kalibracja wag w BAEL

1. Obliczenie tzw. współczynników realizacji R według wzoru:

$$R = (K - N)/K, \tag{17}$$

gdzie K jest oszacowaniem (według wag pierwotnych) liczby mieszkań kwalifikujących się do badania, zaś N oszacowaniem liczby mieszkań kwalifikujących się do badania, lecz nie dających się zbadać z jakiegokolwiek powodu. Współczynniki realizacji wyznacza się dla każdego z 16 województw w sześciu grupach wyróżnionych ze względu na klasę miejscowości.

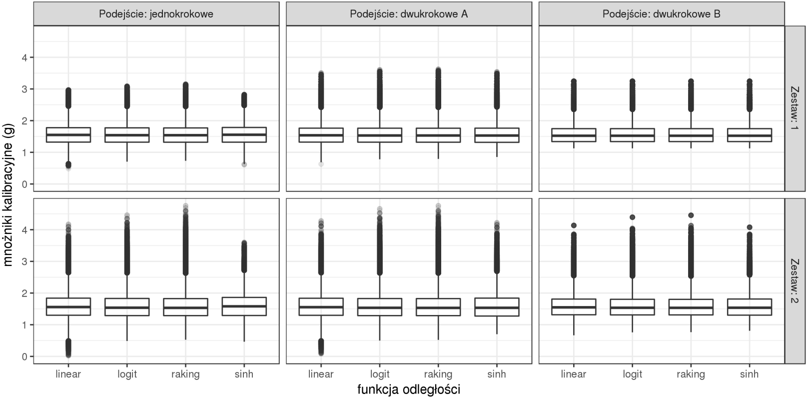
2. Obliczenie wag wtórnych, w drodze dzielenia wag pierwotnych przez R, przy czym współczynnik R bierze się w zależności od województwa i klasy miejscowości, w której znajduje się dane mieszkanie.

3. Wyznaczenie wag finalnych dla danych ludnościowych. Wyliczenia te przeprowadza się tak, aby dostosować wyniki BAEL do bieżących szacunków demograficznych. Dokonuje się tego przez wyliczenie tzw. modyfikatorów, które oblicza się osobno dla każdej z 48 kategorii zdefiniowanych przez miejsce zamieszkania (miasto/wieś) $\times$ płeć $\times$ 12 grup wieku (15–17, 18–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65+), w drodze dzielenia liczby osób w danej kategorii według skorygowanych szacunków demograficznych przez liczbę osób w tejże kategorii wyliczoną z danych BAEL przy zastosowaniu wag wtórnych z etapu drugiego. Wagi finalne otrzymuje się w wyniku mnożenia wag wtórnych przez odpowiednie modyfikatory.

Kalibracja wag w BAEL - podregiony

- Zastosowano cztery funkcje odległości (linear, raking, logit, sinh) oraz trzy metody ich wyznaczania (jednokrokowe, dwukrokowe typu A oraz dwukrokowe typu B) w połączeniu z dwoma zestawami („gwiezdnych”) zmiennych pomocniczych $\mathbf{x}_k^*$.
- **Zestaw 1.** (48 wariantów): płeć (2 warianty: mężczyzna, kobieta) $\times$ miejsce zamieszkania (2 warianty: miasto, wieś) $\times$ grupy wieku (12 wariantów: 15–17, 18–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65+),
- **Zestaw 2.** (121 wariantów): Zestaw 1. (48) + podregiony (73).
- Jako „księżycowy” wektor zmiennych pomocniczych przyjęto zmienne pomocnicze określające przynależność badanego mieszkania do klasy miejscowości w danym województwie (16 województw $\times$ 6 klas miejscowości).

Kalibracja wag w BAEL - podregiony

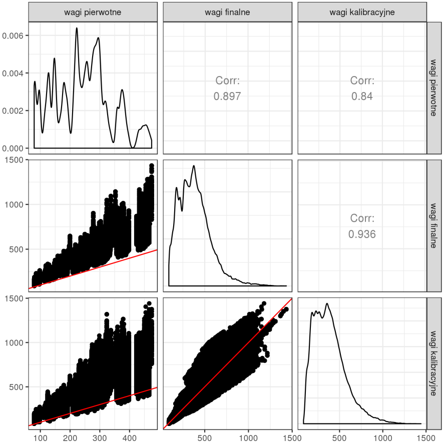


Rys. 1: Porównanie rozkładów mnożników kalibracyjnych (g_k) w zastosowanych podejściach kalibracyjnych (lata 2010–2015)

Mnożniki kalibracyjne (g_k)

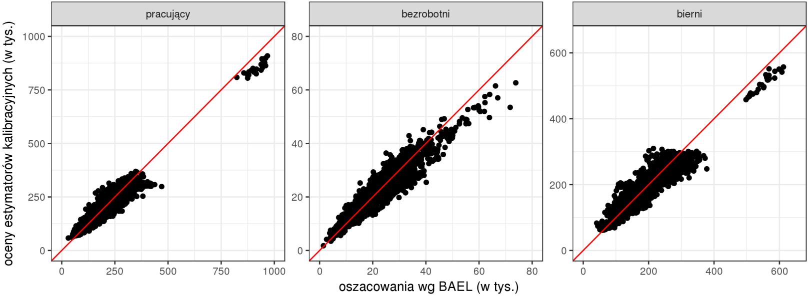
Podejście kalibracyjne	Funkcja odległości	Min	Q1	Q2	Q3	Max	Mean	Sd
Zestaw: 1								
jednokrokowe	linear	0.48	1.32	1.55	1.78	2.97	1.57	0.31
jednokrokowe	logit	0.71	1.32	1.54	1.77	3.09	1.57	0.31
jednokrokowe	raking	0.73	1.32	1.54	1.77	3.15	1.57	0.31
jednokrokowe	sinh	0.60	1.32	1.56	1.78	2.82	1.57	0.31
dwukrokowe A	linear	0.63	1.32	1.54	1.76	3.50	1.57	0.31
dwukrokowe A	logit	0.78	1.32	1.53	1.76	3.60	1.57	0.31
dwukrokowe A	raking	0.79	1.32	1.53	1.76	3.62	1.57	0.31
dwukrokowe A	sinh	0.85	1.32	1.53	1.76	3.54	1.57	0.31
dwukrokowe B	linear	1.13	1.34	1.52	1.75	3.24	1.57	0.29
dwukrokowe B	logit	1.13	1.34	1.52	1.75	3.24	1.57	0.29
dwukrokowe B	raking	1.13	1.34	1.52	1.75	3.24	1.57	0.29
dwukrokowe B	sinh	1.13	1.34	1.52	1.75	3.24	1.57	0.29
Zestaw: 2								
jednokrokowe	linear	0.03	1.30	1.56	1.84	4.17	1.58	0.40
jednokrokowe	logit	0.49	1.29	1.54	1.83	4.46	1.58	0.40
jednokrokowe	raking	0.53	1.29	1.53	1.82	4.76	1.58	0.40
jednokrokowe	sinh	0.46	1.29	1.58	1.86	3.60	1.58	0.40
dwukrokowe A	linear	0.09	1.30	1.55	1.84	4.28	1.58	0.40
dwukrokowe A	logit	0.50	1.29	1.53	1.83	4.66	1.58	0.40
dwukrokowe A	raking	0.52	1.29	1.53	1.82	4.76	1.58	0.40
dwukrokowe A	sinh	0.70	1.27	1.53	1.84	4.22	1.58	0.40
dwukrokowe B	linear	0.67	1.31	1.55	1.81	4.13	1.58	0.37
dwukrokowe B	logit	0.76	1.31	1.54	1.80	4.39	1.58	0.37
dwukrokowe B	raking	0.77	1.31	1.53	1.80	4.45	1.58	0.37
dwukrokowe B	sinh	0.81	1.30	1.54	1.81	4.08	1.58	0.37

Porównanie wag



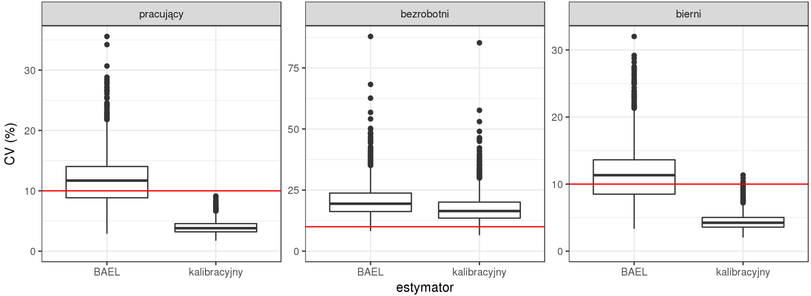
Rys. 2: Porównanie wag pierwotnych, finalnych i kalibracyjnych

Porównanie oszacowań



Rys. 3: Porównanie oszacowań dla pracujących, bezrobotnych i biernych dla wag w BAEL i kalibracyjnych (dwukrokowa procedura typu B w wersji sinh)

Precyzja oszacowań



Rys. 4: Porównanie precyzji oszacowań dla pracujących, bezrobotnych i biernych dla wag w BAEL i kalibracyjnych (dwukrokowa procedura typu B w wersji sinh)

Modele klasy SMO

- Na potrzeby estymacji pośredniej rozważanych w projekcie charakterystyk rynku pracy wykorzystano trzy modele klasy SMO (**wielomianowe modele mieszane**).
- **Model 1** – dla każdego okresu budowany był oddzielny model. Model ten został szczegółowo opisany w pracy Molina i in. (2007).
- **Model 2** – jeden model budowany był dla całego okresu. Model ten został kompleksowo opisany w pracy López-Vizcaíno i in. (2013).
- **Model 3** – jeden model budowany był dla całego okresu z uwzględnieniem autokorelacji w czasie. Model ten został opisany w pracy López-Vizcaíno i in. (2015).

Wykorzystane zmienne pomocnicze

- Jako zmienne pomocnicze wykorzystano udział bezrobotnych zarejestrowanych w urzędach pracy do populacji w danej domenie oraz informację czy dany podregion jest jednocześnie miastem na prawach powiatu.
- W przypadku szacunków dla kwartałów wykorzystano dodatkowo zmienną określającą kwartał, aby uwzględnić addytywny efekt sezonowości.

Ocena jakości oszacowań - ujęcie roczne

Tab. 1: Porównanie względnych błędów szacunków estymatora bezpośredniego, kalibracyjnego (dwukrokowa procedura typu B w wersji sinh) oraz pośredniego (model 1 i 2) dla wszystkich domen za lata 2010–2015 (w %)

Estymator	Min	Q1	Mediana	Średnia	Q3	Max	N/A
Bezpośredni	0.8	5.4	9.9	11.3	14.6	112.9	10.0
Kalibracyjny	0.8	3.5	7.6	9.5	12.6	93.2	10.0
Model 1	0.4	0.9	1.3	3.5	6.8	15.8	
Model 2	0.4	1.4	2.3	4.5	6.7	38.2	

Ocena jakości oszacowań - ujęcie roczne

Tab. 2: Porównanie względnych błędów szacunków estymatora bezpośredniego, kalibracyjnego (dwukrokowa procedura typu B w wersji sinh) oraz pośredniego (model 1 i 2) dla wszystkich domen za lata 2010–2015 według charakterystyk rynku pracy (w %)

Wskaźnik	Estymator	Min	Q1	Mediana	Średnia	Q3	Max	N/A
Aktywni	Bezpośredni	2.0	7.2	9.4	10.1	12.6	41.5	2.0
	Kalibracyjny	1.1	2.6	3.7	5.9	8.9	39.6	2.0
	Model 1	0.5	0.8	1.0	1.0	1.1	1.7	
	Model 2	0.5	1.3	1.7	2.2	2.6	7.4	
Bezrobotni	Bezpośredni	5.8	13.1	16.7	19.1	22.2	112.9	2.0
	Kalibracyjny	4.4	11.2	14.7	17.3	20.3	90.0	2.0
	Model 1	3.6	6.8	8.1	8.2	9.5	15.0	
	Model 2	2.9	6.6	8.3	9.3	10.7	27.7	
Bierni	Bezpośredni	2.4	7.6	9.8	10.3	12.4	41.7	
	Kalibracyjny	1.3	4.0	6.1	6.8	8.7	39.4	
	Model 1	0.4	0.7	0.8	0.8	0.9	1.4	
	Model 2	0.4	0.8	0.9	1.1	1.3	4.1	
Pracujący	Bezpośredni	2.0	7.4	9.5	10.4	13.0	42.9	
	Kalibracyjny	1.2	2.9	4.1	6.4	9.5	40.9	
	Model 1	0.6	1.0	1.2	1.3	1.5	2.5	
	Model 2	0.6	1.5	2.0	2.6	3.0	8.4	
Stopa bezrobocia	Bezpośredni	4.2	10.8	13.7	16.2	18.7	108.4	2.0
	Kalibracyjny	4.3	10.7	13.8	16.2	18.6	93.2	2.0
	Model 1	3.5	6.8	8.2	8.3	9.6	15.8	
	Model 2	2.9	6.5	8.2	9.4	10.4	38.2	
Wsk. zatrudnienia	Bezpośredni	0.8	2.3	3.2	4.4	5.0	21.6	2.0
	Kalibracyjny	0.8	2.3	3.3	4.4	5.0	21.1	2.0
	Model 1	0.6	1.0	1.2	1.3	1.5	2.5	
	Model 2	0.6	1.5	2.0	2.6	3.0	8.4	

Ocena jakości oszacowań - ujęcie kwartalne

Tab. 3: Porównanie względnych błędów szacunków estymatora bezpośredniego, kalibracyjnego (dwukrokowa procedura typu B w wersji sinh) oraz pośredniego (model 2 i model 3) dla wszystkich domen za kwartały 2010-2015 (w %)

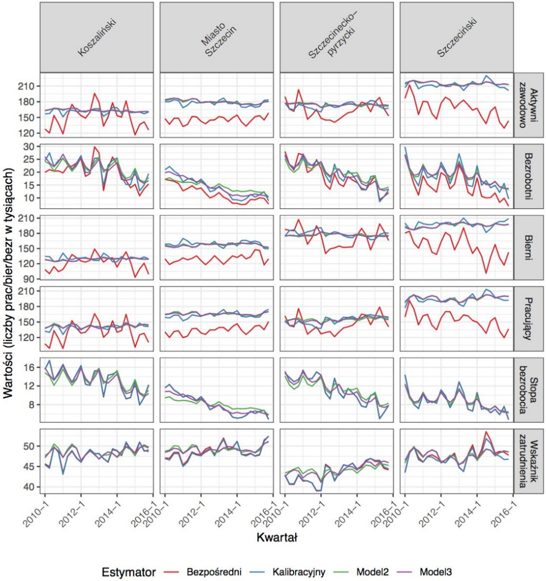
Estymator	Min	Q1	Mediana	Średnia	Q3	Max
Bezpośredni	1.71	7.11	12.29	12.82	16.75	87.87
Kalibracyjny	1.63	3.49	4.60	8.44	13.40	85.28
Model 2	1.10	1.82	2.14	4.21	7.74	13.75
Model 3	0.83	1.44	1.73	3.57	6.65	13.18

Ocena jakości oszacowań - ujęcie kwartalne

Tab. 4: Porównanie względnych błędów szacunków estymatora bezpośredniego, kalibracyjnego (dwukrokowa procedura typu B w wersji sinh) oraz pośredniego (model 2 i 3) dla wszystkich domen za lata 2010–2015 według wskaźników (w %)

Wskaźnik	Estymator	Min	Q1	Mediana	Średnia	Q3	Max
Aktywni	Bezpośredni	2.74	8.75	11.55	11.69	13.83	35.39
	Kalibracyjny	1.63	2.83	3.40	3.54	4.05	9.01
	Model 2	1.10	1.52	1.67	1.67	1.80	2.31
	Model 3	0.83	1.21	1.35	1.36	1.49	2.30
Bezrobotni	Bezpośredni	8.26	16.23	19.42	20.78	23.77	87.87
	Kalibracyjny	6.51	13.53	16.42	17.51	20.08	85.28
	Model 2	4.65	7.78	8.80	8.77	9.71	13.40
	Model 3	4.23	6.73	7.54	7.65	8.45	13.18
Bierni	Bezpośredni	3.33	8.50	11.33	11.48	13.62	32.01
	Kalibracyjny	2.03	3.58	4.24	4.41	5.03	11.34
	Model 2	1.35	1.92	2.10	2.09	2.26	2.96
	Model 3	1.02	1.51	1.68	1.69	1.85	2.81
Pracujący	Bezpośredni	2.89	8.85	11.70	11.84	14.03	35.58
	Kalibracyjny	1.73	3.20	3.81	3.99	4.57	9.13
	Model 2	1.20	1.79	1.97	1.98	2.16	2.91
	Model 3	0.88	1.40	1.55	1.58	1.74	2.61
Stopa bezrobocia	Bezpośredni	6.33	13.26	16.08	17.23	19.83	86.41
	Kalibracyjny	6.34	13.28	16.10	17.23	19.90	85.23
	Model 2	4.39	7.68	8.79	8.75	9.72	13.75
	Model 3	4.12	6.57	7.43	7.55	8.37	13.03
Wsk. zatrudnienia	Bezpośredni	1.71	3.18	3.75	3.92	4.49	8.97
	Kalibracyjny	1.73	3.20	3.81	3.99	4.57	9.13
	Model 2	1.20	1.79	1.97	1.98	2.16	2.91
	Model 3	0.88	1.40	1.55	1.58	1.74	2.61

Porównanie oszacowań - ujęcie kwartalne



Rys. 5: Porównanie oszacowań według typu estymatora i wskaźnika- ujęcie kwartalne

Podsumowanie

- Zadowalające wyniki oszacowań wybranych kategorii rynku pracy dla estymatorów kalibracyjnych biorących pod uwagę poziom podregionów.
- Na ogół obserwowalny zysk na precyzji estymatorów klasy SMO w porównaniu z estymatorami kalibracyjnymi.
- W przypadku wdrożenia prezentowanej metody estymacji opartej na modelach do praktyki badań realizowanych przez Główny Urząd Statystyczny należałoby zapewnić spójność szacunków pośrednich z szacunkami bezpośrednimi na wyższym poziomie agregacji przestrzennej (zasada benchmarkingu).

Literatura

López-Vizcaíno, E., Lombardía, M. J., i Morales, D. (2013). Multinomial-Based Small Area Estimation of Labour Force Indicators. *Statistical modelling*, 13(2):153–178.

López-Vizcaíno, E., Lombardía, M. J., i Morales, D. (2015). Small Area Estimation of Labour Force Indicators Under a Multinomial Model with Correlated Time and Area Effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):535–565.

Molina, I., Saei, A., i Lombardía, M. (2007). Small Area Estimates of Labour Force Participation Under a Multinomial Logit Mixed Model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(4):975–1000.

Särndal, C.-E. i Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. John Wiley & Sons.

Zespół badawczy

Kierownik projektu:
Hanna Strzelecka

Koordynator Merytoryczny w zakresie Modułu II
Marcin Szymkowiak

Zespół autorski:
Maciej Beręsewicz, Iwona Biały, Katarzyna Derucka, Grzegorz Grygiel, Piotr Jastrzębski, Tomasz Józefowski, Tomasz Klimanek, Jacek Kowalewski, Jan Kubacki, Magdalena Łączyńska, Andrzej Młodak, Dorota Malicka, Tomasz Piasecki, Michał Pietrzak, Waldemar Popiński, Małgorzata Saroska, Hanna Strzelecka, Marcin Szymkowiak, Ewa Wieczorek, Kamil Wilak

Dziękuję za uwagę

Marcin Szymkowiak – m.szymkowiak@stat.gov.pl