

Streszczenie raportu końcowego w języku nietechnicznym

*z realizacji pracy badawczej
pt. „Rozszerzenie Badania Aktywności Ekonomicznej
Ludności - wybrane wskaźniki „Europa 2020” oraz wskaźnik
NEET na poziomie województw (NTS 2); podstawowe
agregacje z zakresu rynku pracy na podregiony (NTS 3), miasta
wojewódzkie, grupy podregionów”*

*umowa nr 23/BR-POPT/CBiES/2014 zawarta dnia 8.01.2015 r.
wraz z Aneksem nr 1 z dnia 17.07.2015 roku
pomiędzy Głównym Urzędem Statystycznym z siedzibą w Warszawie
a Centrum Badań i Edukacji Statystycznej z siedzibą w Jachrance*

*Praca badawcza została przeprowadzona w ramach
Projektu Programu Operacyjnego Pomoc Techniczna 2014-2020*

Zespół realizujący:

*Wioletta Zwara
Tomasz Piasecki
Iwona Biały
Waldemar Popiński
Małgorzata Cacko
Dorota Cybart-Napiórkowska
Katarzyna Derucka
Małgorzata Długołęcka
Bartosz Grancow
Grzegorz Gudaszewski
Małgorzata Jaworska
Katarzyna Klamrowska
Zofia Kostrzewa
Marta Łukasiak
Adam Mańkowski
Magdalena Miśko
Jan Nowak
Joanna Stańczak
Hanna Strzelecka
Dorota Szafłys
Marcin Szczepaniak
Beata Ugodzińska
Bogumiła Uljasz
Danuta Waśniewska
Józef Wieczorek
Agnieszka Znajewska*

1. Cel pracy badawczej

Celem zrealizowanej pracy było oszacowanie możliwych danych historycznych oraz opracowanie zmian w metodologii Badania Aktywności Ekonomicznej Ludności pozwalających na uzyskanie reprezentatywnych ocen dotyczących wybranych wskaźników rynku pracy oraz edukacji dla potrzeb określenia obszarów interwencji oraz monitorowania efektów prowadzonych działań na poziomie regionalnym, tj.:

- wybranych wskaźników zawartych w strategii „Europa 2020”, na poziomie województw (NTS 2), tj.:
 - odsetek osób w wieku 18-24 lata z wykształceniem co najwyżej gimnazjalnym, które nie kontynuują nauki i nie dokończają się w ludności ogółem w tej samej grupie wieku,
 - odsetek osób w wieku 30-34 lata posiadających wykształcenie wyższe;
- wskaźnika NEET (young people Neither in Employment nor in Education and Training – osoby młode niepracujące, nieuczące się i niedokształcające się) oraz liczebności tej grupy osób dla kategorii wiekowych 15-24, 15-29 i 25-29 lat na poziomie województw (NTS 2);
- podstawowych agregacji danych z zakresu rynku pracy na poziomie NTS 3, miast wojewódzkich lub w przypadku stwierdzenia braku możliwości prezentowania wyników na poziomie poszczególnych NTS 3, grupowań NTS 3.

2. Charakterystyka źródła danych

W realizacji pracy badawczej zostały wykorzystane wyniki reprezentacyjnego Badania Aktywności Ekonomicznej Ludności (BAEL), które prowadzone jest w Polsce kwartalnie od maja 1992 r. i doskonalone zgodnie z zaleceniami Eurostat. Metodologia badania BAEL oraz jego wyniki są spójne i porównywalne w krajach Unii Europejskiej (jako badanie Labour Force Survey - LFS).

Podstawą metodologii BAEL są definicje dotyczące aktywności ekonomicznej ludności, które opierają się na Rezolucji w sprawie statystyk ludności aktywnej zawodowo, zatrudnienia, bezrobocia i niepełnozatrudnienia, przyjęte na XIII Międzynarodowej Konferencji Statystyków Pracy w październiku 1982 r. (z późniejszymi zmianami) i zalecane przez Międzynarodową Organizację Pracy do stosowania w praktyce przez poszczególne kraje. Przedmiotem badania jest sytuacja w zakresie aktywności ekonomicznej ludności, tzn. fakt wykonywania pracy, pozostawania bezrobotnym lub biernym zawodowo w danym tygodniu.

Podstawowym kryterium podziału ludności na pracujących, bezrobotnych i biernych zawodowo jest praca, czyli wykonywanie, posiadanie bądź poszukiwanie pracy. Przyjęta kolejność wyodrębniania poszczególnych kategorii ludności gwarantuje zaklasyfikowanie każdej osoby tylko do jednej kategorii.

Najniższym poziomem podziału administracyjnego, na którym obecnie są udostępniane dane z BAEL jest województwo. Wynika to z reprezentacyjnego charakteru badania i wielkości próby. Dane na niższych poziomach podziału terytorialnego są obciążone zbyt dużym błędem losowym, podobnie jak zbyt szczegółowe przekroje w ujęciu wojewódzkim.

Szerszy opis metodologii BAEL, w tym doboru próby i uogólniania wyników, oraz definicje kluczowych pojęć i wskaźników zawarte są w publikacjach GUS dostępnych na stronie internetowej, np.

<http://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/aktywnosc-ekonomiczna-ludnosci-polski-w-latach-2010-2012,5,3.html>

3. Metodologia prowadzonych prac

Projekt badawczy był podzielony na dwa etapy. W ramach pierwszego etapu opisano szczegółowo dostępne zasoby informacyjne wraz z metodologią wyznaczania wskaźników wynikowych z BAEL, po czym wstępnie zaproponowano rozwiązania metodologiczne, które mogłyby doprowadzić do realizacji postawionego celu projektu. Drugi etap polegał na: utworzeniu odpowiednich wariantów symulacyjnych na podstawie wniosków z pierwszego etapu oraz przeprowadzeniu symulacyjnych obliczeń w celu przygotowania finalnych propozycji alternatywnych rozwiązań metodologicznych, mogących mieć zastosowanie przy realizacji badania BAEL. Ponadto w drugim etapie dokonano przeliczeń wskaźników według nowej klasyfikacji NTS 2013 wyodrębniającej nowe podregiony (NTS3), a także przeanalizowano wpływ przyszłych zmian w metodologii BAEL wynikających z wdrożenia rewizji badania LFS na poziomie UE.

W pracy badawczej wykorzystano wyniki BAEL dla lat 2010 – 2013 opracowane jako dane miesięczne, kwartalne, półroczne, roczne, półtoraroczne i dwuletnie.

Podstawą wnioskowania na temat możliwości wykorzystania wskaźników uzyskiwanych z badania przy użyciu dotychczasowej metodologii były oceny względnego błędu standardowego szacunku (współczynnik zmienności estymatora – CV) uzyskane jako wynik prac obliczeniowych.

Przeanalizowano łącznie ocenę precyzji dla:

- 6 wskaźników miesięcznych na poziomie NTS 0 (kraj);
- 8 wskaźników na poziomie NTS 2 (województwo) dla wyników kwartalnych, półrocznych, rocznych, półtorarocznych i dwuletnich;
- 160 wskaźników na poziomie NTS 3 (podregion) dla wyników półrocznych, rocznych, półtorarocznych i dwuletnich;
- 58 wskaźników na poziomie miast wojewódzkich dla wyników półrocznych, rocznych, półtorarocznych i dwuletnich;
- 1 wskaźnika na poziomie pogrupowanych podregionów NTS 3 dla wyników półrocznych, rocznych, półtorarocznych i dwuletnich.

W celu umożliwienia syntetycznego wnioskowania dokonano uporządkowania, a następnie agregacji uzyskanych oszacowań precyzji dotyczących poszczególnych wskaźników i przekrojów.

Przeprowadzenie analizy wymagało zatem zdefiniowania następujących elementów:

- określenia zasad oceny precyzji dla pojedynczych uogólnień,
- określenia zasad agregacji ocen precyzji w czasie (dla poszczególnych okresów uogólnienia mieszczących się w okresie analizy, tj. w latach 2010-2013 – np. ocen dla ośmiu okresów półrocznych), tak by uzyskać ocenę syntetyczną na poziomie wskaźnika, poziomu agregacji (przekroju), długości okresu uogólnienia (np. półroczny, dwuletni) i domeny (np. konkretnego województwa, podregionu),

- określenia zasad formułowania wniosków syntetycznych dotyczących danego poziomu agregacji (przekroju) dla danego wskaźnika i długości okresu uogólnienia, tj. agregacji ocen uzyskanych dla poszczególnych domen przekroju.

Efektom przeprowadzonej analizy było nadanie poszczególnym wskaźnikom ocen dotyczących możliwości ich publikowania i wykorzystania. Wskaźniki podzielono na 4 grupy:

- „**akceptowalny**” - wartości wskaźnika w danym przekroju **mogą być publikowane i wykorzystywane bez większych ograniczeń do większości zastosowań**. Dla większości zastosowań nie jest konieczne branie pod uwagę precyzji i jej podawanie przy posługiwaniu się wartościami wskaźnika;
- „**z ostrożnością**”, co oznacza warunkową akceptowalność wskaźnika w danym przekroju. Uogólnienia **mogą być wykorzystywane do niektórych zastosowań, ale trzeba je traktować z ostrożnością**. Analizując dane należy zawsze brać pod uwagę precyzję poszczególnych oszacowań, a publikując wartości wskaźnika należy je opatrywać informacją o precyzji;
- „**do dalszej pracy**” - oszacowania uzyskiwane przy użyciu dotychczasowej metodologii dla takiego wskaźnika **nie są akceptowalne z punktu widzenia precyzji i nie nadają się obecnie do wykorzystania, jako zestaw danych dotyczących całego przekroju**. Wskaźniki/przekroje, które otrzymały ocenę „do dalszej pracy” **zostały włączone do dalszych prac w drugim etapie projektu**, mających na celu umożliwienie ich publikacji i wykorzystania;
- „**do odrzucenia**” - oszacowania wskaźnika w danym przekroju **nie spełniają obecnie wymogów akceptowalności i jest bardzo mało prawdopodobne, by spełniły te wymogi po przeprowadzeniu możliwych do wykonania prac nad metodologią badania**. Rozważanie tak ocenionych wskaźników/przekrojów w drugim etapie projektu uznane zostało za niecelowe i **zostały one wyłączone z dalszych prac**.

4. Rezultaty pracy badawczej

Prace zrealizowane w ramach projektu pozwoliły na osiągnięcie dwóch najważniejszych efektów (i opracowanie powiązanych z nimi produktów), które w znaczącym stopniu rozwijają możliwości statystyki publicznej dotyczące publikacji danych z Badania Aktywności Ekonomicznej Ludności na niższym niż dotąd stopniu agregacji terytorialnej oraz poszerzają zasób dostępnych informacji w tym zakresie.

Pierwszym z głównych efektów, będącym produktem I etapu projektu, jest baza wyliczonych wskaźników obejmująca dane historyczne z lat 2010-2013. Integralną częścią bazy, zawierającą informację o jakości zawartych w niej danych, są tablice z kodami precyzji/jakości, którymi opatrzone wszystkie dane szczegółowe oraz w sposób syntetyczny całe wskaźniki i przekroje.

W II etapie opracowano różne warianty zmian w metodyce Badania Aktywności Ekonomicznej Ludności, które to warianty (alternatywne propozycje metodyki) stanowią produkty tego etapu (wraz z wynikami ich oceny/weryfikacji), pozwalające w znacznym stopniu poszerzyć możliwości zaspokojenia potrzeb informacyjnych objętych zakresem projektu.

Drugim z głównych efektów i produktów projektu jest zestaw rekomendacji dotyczących zmian w sposobie badania, określających zalecaną do wdrożenia nową metodykę losowania próby i uogólniania wyników, która w przypadku wdrożenia powinna w przyszłości znacząco poszerzyć

zestaw informacji z badania dotyczących rynku pracy i edukacji, dostępnych na niższym niż dotąd poziomie agregacji.

Zespół realizujący prace związane z projektem, na podstawie przeprowadzonych analiz i symulacji, zarekomendował następujące rozwiązania dotyczące losowania próby i uogólniania wyników:

- **pozostawienie dwustopniowego schematu losowania próby, przy utrzymaniu dotychczasowych zasad tworzenia jednostek losowania pierwszego stopnia (JPS) i określania liczby mieszkań losowanych z JPS;**
- **utrzymanie liczebności próby badanej zbliżonej do dotychczasowej;**
- **wprowadzenie sposobu podziału na warstwy i alokacji próby w warstwach określanego jako „NTS 2”, zakładającego wydzielenie miast wojewódzkich jako odrębnych warstw z dotychczasowych warstw wojewódzkich oraz algorytm alokacji próby w warstwach ukierunkowany na uzyskanie satysfakcjonujących uogólnień na poziomie miasta wojewódzkiego i województwa z wyłączeniem miast wojewódzkich;**
- **zastosowanie kalibracji wyników na poziomie województwa, w miejsce stosowanej dotąd kalibracji na poziomie kraju, z uwzględnieniem stosowanych dotąd kryteriów demograficznych.**

Rekomendowany wariant metodyki badania powinien pozwolić uzyskać satysfakcjonującą jakość uogólnień na poziomie miasta wojewódzkiego i województwa z wyłączeniem miast wojewódzkich dla większości wskaźników objętych zakresem projektu (w tym wszystkich wskaźników rozważanych w II etapie), bez dającego się przewidzieć zwiększenia kosztów badania i przy stosunkowo niewielkich zmianach organizacyjnych o ograniczonej uciążliwości. Wskazany wariant zmian nie zwiększa w sposób istotny możliwości wykorzystania danych na poziomie NTS 3 (co wymagałoby znaczącego zwiększenia próby), pozwalając uzyskać jedynie wyrównany zbiór wskaźników na tym poziomie o satysfakcjonującej jakości. **Z tego względu nie rekomenduje się rozszerzenia funkcjonalności BAEL o udostępnianie danych na poziomie podregionów – NTS 3. Zastosowanie wskazanego wariantu pozwoli jednak dostarczyć kompletny i użyteczny zestaw wskaźników dla poziomu miast wojewódzkich i województw z wyłączeniem miast wojewódzkich.** Warto podkreślić, że w tym przypadku efekt związany z rekomendowanymi zmianami jest na tyle silny, iż nie można wykluczyć, że po ich wdrożeniu, jakość uogólnień w przypadku wskaźników odrzuconych w I etapie dla przekroju wg miast wojewódzkich, w rzeczywistości będzie miała poziom akceptowalny. Z tego względu zalecane jest przeprowadzenie prac mających na celu ocenę *ex-post* jakości i precyzji danych uzyskanych po realizacji badania zgodnie z nową metodyką w przypadku wskaźników odrzuconych w I etapie dla przekroju wg miast wojewódzkich.

Zastosowanie alternatywnych metod warstwowania i alokacji próby może prowadzić do pogorszenia jakości uogólnień na poziomie kraju. Efekt ten może być istotny w przypadku uogólnień miesięcznych, gdzie podobnie jak w przypadku uogólnień na niskim poziomie agregacji przestrzennej obserwujemy wciąż pewien deficyt precyzji. Syntetyczne oceny jakości uogólnienia dla niektórych spośród rozważanych wskaźników miesięcznych ulegają pogorszeniu, co należy brać pod uwagę podejmując decyzję o ewentualnym wdrożeniu proponowanych rozwiązań. W zasadzie jedyną możliwą drogą poprawy jakości bezpośrednich uogólnień miesięcznych byłoby zwiększenie liczebności próby. Rozważane rozwiązania związane ze zmianami w alokacji i strukturze próby są

w tym przypadku raczej nieskuteczne. Nie jest możliwe uzyskanie poprawy jakości uogólnień w przekrojach terytorialnych i dla całego kraju jednocześnie, bez zwiększenia liczebności próby, a jedynie poprzez zmianę sposobu warstwowania i alokacji próby. W takim przypadku poprawa jakości uogólnień regionalnych uzyskiwana jest kosztem niewielkiego pogorszenia jakości uogólnień dla całego kraju.

Jednocześnie jest prawdopodobne, że faktycznie uzyskane oceny precyzji po realizacji badania okażą się nieco bardziej optymistyczne, niż wynika to z symulacji. Przeprowadzone symulacje mogły być obciążone większym błędem, niż ocena precyzji *ex-post* dokonywana po badaniu, ponieważ w zbiorowości stanowiącej populację symulacyjną (pochodzącej z badania realizowanego na próbie o innym podziale na warstwy i alokacji) występowały bardzo małe zbiory obserwacji dla niektórych warstw, mogące nie odzwierciedlać rozkładu w sposób adekwatny. W związku z tym nie jest wykluczone, że niektóre efekty, które wydały się członkom zespołu realizującego prace nieco zaskakujące (a ten do nich należy), w pewnym stopniu wynikają z błędów symulacji. Należy też podkreślić, że w przypadku rekomendowanego sposobu wdrożenia nowej metodyki jedynie dla prób nowo losowanych, w związku ze schematem rotacji stosowanym w badaniu przez pierwszy rok podpróby o nowej strukturze będą stanowiły maksymalnie połowę całkowitej próby, na której realizowane jest badanie; **wpływ wprowadzonych zmian na wyniki badania będzie zatem też w tym okresie proporcjonalnie ograniczony**. Stanowi to dodatkowe zabezpieczenie, dające możliwość wprowadzenia korekt w przyjętych założeniach przy losowaniu prób na kolejny rok, bez znaczącego uszczerbku dla finalnych wyników badania, gdyby ocena dokonana na podstawie rzeczywistych wyników badania dla nowo losowanych podrób uzasadniała takie działanie. W związku z ww. zastrzeżeniem należy podkreślić, że **konieczne jest prowadzenie prac monitorujących ocenę jakości i precyzji danych uzyskanych po realizacji badania zgodnie z nową metodyką dla uogólnień na poziomie kraju**. Jest to szczególnie istotne w związku z koniecznością zapewnienia reprezentatywności próby losowej wymaganej w Rozporządzeniu Rady (WE) nr 577/98 z dnia 9 marca 1998 r. (z późn. zm.) w sprawie organizacji badania prób losowych dotyczącego siły roboczej we Wspólnocie. **W przypadku zaobserwowania znaczącego pogorszenia, należy wprowadzić korekty w przyjętych założeniach przy losowaniu prób na kolejny rok**.

Rekomendowany wariant został wybrany na podstawie przeprowadzonej analizy efektywności rozważanych rozwiązań, z uwzględnieniem kryteriów organizacyjno-finansowych, które miały znaczący wpływ na ostateczny wybór. W przypadku, gdyby pojawiły się finansowe i organizacyjne możliwości, pozwalające zwiększyć liczebność próby o 50%, byłoby możliwe, na bazie wypracowanych w ramach projektu rozwiązań, wdrożenie metodyki pozwalającej uzyskać oszacowania akceptowalnej jakości dla znaczącej liczby wskaźników na poziomie NTS 3, tworzących kompletny zestaw danych. W takim przypadku, należałoby rozważyć zastosowanie sposobu warstwowania i alokacji próby określanego jako „NTS 3”, który przy założeniu utrzymania dotychczasowej liczebności próby nie znalazł zastosowania ze względu na brak możliwości istotnego i przynoszącego realny zysk informacyjny poszerzenia zasobu dostępnych danych dla poziomu podregionu.

W ramach pracy badawczej została także dokonana weryfikacja czy zastosowanie grupowania podregionów NTS 3 może przyczynić się do uzyskania poprawy jakości wskaźników. Przeprowadzony w projekcie eksperyment wykazał, że zastosowanie metody łączenia podregionów NTS 3, może przynieść oczekiwane rezultaty w postaci poprawy precyzji. Eksperyment został przeprowadzony na przykładzie grupowania opracowanego na potrzeby niniejszego projektu. Wobec braku propozycji grupowań NTS 3 użytecznych dla odbiorców danych statystycznych, nie kontynuowano tych prac.



Unia Europejska
Fundusz Spójności



Jednocześnie pod koniec I etapu projektu Ministerstwo Infrastruktury i Rozwoju zgłosiło zainteresowanie otrzymywaniem danych BAEL w podziale na: miasto wojewódzkie i resztę obszaru województwa (tj. z wyłączeniem obszaru miasta wojewódzkiego), co zostało uwzględnione w II etapie projektu oraz w rekomendacjach.

W zakresie wpływu prac dotyczących modernizacji BAEL na jakość wskaźników będących celem projektu należy stwierdzić, że zmiany planowane do wprowadzenia w EU LFS (w obecnym zakresie) nie będą miały wpływu na jakość rozważanych danych.