

Unia Europejska
Fundusz Spójności



Fundusze
Europejskie
Pomoc Techniczna

Praca badawcza pt.

**„Metoda dekompozycji Produktu Krajowego Brutto (PKB) oraz Wartości
Dodanej Brutto (WDB) w zastosowaniu do analizy struktury różnic
regionalnych”**

Raport końcowy

Praca realizowana w ramach projektu współfinansowanego ze środków Unii Europejskiej w ramach Programu Operacyjnego Pomoc Techniczna 2014-2020.

Wykonawca: Centrum Badań i Edukacji Statystycznej GUS

Dyrektor Mariusz Kraj

Kierownik Projektu: Michał Lewandowski

Opracował zespół badawczy: Maciej Banaś, Dariusz Kotlewski, Joanna Kulczycka,
Dorota Doniec, Grażyna Witkowska i inni.



Część A – dekompozycja wzrostu PKB *per capita* oraz różnic w poziomie WDB *per capita* względem średniej

1 Wprowadzenie

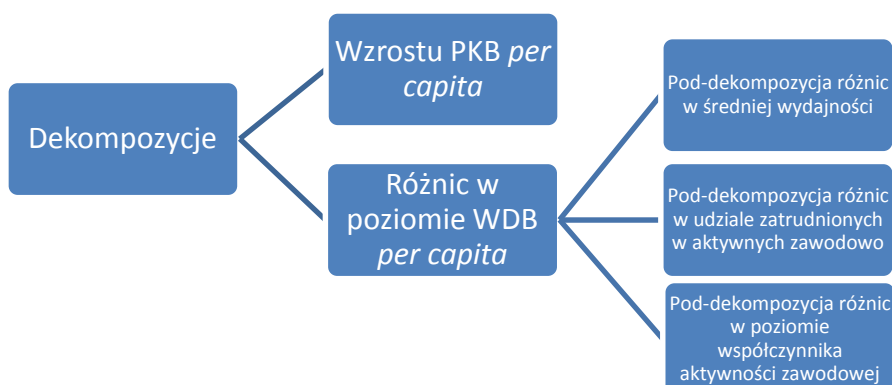
Dekompozycja wzrostu PKB *per capita* i dekompozycja różnic w poziomie WDB *per capita* pozwala na wyodrębnienie czynników obrazujących szerokie spektrum procesów zarówno gospodarczych, jak i społecznych. Przedstawione metody umożliwiają usystematyzowane i jednolite podejście do złożonego zagadnienia identyfikacji źródeł wzrostu lub rozwoju gospodarczego, ponieważ polegają na odpowiednim przetworzeniu ogólnie dostępnych danych, które są regularnie publikowane przez GUS, Eurostat i OECD. Mierniki wykorzystywane w dekompozycji, takie jak wskaźnik zatrudnienia, współczynnik aktywności zawodowej lub poziom PKB na osobogodzinę pracy pracującego, są powszechnie znane, łatwo interpretowalne i często stosowane w analizach społeczno-gospodarczych. Jednocześnie dekompozycja stanowi tylko narzędzie i jako taka nie przesądza o sposobie interpretacji wyników, jednakże pozwala budować cenne obserwacje i umożliwia ich przejrzystą prezentację, która może dotyczyć zarówno struktury, jak i ewolucji struktury wzrostu gospodarczego.

Dekompozycji można dokonać w różnych ujęciach, które wzajemnie się uzupełniają i oparte są na tej samej metodologii:

- W ujęciu dynamicznym analizowane są zmiany strukturalne w czasie dynamiki PKB *per capita*.
- W statycznym ujęciu porównawczym przeprowadzany jest rozkład różnic w poziomie WDB lub PKB dla danej jednostki (makroregionu, województwa, podregionu) w stosunku do średniej dla większego obszaru ją zawierającego (np. dla makroregionu lub województwa może to być Polska, dla podregionu może to być makroregion etc.). W ramach tego ujęcia dekompozycja uwzględnia dodatkowo regionalne zróżnicowanie ze względu na:
 - sektory gospodarki,
 - strukturę wykształcenia (poziom wykształcenia osób pracujących),
 - strukturę wiekową (w grupach wiekowych osób aktywnych zawodowo).

Poniżej przedstawiono schemat obu rodzajów dekompozycji omawianych w niniejszym tekście.

Rysunek 1: Schemat metod dekompozycji



Źródło: opracowanie własne.

Metody dekompozycji mogą być stosowane na różnych poziomach analitycznych, które wzajemnie się uzupełniają i mogą być stosowane łącznie bądź zamiennie. Wyróżniamy prezentację dekompozycji w postaci:

- a. **szeregu czasowego** – pozwala na badanie ewolucji struktury wzrostu/różnic,
- b. **danych przekrojowych** – pozwala na porównanie struktury wzrostu/różnic dla kilku jednostek terytorialnych w danym okresie czasowym,
- c. **danych panelowych** – pozwala na analizę porównawczą ewolucji struktury wzrostu/różnic dla paru jednostek terytorialnych.

Metody dekompozycji mogą być stosowane do analizy jednostek terytorialnych na różnym poziomie agregacji terytorialnej:

- a. makroregionu¹,

¹ Pojęcie makroregionu stosowane w niniejszym opracowaniu nie odpowiada jednostce NTS 1, natomiast odpowiada definicji makroregionu stosowanej przy opracowywaniu strategii ponadregionalnych przez Ministerstwo Infrastruktury i Rozwoju. Wyróżnia się 5 makroregionów: Polska Wschodnia (województwo lubelskie, podkarpackie, podlaskie, świętokrzyskie i warmińsko-mazurskie), Polska Południowa (województwo małopolskie i śląskie), Polska Centralna (województwo mazowieckie i łódzkie), Polska Północna (województwo pomorskie i kujawsko-pomorskie) oraz Polska Zachodnia (województwo zachodnio-pomorskie, wielkopolskie, lubuskie, opolskie i dolnośląskie).

- b. województwa (NTS 2),
- c. podregionu (NTS 3),
- d. etc.

Zaletą przedstawianej metody jest jej wielowymiarowość. Dodatkowo metoda jest otwarta na różnego rodzaju modyfikacje. Zostało to zobrazowane na przykładzie dekompozycji dynamicznej w podrozdziale 2.3.2.

Metodologia stosowana w niniejszej pracy opiera się na nowej, lecz uznanej już w literaturze metodologii opisanej przez Spiezia (2003)², znajdującej swe praktyczne zastosowanie w szeregu publikacji. Bardzo zbliżona metoda stosowana jest na przykład jako element analizy OECD *Regions at a Glance 2009* – jednej z flagowych publikacji OECD w zakresie opracowań regionalnych.

Metoda dekompozycji analizowana w niniejszym opracowaniu polega na rozkładzie danego miernika sytuacji gospodarczej (wzrost PKB *per capita* bądź różnice względem średniej w poziomie WDB *per capita*) na szereg czynników, które po dodaniu do siebie składają się na wartość danego miernika. Zakłada się, że wpływ tych czynników można od siebie oddzielić, zatem rozkład zapisany może być w formie liniowo-addytywnej. Zaletą tego podejścia jest prostota, łatwość prezentacji i możliwość bezpośredniej interpretacji. Wadą jest to, że nie bierze się pod uwagę tych elementów, które wyrażają jednoczesny wpływ paru czynników³ ujętych w dekompozycji, którego nie da się jednoznacznie przyporządkować do żadnego z tych czynników z osobna. W niniejszym opracowaniu po raz pierwszy została poruszona kwestia przybliżenia dekompozycji poprzez porównanie dekompozycji liniowo –addytywnej stosowanej w niniejszej analizie z dekompozycją dokładną, która dodatkowo uwzględnia wszystkie elementy „interakcyjne” pomiędzy poszczególnymi składowymi. W wyniku analizy różnicy pomiędzy tymi dekompozycjami została zaproponowana metoda redukcji błędu. Jest to szczególnie istotne w kontekście dekompozycji różnic względem średniej w poziomie WDB *per capita*, ponieważ różnice te mogą przyjmować duże wartości i błąd przybliżenia jest wówczas duży. Dzięki algorytmowi redukcji błędu dekompozycja różnic od średniej staje się możliwa bez konieczności

² Vincenzo Spiezia, *Identifying the determinants of regional performances*, 6th Session of the Working Party on Territorial Indicators, 2003.

³ Te elementy nazywać będziemy elementami „interakcyjnymi”.

akceptowania dużego błędu, co ma miejsce szczególnie w przypadku dużych różnic od średniej przy jednoczesnym zachowaniu zalet dekompozycji przybliżonej, tj. prostoty i łatwości prezentacji i interpretacji. Kwestia błędu przybliżenia staje się istotna również w przypadku dekompozycji dynamiki PKB *per capita* w sytuacji, kiedy zmiany w czasie są odpowiednio duże, co może mieć miejsce szczególnie jeśli kolejne analizowane okresy są od siebie mocno odległe w czasie. O wadze problemu przybliżenia świadczy fakt, iż z powodu nie uwzględnienia w pracach Spiezia problemu błędu przybliżenia, niemożliwe okazało się zastosowanie tej metodologii w ujęciu statyczno-porównawczym tj. w rozkładzie różnic w poziomie WDB *per capita* dla regionu w stosunku do średniej dla grupy regionów. W związku z tym ograniczono się do wykorzystania dekompozycji w ujęciu dynamicznym, tj. rozkładu wzrostu (w okresie rocznym) WDB *per capita*. W niniejszej publikacji, dzięki poruszeniu problemu błędu przybliżenia rozkładu oraz opracowaniu sposobu uwzględnienia go w rozkładzie, możliwe stało się wykorzystanie metodologii dekompozycji w pełnym zakresie, tj. w obu ujęciach, dynamicznym i statyczno-porównawczym.

Praca składa się z czterech rozdziałów. Rozdział 1 wprowadza idee podstawowych metod dekompozycji. W rozdziale 2 omówiona została metodologia różnych rodzajów dekompozycji. Najpierw przedstawiono ogólną metodę dekompozycji w zapisie matematycznym wraz z autorskim rozszerzeniem tej metody na analizę błędu przybliżenia. Następnie wprowadzono zmienne i wskaźniki wykorzystywane w dekompozycjach. W kolejnych dwóch podrozdziałach omówiono metodologię dekompozycji wzrostu PKB *per capita* oraz różnic w poziomie WDB *per capita* względem średniej. W kolejnej części wprowadzono metodologię dekompozycji według współczynnika Theila. Wyzwania praktycznej implementacji przedstawionych metod zostały opisane w następnym podrozdziale. W rozdziale 3 przedstawiono opis techniczny bazy danych, plików wynikowych oraz narzędzia programu MS Excel z makrami VBA. Rozdział 4 zawiera wyniki i ich interpretacje. W aneksie na końcu raportu zebrano wybrane wykresy.

1.1 Idea dekompozycji wzrostu na przykładzie

Zacznijmy od przykładu dekompozycji wzrostu PKB *per capita*. Realne PKB *per capita* można zapisać jako iloczyn wskaźników dotyczących:

- Efektywności gospodarki
 - Średnia wydajność pracy
- Rynku pracy

- Wskaźnik zatrudnienia
- Faktyczny roczny czas pracy
- Demografii
 - Udział populacji 15-64 w całości populacji

Wybrane wskaźniki ilustrują syntetyczny sposób opisu złożonej struktury gospodarki i jako takie powinny być traktowane jako bardzo wycinkowy obraz całości. Są to jednak wskaźniki powszechnie wykorzystywane oraz oficjalnie publikowane przez GUS. Przykładowy sposób zapisu realnego PKB *per capita* na składniki opisane powyżej może być następujący:

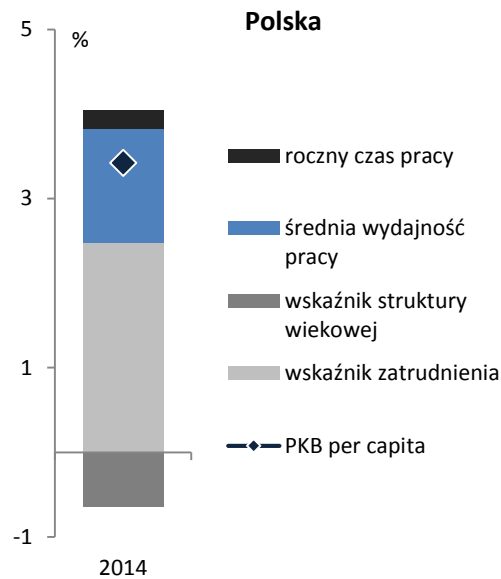
$$\underbrace{\frac{\text{PKB}}{\text{Populacja}}}_{\text{PKB per capita}} = \underbrace{\frac{\text{PKB}}{\text{Pracujący} \times \text{Czas pracy}}}_{\text{wydajność pracy}} \times \underbrace{\frac{\text{Pracujący}}{\text{Populacja 15-64}}}_{\text{wsk. zatrudnienia}} \times \text{Czas pracy} \times \underbrace{\frac{\text{Populacja 15-64}}{\text{Populacja}}}_{\text{wsk. str. wiekowej}}$$

Powyzsza postać multiplikatywna wyrażenia może zostać przekształcona w postać addytywną przy przejściu na **względne zmiany w czasie** odpowiednich wskaźników – poniżej podajemy przykład dla realnego PKB *per capita* Polski w 2014 r.:

$$\underbrace{\Delta_t^{\%} \frac{\text{PKB}}{\text{Populacja}}}_{+3,42\%} \approx \underbrace{\Delta_t^{\%} \frac{\text{PKB}}{\text{Pracujący} \times \text{Czas pracy}}}_{+1,36\%} + \underbrace{\Delta_t^{\%} \frac{\text{Pracujący}}{\text{Populacja 15-64}}}_{+2,47\%} + \underbrace{\Delta_t^{\%} \text{Czas pracy}}_{+0,22\%} + \underbrace{\Delta_t^{\%} \frac{\text{Populacja 15-64}}{\text{Populacja}}}_{-0,64\%}$$

Powyzsza zależność zachodzi w sposób przybliżony przy założeniu małych zmian względnych poszczególnych wskaźników. Wykres 1 przedstawia tę przykładową dekompozycję:

Wykres 1: Dekompozycja wzrostu realnego PKB per capita Polski w 2014 r.



Źródło: opracowanie własne na podstawie danych GUS.

Powyższy przykład obrazuje, w jaki sposób zmiany w czasie PKB *per capita* mogą zostać rozbite na addytywne składowe zmiany dotyczące różnych obszarów gospodarki.

1.2 Idea dekompozycji różnic na przykładzie

Drugi przykład dotyczy dekompozycji różnic w poziomie WDB *per capita*. WDB *per capita* można rozbić jako iloczyn wybranych wskaźników dotyczących:

- efektywności gospodarki:
 - średnia wydajność pracy,
- rynku pracy:
 - wskaźnik zatrudnienia zapisany jako iloczyn:
 - udziału pracujących w liczbie aktywnych zawodowo,
 - oraz współczynnika aktywności zawodowej,
 - faktyczny roczny czas pracy,
- demografii:
 - udział populacji w wieku 15-64 lat w całości populacji.

Powyższy podział jest podobny do podziału wykorzystanego przy ilustracji dekompozycji wzrostu, różni się jednak od niego, aby uwzględnić różną specyfikę dwóch metod dekompozycji. Przykładowy sposób zapisu WDB *per capita* na składniki opisane powyżej może być następujący:

$$\frac{\text{WDB}}{\text{Populacja}} = \frac{\text{WDB}}{\text{Pracujący} \times \text{Czas pracy}} \times \frac{\text{Pracujący}}{\text{Aktywni zawodowo}} \times \frac{\text{Aktywni zawodowo}}{\text{Populacja 15-64}} \times \text{Czas pracy}$$

WDB *per capita*
wydajność pracy
wskaźnik zatrudnienia
wsp. akt. zawodowej

$$\times \frac{\text{Populacja 15-64}}{\text{Populacja}} \times \text{Różnice w metodologii}$$

wskaźnik str. wiekowej

Liczba przeciętnie pracujących w ciągu roku jest liczona inaczej na potrzeby wydajności pracy (zgodna z Rachunkami Narodowymi) i inaczej w zastosowaniu do wskaźnika zatrudnienia (według BAEL): w wyrażeniu powyżej zaznaczone kolejno kolorem czerwonym i zielonym.

Populacja 15-64 jest liczona inaczej na potrzeby wskaźnika zatrudnienia i współczynnika aktywności zawodowej (według BAEL) i inaczej na potrzeby badań demograficznych i wykorzystywanej do liczenia wskaźnika struktury wiekowej: w wyrażeniu powyżej zaznaczone kolejno kolorem zielonym i niebieskim.

Stąd składnik „Różnice w metodologii” dzieli się na dwie składowe: te wynikające z „przejścia” z jednej liczby pracujących na drugą i tę wynikającą z „przejścia” jednej populacji 15-64 na drugą.

Powyższa postać multiplikatywna wyrażenia może zostać przekształcona w postać addytywną przy przejściu na **różnice względem średniej** odpowiednich wskaźników – poniżej podajemy przykład dla WDB *per capita* według parytetu siły nabywczej Polski w stosunku do Unii Europejskiej w roku 2014:

$$\underbrace{\Delta_{\text{PL vs UE}}^{\%} \frac{\text{WDB}}{\text{Populacja}}}_{-32,1\%} = \underbrace{\Delta_{\text{PL vs UE}}^{\%} \frac{\text{WDB}}{\text{Pracujący} \times \text{Czas pracy}}}_{-41,2\%} + \underbrace{\Delta_{\text{PL vs UE}}^{\%} \frac{\text{Pracujący}}{\text{Aktywni zawodowo}} + \Delta_{\text{PL vs UE}}^{\%} \frac{\text{Aktywni zawodowo}}{\text{Populacja 15-64}}}_{-4,7\%}$$

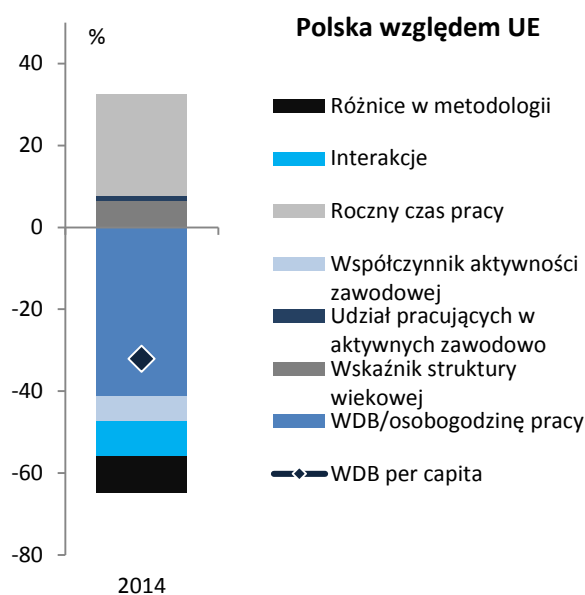
$$+ \underbrace{\Delta_{\text{PL vs UE}}^{\%} \text{Czas pracy}}_{+24,9\%} + \underbrace{\Delta_{\text{PL vs UE}}^{\%} \frac{\text{Populacja 15-64}}{\text{Populacja}}}_{+6,4\%} + \underbrace{\text{Różnice w metodologii}}_{-8,8\%} + \underbrace{\text{Interakcje}}_{-8,6\%}$$

Powyższa zależność zachodzi w sposób ścisły z uwagi na uwzględnienie elementu „Interakcje”, który reprezentuje przybliżenie będące rezultatem przejścia z formy multiplikatywnej do formy addytywnej. Powodem umieszczenia dodatkowego elementu w dekompozycji różnic przy jednoczesnym nie uwzględnieniu podobnego elementu w dekompozycji wzrostu jest fakt, iż

roczne zmiany w czasie poszczególnych wskaźników są najczęściej małe (<10%) przez co dokładność przybliżenia w dekompozycji może być uznana za wystarczającą (błąd najczęściej <0,1%). Z drugiej strony różnice w poziomie wskaźników dla Polski w stosunku do średniej dla UE są najczęściej znacznie większe i stąd dokładność przybliżenia jest mniejsza, co sprawia, że jest konieczne, aby uwzględnić błąd (reprezentowany za pomocą elementu „Interkacje”⁴).

Wykres 2 przedstawia przykładową dekompozycję różnic:

Wykres 2: Dekompozycja różnic w poziomie realnego PKB per capita według parytetu siły nabywczej Polski w stosunku do Unii Europejskiej w 2014 r.



Źródło: opracowanie własne na podstawie danych GUS.

Powyższy przykład pokazuje w jaki sposób różnice w poziomie WDB *per capita* mogą zostać rozbite na addytywne składowe różnice dotyczące różnych obszarów gospodarki.

⁴ W dalszej części opracowania znajduje się dyskusja na temat błędu przybliżenia oraz sposobów radzenia sobie z nim przy wykorzystaniu algorytmu redukcji błędów poprzez uwzględnienie części elementów interakcyjnych z dekompozycji pełnej – stąd nazwa „Interakcje”.

2 Metodologia

W niniejszej części przedstawiony zostanie zarys teoretyczny metod.

2.1 Podstawy matematyczne metod dekompozycji

Niech dany wskaźnik X , którego chcemy opisywać (np. PKB *per capita*), będzie zapisany jako iloczyn N składowych:

$$X = X_1 \times X_2 \times \dots \times X_N, \quad (*)$$

gdzie zarówno X , jak i wszystkie zmienne X_i są dodatnie. Dla potrzeb implementacji ważne jest, aby poszczególne składowe miały interpretację ekonomiczną i podobnie jak w przykładzie powyżej dotyczyły poszczególnych aspektów mających wpływ na strukturę X .

Większość wskaźników ekonomicznych daje się zapisać w formie (*) opisanej powyżej, czyli w postaci multiplikatywnej. Często pojawia się potrzeba sprowadzenia wyrażenia do postaci addytywnej, gdyż postać taka jest łatwiejsza w interpretacji. W tym celu, powyższe wyrażenie wystarczy zlogarytmować:

$$\log X = \log X_1 + \log X_2 + \dots + \log X_N$$

Następnie korzystając z przybliżenia $\Delta \log x \cong \Delta x/x$ dla małych Δx możemy zapisać:

$$\frac{\Delta X}{X} \cong \frac{\Delta X_1}{X_1} + \frac{\Delta X_2}{X_2} + \dots + \frac{\Delta X_N}{X_N}, \quad (**)$$

gdzie $\Delta X_i/X_i$ oznacza zmianę procentową składowej X_i .

W wyniku prostego przekształcenia wyrażenia multiplikatywnego otrzymujemy postać addytywną, gdzie poszczególne elementy są zmianami względnymi początkowych zmiennych, przy czym „względność” tych zmian dotyczyć może zarówno czasu, jak i porównań między jednostkami. W szczególności, jeśli X oznacza PKB *per capita* dla Polski w roku 2013, to $\frac{\Delta X}{X}$ może oznaczać wzrost PKB *per capita* w 2014 roku w stosunku do roku 2013, tj. ogólnie dynamikę tej zmiennej, ale może również oznaczać różnicę w poziomie PKB *per capita* dla Polski względem średniej dla całej Unii Europejskiej. W pierwszym przypadku ΔX oznacza zmianę w czasie zmiennej X , natomiast w drugim przypadku ΔX oznacza różnicę względem średniego poziomu zmiennej X . Te dwa przypadki odpowiadają dwom rodzajom dekompozycji przedstawionym na Rysunku Rysunek 1, tj. dekompozycji wzrostu PKB *per capita* oraz dekompozycji różnic względem

średniej w poziomie WDB *per capita*. Powód, dla którego w jednym przypadku mowa jest o PKB *per capita*, a w drugim o WDB *per capita*, jest taki, że dekompozycja różnic względem średniej umożliwia dodatkowe pod-dekompozycje, które są opisane w dalszej części pracy, a te dekompozycje wymagają danych szczegółowych, które dostępne są tylko dla WDB *per capita*, a nie są dostępne dla PKB *per capita*. Jednak ponieważ PKB i WDB różnią się zasadniczo tym, że PKB zawiera podatki od produktów a nie zawiera dotacji do produktów, a WDB odwrotnie, to nie ma to większego znaczenia dla wyjaśnienia struktury wzrostu gospodarczego bądź struktury różnic względem średniej w poziomie rozwoju gospodarczego.

Własność przybliżenia logarytmu przedstawiona powyżej (**), która umożliwia przedstawienie składowych zmian zmiennej X w postaci addytywnej będzie wykorzystywana w metodach dekompozycji przedstawionych w niniejszej pracy.

2.1.1 Rozszerzenie: dekompozycja błędu przybliżenia

Przekształcenie w postać addytywną wykorzystywane w metodzie dekompozycji a opisane w rozdziale 2.1 nie zachodzi w sposób ścisły, ponieważ przyjmuje się przybliżenie liniowe funkcji logarytmicznej. Zatem opisana własność zachodzi tylko wówczas, gdy przyrosty zmiennych są małe⁵. W praktyce oznacza to, że dla małych przyrostów poszczególnych zmiennych (ΔX_i), mały jest również błąd przybliżenia. Nawiązując do dekompozycji opisanych w niniejszej pracy, małych wartości przyrostów ΔX_i można się spodziewać w przypadku dekompozycji dynamicznej, ponieważ wówczas ΔX_i oznacza zmianę zmiennej X_i w czasie, a roczne przyrosty (bądź spadki) zmiennych rzadko przekraczają 10%. Na przykład, w 2007 roku wzrost realnego PKB *per capita* w poszczególnych województwach Polski oscylował w granicach od trzech do dziewięciu procent w zależności od województwa. Z kolei zmienna ΔX_i dla dekompozycji różnic regionalnych oznacza różnicę w stosunku do średniej, a ta zaś dla regionów bardzo zróżnicowanych może okazać się wysoka. Na przykład, poziom WDB *per capita* w województwie mazowieckim w 2004 roku był ponad 50% wyższy od średniej dla Polski. Dlatego problem przybliżenia jest szczególnie istotny w przypadku dekompozycji różnic regionalnych, choć dla dekompozycji dynamicznej również występuje.

Poniższa analiza dotycząca błędu przybliżenia ma na celu wskazanie sposobu, w jaki z błędem

⁵ Formalnie tylko wówczas, gdy przyrosty absolutne poszczególnych zmiennych dążą do zera.

przybliżenia można postępować.

Wartość ΔX można obliczyć stosując wyrażenie (**). Jednak jest to tylko przybliżona wartość w przypadku, gdy przyrosty zmiennych są małe. Dokładny wzór uzyskujemy w wyniku zastosowania definicji (*) dla nowej wartości zmiennej X i przemnożenia odpowiednich składowych:

$$\begin{aligned}
 (X + \Delta X) &= (X_1 + \Delta X_1) \times (X_2 + \Delta X_2) \times \dots \times (X_N + \Delta X_N) \\
 &= X_1 X_2 \dots X_N + \\
 &+ \Delta X_1 X_2 \dots X_N + X_1 \Delta X_2 \dots X_N + \dots + X_1 X_2 \dots \Delta X_N + \\
 &+ \Delta X_1 \Delta X_2 \dots X_N + \Delta X_1 X_2 \Delta X_3 \dots X_N + \dots + X_1 X_2 \dots \Delta X_{N-1} \Delta X_N + \\
 &+ \dots + \\
 &+ \Delta X_1 \Delta X_2 \dots \Delta X_N
 \end{aligned}$$

W wyniku prostego algebraicznego przekształceniu, otrzymujemy postać dokładnego rozkładu przyrostów względnych zmiennej X :

$$\begin{aligned}
 \frac{\Delta X}{X} &= \overbrace{\left[\frac{\Delta X_1}{X_1} + \frac{\Delta X_2}{X_2} + \dots + \frac{\Delta X_N}{X_N} \right]}^{\text{część ujęta w dekompozycji}} + \\
 &+ \frac{\Delta X_1}{X_1} \frac{\Delta X_2}{X_2} + \frac{\Delta X_1}{X_1} \frac{\Delta X_3}{X_3} + \dots + \frac{\Delta X_{N-1}}{X_{N-1}} \frac{\Delta X_N}{X_N} + \\
 &+ \dots + \\
 &+ \frac{\Delta X_1}{X_1} \frac{\Delta X_2}{X_2} \dots \frac{\Delta X_N}{X_N}, \quad (***)
 \end{aligned}$$

Jak widać, dokładne rozwinięcie przyrostów zmiennej X jest sumą części ujętej w dekompozycji opisanej wzorem (**) oraz reszty. Zatem błąd przybliżenia zdefiniować można następująco:

$$\begin{aligned}
 &\frac{\Delta X_1}{X_1} \frac{\Delta X_2}{X_2} + \frac{\Delta X_1}{X_1} \frac{\Delta X_3}{X_3} + \dots + \frac{\Delta X_{N-1}}{X_{N-1}} \frac{\Delta X_N}{X_N} + \\
 &+ \dots + \\
 &+ \frac{\Delta X_1}{X_1} \frac{\Delta X_2}{X_2} \dots \frac{\Delta X_N}{X_N}
 \end{aligned}$$

Z powyższego wyprowadzenia widać, że błąd przybliżenia można podzielić na czynniki drugiego rzędu, trzeciego i tak aż do N-tego rzędu. Jeśli przyrosty poszczególnych zmiennych są małe, tym bardziej małe będą ich iloczyny i tym samym błąd przybliżenia można uznać za nieistotny. W sytuacji, gdy przyrosty poszczególnych zmiennych nie są wystarczająco małe, w oparciu o istotne elementy wyższego rzędu, możliwe jest wprowadzenie odpowiednich poprawek. Zaletą takiego

rozkładu błędu jest to, iż znamy jego strukturę, a jednocześnie decydując się na zastosowanie dekompozycji w praktyce mamy pełną elastyczność w zakresie doboru ilości składowych, na których nam zależy. Możliwe jest bowiem wprowadzenie dodatkowych elementów interakcyjnych obok głównych składowych dekompozycji w celu obniżenia błędu przybliżenia. Jednocześnie, nie jest koniecznym uwzględnienie wszystkich elementów podobnie jak w wyrażeniu (***) – uczyniłoby to zresztą dekompozycję mało przejrzystą i trudną w interpretacji, a dodatkowo praktycznie uniemożliwiłoby przejrzystą prezentację graficzną składowych dekompozycji. Jednakże dzięki znajomości struktury błędu przybliżenia, można zdecydować się na włączenie do dekompozycji tych elementów interakcyjnych drugiego i ewentualnie wyższych rzędów, które są relatywnie wysokie co do wartości w porównaniu do innych elementów błędu i jako takie w głównej mierze odpowiadają za wysokość samego błędu przybliżenia. Zmienne drugiego i wyższych rzędów jest co prawda trudniej interpretować, jednak nie jest to niemożliwe. Na przykład zmienna drugiego rzędu taka jak $\frac{\Delta X_1}{X_1} \frac{\Delta X_2}{X_2}$ reprezentuje wpływ jednoczesnej zmiany zmiennej X_1 i zmiennej X_2 , którego nie da się przyporządkować wyłącznie jednej z tych zmiennych. Tego typu zmienne interakcyjne mają zatem podobną interpretację do zmiennych multiplikatywnych używanych w ekonometrii często, gdy liniowa postać modelu ekonometrycznego zostanie negatywnie zweryfikowana i konieczne jest włączenie nieliniowych elementów.

Zatem istnieje prosta zależność pomiędzy wielkością błędu przybliżenia, na którą musimy się zgodzić a prostotą i łatwością w interpretacji składowych rozkładu. Gdy więc błąd przybliżenia w dekompozycji liniowej⁶ jest nieakceptowalny, wówczas trzeba znaleźć kompromisowe rozwiązanie pomiędzy dekompozycją pełną a dekompozycją liniową. Można to na przykład zrobić dla danego poziomu akceptowalnego błędu przybliżenia i dobrać minimalną ilość składowych wyższego rzędu w kolejności wielkości, która spowoduje obniżenie błędu przybliżenia poniżej zadanej wielkości.

Reasumując, powyższa metoda uwzględnienia błędu przybliżenia dekompozycji umożliwia dokonanie dekompozycji i jej sensowną interpretację w sytuacji, gdy przyrosty bądź różnice dla poszczególnych zmiennych są duże. W publikacji OECD *Regions at a Glance 2009*

⁶ Tj. takim, w którym występują wyłącznie zmienne pierwszego rzędu.

prawdopodobnie właśnie duże wartości różnic dla zmiennych, a co za tym idzie duży błąd przybliżenia dekompozycji, spowodowały rezygnację z publikowania dekompozycji różnic regionalnych i dlatego porzeczano na dekompozycji wzrostu PKB *per capita*. Wyżej omówiona metoda pozwala na rozwiązanie tego problemu przy użyciu metody rozkładu błędu opisanego powyżej. Praktyczne zastosowanie tej metody zostanie zaprezentowane w dalszej części pracy.

2.2 Definicje zmiennych i wskaźników wykorzystywanych w dekompozycji wzrostu oraz różnic

Aby wprowadzić zapis formalny dekompozycji, potrzebne są definicje i notacja dla poszczególnych zmiennych. Wprowadźmy następujące oznaczenia:

- DW – dekompozycja wzrostu PKB *per capita*
- DR – dekompozycja różnic w poziomie WDB *per capita*
- PDR1, PDR2, PDR3 – pod-dekompozycja różnic składowych dekompozycji DR

Stosowane są również następujące subskrypty przy zmiennych:

- Subskrypt „i” oznacza daną jednostkę terytorialną (np. województwo),
- Subskrypt „s” oznacza grupy sekcji według podziału PKD 2007
- Subskrypt „w” oznacza grupy wiekowe
- Subskrypt „e” oznacza grupy wykształcenia

Tablica 1 zawiera spis zmiennych wykorzystanych w metodach dekompozycji. Źródła danych podane są w Rozdziale 5 - Aneksie.

Tablica 1: Spis zmiennych wykorzystywanych w dekompozycji

Lp.	Oznaczenie	Zmienna	Metodologia	Rodzaj dekompozycji*	Liczba podkategorii j	Uwagi
1	PKB_i	PKB ogółem w cenach stałych	Rachunki regionalne PKD 2007, ESA 2010	DW		wyliczone z PKB ogółem w cenach bieżących przemnożone przez dynamikę PKB ogółem w cenach stałych
2	WDB_i	WDB ogółem w cenach bieżących	Rachunki regionalne, PKD 2007, ESA 2010	DR		Alternatywnie w cenach stałych: można policzyć wykorzystując dynamikę PKB ogółem w cenach stałych oraz deflatory dla poszczególnych grup sekcji na poziomie krajowym (założenie byłoby takie, że nie ma interakcji między grupami sekcji a województwami, tj. deflatory dla tych dwóch przekrojów są niezależne od siebie
3	WDB_is	WDB ogółem w cenach bieżących według grup sekcji s	Rachunki regionalne, PKD 2007, ESA 2010	PDR1	7 grup sekcji: A, BDE, C, F, GHIJ, KL, MNOPQRST	
4	PP_i	Przeciętna liczba pracujących w miejscu pracy	Badania przedsiębiorstw, ale osobna kategoria zgodna z Rachunkami regionalnymi	DW, DR		wyliczone jako iloraz WDB ogółem w cenach bieżących i WDB na 1 pracującego w cenach bieżących
5	PP_is	Przeciętnie liczba pracujący w miejscu pracy według grup sekcji s	Badania przedsiębiorstw, ale osobna kategoria zgodna z Rachunkami regionalnymi	PDR1	7 grup sekcji: A, BDE, C, F, GHIJ, KL, MNOPQRST	wyliczone jako iloraz WDB ogółem w cenach bieżących według grup sekcji s i WDB na 1 pracującego w cenach bieżących według grup sekcji s

Lp.	Oznaczenie	Zmienna	Metodologia	Rodzaj dekompozycji*	Liczba podkategorii j	Uwagi
6	N_i	Populacja ogółem	Stan ludności	DW, DR		wyliczone jako iloraz PKB lub WDB ogółem w cenach bieżących i PKB lub WDB per capita w cenach bieżących, alternatywnie ludność ogółem (faktyczne miejsce zamieszkania) stan na 30 VI
7	N15+_i	Populacja 15+	Stan ludności	DW, DR		wyliczona jako różnica ludności ogółem i ludności w wieku przedprodukcyjnym (14 lat i mniej)
8	NB15+_i	Populacja 15+	BAEL	DW, DR		Alternatywnie te same kategorie ale w wieku produkcyjnym 15-59/64
9	NB15+_iw	Populacja 15+ według grup wiekowych w	BAEL	PDR3	4 grupy wiekowe: 15-29, 30-39, 40-49, 50+	
10	AZ_i	Aktywni zawodowo 15+	BAEL	DW, DR		Alternatywnie te same zmienne ale w wieku produkcyjnym 15-59/64
11	AZ_iw	Aktywni zawodowo 15+ według grup wiekowych	BAEL	PDR3	4 grupy wiekowe: 15-29, 30-39, 40-49, 50+	
12	AZ_ie	Aktywni zawodowo 15+ według grup wykształcenia (edukacji)	BAEL	PDR2	5 grup wykształcenia: wyższe, policealne oraz średnie zawod., średnie ogólnokoszt., zasad. zawod., gimnazj., podstaw. i niższe	
13	PZ_i	Pracujący w miejscu zamieszkania 15+	BAEL	DW, DR		Alternatywnie te same zmienne ale w wieku produkcyjnym 15-59/64

Lp.	Oznaczenie	Zmienna	Metodologia	Rodzaj dekompozycji*	Liczba podkategorii j	Uwagi
14	PZ_ie	Pracujący w miejscu zamieszkania 15+ według grup wykształcenia (edukacji)	BAEL	PDR2	5 grup wykształcenia: patrz powyżej	

*Patrz objaśnienia strona 18.

Na podstawie tych podstawowych zmiennych budowane są wskaźniki, które znajdują zastosowanie w omawianych metodach dekompozycji. Tablica 2 przedstawia te wskaźniki.

Poniżej przedstawiono krótką dyskusję na temat zastosowania tych wskaźników:

1. Zmienną dekomponowaną zarówno w dekompozycji wzrostu, jak i w dekompozycji różnic może być PKB *per capita*. W przypadku dekompozycji różnic zdecydowano się jednak na dekompozycję WDB *per capita*. Powodem jest to, że dla WDB dostępne są szczegółowe dane w rozbiciu na sekcje PKD a dla PKB takich danych brakuje. Te dane wykorzystywane są do pod-dekompozycji różnic w wydajności pracy, jednej z głównych składowych dekompozycji różnic.
2. Średnia wydajność pracy może być zdefiniowana albo jako WDB bądź PKB na 1 pracującego lub też jako WDB bądź PKB na osobogodzinę pracy osoby pracującej. Ten drugi rodzaj wydajności, który dodatkowo uwzględnia faktyczny roczny czas pracy został zastosowany w przykładowej dekompozycji różnic dla Polski względem Unii Europejskiej w 2014 r. we wprowadzeniu do niniejszego Raportu. W pozostałych dekompozycjach zastosowano jednak pierwszą definicję wydajności pracy (WDB lub PKB na 1 pracującego), ponieważ dane dotyczące czasu pracy w odpowiednich sekcjach PKD dla makroregionów, województw i podregionów nie są dostępne.

Tablica 2: Spis wskaźników stosowanych w obu rodzajach dekompozycji

Lp.	Wskaźnik	Opis	Dwie metodologie*	Rodzaj dekompozycji*	Uwagi
1	PKB _i /N _i	PKB <i>per capita</i>	ESA/demo	DW	Zmienna dekomponowana
2	WDB _i /N _i	WDB <i>per capita</i>	ESA/demo	DR	Zmienna dekomponowana
3	PKB _i /PP _i	PKB na 1 pracującego	Nd.	DW	Alternatywnie: PKB _i /(PP _i *H _i), gdzie H _i oznacza faktyczny roczny czas pracy osoby pracującej
4	WDB _i /PP _i	WDB na 1 pracującego	Nd.	DR	
5	WDB _{is} /PP _{is}	WDB na 1 pracującego w grupie sekcji s	Nd.	PDR1	
6	PP _{is} /PP _i	Udział pracujących w grupie sekcji s w ogóle pracujących	Nd.	PDR1	
7	PP _i /PZ _i *(PZ/PP)	Współczynnik dojazdów do pracy netto	ESA/BAEL	DW/DR	PZ/PP oznacza iloraz pracujących ogółem w Polsce według BAEL i według ESA, uwzględnienie tego elementu jest konieczne, aby wyeliminować błąd doboru próby; sam współczynnik dojazdów do pracy jest zatem ilorazem udziału liczby pracujących w miejscu pracy w danym województwie oraz udziału liczby pracujących w miejscu zamieszkania w danym województwie
8	PZ _i /AZ _i	Udział pracujących w aktywnych zawodowo	Nd.	DW/DR	
9	PZ _{ie} /AZ _{ie}	Udział pracujących w aktywnych zawodowo w grupie wykształcenia e	Nd.	PDR2	

Lp.	Wskaźnik	Opis	Dwie metodologie*	Rodzaj dekompozycji*	Uwagi
10	AZ_ie/AZ_i	Udział aktywnych zawodowo w grupie wykształcenia e w ogóle akt. zawod.	Nd.	PDR2	
11	AZ_i/NB15+_i	Współczynnik aktywności zawodowej	Nd.	DW/DR	Alternatywnie te same wskaźniki dla ludności w wieku produkcyjnym
12	NB15+_i/N15+_i	Wskaźnik różnicy w metodologii BAEL/demo	BAEL/demo	DW/DR	
13	N15_i/N_i	Wskaźnik struktury wiekowej	Nd.	DW/DR	
14	AZ_iw/NB15+_iw	Współczynnik aktywności zawodowej dla grupy wiekowej w	Nd.	PDR3	
15	NB15+_iw/NB15+_i	Udział populacji 15+ w grupie wiekowej w całości populacji 15+	Nd.	PDR3	

*BAEL: Badanie Aktywności Ekonomicznej Ludności, ESA: European System of National and Regional Accounts in the European Union (ESA 2010), demo: badania demograficzne

3. Następnym wskaźnikiem stosowanym w niniejszej pracy jest współczynnik dojazdów do pracy. Jest on zdefiniowany jako iloraz liczby pracujących w miejscu pracy w jednostce terytorialnej i i liczby pracujących w miejscu zamieszkania w tejże jednostce :

$$\frac{PP_i}{PZ_i} = \frac{PZ_i + DN_i}{PZ_i} = 1 + \frac{DN_i}{PZ_i},$$

gdzie DN_i oznacza dojazdy do pracy netto i jest to różnica pomiędzy liczbą pracujących w miejscu pracy w jednostce i i liczbą pracujących w miejscu zamieszkania w jednostce i ⁷. Zauważmy, że w momencie kiedy rozpatrujemy różnice lub dynamikę współczynnika dojazdów do pracy, wówczas zachodzi następująca relacja:

$$\Delta \frac{PP_i}{PZ_i} = \Delta \frac{DN_i}{PZ_i}$$

Współczynnik dojazdów do pracy umożliwia uwzględnienie wpływu osób, które zamieszkują w jednej jednostce terytorialnej a pracują w drugiej, wytwarzając tam wartość dodaną. Wpływ ten jest szczególnie znaczący, im mniejsze obszary analizujemy, ponieważ dojazdy do pracy rzadko występują na znacznych odległościach. W szczególności, jeśli by zignorować grono osób zamieszkujących teren Rzeczypospolitej Polskiej a pracujących za granicą, to na terytorium całej Polski liczba pracujących w miejscu zamieszkania jest równa liczbie pracujących w miejscu pracy i uwzględnianie dojazdów do pracy nie ma wówczas sensu. W przypadku mniejszych jednostek terytorialnych, takich jak województwa albo, nawet bardziej, powiaty lub gminy, dojazdów do pracy nie powinno się ignorować⁸, mogą bowiem w znaczny sposób wpływać na wyniki danego regionu. W podrozdziale 2.6 dotyczącym wyzwań praktycznej implementacji poruszono problem różnych prób (*sampling error*) przy liczeniu współczynnika dojazdów netto i wyjaśniono sens stosowanej poprawki,

⁷ Ponieważ aktywni zawodowo to pracujący i bezrobotni, a skoro bezrobotni nie mają pracy to w ich przypadku rozróżnienie pomiędzy miejscem pracy i miejscem zamieszkania jest bezprzedmiotowe, zatem dojazdy netto można alternatywnie zdefiniować jako różnicę pomiędzy liczbą aktywnych zawodowo w miejscu pracy a liczbą aktywnych zawodowo w miejscu zamieszkania w danym województwie.

⁸ Por. Kruszka K. (red.), *Dojazdy do pracy w Polsce. Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem*. Główny Urząd Statystyczny i Urząd Statystyczny w Poznaniu, Poznań 2010.

tj. zastosowanie wzoru $\left(\frac{PP_i}{PZ_i}\right) / \left(\frac{PP}{PZ}\right)$ zamiast prostego $\frac{PP_i}{PZ_i}$, gdzie zmienne z subskryptem dotyczą danej jednostki terytorialnej (makroregionu, województwa lub podregionu) a zmienne bez subskryptu to zmienne dla Polski.

4. Kolejnym wskaźnikiem używanym w dekompozycjach analizowanych w niniejszym opracowaniu jest udział pracujących w liczbie aktywnych zawodowo, zdefiniowany jako iloraz liczby pracujących w miejscu zamieszkania i liczby aktywnych zawodowo w miejscu zamieszkania: $\frac{PZ_i}{AZ_i}$.

5. Współczynnik aktywności zawodowej definiujemy na dwa możliwe sposoby⁹ jako stosunek:

a. Liczby aktywnych zawodowo do populacji w wieku od 15 roku życia: $\frac{AZ_i}{NB_i^{15+}}$

b. Liczby aktywnych zawodowo do liczby populacji ludzi w wieku produkcyjnym¹⁰: $\frac{AZ_i}{NB_i^{15-64}}$

Podobnie na dwa sposoby można zdefiniować wskaźnik zatrudnienia ($\frac{PZ_i}{NB_i^{15+}}$ lub

$\frac{PZ_i}{NB_i^{15-64}}$), który jest iloczynem udziału pracujących w aktywnych zawodowo (por. punkt

4. Powyżej) oraz współczynnika aktywności zawodowej.

Odpowiednie definicje przy uwzględnieniu wieku 15-64 zastosowano w przykładowych dekompozycjach dla Polski w 2014 r. na początku niniejszego Raportu. W dekompozycjach dla województw zastosowano jednak wskaźniki uwzględniające wiek 15 lat i więcej, ponieważ ta kategoria, w odróżnieniu od tej uwzględniającej wiek 15-64 lat, jest spójna z grupami wiekowymi stosowanymi przy pod-dekompozycji różnic w poziomie współczynnika aktywności zawodowej. Obie definicje umożliwiają oddzielenie wpływu zjawisk zachodzących na rynku pracy od zjawisk czysto demograficznych, które determinują popyt na pracę na rynku pracy – udział osób, które mogą pracować ze względu na wiek w ogólnej liczbie ludności. Jednak pierwsza definicja współczynnika aktywności zawodowej przedstawiona powyżej umożliwia

⁹ Te dwa sposoby odpowiadają odpowiednio tradycyjnej metodologii GUS oraz metodologii EUROSTATu.

¹⁰ Główny Urząd Statystyczny przyjmuje za wiek produkcyjny wiek 18-59 dla kobiet i 18-64 dla mężczyzn. Wiek 15-64 jest natomiast stosowany w Eurostacie.

uchwycenie tylko tych zjawisk demograficznych, które mają wpływ na proporcje liczby ludzi w wieku 15 lat i więcej do tych, którzy mają mniej niż 15 lat. Jednak wiadomo, że nie tylko osoby młode nie są zdolne do pracy z racji wieku, ale również osoby w odpowiednio podeszłym wieku bądź nie są zdolne do pracy, bądź osiągnęły już wiek emerytalny i nie muszą ani nie chcą już pracować. Druga definicja współczynnika aktywności zawodowej uwzględnia wpływ czynników demograficznych zarówno w grupie osób, które nie osiągnęły jeszcze odpowiedniego wieku (umownie 15 lat), jak i w grupie osób, które przekroczyły już wiek 64 lat.

6. Ponieważ populację ogółem oraz w wieku 15+ lub w wieku 15-64 można liczyć na dwa sposoby, tj. według metodologii BAEL oraz według badań demograficznych, to potrzebny jest wskaźnik „przejścia metodologii”, który definiowany jest również na dwa sposoby w zależności od definicji wskaźnika zatrudnienia/współczynnika aktywności zawodowej i wskaźnika struktury wiekowej jako $\frac{NB_i^{15+}}{N_i^{15+}}$ lub $\frac{NB_i^{15-64}}{N_i^{15-64}}$. Więcej na ten temat w podrozdziale 2.6 dot. wyzwań praktycznej implementacji.

7. Uwzględniając podwójną definicję współczynnika aktywności zawodowej, również w dwojaki sposób zdefiniowany jest współczynnik struktury wiekowej, jako udział:

a. Populacji od 15 roku życia wzwyż w całkowitej populacji: $\frac{N_i^{15+}}{N_i}$

b. Populacji w wieku produkcyjnym w całkowitej populacji: $\frac{N_i^{15-64}}{N_i}$

Współczynnik struktury wiekowej przybliża wpływ struktury wiekowej na wysokość PKB (bądź WDB) *per capita*. Przykładowo, im więcej osób w wieku produkcyjnym w stosunku do całkowitej liczby ludności, tym więcej osób teoretycznie zdolnych do pracy i wytwarzania PKB lub WDB.

Po wprowadzeniu podstawowych pojęć i definicji zmiennych zostaną omówione oba rodzaje dekompozycji.

2.3 Dekompozycja wzrostu PKB *per capita*

Pierwsza z omawianych dekompozycji to dekompozycja wzrostu PKB *per capita*. Poniżej omówiono ogólną charakterystykę dekompozycji, a następnie wprowadzono charakterystykę formalną metody.

2.3.1 Charakterystyka ogólna

Głównym celem tego rodzaju dekompozycji jest zidentyfikowanie struktury wzrostu PKB *per capita* dla danej jednostki poprzez rozkład na szereg składników, reprezentujących względne zmiany w czasie poziomów następujących mierników:

- średniej wydajności pracy
- wskaźnika dojazdów do pracy
- wskaźnika zatrudnienia
- współczynnika aktywności zawodowej
- wskaźnika struktury wiekowej.

Dzięki takiemu rozkładowi wzrostu PKB *per capita*, dekompozycja umożliwia oddzielenie wpływu:

1. średniej wydajności pracy, tj. czynnika umożliwiającego trwały wzrost PKB *per capita* i konkurencyjności gospodarki;
2. zmian sytuacji na rynku pracy mierzonych za pomocą:
 - a. dynamiki wskaźnika zatrudnienia;
 - b. dynamiki współczynnika aktywności zawodowej;
 - c. zmianami wskaźnika dojazdów do pracy;
3. wpływu czynników demograficznych mierzonych udziałem osób w wieku produkcyjnym w całości populacji.

Zaletą tego rodzaju dekompozycji jest wyodrębnienie czynników, które są elementami odrębnych obszarów polityki państwa. Na przykład, wydajność pracy jest uzależniona od poziomu i struktury inwestycji, poziomu innowacyjności w gospodarce oraz innych czynników wpływających na jakość pracy. Wskaźnik zatrudnienia i współczynnik aktywności zawodowej mierzy skuteczność polityki rynku pracy. Wskaźniki struktury wiekowej podlegają natomiast wpływowi zjawisk i polityki demograficznej.

2.3.2 Charakterystyka formalna

Zapis formalny dekompozycji dynamicznej przedstawia się następująco. Zgodnie z opisem w

rozdziale 2.1, punktem wyjścia jest tożsamość w postaci iloczynu¹¹:

$$\frac{PKB_i}{N_i} \equiv \underbrace{\frac{PKB_i}{PP_i}}_{\text{wydajność}} \times \underbrace{\frac{PP_i/PP}{PZ_i/PZ} \times \frac{PZ_i}{AZ_i} \times \frac{AZ_i}{NB_i^{15+}} \times \frac{N_i^{15+}}{N_i}}_{\text{intensywność}} \times \underbrace{\frac{NB_i^{15+}}{N_i^{15+}} \times \frac{PP}{PZ}}_{\text{przejście metodologii}}$$

Wykorzystując przybliżenie liniowe, otrzymujemy następujące wyrażenie:

$$\frac{\Delta \frac{PKB_i}{N_i}}{\frac{PKB_i}{N_i}} \cong \underbrace{\frac{\Delta \frac{PKB_i}{PP_i}}{\frac{PKB_i}{PP_i}} + \frac{\Delta \left(\frac{PP_i}{PZ_i}\right) / \left(\frac{PP}{PZ}\right)}{\left(\frac{PP_i}{PZ_i}\right) / \left(\frac{PP}{PZ}\right)} + \frac{\Delta \frac{PZ_i}{AZ_i}}{\frac{PZ_i}{AZ_i}} + \frac{\Delta \frac{AZ_i}{NB_i^{15+}}}{\frac{AZ_i}{NB_i^{15+}}} + \frac{\Delta \frac{NB_i^{15+}}{N_i^{15+}}}{\frac{NB_i^{15+}}{N_i^{15+}}} + \frac{\Delta \frac{N_i^{15+}}{N_i}}{\frac{N_i^{15+}}{N_i}} + \frac{\Delta \frac{PP}{PZ}}{\frac{PP}{PZ}}}_{\text{część ujęta w dekompozycji}} + \underbrace{\text{błąd przybliżenia}}_{\text{część nieujęta w dekompozycji}}$$

W zależności od potrzeb, dostępności danych i przyjętych założeń dotyczących metodologii powyższą dekompozycję można modyfikować. Na przykład jeśli wskaźnik typu $\frac{PZ_i}{AZ_i}$ jest niepożądany z uwagi na inną interpretację niż wskaźnik zatrudnienia liczony według metodologii GUS bądź Eurostat, wówczas dwa składniki, tj. $\frac{AZ_i}{NB_i^{15+}}$ i $\frac{PZ_i}{AZ_i}$ możemy scalić w jeden $\frac{PZ_i}{N_i^{15+}}$, uzyskując wskaźnik zatrudnienia podawany przez GUS i Eurostat. Również liczbę osób w wieku 15+ możemy zastąpić liczbą osób w wieku produkcyjnej, wówczas wskaźnik zatrudnienia nie będzie liczony według metodologii GUS, a według metodologii Eurostatu. Dodatkowo uniknąć można odrębnego ujęcia dynamiki dojazdów do pracy, poprzez uwzględnienie liczby pracujących z jednego tylko źródła (oznaczonych przez P).

W następnym rozdziale zostaną omówione kwestie związane ze spójnością dostępnych danych – co umożliwi opis dodatkowych możliwych modyfikacji powyższej metody w zależności od potrzeb i kryteriów oceny jakości metody przyjętych przez osoby, które zechcą wykorzystać metodę.

Następny podrozdział opisuje drugi rodzaj dekompozycji, zwanej dekompozycją różnic.

2.4 Dekompozycja różnic w poziomie WDB *per capita* względem średniej

Podobnie jak w poprzednim podrozdziale, tak i tutaj zachowano podział na dwie części. W pierwszej części zostanie omówiona ogólna charakterystyka metody wraz z dyskusją dotyczącą identyfikacji niewykorzystanych zasobów regionu. Następnie wprowadzony

¹¹ Wyrażenie $\frac{PP_i/PP}{PZ_i/PZ}$ uwzględnia dodatkowo błąd próby, który omówiony jest w podrozdziale 2.6.

zostanie szczegółowy opis formalny tych metod, tj. z wykorzystaniem wzorów. Dyskusja dotycząca implementacji, podobnie jak w przypadku dekompozycji wzrostu PKB *per capita*, została odroczone do rozdziału 2.6.

2.4.1 Charakterystyka ogólna

Dana jest jednostka większa (na przykład Polska) i jednostka mniejsza w niej zawarta (na przykład województwo mazowieckie). Głównym celem dekompozycji różnic regionalnych jest uchwycenie czynników, które determinują zróżnicowanie poziomu WDB *per capita* dla mniejszej jednostki w relacji do średniego poziomu dla jednostki większej. Czynnikiemami tymi są:

- różnice w średniej wydajności pracy;
- różnice w poziomie wskaźnika zatrudnienia;
- różnice w poziomie wskaźnika dojazdów do pracy;
- różnice w poziomie współczynnika aktywności zawodowej.

Każda z wymienionych składowych może być interpretowana jako miernik poszczególnych determinant sukcesu gospodarczego na poziomie regionalnym.

Średnia wydajność pracy jest wskaźnikiem wydajności regionalnej produkcji. Wskaźnik zatrudnienia mierzy efektywność lokalnego rynku pracy. Współczynnik aktywności zawodowej podsumowuje charakterystyki regionalnej siły roboczej, natomiast wskaźnik dojazdów do pracy może być traktowany jako miernik efektów lokalizacji geograficznej.

Te cztery ważne elementy wpływające na poziom rozwoju gospodarczego danej jednostki, a więc wydajność regionalnej produkcji, efektywność lokalnego rynku pracy, charakterystyka regionalnej siły roboczej oraz lokalizacja geograficzna mogą być przyporządkowane albo do naturalnego „wyposażenia” albo do niewykorzystanych zasobów.

Naturalne „wyposażenie” odnosi się do takich charakterystyk danej jednostki, które albo nie mogą być zmienione albo mogą być zmienione tylko w długim horyzoncie czasowym, na przykład lokalizacja geograficzna, zasoby naturalne, demografia, typ miejski lub wiejski.

Niewykorzystane zasoby odnoszą się natomiast do wszystkich tych charakterystyk danej jednostki, które mogłyby być bardziej efektywnie wykorzystywane i alokowane celem podniesienia poziomu WDB *per capita*, na przykład transport, technologia i infrastruktura, instytucje i regulacje rynku pracy, kapitał ludzki i społeczny.

Powyższy podział, który pozwala identyfikować źródła słabości lub siły danej jednostki i przez to rekomendować odpowiedni program naprawczy, jest możliwy w ramach prezentowanej metodologii dzięki dodatkowym pod-dekompozycjom składowych różnic powstałych w wyniku głównej dekompozycji:

- 1) Różnice w średniej wydajności pracy zostaną rozbite na następujące elementy:
 - a) specjalizacja sektorowa;
 - b) różnice w poziomie rozwoju infrastruktury i technologii;
- 2) Z kolei różnice w poziomie wskaźnika zatrudnienia zostaną zdekomponowane na:
 - a) profil wykształcenia;
 - b) różnice w efektywności rynku pracy;
- 3) Różnice w poziomie współczynnika aktywności zawodowej ulegną rozbiciu na:
 - a) profil struktury wiekowej;
 - b) różnice w poziomie uczestnictwa w rynku pracy.

Powyższe elementy można przyporządkować teraz do dwóch kategorii zasobów omówionych powyżej. Do naturalnego wyposażenia regionu zaliczyć można:

- specjalizację;
- strukturę demograficzną;
- dojazd do pracy do innego regionu/kraju niż miejsce zamieszkania.

Do niewykorzystanych zasobów zaliczyć można:

- technologię i infrastrukturę;
- wykształcenie/umiejętności;
- zatrudnienie/rynek pracy;
- uczestnictwo w rynku pracy.

Rysunek 2: Naturalne wyposażenie i niewykorzystane zasoby: podział czynników

Naturalne wyposażenie	Niewykorzystane zasoby
Specjalizacja	Technologia i Infrastruktura
Struktura demograficzna	Wykształcenie/umiejętności
Dojazdy do pracy	Zatrudnienie/rynek pracy
	Uczestnictwo w rynku pracy

Dekompozycja różnic regionalnych zasadniczo ma charakter statyczny, choć można nadać jej również charakter dynamiczny poprzez porównywanie różnic w czasie dla różnych okresów. Ujęcie „dynamiczne” pozwala na uchwycenie zmian w poziomie wykorzystania zasobów i zmian w strukturze wiekowej, profilu wykształcenia oraz specjalizacji sektorowej danego regionu.

2.4.2 Charakterystyka formalna dekompozycji podstawowej

Dekompozycja może zostać zapisana za pomocą poniższej relacji i przekształcenia, według ogólnej metody przedstawionej w rozdziale 2.1:

$$\frac{WDB_i}{N_i} = \frac{WDB_i}{PP_i} \cdot \frac{PP_i}{PZ_i} \cdot \frac{PZ_i}{AZ_i} \cdot \frac{AZ_i}{NB_i^{15+}} \cdot \frac{NB_i^{15+}}{N_i^{15+}} \cdot \frac{N_i^{15+}}{N_i}$$

$$\begin{aligned} & \frac{\frac{WDB_i}{N_i} - \frac{WDB}{N}}{\frac{WDB}{N}} \\ \frac{\Delta X}{X}: & \underbrace{WDB \text{ per capita}} \\ & = \underbrace{\frac{\frac{WDB_i}{PP_i} - \frac{WDB}{P}}{\frac{WDB}{P}}}_{\text{Produktywność}} + \underbrace{\frac{\frac{PP_i}{PZ_i} - \frac{P}{P}}{\frac{P}{P}}}_{\text{Wsp.dojazdów}} + \underbrace{\frac{\frac{PZ_i}{AZ_i} - \frac{P}{A}}{\frac{P}{A}}}_{\text{Wsk.zatrudnienia}} + \underbrace{\frac{\frac{AZ_i}{NB_i^{15+}} - \frac{A}{NB^{15+}}}{\frac{A}{NB^{15+}}}}_{\text{Wsp.aktywności}} + \underbrace{\frac{\frac{N_i^{15+}}{N_i} - \frac{N^{15+}}{N}}{\frac{N^{15+}}{N}}}_{\text{Wsk.demogr.}} \\ & + \underbrace{\frac{\frac{NB_i^{15+}}{N_i^{15+}} - \frac{NB^{15+}}{N^{15+}}}{\frac{NB^{15+}}{N^{15+}}}}_{\text{Wsk.zmiany metodologii}} + \text{Interakcje} \end{aligned}$$

W przypadku tej dekompozycji znajduje zastosowanie ta sama metoda, co w przypadku dekompozycji dynamicznej omówionej w podrozdziale 2.3.2 z tą różnicą, że tym razem rozkład dotyczy różnic w stosunku do średniej, a poprzednio rozkład dotyczył zmian w stosunku do roku poprzedniego. Matematycznie jest to jednak to samo przekształcenie. Należy jednak uwzględnić fakt, że różnice w stosunku do średniej dla wybranych wskaźników mogą być znacznie wyższe niż różnice w czasie w stosunku do roku poprzedniego, w szczególności w przypadku jednostek terytorialnych, które bardzo odstają od poziomu średniego (dla większego obszaru) pod względem poziomu danego wskaźnika. Wynika z tego,

iz problem błędu przybliżenia odgrywa w tym wypadku znaczącą rolę i dla niektórych jednostek terytorialnych nie powinno się go pomijać¹². Przykład algorytmu radzenia sobie z błędem przybliżenia znaleźć można w Rozdziale 4.4 niniejszego opracowania.

2.4.3 Charakterystyka formalna pod-dekompozycji wybranych składowych

Oprócz dekompozycji przedstawionej powyżej, dokonywanych jest szereg pod-dekompozycji wybranych różnic dotyczących:

- Średniej wydajności pracy,
- Udziału zatrudnionych w liczbie aktywnych zawodowo,
- Współczynnika aktywności zawodowej.

Celem tych pod-dekompozycji jest uchwycenie dodatkowych zależności i struktury pewnych zjawisk, o czym mowa poniżej.

2.4.3.1 Pod-dekompozycja różnic w wydajności pracy

Różnice w średniej wydajności pracy mogą wynikać z tego, iż w danym regionie dany sektor gospodarki ma większy udział niż średnio dla kraju, a w tym sektorze średnia krajowa wydajność pracy odbiega od średniej produktywności w całej gospodarce. Tę składową różnic w produktywności określa się mianem specjalizacji sektorowej. Pozostała składowa różnic w produktywności bierze się z tego, iż w poszczególnych sektorach średnia wydajność pracy w danym regionie jest wyższa (lub niższa) niż średnia krajowa wydajność pracy w tych sektorach. Jest to więc różnica w produktywności, która wynika z rzeczywistej przewagi danego regionu w intensywności wykorzystania zasobów zarówno ludzkich, jak i kapitału, infrastruktury, etc. Jeśli przez j oznaczymy dany sektor gospodarki, a przez i daną jednostkę, to powyższy podział uwzględnić możemy w dodatkowej dekompozycji różnic średniej wydajności pracy:

$$\left(\frac{WDB_i}{PP_i} - \frac{WDB}{P}\right) = \underbrace{\sum_j \frac{PP_{ij}}{PP_i} \left(\frac{WDB_{ij}}{PP_{ij}} - \frac{WDB_j}{P_j}\right)}_{\text{pozostałe różnice}} + \underbrace{\sum_j \frac{WDB_j}{P_j} \left(\frac{PP_{ij}}{PP_i} - \frac{P_j}{P}\right)}_{\text{specjalizacja sektorowa}}$$

Dodatnia specjalizacja sektorowa dla sektora j i regionu¹³ i oznacza, że udział pracujących w

¹² Patrz rozdział 4.4, gdzie znajduje się dyskusja formalna na temat błędu przybliżenia.

¹³ Słowo region ma w niniejszej pracy podwójne znaczenie. Po pierwsze oznacza jednostkę terytorialną NTS-1

tym sektorze w tym regionie jest większy niż udział pracujących w tym sektorze dla całego kraju. Ujemna specjalizacja sektorowa sektora j w regionie i oznacza natomiast mniejszy udział pracujących w tym sektorze w tym regionie niż średnio w całym kraju.

Rozróżnienie na specjalizacje sektorowe oraz pozostałe różnice średniej wydajności pracy jest kluczowe w identyfikacji niewykorzystanych zasobów. Jeśli na przykład średnia wydajność pracy w danym województwie jest niższa niż w kraju ze względu na fakt, iż to województwo jest wyspecjalizowane w rolnictwie, a w rolnictwie średnia wydajność pracy jest, przypuśćmy, niższa niż w innych sektorach, to nie oznacza to jeszcze, że zasoby tego województwa w zakresie rolnictwa są niewykorzystane. Jeśli natomiast średnia wydajność pracy w danym województwie w rolnictwie jest niższa niż średnio dla kraju w sektorze rolniczym, to wówczas oznacza to, że to województwo ma niewykorzystane zasoby w rolnictwie. Oczywiście te dwa zjawiska mogą się ze sobą łączyć, tj. dane województwo może wykazywać jednocześnie dodatnią specjalizację sektorową w rolnictwie oraz niższą niż dla kraju wydajność pracy rolnictwa. Wystąpienie takiego zjawiska może być nawet spotęgowane wówczas, gdy krańcowa produktywność pracy jest malejąca – jeśli w danym regionie w rolnictwie pracuje mało osób, to każda z nich ma wysoką wydajność, natomiast jeśli liczba osób pracujących wzrośnie, wydajność na osobę spada¹⁴. Niemniej występowanie tego zjawiska nie jest automatyczne. Ważne jest to, że w wyniku pod-dekompozycji średniej wydajności pracy, możliwe jest zidentyfikowanie niewykorzystanych zasobów gospodarki danego regionu poprzez oddzielenie specjalizacji sektorowej od różnic w wydajności pracy wynikających z różnego poziomu rozwoju technologii, infrastruktury oraz różnego poziomu jakości pracy.

według nomenklatury NUTS, a po drugie oznacza po prostu jednostkę terytorialną o niesprecyzowanej definicji. Jeśli nie jest wyraźnie podkreślone, że chodzi o pierwsze znaczenie tego słowa, region należy rozumieć według drugiej definicji.

¹⁴ Taki proces można rozumieć w dwojaki sposób:

- 1) Na początku w danym sektorze pracują tylko ci najbardziej wydajni, potem dopiero dochodzą ci mniej wydajni. Jeśli jest kryzys, to znowu pierwsi bankrutują ci mniej wydajni, a pozostają ci bardziej wydajni.
- 2) Malejące korzyści skali – im więcej osób tym większy problem organizacji pracy, nowi pracownicy zajmują się już mniej rentownymi zajęciami, ponieważ te bardziej rentowne są już obsadzone przez poprzednich pracowników, etc.

2.4.3.2 Pod-dekompozycja różnic w udziale pracujących w aktywnych zawodowo

To samo dotyczy udziału zatrudnionych w liczbie aktywnych zawodowo. Różnice w poziomie tego udziału w stosunku do średniej krajowej można z kolei rozbić na różnice w profilu wykształcenia pracujących w danym województwie w stosunku do średniego profilu wykształcenia pracujących w kraju oraz różnice w rzeczywistym udziale liczby osób pracujących do liczby osób aktywnych zawodowo, odzwierciedlające rzeczywistą sytuację na rynku pracy. Jeśli przez j oznaczymy odpowiednią grupę wykształcenia a przez i region, otrzymamy następującą dekompozycję różnic w udziale zatrudnionych w liczbie aktywnych zawodowo:

$$\left(\frac{PZ_i}{AZ_i} - \frac{P}{A}\right) = \underbrace{\sum_j \frac{AZ_{ij}}{AZ_i} \left(\frac{PZ_{ij}}{AZ_{ij}} - \frac{P_j}{A_j}\right)}_{\text{efektywność rynku pracy}} + \underbrace{\sum_j \frac{P_j}{A_j} \left(\frac{AZ_{ij}}{AZ_i} - \frac{A_j}{A}\right)}_{\text{profil wykształcenia}}$$

Podobnie jak w przypadku specjalizacji sektorowej, dodatni profil wykształcenia dla grupy wykształcenia j w regionie i oznacza, że w tym regionie jest relatywnie więcej osób z grupy wykształcenia j niż średnio w kraju. Składnik „efektywność rynku pracy” dla regionu i dla grupy wykształcenia j oznacza z kolei, że w tym regionie i dla tej grupy wykształcenia udział liczby osób pracujących do liczby osób aktywnych zawodowo różni się od tego udziału dla tej grupy wykształcenia dla całego kraju. Rozbicie różnic w poziomie udziału zatrudnionych w liczbie aktywnych zawodowo ma podobny cel do rozbicia różnic w średniej wydajności pracy. Na przykład, jeśli w danym województwie jest więcej osób z wyższym wykształceniem niż średnio w kraju, a udział zatrudnionych w liczbie aktywnych zawodowo dla osób z wyższym wykształceniem jest ogólnie wyższy niż dla osób należących do innych grup wykształcenia. Znajdzie to odbicie w dodatnim składniku profilu wykształcenia dla tej grupy wykształcenia, jednak nie świadczy to o tym, że rynek pracy w tym województwie funkcjonuje lepiej niż średnio dla kraju. Świadczy to tylko o większej liczbie osób z wyższym wykształceniem, dla których udział zatrudnionych w liczbie aktywnych zawodowo okazuje się być ogólnie wyższy. Aby zidentyfikować pole do poprawy rynku pracy należy raczej zwrócić uwagę na pozostały składnik różnic w udziale zatrudnionych w aktywnych zawodowo, tj. „efektywność rynku pracy”. Jeśli ten składnik przyjmuje wartość dodatnią dla danego województwa i dla danej grupy wykształcenia, oznacza to, że dla tej grupy, udział zatrudnionych w liczbie aktywnych zawodowo jest wyższy w danym województwie niż średnio w całym kraju. Ten składnik świadczy najprawdopodobniej o lepszym funkcjonowaniu rynku pracy w danym

województwie.

2.4.3.3 Pod-dekompozycja różnic we współczynniku aktywności zawodowej

Podobnie jak w przypadku różnic w średniej wydajności pracy, jak i różnic w udziale zatrudnionych w liczbie aktywnych zawodowo, można dokonać pod-dekompozycji różnic w poziomie współczynnika aktywności zawodowej. Tym razem jednak rozkład opiera się nie na sektorach gospodarki ani na grupach wykształcenia, ale na grupach wiekowych ludności. W ten sposób możliwe jest oddzielenie różnic w poziomie współczynnika aktywności zawodowej wynikających ze struktury wiekowej oraz tych wynikających z różnego poziomu udziału w rynku pracy. Jeśli przez j oznaczymy grupę wiekową, a przez i województwo, to otrzymamy następującą dekompozycję:

$$\left(\frac{AZ_i}{NB_i^{15+}} - \frac{A}{NB^{15+}} \right) = \underbrace{\sum_j \frac{NB_{ij}^{15+}}{NB_i^{15+}} \left(\frac{AZ_{ij}}{NB_{ij}^{15+}} - \frac{A_j}{NB_j^{15+}} \right)}_{\text{poziom uczestnictwa w rynku pracy}} + \underbrace{\sum_j \frac{A_j}{NB_j^{15+}} \left(\frac{NB_{ij}^{15+}}{NB_i^{15+}} - \frac{NB_j^{15+}}{NB^{15+}} \right)}_{\text{profil struktury wiekowej}}$$

Podobnie jak wcześniej, dodatni profil struktury wiekowej dla regionu i dla grupy wiekowej j oznacza, że w tym regionie jest relatywnie więcej osób z grupy wiekowej j niż średnio w kraju. Poziom uczestnictwa w rynku pracy wskazuje na różnice dla regionu i i grupy wiekowej j w udziale liczby osób aktywnych zawodowo w stosunku do liczby osób od 15 roku życia w stosunku do średniej dla kraju w tej grupie wiekowej.

Reasumując, idea tych trzech pod-dekompozycji polega na tym, aby porównywać regiony w podobnych warunkach – ich produktywność, efektywność rynku pracy oraz poziom uczestnictwa w rynku pracy nie zależy wyłącznie od ich bezpośrednich starań, ale również obiektywnych, trwalszych uwarunkowań tj. od specjalizacji sektorowej, profilu wykształcenia oraz profilu struktury wiekowej.

2.5 Dekompozycja według współczynnika Theila – podstawy metodologiczne

Współczynnik Theila należy do miar opartych na entropii i służy do mierzenia nierówności (zróżnicowania) poziomu dochodu, bądź innych mierzalnych zmiennych przekrojowych. Załóżmy, że interesuje nas pomiar zróżnicowania poziomu PKB *per capita* w 66 podregionach (jednostka terytorialna NTS3) Polski. Zdefiniujmy następujące wielkości: N_i populacja w podregionie i , N populacja Polski, PKB_i poziom PKB w podregionie i , PKB poziom PKB w Polsce. A zatem $\frac{PKB_i}{PKB}$, $\frac{N_i}{N}$ oznaczają kolejno udział PKB podregionu i w PKB ogółem dla Polski oraz udział populacji w podregionie i w populacji Polski ogółem. Współczynnik Theila

definiowany jest wówczas następująco:

$$T = \sum_{i=1}^{66} \frac{PKB_i}{PKB} \ln \frac{\frac{PKB_i}{PKB}}{\frac{N_i}{N}} \quad (****)$$

Współczynnik ten przyjmuje wartość 0 w hipotetycznej sytuacji, gdy poziom udział PKB każdego podregionu w PKB dla Polski byłby dokładnie równy udziałowi populacji tego podregionu w populacji Polski – oznaczałoby to brak zróżnicowania pod względem PKB *per capita* w podregionach Polski. Z kolei w innej hipotetycznej sytuacji, kiedy jeden podregion wytwarzałby prawie cały PKB Polski (maksymalne możliwe zróżnicowanie PKB *per capita*) a inne podregiony wytwarzałyby bardzo niewielkie ilości PKB, wartość współczynnika Theila dążyłaby do maksymalnej wartości $\ln 66$. A zatem współczynnik Theila przyjmuje wartości od najmniejszej (0) do największej ($\ln L_j$), gdzie L_j to liczba jednostek, których zróżnicowanie badamy (w naszym przykładzie są to podregiony). Najmniejsza wartość oznacza najmniejsze możliwe zróżnicowanie, a największa oznacza największe zróżnicowanie. Wartości współczynnika Theila można znormalizować do przedziału $[0,1]$, poprzez podzielenie wartości wyrażenia (****) powyżej przez wartość najwyższą, tj. $\ln 66$. W niniejszym opracowaniu zdecydowano się jednak na utrzymanie nieznormalizowanej wersji.

Zauważmy, że współczynnik Theila stanowi sumę ważoną logarytmu stosunku PKB *per capita* w danym podregionie do PKB *per capita* w Polsce z wagami równymi udziałom PKB w danym podregionie w PKB dla Polski. A zatem w podregionach biedniejszych niż średnia dla Polski wartość logarytmu stosunku PKB *per capita* danego w danym podregionie do PKB *per capita* dla Polski jest ujemna a w podregionach bogatszych niż Polska wartość ta jest dodatnia.

Podstawowa wersja współczynnika Theila, zwana również współczynnikiem Theila z udziałami PKB jest bardziej wrażliwa na zmiany w bogatych podregionach a mniej na zmiany w biednych podregionach. Istnieje wersja druga współczynnika Theila, zwana współczynnikiem Theila z udziałami populacji, która jest bardziej wrażliwa na zmiany w biedniejszych podregionach. Można ją zapisać następująco:

$$T' = - \sum_{i=1}^{66} \frac{N_i}{N} \ln \frac{\frac{PKB_i}{PKB}}{\frac{N_i}{N}} \quad (*****)$$

Zarówno współczynnik Theila z udziałami PKB T jak i druga wersja współczynnika Theila z udziałami populacji T' może być rozszerzona o dwa elementy:

- a) Dekompozycja współczynnika według podgrup;
- b) Dekompozycja współczynnika dla PKB *per capita* na składowe.

Zacznijmy od pierwszego elementu. Chcemy zbadać zróżnicowanie poziomu PKB *per capita* w podregionach Polski, ale to zróżnicowanie chcemy rozbić na dwie składowe:

- Zróżnicowanie podregionów względem makroregionów (nazywane zróżnicowaniem „wewnątrz”);
- Zróżnicowanie makroregionów względem Polski (nazywane zróżnicowaniem „pomiędzy”).

Aby zobrazować metodę dekompozycji współczynnika Theila według podgrup czyli w naszym przykładzie makroregionów, wprowadźmy dodatkowe oznaczenia: dana zmienna z subskrypcją k oznaczać będzie jeden z pięciu makroregionów Polski (Polska Wschodnia, Zachodnia, Południowa, Centralna oraz Polska Północna) natomiast n_k oznaczać będzie liczbę podregionów w makroregionie k . A zatem 66 podregionów (oznaczonych jako i) dzieli się na 5 makroregionów k , z których każdy ma n_k podregionów. Poniższe wyprowadzenie ilustruje metodę dekompozycji na podgrupy:

$$\begin{aligned}
 T &= \sum_{i=1}^{66} \frac{PKB_i}{PKB} \ln \frac{PKB_i}{\frac{N_i}{N}} \\
 &= \sum_{k=1}^5 \sum_{i=1}^{n_k} \frac{PKB_k}{PKB} \frac{PKB_i}{PKB_k} \ln \frac{PKB_k}{PKB} \frac{PKB_i}{PKB_k} \\
 &= \sum_{k=1}^5 \frac{PKB_k}{PKB} \left[\sum_{i=1}^{n_k} \frac{PKB_i}{PKB_k} \left(\ln \frac{PKB_i}{PKB_k} + \ln \frac{PKB_k}{PKB} \right) \right] \\
 &= \underbrace{\sum_{k=1}^5 \frac{PKB_k}{PKB} T_k}_{\text{zróżnicowanie wewnątrz}} + \underbrace{\sum_{k=1}^5 \frac{PKB_k}{PKB} \ln \frac{PKB_k}{PKB}}_{\text{zróżnicowanie pomiędzy}},
 \end{aligned}$$

gdzie T_k to współczynniki Theila dla podregionów względem makroregionów.

Współczynnik Theila czyli zróżnicowanie podregionów Polski względem PKB *per capita* rozdzielony został na dwie składowe:

- Zróżnicowanie podregionów względem makroregionów;

- Zróżnicowanie makroregionów względem Polski.

Dzięki takiej dekompozycji wiadomo jak bardzo makroregiony są homogeniczne względem siebie i (na poziomie podregionów) wewnątrz siebie.

Podobnie dokonuje się dekompozycji wersji współczynnika Theila z udziałami populacji:

$$\begin{aligned}
 T' &= - \sum_{i=1}^{66} \frac{N_i}{N} \ln \frac{\frac{PKB_i}{PKB}}{\frac{N_i}{N}} \\
 &= - \sum_{k=1}^5 \sum_{i=1}^{n_k} \frac{N_k}{N} \frac{N_i}{N_k} \ln \frac{\frac{PKB_k}{PKB} \frac{PKB_i}{PKB_k}}{\frac{N_k}{N} \frac{N_i}{N_k}} \\
 &= - \sum_{k=1}^5 \frac{N_k}{N} \left[\sum_{i=1}^{n_k} \frac{N_i}{N_k} \left(\ln \frac{\frac{PKB_i}{PKB_k}}{\frac{N_i}{N_k}} + \ln \frac{\frac{PKB_k}{PKB}}{\frac{N_k}{N}} \right) \right] \\
 &= \underbrace{\sum_{k=1}^5 \frac{N_k}{N} T'_k}_{\text{zróżnicowanie wewnątrz}} - \underbrace{\sum_{k=1}^5 \frac{N_k}{N} \ln \frac{\frac{PKB_k}{PKB}}{\frac{N_k}{N}}}_{\text{zróżnicowanie pomiędzy}},
 \end{aligned}$$

gdzie T'_k to współczynniki Theila drugiego rodzaju dla podregionów względem makroregionów.

Teraz przejdziemy do omówienia drugiego rozszerzenia współczynnika Theila, który polega na jego dekompozycji na poszczególne składowe ekonomiczne. Wiadomo, że PKB *per capita* dla podregionów rozbić możemy na składowe multiplikatywne według następującego wzoru:

$$\frac{PKB_i}{N_i} = \frac{PKB_i}{PP_i} \frac{PP_i}{N_i^{15+}} \frac{N_i^{15+}}{N_i}$$

Gdzie P_i to liczba pracujących w podregionie i , a

Możemy rozbić współczynnik Theila z udziałami PKB dla PKB *per capita* na trzy składowe:

$$\begin{aligned}
 T &= \sum_{i=1}^{66} \frac{PKB_i}{PKB} \ln \frac{\frac{PKB_i}{PKB}}{\frac{N_i}{N}} \\
 &= \sum_{i=1}^{66} \frac{PKB_i}{PKB} \left(\ln \frac{\frac{PKB_i}{PKB}}{\frac{PP_i}{PP}} + \ln \frac{\frac{PP_i}{PP}}{\frac{N_i^{15+}}{N^{15+}}} + \ln \frac{\frac{N_i^{15+}}{N^{15+}}}{\frac{N_i}{N}} \right)
 \end{aligned}$$

$$= \underbrace{\sum_{i=1}^{66} \frac{PKB_i}{PKB} \ln \frac{PKB_i}{PP_i}}_{\text{zróż. wydajności}} + \underbrace{\sum_{i=1}^{66} \frac{PKB_i}{PKB} \ln \frac{PP_i}{N_i^{15+}}}_{\text{zróż. wsk. zatrudnienia}} + \underbrace{\sum_{i=1}^{66} \frac{PKB_i}{PKB} \ln \frac{N_i^{15+}}{N}}_{\text{zróż. wsk. str. wiekowej}}$$

A zatem całkowite zróżnicowanie PKB *per capita* w podregionach względem Polski można rozbić na zróżnicowanie dotyczące wydajności pracy, zróżnicowanie dotyczące wskaźnika zatrudnienia oraz zróżnicowanie dotyczące wskaźnika struktury wiekowej.

Podobnie postępujemy ze współczynnikiem Theila z udziałami populacji:

$$\begin{aligned} T' &= - \sum_{i=1}^{66} \frac{N_i}{N} \ln \frac{PKB_i}{N_i} \\ &= - \sum_{i=1}^{66} \frac{N_i}{N} \left(\ln \frac{PKB_i}{PP_i} + \ln \frac{PP_i}{N_i^{15+}} + \ln \frac{N_i^{15+}}{N} \right) \\ &= \underbrace{- \sum_{i=1}^{66} \frac{N_i}{N} \ln \frac{PKB_i}{PP_i}}_{\text{zróż. wydajności}} - \underbrace{\sum_{i=1}^{66} \frac{N_i}{N} \ln \frac{PP_i}{N_i^{15+}}}_{\text{zróż. wsk. zatrudnienia}} - \underbrace{\sum_{i=1}^{66} \frac{N_i}{N} \ln \frac{N_i^{15+}}{N}}_{\text{zróż. wsk. str. wiekowej}} \end{aligned}$$

Można połączyć dekompozycję współczynnika Theila według podgrup oraz dekompozycję współczynnika Theila dla PKB *per capita* na składowe w jedno. Wówczas otrzymamy informację jak rozbić zróżnicowanie PKB *per capita* podregionów Polski na zróżnicowanie poszczególnych składowych (wydajności, wskaźnika zatrudnienia i wskaźnika struktury wiekowej) i jednocześnie dla każdego z tych składowych na zróżnicowanie wewnątrz makroregionów i zróżnicowanie pomiędzy makroregionami.

Metoda ma również tę zaletę, że dla każdego podregionu (makroregionu) i dla każdej składowej możemy sprawdzić, czy ten podregion przyczynia się do zmniejszenia czy też zwiększenia zróżnicowania pod względem tej składowej.

2.6 Wyzwania praktycznej implementacji dekompozycji

Niniejszy rozdział dotyczy praktycznej implementacji metod dekompozycji na podstawie dostępnych danych GUS. Związek między teorią a praktyką jest w tym wypadku bardzo istotny, ponieważ dane, które są dostępne nie zawsze w pełni przystają do teoretycznych definicji z uwagi na rozmaite problemy związane z ich pozyskiwaniem i przetwarzaniem.

Należy pamiętać o tych ograniczeniach, aby później nadać właściwą interpretację wynikom. Dokładne wyjaśnienia metodologiczne dostępne są w odpowiednich publikacjach, natomiast pewne ogólne definicje i pojęcia odnoszące się do definicji podanych poniżej prezentowane są w słowniku pojęć w aneksie do niniejszego opracowania. Problemy związane z zastosowaniem istniejących i dostępnych danych do metody dekompozycji można podzielić na trzy grupy:

- a) Problem łączenia danych pochodzących w różnych źródłach i obliczanych według różnych metodologii (np. we współczynniku dojazdów do pracy wykorzystana jest zarówno liczba pracujących w miejscu pracy (formularze administracyjne) oraz liczba pracujących w miejscu zamieszkania (BAEL))
- b) Problem niedokładnej mierzalności pewnych wielkości (np. wskaźnik zatrudnienia powinien informować o udziale osób pracujących w ogóle populacji zdolnych do pracy, jednak ogólna liczba osób zdolnych do pracy jest wielkością niemierzalną, dlatego w zamian stosuje się liczbę osób w wieku produkcyjnym bądź liczbę osób od 15 roku życia)
- c) Problem niedokładności dostępnych danych.

Niniejszy rozdział ma na celu zidentyfikowanie problemów implementacji metod dekompozycji w praktyce na podstawie danych GUS oraz zarysowanie sposobów radzenia sobie z nimi. Poprzez odpowiednie modyfikacje bądź dodatkowe zabiegi metodologiczne, części spośród zidentyfikowanych problemów daje się zapobiec. Części problemów w pełni nie da się rozwiązać. Chodzi o to, aby wskazać minimalny zbiór niedoskonałości, które trzeba zaakceptować, jeśli chce się stosować metody dekompozycji.

2.6.1 Dane pochodzące z różnych metodologii

Zasadniczy problem dotyczący spójności metodologii danych w metodach dekompozycji jest następujący. W dekompozycji stosuje się zasadniczo dwa rodzaje danych, których metodologia znacznie się między sobą różni:

- Dane pozyskiwane z rachunków narodowych bądź regionalnych według formularzy sprawozdawczych oraz spójne z nimi dane dotyczące liczby ludności i dane pochodzące z badań przedsiębiorstw (*PKB, WDB, PP, N, N¹⁵⁺, N¹⁵⁻⁶⁴*)

- Dane pochodzące z Badania Aktywności Ekonomicznej Ludności (BAEL), które są badaniami reprezentacyjnymi (na podstawie próby gospodarstw domowych, wyniki ekstrapoluje się na całość populacji) (PZ, AZ oraz NB^{15+}, NB^{15-64})

Specyfika metody dekompozycji polega na tym, że trzeba przejść z danych z jednej grupy na dane z drugiej grupy za pomocą wskaźnika/współczynnika ilorazowego. Moment przejścia w dekompozycjach przedstawionych w podrozdziałach 2.3.2 oraz 2.4.2 następuje w dwóch miejscach:

- Wskaźnik dojazdów do pracy $\frac{PP_i}{PZ_i}$, gdzie następuje „przejście” z metodologii zgodnej z Rachunkami Narodowymi (PP_i) – publikuje się WDB na 1 pracującego, czyli $\frac{WDB_i}{PP_i}$ – na metodologię BAEL (PZ_i);
- Wskaźnik $\frac{NB_i^{15+}}{N_i^{15+}}$, gdzie następuje „przejście” z metodologii BAEL na metodologię badań demograficznych (zgodną z Rachunkami Narodowymi) – publikuje się przecież WDB i PKB *per capita*, czyli $\frac{WDB_i}{N_i}, \frac{PKB_i}{N_i}$, a zmienna N_i jest zgodna z N_i^{15+} .

2.6.1.1 Błąd różnych prób we współczynniku dojazdów do pracy

We współczynniku dojazdów do pracy netto oprócz liczby pracujących w miejscu pracy (PP) wykorzystuje się liczbę pracujących w miejscu zamieszkania (PZ), która to liczba pozyskiwana jest na podstawie Badania Aktywności Ekonomicznej Ludności (BAEL) prowadzonego przez GUS. Tym razem jest to średnioroczna liczba pracujących, jednakże z uwagi na odmienną metodologię, różnice w zakresie zliczanych osób i fakt, że BAEL jest badaniem reprezentatywnym, liczba pracujących w miejscu pracy i liczba pracujących w miejscu zamieszkania może się między sobą znacznie różnić i nie wszystkie z tych różnic wynikają li tylko z odmiennego zakresu i okresu zliczania. Występuje tutaj problem różnych prób (ang. *sampling error*), tj. różnych metodologii. Błąd ten dla pracujących w miejscu pracy oraz pracujących w miejscu zamieszkania jest obserwowany na przykład na poziomie krajowym – liczba pracujących w miejscu pracy i w miejscu zamieszkania powinna być teoretycznie dla całego kraju taka sama, jeśli założy się, że poziom dojazdów do pracy za granicą jest pomijalny, jednakże występują nieraz znaczne różnice wynikające z różnych metodologii – liczba pracujących w miejscu pracy dla całego kraju według rachunków regionalnych nie będzie równa liczbie pracujących w miejscu zamieszkania dla całego kraju na podstawie BAEL. Różnica między nimi stanowi błąd różnych próbek (ang. *sampling error*). Możliwym

rozwiązaniem tego problemu jest zastosowanie udziałów pracujących zamiast poziomów, tj. przy obliczaniu współczynnika dojazdów do pracy zamiast samej liczby pracujących w miejscu pracy w danym regionie, weźmiemy udział tychże w ogóle pracujących w kraju mierzonych według rachunków regionalnych, a zamiast samej liczby pracujących w miejscu zamieszkania w danym regionie weźmiemy udział tychże w ogólnej liczbie pracujących w kraju na podstawie BAEL, wówczas błąd próbki zostanie skorygowany. Można zilustrować to formalnie – jeśli oznaczymy przez $PP_i(\text{Rach. reg.})$, $PP(\text{Rach. reg.})$ liczbę pracujących w miejscu pracy (według Rachunków regionalnych) kolejno w jednostce i oraz w całym kraju a przez $PZ_i(\text{BAEL})$, $PZ(\text{BAEL})$ liczbę pracujących w miejscu zamieszkania (według BAEL) kolejno w jednostce i oraz dla całego kraju, to współczynnik dojazdów do pracy netto powinniśmy policzyć według następującego wzoru:

$$DN_i = \frac{\frac{PP_i(\text{Rach. reg.})}{PP(\text{Rach. reg.})}}{\frac{PZ_i(\text{BAEL})}{PZ(\text{BAEL})}}$$

2.6.1.2 Przeciętna liczba pracujących

Liczba pracujących według faktycznego miejsca pracy (PP) pozyskiwana jest przez GUS na podstawie sprawozdań dostarczanych przez jednostki zobowiązane do ich terminowego nadsyłania – liczba ta jest podawana według stanu na 31 XII każdego roku. Jednak PKB bądź WDB to produkt wytworzony w ciągu całego roku, należałoby zatem ten produkt podzielić przez średnioroczną liczbę pracujących. Możliwym rozwiązaniem tego problemu jest zastosowanie średniej arytmetycznej z dwóch kolejnych okresów. Ten sposób nie uwzględnia jednak efektu sezonowego – liczba pracujących na koniec grudnia może się różnić systematycznie i zgodnie z określonym wzorcem od liczby pracujących według stanu na inny miesiąc roku. Drugim rozwiązaniem, które można zastosować to wykorzystanie średniorocznej liczby pracujących pozyskanej w wyniku podzielenia wartości WDB (PKB) przez WDB (PKB) na jednego pracującego. Taka wartość musi być w ten prosty sposób wyliczona ponieważ nie jest publikowana osobno przez GUS. Jest to liczba pracujących policzona na podstawie przeciętnej liczby zatrudnionych w roku, do których dołączone są pozostałe grupy pracujących – patrz aneks do niniejszego podrozdziału. Na potrzeby pod-dekompozycji różnic w poziomie wydajności pracy potrzebne są również dane dotyczące pracujących w poszczególnych sektorach zgodne z danymi dotyczącymi sektorowego WDB:

- **W układzie NTS 2** liczba pracujących obliczona z WDB na jednego pracującego jest dostępna według 7 grup sekcji PKD, zatem przy nieznaczonej utracie szczegółowości (7 grup sekcji zamiast 21 sekcji PKD) można wykorzystać przeciętną liczbę pracujących zamiast liczbę pracujących według stanu na 31 XII danego roku.
- **W układzie NTS 3** przeciętną liczbę pracujących z podziałem na 7 grup NACE zgodnym z podziałem WDB podaje EUROSTAT.

Jeśli nie jest możliwe, aby uzyskać dobrej jakości przeciętną liczbę pracujących w roku, bądź interesuje nas większa szczegółowość sektorowa (według sekcji PKD), wówczas zmuszeni jesteśmy do wykorzystania liczby pracujących według stanu na 31 XII. Oznacza to wówczas, że średnia wydajność pracy liczona jest z pewnym błędem, wynikającym z faktu, iż PKB w cenach bieżących tworzony jest w ciągu całego roku rozliczeniowego, natomiast liczba pracujących dotyczy stanu na ostatni dzień roku kalendarzowego. Dynamika (bądź różnice w poziomie) takiej zmiennej to w przybliżeniu suma dynamiki (różnicy w poziomie) zmiennej właściwej (tj. średniej wydajności pracy liczonej przy uwzględnieniu średniorocznej liczby pracujących) oraz dynamika (różnica w poziomie) ilorazu średniorocznej liczby pracujących oraz liczby pracujących z dnia 31 XII danego roku, który to wskaźnik informować może o zmianach w liczbie pracujących pod koniec roku w stosunku do średniej dla całego roku. Aby nadać właściwą interpretację średniej wydajności pracy liczonej na podstawie liczby pracujących według stanu z 31 XII danego roku, należy o tym pamiętać i uwzględnić to w proponowanych wnioskach.

Aneks: Przeciętna roczna liczba pracujących w miejscu pracy

Pracujący składają się z:

1. Zatrudnieni na podstawie stosunku pracy
2. Pracodawcy i pracujący na własny rachunek
3. Osoby wykonujące pracę nakładczą
4. Agenci

Ponieważ liczba pracujących w gospodarce narodowej wyraża stan na 31 XII każdego roku, w celu obliczenia średniorocznej wydajności pracy, potrzeba innego wskaźnika liczby pracujących, który oddawałby przeciętną liczbę pracujących w ciągu roku. Taka wielkość obliczana jest następująco. Startujemy z przeciętnego zatrudnienia w ciągu roku i oszacowujemy dodatkowo średnią ilość pracujących, którzy nie są zatrudnieni na podstawie

stosunku pracy. Następnie oblicza się średnią liczbę pracujących jako średnią ważoną średniego zatrudnienia i oszacowanych średnich dla pozostałych kategorii. Daje to pozycję, która widnieje w roczniku statystycznym pod nazwą „Pracujący (przeciętnie w roku) według statusu zatrudnienia”. Jednak pozycja ta nie uwzględnia jeszcze jednej pozycji pracujących – mianowicie osób pracujących w jednostkach budżetowych prowadzących działalność w zakresie obrony narodowej i bezpieczeństwa publicznego, tj. ok. 292 tysiące osób. Pozostaje dodać tę wielkość i w ten sposób otrzymujemy średnią liczbę pracujących, którą można użyć do obliczenia wartości dodanej brutto na jednego pracującego, czyli średniorocznej wydajności pracy.

2.6.2 Problemy niedokładnej mierzalności wybranych wielkości

Pomiar wielkości opisujących zjawiska społeczno-gospodarcze jest utrudniony ograniczonym dostępem do informacji oraz technicznymi uwarunkowaniami pozyskiwania informacji. Stąd w wielu przypadkach przyjmuje się konwencję, według której świadomie dokonywany jest pomiar przybliżony danej wielkości, bądź też mierzona jest wielkość, która jest ściśle skorelowana z wielkością, którą chcemy mierzyć. Przykładem pierwszego rodzaju postępowania jest umowne przyjęcie wieku produkcyjnego w przedziale 18-64 lata, przy założeniu, że większość osób w tym czasie może bądź powinna pracować. Przykładem drugiego rodzaju postępowania jest przyjęcie wyniku testu IQ za wielkość w przybliżeniu opisującą zdolności intelektualne danej osoby.

Problemy niedokładnej mierzalności pewnych wielkości, które są specyficzne dla metod dekompozycji prezentowanych w niniejszym opracowaniu są omówione poniżej.

2.6.2.1 Udział zatrudnionych w liczbie aktywnych zawodowo

Liczba aktywnych zawodowo pochodzi ze wspomnianego powyżej badania BAEL i obejmuje osoby pracujące zawodowo oraz zarejestrowanych bezrobotnych. Wskaźniki zatrudnienia używane przez GUS oraz Eurostat zdefiniowane są różnie, bowiem liczony jest udział pracujących w grupie wszystkich osób z przedziału wiekowego, odpowiednio od 15 lat w przypadku GUS i w wieku 15-64 lat według EU Labour Force Survey (publikowany przez Eurostat). W przypadku dekompozycji prezentowanych w niniejszym opracowaniu wykorzystano udział zatrudnionych w liczbie aktywnych zawodowo zamiast wskaźnika zatrudnienia zdefiniowanego przez GUS lub przez EU Labour Force Survey, gdyż pozwala to na oddzielenie dwóch odrębnych zjawisk na rynku pracy – dynamiki zatrudnienia wśród osób

aktywnych zawodowo¹⁵ oraz dynamiki osób aktywnych zawodowo wśród osób umownie zdolnych do pracy ze względu na wiek¹⁶. Jeśli taka definicja wskaźnika zatrudnienia jest nie do zaakceptowania, wówczas można te dwa wskaźniki scalić w jeden (wskaźnik zatrudnienia według metodologii GUS lub EUROSTAT):

$$\frac{PZ_i}{N_i^{15+}} = \frac{PZ_i}{AZ_i} \times \frac{AZ_i}{N_i^{15+}}, \quad \frac{PZ_i}{N_i^{15-64}} = \frac{PZ_i}{AZ_i} \times \frac{AZ_i}{N_i^{15-64}}$$

Liczba ludności w wieku produkcyjnym bądź ludności w wieku od 15 roku życia powinna być pozyskana z badania BAEL w celu zachowania spójności metodologicznej. Wówczas „przejście” z powrotem na metodologię rachunków narodowych następowałoby przy wskaźniku struktury wiekowej.

2.6.2.2 Współczynnik aktywności zawodowej

Współczynnik aktywności zawodowej liczyć można na dwa możliwe sposoby¹⁷ jako stosunek:

- Liczby aktywnych zawodowo do liczby populacji od 15 roku życia: $\frac{AZ_i}{N_i^{15+}}$
- Liczby aktywnych zawodowo do liczby populacji ludzi w wieku produkcyjnym:

$$\frac{AZ_i}{N_i^{15-64}}$$

Wiek produkcyjny zdefiniowany jest według Eurostatu jako 15-64 lat. Obie definicje umożliwiają w pewnym sensie oddzielenie wpływu zjawisk zachodzących na rynku pracy (współczynnik aktywności zawodowej informuje, aczkolwiek niebezpośrednio, o tym, jak dużo osób jest zniechęconych, tj. ani nie pracuje ani aktywnie nie poszukuje pracy – zarejestrowani bezrobotni) od zjawisk czysto demograficznych, które determinują popyt na pracę na rynku pracy – udział osób, które mogą pracować ze względu na wiek w ogólnej liczbie ludności. Jednak pierwsza definicja współczynnika aktywności zawodowej przedstawiona powyżej umożliwia uchwycenie tylko tych zjawisk demograficznych, które mają wpływ na proporcje liczby ludzi w wieku od 15 lat do tych poniżej tego wieku. Jednak

¹⁵ Symetrycznie, można mówić o dynamice stopy bezrobocia, bowiem stopa bezrobocia jest liczbą bezrobotnych zarejestrowanych w stosunku do liczby osób aktywnych zawodowo, natomiast ludność aktywna zawodowo to osoby pracujące zawodowo oraz bezrobotni zarejestrowani.

¹⁶ Patrz definicja współczynnika aktywności zawodowej poniżej.

¹⁷ Te dwa sposoby odpowiadają odpowiednio tradycyjnej metodologii GUS oraz metodologii EUROSTATu.

wiadomo, że nie tylko osoby zbyt młode nie są zdolne do pracy z racji wieku, ale również osoby w odpowiednio podeszłym wieku bądź nie są zdolne do pracy, bądź osiągnęły już wiek emerytalny i nie muszą ani nie chcą już pracować. Druga definicja współczynnika aktywności zawodowej uwzględnia wpływ czynników demograficznych zarówno w grupie osób, które nie osiągnęły jeszcze odpowiedniego wieku (umownie 15 lat), jak i w grupie osób, które przekroczyły już wiek 64 lat.

2.6.2.3 Wskaźnik struktury wiekowej

Uwzględniając podwójną definicję współczynnika aktywności zawodowej, również w dwojaki sposób zdefiniowany jest współczynnik struktury wiekowej, jako udział:

- Populacji od 15 roku życia w całkowitej populacji: $\frac{N_i^{15+}}{N_i}$
- Populacji w wieku produkcyjnym w całkowitej populacji: $\frac{N_i^{15-64}}{N_i}$

Współczynnik struktury wiekowej, o czym była mowa powyżej, pomyślany jest jako miara wpływu demograficznego na wysokość PKB (bądź WDB) *per capita*. Przykładowo, im więcej osób w wieku produkcyjnym w stosunku do całkowitej liczby ludności, tym więcej osób teoretycznie zdolnych do pracy i wytwarzania PKB lub WDB.

2.6.2.4 Efekt kalendarza

Średnią liczbę przepracowanych godzin w ciągu roku uzyskujemy jako iloczyn liczby pracujących i średniego tygodniowego czasu pracy pomnożony przez 52 tygodnie. Tak więc umownie przyjmujemy $52 \times 7 = 364$ dni w roku. Takie ujęcie nie uwzględnia efektu kalendarza. Zakłada się, że w każdym roku liczba dni przepracowanych jest taka sama. Możliwym sposobem na zaradzenie trudności, jest wykorzystanie średniej liczby faktycznie przepracowanych godzin w ciągu roku na jednego pracownika. Wówczas uwzględniony jest efekt kalendarza, jednak utracona zostaje informacja o dynamice średniego tygodniowego czasu pracy. Jeśli ta dynamika powinna być uwzględniona, wówczas można rozbić faktycznie przepracowane godziny w ciągu roku na jednego pracownika jako iloczyn średniego tygodniowego czasu pracy oraz liczbę tygodni w roku faktycznie przepracowaną przez przeciętnego pracownika (pozyskaną jako iloraz pozostałych dwóch wielkości). Dane dotyczące średniego tygodniowego czasu pracy dostępne są w EUROSTAT, który otrzymuje je od GUS.

2.6.2.5 Ceny bieżące a ceny stałe

W niniejszym Raporcie tam gdzie to było możliwe stosowano wartości PKB i WDB w cenach stałych. W sytuacjach, gdy jest to niemożliwe i trzeba zastosować wartości w cenach bieżących, wówczas należy pamiętać o właściwej interpretacji wyników. Na przykład, jeżeli wydajność pracy mierzymy jako osobogodzinę pracy osoby pracującej w cenach bieżących, to na wzrost tej wielkości może mieć wpływ zarówno rzeczywisty wzrost wartości PKB (lub WDB), jak i wzrost poziomu cen w gospodarce. W szczególności, dynamika średniej wydajności pracy przy użyciu PKB (WDB) w cenach bieżących jest w przybliżeniu¹⁸ równa sumie dynamiki średniej wydajności pracy przy użyciu PKB (WDB) w cenach stałych i inflacji. Powyższy problem dotyczy dekompozycji dynamicznej oraz porównywania wyników dekompozycji różnic regionalnych w czasie, nie dotyczy jednak statyczno-porównawczego ujęcia dekompozycji różnic regionalnych, ponieważ ujęcie owo nie zawiera porównań międzyokresowych.

2.6.3 Problem niedokładności pomiaru dostępnych zmiennych

Problem niedokładności dostępnych danych wynika z technicznych uwarunkowań procesu pozyskiwania informacji i jest zjawiskiem występującym we wszystkich badaniach statystycznych. Błędy tego rodzaju są charakterystyczne dla badań reprezentacyjnych i spisywane są w formie tabel precyzji poszczególnych badań. W odniesieniu do metod przedstawionych w niniejszym opracowaniu, poszczególne składowe dekompozycji oraz pod-dekompozycji powinny być zestawione z tabelami precyzji dla tych wielkości w celu stwierdzenia, czy błąd jest duży w stosunku do samej składowej dekompozycji. Na podstawie wstępnych obliczeń można stwierdzić, że błąd niedokładności danych jest poważnym problemem dla mniej licznych i „skrajnych” grup populacji: na przykład wskaźniki zatrudnienia lub współczynniki aktywności zawodowej dla osób w wieku 15-17, 18-19 oraz 20-24 wykazują dużo większe wartości oszacowania względnych błędów standardowych niż te same wskaźniki dla pozostałych grup wiekowych. Wynika z tego, że problem niedokładności dostępnych danych ma szczególne znaczenie w przypadku pod-dekompozycji

¹⁸ Przybliżenie, o którym mowa, wynika z dwóch faktów. Po pierwsze, wynika z przybliżenia liniowego dekompozycji dynamiki PKB w cenach bieżących na sumę dynamiki PKB w cenach stałych i dynamiki poziomu cen w gospodarce. Po drugie, wynika z tego, że inflacja może być liczona dla nieco odmiennego koszyka dóbr niż składowe PKB.

poszczególnych składowych głównych i występuje dla pewnych grup wieku, wykształcenia (w mniejszej skali) lub sektorów gospodarczych (również mniejsza skala zjawiska).