

rozprawy i studia

ANDRZEJ ROSNER¹

PROBLEM POMIARU POZIOMU ROZWOJU SPOŁECZNO-GOSPODARCZEGO. SKALA POMIAROWA I JEJ WŁAŚCIWOŚCI

Streszczenie: Przedmiotem analizy są metodologiczne ograniczenia pomiaru zmiennych określanych w terminach teoretycznych (na wysokim poziomie abstrakcji), takich jak np. poziom rozwoju społeczno-gospodarczego. Broniona jest teza, że nawet jeśli wszystkie składowe elementy pomiaru (wskaźniki empiryczne uwikłane w algorytmie tworzenia skali pomiarowej takiego pojęcia) mają charakter pomiaru na skalach co najmniej interwałowych, to i tak syntetyczna skala pomiarowa ma charakter słabszy, porządkowy. Zdania korzystające z wyników takiego pomiaru mają charakter relatywizacji przestrzennej i historycznej. Dodatkowo rozważana jest zależność między prostotą (elegancją formalną) algorytmu pomiaru i agregacją przestrzenną analizowanych jednostek. Zależność ta powoduje, że im mniejsze jednostki przestrzenne, tym bardziej złożony algorytm, co jest wynikiem różnych możliwych dróg prowadzących do realizacji tego samego celu (np. postępu w osiągniętym poziomie rozwoju).

Słowa kluczowe: rozwój społeczno-gospodarczy, zróżnicowanie przestrzenne rozwoju, skala pomiarowa, operacjonalizacja terminów teoretycznych

CO PODLEGA POMIAROWI? KWESTIA TREŚCI POJĘĆ TEORETYCZNYCH

Problem pomiaru poziomu rozwoju społeczno-gospodarczego staje się istotny wówczas, gdy przedmiotem rozważań jest porównanie dwóch lub większej liczby jednostek pod względem pewnej, trudnej do jednoznacznego zdefiniowania, cechy, jaką jest właśnie to, co potocznie nazywamy „poziomem rozwoju” lub porównanie „poziomów rozwoju” tej samej jednostki, ale w dwóch punktach czasowych. Badanie takie napotyka jednak szereg trudności, jedna z nich związana jest z faktem, że rzadko

¹ Autor jest pracownikiem naukowym Instytutu Rozwoju Wsi i Rolnictwa PAN.

używa się obecnie pojęcia „rozwój” w podstawowym (dla dalszych rozważań) sensie ekonomicznym inaczej, niż uzupełniając je przymiotnikiem, takim jak „zrównoważony”, „ekologiczny (ekorozwój)”, „społeczny”, „endogeny”² itp. Tym samym nabierają one cech interdyscyplinarnych. Ponadto samo pojęcie „rozwój” zawiera w sobie element dynamiki, niezależnie od tego, jakim przymiotnikiem jest opatrzony – oznacza zmianę pewnego stanu rzeczy w inny, bardziej korzystny z punktu widzenia pewnych kryteriów. Pojęciem pierwotnym logicznie jest to, co można nazwać „stanem rzeczy” z danego punktu widzenia, a więc rozumianym statycznie pewnym „poziomem rozwoju”. Pierwotnym, bo badanie „rozwoju” (więc dynamiki) zakłada porównanie dwóch jego poziomów rozdzielonych czasem.

Porównanie jednostek między sobą może dotyczyć zarówno osiągniętego poziomu rozwoju, samego rozwoju, a więc dynamiki zjawiska dokonujących się zmian, jak i kryteriów decydujących o tym, że zmianę rozumie się jako rozwój, a zatem, że jej kierunek (oceniany ze względu na dane kryteria) jest korzystny, to znaczy – zmierza ona do doskonalszego ich wypełniania. Łatwo można wyobrazić sobie sytuację w której dynamika zjawisk jest wysoka, jednak w relatywnie małym stopniu dokonuje się poprawa ze względu na przyjęte kryteria określające „poziom rozwoju” jak i taką, w której relatywnie mniejsza dynamika zjawisk składających się na rozwój wywiera istotny wpływ na spełnianie założonych jego kryteriów. Jest to oczywista konsekwencja faktu, że nie każda zmiana prowadzi do poprawy spełniania kryteriów rozwoju.

Wieloznaczność pojęć „rozwój” i „poziom rozwoju” sprzyja temu, że w różnych rozważaniach na jego temat przypisuje im się odmienną konotację, częściowo przynajmniej związaną z różnymi kryteriami oceny. I tak np. w ekspertyzie zatytułowanej *Błąd pomiaru. Dlaczego PKB nie wystarcza* autorzy [Stiglitz, Sen, Fitoussi 2013] skłaniają się do utożsamiania „rozwoju” ze „wzrostem dobrobytu” lub „jakości życia”. Warto zauważyć, że pojęcia te są równie niejasne i wieloznaczne jak pojęcie, które w założeniu mają precyzować, może zachodzić więc sytuacja określana jako *ignotum per ignotum*. Ponadto osiągnięty poziom „dobrobytu” czy „jakości życia” są również pojęciami języka teoretycznego (abstrakcyjnego), problem polega więc na tym, czy ich operacyjne definicje są tożsame. Gdyby tak było – byłyby to pojęcia tożsame (przynajmniej ekstensjonalnie), mogłyby różnić się pod względem konotacji, kładąc akcent na różne cechy tego samego zjawiska. Aby sprawa stała się jaśniejsza, byłaby ona podobna do dwóch pojęć: kwadratu i czworoboku foremnego. Oba wyznaczają ten sam zbiór desygnatów, jednak jedno akcentuje pewne ich cechy, drugie – inne.

W przypadku takich pojęć, jak „rozwój” czy „poziom rozwoju” dopiero sformułowanie definicji operacyjnych zapewniłoby możliwość dokonania pomiaru i porównań między jednostkami. Co więcej, operacjonalizacja ta powinna być powszechnie uznana, wówczas będzie miała charakter definicji sprawozdawczej, będzie informowała, co w języku nauki rozumiemy przez „rozwój”. W przeciwnym wypadku definicja pojęcia ma cechy tzw. definicji projektującej, zrelatywizowanej

² Podobna sytuacja występuje również w innych dyscyplinach naukowych, ogólne pojęcie „rozwój” opatrywane jest przymiotnikami, często złożonymi, takimi jak np. psychofizyczny.

do danego badania (analizy), to znaczy, że przybiera ona postać „w tym opracowaniu przez pojęcie X rozumie się to i to”. Pozostaje jednak nierozstrzygnięte pytanie, czy taką definicję sprawozdawczą można sformułować i czy jest to potrzebne.

Co więcej, gdyby utożsamić „rozwój” ze „wzrostem dobrobytu” oznaczałoby to, że niebudząca zastrzeżeń operacjonalizacja jednego z tych pojęć jest (przynajmniej w sensie ekstensjonalnym) również operacjonalizacją drugiego. Pozostaje jednak niejasne, czy „dobrobyt” i „poziom rozwoju”, a także „rozwój” i „wzrost dobrobytu” są (parami) tym samym pojęciem. Wydaje się, że na gruncie języka potocznego – nie. Podobnie jakość życia i jej poprawa nie musi wiązać się z poziomem rozwoju i samym rozwojem. Wydaje się, że jest pojęciem szerszym o czynniki pozaekonomiczne wpływające na nią, aczkolwiek prawdopodobnie węższym w wymiarze czysto ekonomicznym, bo nie zawsze poprawa poziomu rozwoju ekonomicznego musi wiązać się z poprawą jakości życia. Do relacji między tymi pojęciami wrócimy jeszcze w dalszej części artykułu.

Warto zauważyć, że w naukach społecznych trudności z operacjonalizacją pojęć ogólnych (teoretycznych) występują często, wiele z nich, choć przydatne dla opisu zjawisk społecznych i psychologicznych, jest – w sensie przekładu operacyjnego – niejednoznacznych, a mimo to są one powszechnie używane. Czasem w stosunku do tego rodzaju pojęć używa się określenia, że są one „płodne”, co oznacza, że mimo ich niejednoznaczności pozwalają na przekonujący opis i wyjaśnianie mechaniki zjawisk i procesów społecznych. Do pojęć tego rodzaju zaliczyć można np. pojęcie prestiżu, pozycji społecznej itp. Na gruncie psychologii trudności takie występują np. w stosunku do pojęcia „inteligencja”, co prowadzi do paradoksów w rodzaju stwierdzenia, że jest to ta cecha, którą mierzy test IQ. Mimo to wymienione pojęcia są niezwykle użyteczne i powszechnie używane.

Dodatkowa trudność wynika z tego, że próby operacjonalizacji pojęcia „poziom rozwoju” nie mogą mieć charakteru uniwersalnego, co wiąże się z faktem, że zdania korzystające z tych pojęć mają charakter twierdzeń o zrelatywizowanym zasięgu historycznym i przestrzennym [Nowak 1965, s. 43 i następne]. Jako przykład potrzeby takiej relatywizacji przytoczyć można pytanie zadane przez jednego z uczestników seminarium IRWiR PAN i Instytutu Rozwoju Wsi Chińskiej Akademii Nauk po wysłuchaniu referatu na temat marginesu biedy na wsi w Polsce: „jaki procent Polaków żyje za mniej niż jednego dolara dziennie”. Z perspektywy badacza chińskiej wsi pytanie to miało istotne znaczenie, co więcej, odnosiło się do uznanych miar stosowanych w porównaniach międzynarodowych. Było jednak nieadekwatne do poziomu rozwoju społeczno-ekonomicznego w krajach UE, nawet najbardziej ubogich, między innymi ze względu na występowanie systemów świadczeń pomocy społecznej. Byłoby ono trafne w odniesieniu do wielu krajów Afryki, biednych krajów Azji itp.

„Rozwój” (niezależnie od tego, z jakim przymiotnikiem) jest procesem w trakcie którego zmieniają się cele, które są (czy też powinny być) osiągane. W pewnym okresie celem takim może być dostarczenie wszystkim członkom społeczeństwa możliwości fizycznego przetrwania, a więc np. uniknięcie głodu rozumianego jako powszechna dostępność żywności w ilości zapewniającej minimalną normę wartości kalorycznych. Całkiem inny etap „rozwój” stawia sobie za cel np. dostęp do elementarnej edukacji,

jeszcze inny – do zdrowego środowiska. W praktyce oznacza to, że nie istnieje lub jest bardzo ograniczona możliwość uniwersalnej operacjonalizacji pojęcia „rozwój”, która pozwoliłaby na wykorzystanie do pomiarów całej skali, a nie tylko pewnej jej sekcji czy odcinka. Wyznaczenie takiego wykorzystywanego przedziału skali, to właściwie relatywizacja przestrzenna i historyczna zdań wykorzystujących wyniki pomiaru.

UWAGI O LOGICZNEJ CHARAKTERYSTYCE SKAL POMIAROWYCH

Każdy pomiar musi dokonywać się na jakiejś skali, a więc na pewnym konstrukcie logicznym „nabudowanym” na osi liczbowej, choć czasem oś ta przybiera postać słownych opisów kolejnych punktów. Tak dzieje się np. wówczas, gdy respondentowi dajemy do wyboru jedną z pięciu odpowiedzi: złe, raczej złe, średnie, raczej dobre, dobre. W rzeczywistości posługujemy się wówczas pięciopunktową skalą do pomiaru pewnej cechy, o której wnioskujemy z odpowiedzi respondenta. Jest to jednak tylko językowa stylizacja odcinka osi liczbowej.

W rozważaniach logiczno-metodologicznych przyjmuje się, że pomiar to system składający się ze zbioru elementów poddawanych pomiarowi (polu pomiaru) i pewnego zbioru relacji (czasem funkcji) opisanych na polu pomiaru [Suppes, Zinnes 1962]. W rozważaniach metodologicznych z zakresu nauk społecznych analizie poddaje się przede wszystkim relacje (i funkcje) tworzące daną skalę pomiarową. Analiza tych relacji pozwala wyróżnić cztery podstawowe skale do pomiaru zmiennych³.

Najsłabsza logicznie skala pomiarowa to tzw. skala nominalna. Jej konstrukcja logiczna podobna jest do podziału logicznego zbioru z tym, że o ile w przypadku podziału wymagane jest spełnienie trzech warunków (rozłączność podzbiorów, wyczerpywanie zbioru przez podzbiory oraz niepustość podzbiorów), o tyle w przypadku skali nominalnej zazwyczaj pomija się wymóg niepustości podzbiorów. Skala nominalna oparta jest na jednej tylko relacji: równości. Dzieli ona zbiór na klasy przedmiotów pod danym względem równych i różnych między wyodrębnionymi klasami. Jako przykład posłużyć może pomiar linii tramwajowych na takie, które oznaczone są numerem 1, numerem 2 itd.

Skalą „mocniejszą” logicznie jest skala porządkowa. Wyznaczają ją dwie relacje, równości i większości (pod danym względem). Pomiar na takiej skali dokonuje nie tylko podziału zbioru na klasy elementów równych pod badanym względem, ale klasy te dodatkowo porządkuje. Skale tego rodzaju są niezwykle rozpowszechnione w naukach społecznych, ale wykorzystywane są również w innych celach, np. taki charakter ma skala twardości minerałów Mohsa⁴. W socjologii skale porządkowe wykorzystywane są do pomiaru np. dystansu społecznego, prestiżu itp. Najważniejszą

³ Pomijam tu problem tzw. skal złożonych, a więc takich, które powstają w wyniku operacji dokonywanych na skalach prostych.

⁴ Mowa o dziesięciopunktowej skali twardości minerałów. Została ona zdefiniowana w ten sposób, że jeśli minerałem A można zarysować minerał B i odwrotnie – są one tej samej twardości. Jeśli minerał X zarysowuje minerał Y, ale nie odwrotnie – minerał X jest twardszy od Y.

cechą skal porządkowych jest to, że pozwalają one określić porządek (według stopnia spełniania danego kryterium) elementów zbioru poddawanych pomiarowi. Jednocześnie nie mówią one o tym porządku nic ponad to, a w szczególności punkty nazwane na skali kolejnymi np. numerami nie muszą być rozłożone w równych odległościach. Szczególnym przykładem skali porządkowej jest szereg rangowy, porządkuje on zbiór, ale nie informuje o dystansach między poszczególnymi punktami szeregu. Na jego podstawie można określić, czy między pozycją X i Y występują pozycje pośrednie, ale występowanie ich nie oznacza, że dystans między tymi pozycjami wzrasta.

Warto też zwrócić uwagę na to, że na podstawie wyników pomiaru na skali porządkowej nie można nic powiedzieć o charakterze rozkładu tych wyników, w szczególności, czy jest on zbliżony do rozkładu normalnego, biegunowego, czy jest symetryczny itp.

Kolejnym, trzecim typem skal pomiarowych jest tzw. skala interwałowa. Różni się ona od porządkowej tym, że na tworzącej ją osi liczbowej zachowane są stałe odległości między punktami, między których nazwami arytmetyczna różnica jest ta sama. Skale takie mogą być wyrażone w jednostkach (np. w stopniach Celsjusza, w dolarach itp.). Skala interwałowa może być opisana na pewnym segmencie osi liczbowej, jej charakter nie wymaga występowania punktu początkowego (zerowego) o charakterze wyróżnionym w sposób absolutny. Z tego względu pomiar dokonywany na skali interwałowej może przybierać wartości ujemne i dodatnie. Ten typ skal jest przydatny np. do określania odchylenia się wartości dla poszczególnych elementów zbioru od średniej dla całego zbioru.

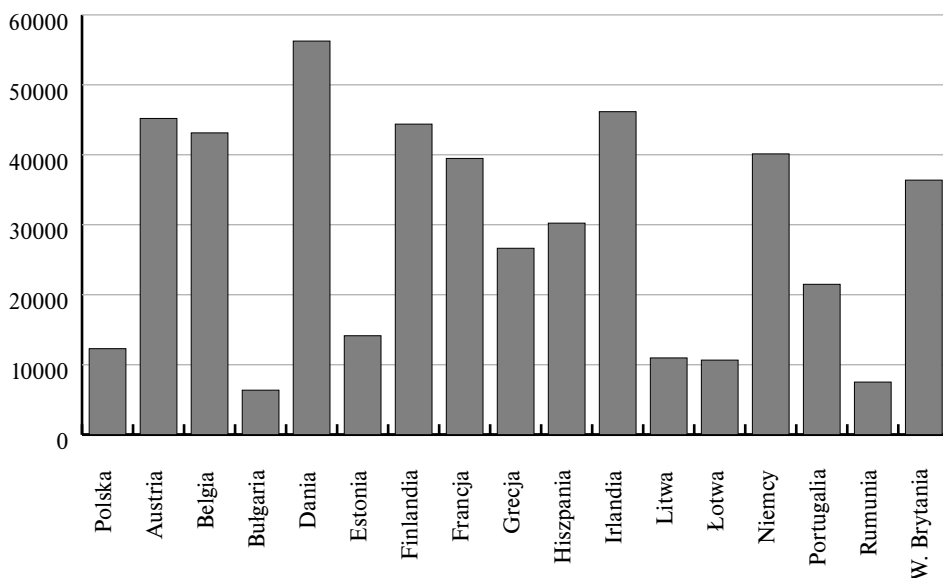
Wreszcie najmocniejszy typ skali pomiarowej, tzw. skala absolutna, różni się od interwałowej tym, że punkt zerowy osi liczbowej jest jej końcem, więc wynik pomiaru nie może być ujemny. Jako przykład posłużyć może pomiar temperatury, jeśli wykorzystuje skalę Celsjusza – dokonywany jest na skali interwałowej, jeśli na skali Kalwina – skalę absolutną. W naukach społecznych, zwłaszcza ekonomicznych, zarówno skale interwałowe, jak też absolutne są często wykorzystywane.

Warto jednak zauważyć, że charakter skali pomiarowej decyduje o tym, jakie operacje arytmetyczne mogą być wykonywane na wynikach pomiarów. Skale nominalne wyłącznie pozwalają na klasyfikacje elementów podlegających pomiarowi na klasy różniące się pod danym względem. Skale porządkowe – dodatkowo pozwalają porządkować wyróżnione klasy według kryterium definiowanego przez skalę. Skala interwałowa dopuszcza operacje dodawania i odejmowania dokonywane na wynikach pomiarów, a absolutna również mnożenia i dzielenia. Także wnioski, jakie można formułować opierając się na wynikach pomiaru, zależą od typu skali; ogólnie można powiedzieć, że im mocniejszy typ skali, tym więcej informacji dostarcza pomiar. Jednak, opierając się na tym stwierdzeniu, nie można postulować, aby konstruować skale mocniejsze, typ skali zależny jest w ogromnym stopniu od tego, co poddawane jest pomiarowi⁵.

⁵ Szerzej o błędach w formułowaniu wniosków nieadekwatnych do typu skali pomiarowej pisał T. Pawłowski w pracy: *Rodzaje skal a struktura zdań stwierdzających związki między wielkościami*. [w:] *Logiczna Teoria Nauki*. PWN, Warszawa 1966.

PROBLEM ZAKRESU TREŚCI MIERZONEJ WIELKOŚCI A ELEGANCJI FORMALNEJ ALGORYTMU POMIARU

Poszukiwanie uniwersalnej skali „poziomu rozwoju”, co więcej, eleganckiej, bo opartej na powszechnie uznanym wskaźniku i formalnie prostym algorytmie, czasem dokonuje się na gruncie ekonomii. Taką wartość wskaźnikową przypisuje się PKB *per capita*, co jednak bywa ostro krytykowane. Nie znaczy to, że algorytm obliczania PKB jest prosty, ale miara ta jest stosowana również dla innych celów, tzw. rachunki narodowe przeprowadzane są przez wszystkie kraje w ich urzędach statystycznych. Miara „poziomu rozwoju” oparta na PKB korzysta więc z wyliczeń przeprowadzanych standardowo przez krajowe urzędy statystyczne i z tego powodu są one dostępne. Rezultat takiej miary dla wybranej grupy krajów zawiera wykres 1.

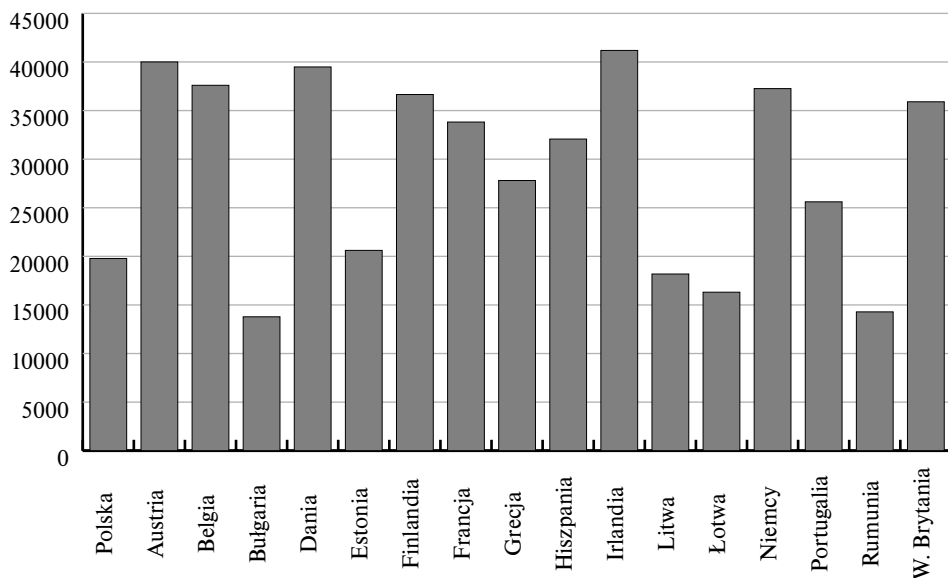


WYKRES 1. PKB *per capita* w wybranych krajach UE w 2010 roku (w dol. USA, ceny bieżące)
GRAPH 1. GDP *per capita* in chosen EU countries in 2010 (USA \$'s, current price)

Źródło: Dane według Rocznika Statystyki Międzynarodowej 2012. GUS, Warszawa 2012.

Source: Data according to *Rocznik Statystyki Międzynarodowej (Annual International Statistics) 2010*. Central Statistical Office, Warsaw, 2012.

Relatywnie lepszym wskaźnikiem jest PKB *per capita* uwzględniający wartość nabywczą pieniądza w poszczególnych jednostkach poddawanych pomiarowi. Wskaźnik ten (PKB *per capita* według PPP) znosi jednak tylko część zastrzeżeń formułowanych w stosunku do miary PKB *per capita*. Wykres 2 przedstawia wartość wskaźnika PPP dla tych samych krajów, które zestawione zostały na wykresie 1.



WYKRES 2. PKB według PPP *per capita* w wybranych krajach UE w 2010 roku (w dol. międzynarodowych, ceny bieżące)

GRAPH 2. GDP according to PPP *per capita* in chosen EU countries in 2010 (international \$'s, current price)

Źródło: Dane według Rocznika Statystyki Międzynarodowej 2012. GUS, Warszawa 2012.

Source: Data according to *Rocznik Statystyki Międzynarodowej (Annual International Statistics)* 2010. Central Statistical Office, Warsaw, 2012.

Zastrzeżenia dotyczące miar opartych na PKB wynikają między innymi z tego, że dotyczy on wyłącznie rozwoju gospodarczego. Zwraca się również uwagę, że rachunek PKB wykazuje wzrost wskaźnika w sytuacji wystąpienia klęsk naturalnych, nie oznacza on jednak wówczas wzrostu gospodarczego. I wreszcie, jest on przydatny dla dużych agregatów (krajów), w miarę zaś schodzenia do mniejszych jednostek, precyzja miar opartych na PKB gwałtownie spada. (Pomijam tu szczegółowe kwestie dotyczące niedokładności związanych z tzw. rachunkami narodowymi, które również prowadzą do wniosku, że jest to miara nieprecyzyjna i podatna do manipulacji statystycznych)⁶. Obok takich zastrzeżeń wskaźniki oparte na PKB mają również pewne zalety formalne. Skale pomiarowe wykorzystywane w przypadku tych wskaźników zdawałoby się, że są stosunkowo silne, na pierwszy rzut oka mają one charakter co najmniej interwałowy. Pozwala to nie tylko na porównywanie analizowanych jednostek (do tego wystarczyłyby skale porządkowe), ale również na określanie dystansu między nimi wzdłuż danej skali (na co skala porządkowa nie pozwala).

⁶ Problem ten jest szeroko omówiony [w:] Joseph E. Stiglitz, Amartya Sen, Jean-Paul Fitoussi: *Błąd pomiaru. Dlaczego PKB nie wystarcza. Raport Komisji ds. Pomiaru Wydajności Ekonomicznej i Postępu Społecznego*. PTE 2013.

Określenie „na pierwszy rzut oka”, użyte wcześniej, odnosi się do kwestii bardzo ważnej, choć rzadko podejmowanej. Skala empiryczna, np. mówiąca o PKB w dolarach na mieszkańca, ma niewątpliwie charakter co najmniej interwałowy, wydaje się, że można by nawet bronić tezy, że jest to tzw. skala absolutna. Interwałem jest jeden dolar, punktem zerowym – brak PKB. Jednak jeśli skalę taką potraktować jako miarę rozwoju gospodarczego, powstaje pytanie, czy nie straci ona swojej logicznej własności równych interwałów. Różnica o x dolarów dla jednostek o niskim wskaźniku PKB ma większe znaczenie niż dla jednostek o wysokich wartościach tego wskaźnika. Z tego względu czasem można spotkać w analizach przeliczenia różnic między krajami nie w liczbie interwałów skali (w dolarach), ale w miarach względnych (np. wówczas, gdy mówimy, że PKB *per capita* w jednostce A jest wyższe niż w B o $y\%$). Takie postawienie sprawy prowadzi do konieczności rozważenia dwóch kwestii⁷:

- 1) metodologicznej, czy między typem interwałowym skali i porządkowym występuje jakiś typ pośredni, na co wskazywałoby używanie względnych (procentowych) dystansów między jednostkami skali, oraz
- 2) czy mimo iż operacyjna, wskaźnikowa, skala oparta na jednym tylko wskaźniku ma charakter co najmniej interwałowy, to pomiar *indicatum* (poziom rozwoju) mierzony w oparciu o taką definicję wykazuje charakter słabszy logicznie, np. porządkowy. Inaczej mówiąc, czy istnieje absolutna miara poziomu rozwoju, czy też samo pojęcie „poziom rozwoju” nie powoduje, że skala mierząca zróżnicowanie ma charakter słabszy, niż skala mierząca wskaźnik występujący w definicji operacyjnej. Osobiście opowiadam się za tą drugą możliwością. Sprowadza się ona do stwierdzenia, że ta sama skala, mając charakter wskaźnikowy dla różnych pojęć, może mieć różny charakter logiczny, zależny od konotacji tych pojęć.

Obie te kwestie mają fundamentalne znaczenie dla problemu pomiaru w naukach społecznych, mają też bardzo dużą literaturę na różnym poziomie abstrakcji (od czysto logicznych rozważań dotyczących typów skali, do kwestionowania szczegółowych pomiarów ze względu na operacje arytmetyczne dokonywane na skalach ich niedopuszczających).

Przyjęcie, że miarą rozwoju jest wskaźnik PKB *per capita* lub PPP również liczony na mieszkańca ma jednak jeszcze inne konsekwencje. Można je prześledzić na przykładzie tej samej grupy krajów UE, które wymienione zostały na wykresach 1 i 2.

Dane zestawione w tabeli 1 ujawniają, że skale pomiarowe poziomu rozwoju gospodarczego według PKB i PPP liczonych na mieszkańca nie zawsze dają takie samo uporządkowanie jednostek. Gdyby użycie wskaźnika PPP poprawiało tylko dokładność pomiaru (i relacji między jednostkami poddanymi pomiarowi) porządki rangowe powinny pozostać niezmienione, okazuje się jednak, że zmiany występują. Np. według miary opartej na PKB Wielka Brytania jest lepiej rozwinięta od Francji, ale wskaźnik PPP ujawnia, że jest odwrotnie. Jednocześnie nie jest możliwe, aby jakiś kraj był zarazem lepiej i gorzej rozwinięty od innego. Byłoby to sprzeczne z podstawowymi prawami logiki, w szczególności z zasadą niesprzeczności (nieprawda, że p i $nie p$ zarazem).

⁷ Można również rozważyć trzecią możliwość, czy nie jest to wówczas skala pomiarowa złożona, a więc skala składająca się z co najmniej dwóch skal i relacji między nimi. Kwestia ta jako ściśle logiczna wykracza poza ramy artykułu.

TABELA 1. Poziom rozwoju wybranych krajów UE według PKB i PKB według PPP *per capita* w 2010 r.

TABLE 1. Chosen EU countries development level according to GDP and GDP according PPP *per capita* in 2010

	PKB <i>per capita</i>	PKB według PPP <i>per capita</i>	Ranga według PKB	Ranga PKB, według PPP <i>per capita</i>	Różnica rang (PPP-PKB)
Polska	12298	19783	13	13	0
Austria	45191	40005	3	2	-1
Belgia	43127	37600	5	4	-1
Bułgaria	6365	13780	17	17	0
Dania	56255	39489	1	3	+2
Estonia	14146	20615	12	12	0
Finlandia	44380	36651	4	6	+2
Francja	39475	33820	7	8	+1
Grecja	26631	27805	10	10	0
Hiszpania	30222	32070	9	9	0
Irlandia	46160	41188	2	1	-1
Litwa	10975	18184	14	14	0
Łotwa	10663	16312	15	15	0
Niemcy	40123	37260	6	5	-1
Portugalia	21486	25610	11	11	0
Rumunia	7522	14287	16	16	0
W. Brytania	36373	35904	8	7	-1

Źródło: Obliczenia własne na podstawie danych GUS. Rocznik Statystyki Międzynarodowej 2012. GUS, Warszawa 2012

Source: Own calculations based on data provided by Central Statistical Office.

Aby rozwiązać ten problem, można postąpić dwojako: albo przyjąć, że skale rozwoju według PKB i według PPP mierzą inne pojęcie rozwoju; obie są prawidłowe, ale dotyczą innych cech jednostek. Musimy wówczas pogodzić się z tym, że pojęcie „rozwój” jest wieloznaczne. Można też postąpić inaczej – przyjąć, że zarówno wartość PKB jak i PPP są wskaźnikami empirycznymi dla tego samego pojęcia „rozwoju”, ale ich cechy metodologiczne jako wskaźników są różne. W szczególności nie są to wskaźniki definicyjne, ale statystyczne i charakteryzują się inną „mocą wskaźnikową”⁸. Oznacza to, że zarówno jedna, jak i druga skala daje tylko przybliżony wynik uporządkowania jednostek (krajów) według poziomu rozwoju gospodarczego, przy czym ten „poziom rozwoju gospodarczego” traktowany jest jako pojęcie języka teoretycznego.

Zastrzeżenie dotyczące miar poziomu rozwoju opartych na PKB prowadziły do prób budowy miar bardziej złożonych, wykorzystujących pewne zbiory wskaźników oraz algorytm łączenia ich w jedną miarę syntetyczną (a więc mniej eleganckich formalnie), jednak poszukiwanie takich miar odpowiadało także na postulat szerszego niż wąsko ekonomicznie rozumienia pojęcia „rozwój”. Rozszerzenie treści pojęcia „rozwój” odbywało się więc kosztem elegancji metodologicznej. Poszukiwano przy tym możliwie małego zbioru wskaźników, takich aby miały względnie uniwersalne znaczenie, to znaczy takich, które, po pierwsze, pozwalały na relatywizację pomiaru

⁸ Relacje między wskaźnikami i *indicatum* szczegółowo analizował Stefan Nowak w pracy: *Studia z metodologii nauk społecznych*. PWN, Warszawa 1965.

do względnie dużego przedziału historycznego i przestrzennego, po drugie – uwzględniały fakt, że rozwój może dokonywać się według różnych scenariuszy, inaczej mówiąc, ten sam efekt może być uzyskiwany na różne sposoby. Na przykład, nie jest uniwersalny w tym rozumieniu wskaźnik udziału ludności miejskiej w danym społeczeństwie, choć przeważnie wzrost udziału ludności w miastach wiąże się ze wzrostem poziomu rozwoju gospodarczego. Może się on jednak dokonywać zarówno poprzez przenoszenie ludności ze wsi do miast, jak i poprzez prowadzenie polityki sprzyjającej powstawaniu pozarolniczych miejsc pracy na obszarach wiejskich i sprzyjaniu osiedlania się na wsi. Ponadto klasyfikacja miejscowości na wiejskie i miejskie jest różna w poszczególnych krajach, a próby ominięcia tej bariery (np. podejmowane przez OECD) budzą wiele wątpliwości. Natomiast dobrym wskaźnikiem jest np. miara średniej długości życia w danej populacji. Co prawda, wydłużanie życia nie jest celem rozwoju, ale jest następstwem szeregu dokonujących się zmian (poprawy standardów życia, dostępności usług medycznych, rozwoju usług publicznych, niskiego wskaźnika umieralności niemowląt, redukcji stresu itp.), które takim celem są. Średnia długość życia jest więc wskaźnikiem pośrednio wskazującym na poziom rozwoju społeczno-gospodarczego niezależnie od tego, jaki model rozwoju przez analizowane jednostki jest realizowany.

Pomijamy tu fakt, że wydłużanie życia nie musi łączyć się z jego jakością, współczesna medycyna opracowała metody podtrzymywania życia bez perspektyw wyleczenia schorzeń decydujących o tym, że jakość tego życia staje się wegetacją. Nawet środowiska stawiające w centrum swoich ideologicznych wartości życie ludzkie dorobiły się terminu „uporczywej terapii”, co można rozumieć jako podtrzymywanie życia bez względu na jego jakość.

Wychodząc z przesłanek wskazujących na wady i jednostronność miar opartych na PKB lub wskaźniki od niego pochodne, został zbudowany w środowisku ekonomistów i polityków społecznych związanych z agendami ONZ wskaźnik zwany skrótowo HDI (*Human Development Index*). W odróżnieniu od omówionych wcześniej miar opartych na PKB dotyczy on rozwoju społeczno-gospodarczego, a nie tylko komponentu gospodarczego. Z tego względu należy go traktować raczej jako odpowiedź na krytykę wskaźników opartych na PKB podnoszącą zawężony do gospodarki zakres spraw badanych przez nie, a nie jako ich bezpośrednią konkurencję.

Podobnie jak wskaźniki oparte na miarach związanych z PKB, HDI stał się miarą wykorzystywaną dla analiz zróżnicowania poziomu rozwoju w układzie regionów świata i krajów. Konstrukcja indeksu zakłada, że jest on miarą syntetyczną, składającą się z trzech komponentów: ekonomicznego, edukacyjnego oraz demograficznego. Komponenty te ulegały modyfikacjom, dotyczyło to przede wszystkim edukacyjnego. W starszych wersjach algorytmu HDI (sprzed 2010 roku) brał on pod uwagę zjawisko analfabetyzmu i wskaźnik skolaryzacji, nadając im odpowiednio wagi 2/3 i 1/3, obecnie zrezygnowano z takiej konstrukcji na rzecz odpowiednio uogólnionych dwóch wskaźników prostszych: średniej liczby lat nauki szkolnej osób w wieku 25 lat i więcej oraz przewidywanej średniej liczby lat nauki dzieci w wieku 5 lat. Wskaźniki te mają taką samą wagę. Komponent demograficzny, interpretowany jako wskaźnik zdrowotności i pośrednio jakości życia, to przeciętna długość życia osoby w wieku 0 lat (według danych tablic trwania życia). Trzecim składnikiem

HDI jest wskaźnik ekonomiczny korzystający z miary PPP *per capita*. Pomijając szczegóły statystyczne, wszystkie trzy wskaźniki są standaryzowane w taki sposób, aby ich wartość mieściła się w przedziale 0–1, a końcowa wartość HDI to średnia geometryczna z trzech wartości odpowiadających tym komponentom.

Warto dodać, że w oparciu o HDI powstał szereg podobnych wskaźników, uwzględniających np. umieralność niemowląt, które jednak nie stanowią standardu równie często przywoływanego, jak podstawowa wersja HDI opracowana przez UNDP.

Ze względu na swoją konstrukcję skala HDI ma charakter porządkowy, odwzorowuje ona relację o charakterze porządkowym oraz równościową. Na podstawie wskaźnika można więc stwierdzić, który kraj z danej pary jest wyżej, a który niżej rozwinięty w sensie społeczno-gospodarczym lub że są one na tym samym poziomie rozwoju. Nie można jednak nic powiedzieć o dystansie rozwojowym między nimi. Jednak, jak się zdaje, nie tylko względy formalne decydują o tym, że skala HDI ma charakter porządkowy, również analiza konotacji pojęcia *human development* przy przyjęciu silniejszej miary niż porządkowa budzi sprzeciw intuicji językowych. Warto dodać, że trzy składowe HDI (ekonomiczny, edukacyjny i demograficzny) traktowane są jako wskaźniki empiryczne (w rozumieniu typologii wskaźników), a nie definicyjne. Nie jest więc od nich wymagane, aby wpływały na końcowy wynik pomiaru indeksu inaczej niż tylko statystycznie.

TABELA 2. Rangi wybranych krajów według trzech wskaźników (PKB, PPP i HDI)
TABLE 2. Chosen countries rank accordingly to three indicators (GDP, PPP and HDI)

Kraje	Ranga według PKB	Ranga według PPP	Ranga w zbiorze wybranych krajów HDI ^b	Ranga w zbiorze uniwersalnym ^a HDI
Polska	13	13	11	35
Austria	3	2	6	21
Belgia	5	4	6	21
Bułgaria	17	17	15	58
Dania	1	3	2	10
Estonia	12	12	10	33
Finlandia	4	6	7	24
Francja	7	8	5	20
Grecja	10	10	9	29
Hiszpania	9	9	8	27
Irlandia	2	1	3	11
Litwa	14	14	11	35
Łotwa	15	15	13	48
Niemcy	6	5	1	6
Portugalia	11	11	12	41
Rumunia	16	16	14	54
W. Brytania	8	7	4	14

^a w zbiorze krajów, dla których UNDP opublikowało wyniki analizy; ^b w przypadku takiego samego wskaźnika HDI uzyskiwanego przez dwa kraje nadano im tę samą pozycję w szeregu rangowym

Źródło: Obliczenia własne na podstawie danych GUS. *Rocznik Statystyki Międzynarodowej 2012*. GUS Warszawa 2012 oraz *UNDP Human Development Report 2012* (w internecie).

Source: own calculations based on data provided by Central Statistical Office and *UNDP Human Development Report 2012*.

Tabela 2 ujawnia, że mimo iż skale, o których mowa, a więc oparte na PKB *per capita*, PPP *per capita* i HDI mierzą różne wielkości, to jednak są one bardzo silnie skorelowane. Między skalami uwzględniającymi PKB i PPP (dla szeregu rangowego krajów wybranych dla ilustracji w tabelach) $r = 0,98$, analogicznie liczony współczynnik korelacji dla szeregów rangowych według HDI i PKB *per capita* $r = 0,89$, a HDI i PPP *per capita* $r = 0,90$. Oznacza to, że mimo różnic w konotacji pojęć „poziom rozwoju gospodarczego” i „poziom rozwoju społeczno-gospodarczego” empirycznie utworzone szeregi rangowe państw są bardzo zbliżone do siebie. A zatem, co prawda konotacje (treść) przypisane do tych pojęć są różne, ale denotacje (ekstensje) bardzo zbliżone. Z punktu widzenia praktyki analizy zróżnicowań (porządków, gdyż wszystkie omawiane skale mają charakter porządkowy) poziomu rozwoju, różnice między wymienionymi wskaźnikami są nieznaczne, jednak różnią się zasadniczo zarówno, jeśli pod uwagę weźmiemy konotację pojęć używanych przy ich operacjonalizacji, jak i formalną elegancję algorytmów.

Na marginesie warto zauważyć, że niezależnie od bieżącej pozycji na skali (rangi) miara HDI dla Polski wzrasta od wielu lat. Jest to jednak mylące, biorąc pod uwagę logiczne właściwości skali porządkowej. Według autorów raportu UNDP Polska zaliczona została do krajów o bardzo wysokim poziomie HDI. Bardziej rygorystyczna metodologicznie interpretacja wyników pomiaru ujawnia, że wynik pomiaru w 2013 roku lokował Polskę na 35. pozycji na 187 państw poddawanych analizie, podczas gdy rok wcześniej zajmowała 34. pozycję.

Wskaźnik rozwoju społeczno-gospodarczego HDI, podobnie jak wcześniej omawiane miary rozwoju gospodarczego, doczekał się krytyki z różnych stron. Najważniejsze argumenty osób polemizujących z jego autorami dotyczą faktu, że stanowi on odbicie „zachodniego modelu rozwoju”, nie uwzględniając szeregu czynników o charakterze kulturowym, związanych z istniejącymi różnicami, np. skal wartości itp. Z innych pozycji HDI krytykowane jest za to, że nie uwzględnia różnic związanych ze środowiskiem naturalnym. Polemizujący z nim wskazują na niedokładności danych pochodzących ze statystyk krajowych, a także na niekontrolowany wpływ, jaki mogą wywierać różnice metodologiczne poszczególnych (krajowych) źródeł informacji. Poddawane są ocenie szczegóły algorytmu oraz trafność doboru wskaźników empirycznych dla trzech komponentów (zmiany dokonane w 2010 roku w komponencie edukacyjnym były rezultatem tego rodzaju krytyki).

Oddzielne miejsce zajmuje polemika z UNDP na temat tego, czy agenda ONZ powinna w ogóle zajmować się tego rodzaju problematyką, gdyż może ona prowadzić do stygmatyzacji niektórych społeczeństw jako nisko rozwiniętych, przy czym poziom rozwoju oceniany jest według kategorii świata zachodniego, a nie skali wartości specyficznych dla nich właśnie. W szczególności wielu krytyków podważa wyznaczone przez autorów raportu „koszyki” krajów bardzo wysoko, wysoko, nisko i bardzo nisko rozwiniętych, wskazując, że granice między nimi są umowne, ale mogą zostać wykorzystane dla celów polityki międzynarodowej (np. jako kryteria krajów wymagających wsparcia, czy też takich, które powinny udzielać wsparcia). Na tego rodzaju krytykę autorzy HDI odpowiadają, że indeks jest kategorią analityczną, pod pewnymi względami lepszą (bogatszą treściowo) od miar opartych na PKB, a jego konstrukcja będzie ewoluować w miarę zmieniania się rzeczywistości, którą opisuje. Ta ostatnia uwaga

oznacza, że autorzy HDI zdają sobie sprawę, iż miara nie ma charakteru absolutnego, ale jest zrelatywizowana historycznie. Wyznaczone koszyki poziomów rozwoju to nic innego, jak próba skrócenia skali do pięciu punktów przez arbitralnie wyznaczone granice przedziałów na skali o nieokreślonej, ale bardzo dużej, liczbie punktów.

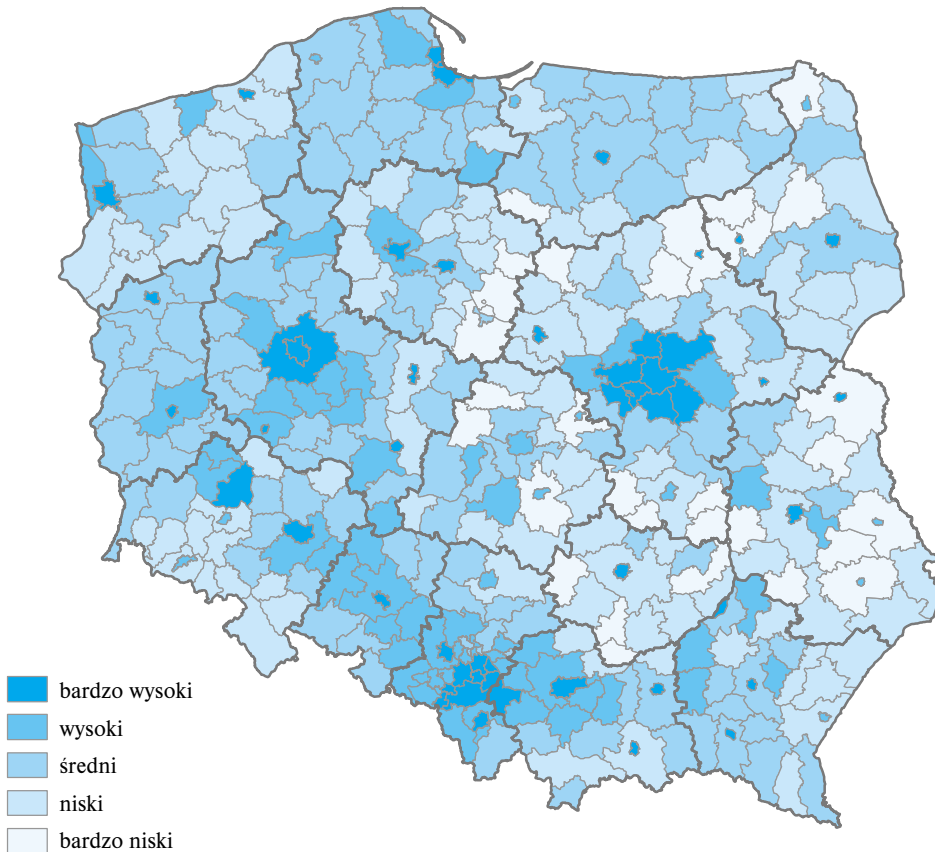
Pewną odpowiedzią na krytyki, z jakimi spotkało się opracowanie HDI, jest przygotowywany przez OECD BLI (*Better Life Index*), w założeniu ma on uwzględniać szereg dodatkowych wskaźników, takich jak satysfakcja życiowa, poczucie bezpieczeństwa ekonomicznego, poczucie sprawstwa, problem bezpieczeństwa pracy i zatrudnienia, środowiska naturalnego itp. Wskaźnik ten jest jednak obecnie na etapie opracowywania i jego ostateczne założenia metodologiczne, algorytmy i zakres merytoryczny są trudne do oceny. Wstępne zapowiedzi wskazują, że autorzy zamierzają wykonać kolejny krok w kierunku rozszerzenia zakresu treści mierzonej wielkości kosztem jej elegancji formalnej.

FORMALNA ELEGANCJA ALGORYTMU OPERACJONALIZACJI POJĘCIA „POZIOM ROZWOJU” A PRZESTRZENNA AGREGACJA JEDNOSTEK ANALIZY

Wymienione wcześniej wskaźniki poziomu rozwoju (PKB, PPP i HDI) mają tę wspólną cechę, że informują o zróżnicowaniu tego poziomu według relatywnie bardzo dużych jednostek – krajów. Doświadczenie wskazuje, że miara tego rodzaju dla kraju stanowi uśrednienie niekiedy bardzo silnie zróżnicowanej wewnątrznie przestrzeni. Wystarczy wspomnieć o różnicach, np. edukacyjnych między miastem i wsią w Chinach, czy też struktur gospodarczych Włoch północnych i południowych. Z tego względu duże znaczenie ma konstrukcja miar poziomu rozwoju, która mogłaby służyć ujawnianiu zróżnicowań przestrzennych wewnątrz krajów, w układach regionów i lokalnych.

Najbardziej dojrzałą spośród znanych prób tego rodzaju jest koncepcja wskaźnika HDI zmodyfikowana dla celów analizy regionów i subregionów (według niektórych wypowiedzi Autorów również układów lokalnych). Wskaźnik ten nazwany został LHDI (*Local Human Development Index*). Rozkład przestrzenny wskaźnika według powiatów przedstawia rysunek 1.

Zgodnie z oczekiwaniami LHDI pod względem formalnym jest mniej „elegancki” niż HDI, jego algorytm składa się z większej liczby podstawowych wskaźników empirycznych. Jest to konsekwencja potrzeby uwzględnienia dużej różnorodności jednostek analizy, która jednak nie musi prowadzić do hierarchicznych różnic między nimi z punktu widzenia celu tej analizy. Dla przykładu komponent „zamożność” jest zbudowany z kilku zmiennych, co uwzględnia różnice struktur gospodarczych między jednostkami (powiatami). Poza dochodami opodatkowanymi przez PIT, także bierze pod uwagę to, że rolnicy nie są płatnikami tego podatku. Wskaźnik musiał uwzględniać w jakiś sposób dochody tej grupy ludności, wprowadzono więc korektę związaną z obszarem użytków rolnych wykorzystywanych w jednostce (powiecie) skorygowanym o ich przydatność agrotechniczną, a więc w wymiarze tzw. powierzchni przeliczeniowej. Dodatkowo uwzględniono dochody ludności trafiające za pośrednictwem systemu pomocy społecznej.



RYSUNEK 1. Miara LHDI

FIGURE 1. LHDI measurement

Źródło: Arak P., Ivanov A., Peleah M., Płoszaj A., Rakocy K., Rok J., Wyszowski K., 2012, *Krajowy Raport o Rozwoju Społecznym. Polska 2012. Rozwój regionalny i lokalny*, Ministerstwo Rozwoju Regionalnego – Biuro Projektowe UNDP w Polsce, Warszawa.

Source: Arak P., Ivanov A., Peleah M., Płoszaj A., Rakocy K., Rok J., Wyszowski K., 2012: *Krajowy Raport o Rozwoju Społecznym. Polska 2012. Rozwój regionalny i lokalny*, (Domestic Rapport on Social Development. Poland 2012), Ministry of Regional Development – Project Office UNDP in Poland, Warsaw.

Komponent demograficzny, w HDI biorący pod uwagę jeden wskaźnik, w LHDI opiera się na dwóch: długości trwania życia oraz strukturze przyczyn zgonów. Komponent edukacyjny – wykorzystuje dwa wskaźniki, skolaryzację w wieku przedszkolnym oraz wyniki testów z części przedmiotów egzaminacyjnych.

Każdy z komponentów obliczany jest oddzielnie, a wskaźnik LHDI (podobnie jak HDI) powstaje jako średnia geometryczna wyników pomiaru trzech komponentów. Można więc powiedzieć, że w zasadzie każdy z nich ma taką samą wagę dla miary syntetycznej, a przynajmniej jest (od strony rachunkowej) traktowany tak samo.

Rezultaty uzyskane z analizy i nałożenie ich na mapę powiatów daje bardzo ciekawy obraz kraju. Potwierdza się znane z innych badań i potocznych obserwacji, że

zróźnicowanie przestrzenne powstaje obecnie w wyniku nałożenia się dwóch porządków: historycznego (zwłaszcza z okresu rozbiorowego) oraz związanego z siecią miejską, a w szczególności siecią dużych miast. Rezultat taki uzyskiwany jest przez wielu autorów zajmujących się zróźnicowaniem przestrzennym kraju niezależnie od tego, jaką wykorzystują metodologię w swoich analizach oraz niezależnie od tego, czy dotyczą one poziomu rozwoju gospodarczego, społeczno-gospodarczego czy też rozwiniętych funkcji społecznych i gospodarczych.

Jednak wskaźnik LHDI posiada również pewne cechy, na które warto zwrócić uwagę. Czasem jest on rozumiany jako wskaźnik jakości życia, co wydaje się interpretacją nadmiernie szeroką. Na jakość życia wpływa szereg czynników niebranych pod uwagę przez LHDI, np. jakość środowiska naturalnego, poczucie bezpieczeństwa fizycznego i ekonomicznego, poziom stresu, poczucie związane ze sprawstwem lub alienacją władzy itp. Wskaźnik ujawnia poziom rozwoju społecznego i gospodarczego, co więcej – jest miarą silnie uśredniającą ten poziom. Uśrednienie to dotyczy nie tylko faktu, że powiat jest jednostką niejednorodną, że obok powiatów grodzkich występują także powiaty ziemskie, a więc posiadające całkiem inną strukturę społeczną i gospodarczą. Uśrednienie to dotyczy także faktu, że wykorzystywane wskaźniki nie zawierają elementu związanego z wewnętrznym zróźnicowaniem społecznym i gospodarczym, a więc np. informacja o systemie pomocy społecznej uwikłana jest wyłącznie w dane dotyczące zamożności ludności, a nie jest wykorzystywana jako wskaźnik rozwarstwienia społeczności lokalnej.

Uśredniająca charakter wskaźnika w pierwszym sensie (dotyczącym zróźnicowania przestrzeni, której dotyczy) jest nieunikniony, jego znaczenie zwiększa się wraz z powiększaniem się jednostek przestrzennych poddawanych analizie. W tym sensie dane dla krajów są uśrednieniem danych dla regionów kraju, a dane dla regionów – uśrednieniem informacji dla układów lokalnych. Nawet jeśli jednostką analizy pozostaje gmina, miary tworzone dla niej nie uwzględniają faktu, że jest to jednostka często niejednorodna, może składać się np. ze strefy rolniczej i np. strefy o rozwiniętej funkcji turystycznej. Oczywiście im mniejsza jednostka analizy, tym rozkład przestrzenny zjawisk poddawanych analizie ujawnia większe zróźnicowanie przestrzeni. Tym samym jest on bardziej zbliżony do rzeczywistości społeczno-gospodarczej.

Wykorzystanie małych jednostek analizy ma jednak również inne konsekwencje, dane dotyczą niewielkiej populacji, małych ilości zdarzeń, tym samym w większym stopniu mogą być przypadkowe. LHDI korzysta z informacji o długości życia dla populacji powiatowych. Sama miara „przewidywana długość życia osoby w wieku 0 lat” powinna być odnoszona do odpowiednio dużych populacji, powiat jest jednostką, dla której może ona być bardzo mało precyzyjna i w sposób przypadkowy zmienna w czasie. Podobnie pewne wątpliwości może budzić korzystanie przez LHDI z danych o strukturze zgonów według przyczyn. O ile dla dużych populacji (krajowej, wojewódzkich) przyczyny zgonów wykazują pewne, względnie stałe, rozkłady oraz charakterystyczne trendy zmian, o tyle dla powiatów dane mogą tracić taki charakter, a dla gmin korzystanie z nich byłoby absurdem. W przypadku gminy poważniejszy wypadek komunikacyjny, np. z udziałem autobusu, może

w sposób zasadniczy zakłócić zarówno dane o długości życia, jak i przyczynach zgonów.

Jeszcze poważniejsze konsekwencje, jak się zdaje, ma jednak drugie rozumienie „uśredniania” danych. Taką samą wartość np. wskaźnika zamożności uzyskuje jednostka, w której społeczność ma relatywnie duży biegun ludności zamożnej i również duży ubogiej, jak i taka, w której nie występują ani ludzie wybitnie zamożni, ani szczególnie ubodzy, a więc mająca charakter społeczności bardziej spójnej wewnętrznie ze względu na poziom zamożności. To samo dotyczy wskaźników wykształcenia, łatwo można sobie wyobrazić jednostki przestrzenne (gminy, powiaty) w których jest relatywnie bardzo mało osób zarówno z wykształceniem wyższym, jak i podstawowym oraz te, w których występuje relatywnie liczna elita osób dobrze wykształconych (mających wykształcenie wyższe) i duża grupa osób z wykształceniem podstawowym lub nawet niepełnym podstawowym. W obu sytuacjach miara oparta na średniej liczbie lat nauki może być bardzo zbliżona, choć z punktu widzenia społecznego są to sytuacje całkowicie różne.

Reasumując, inicjatywa utworzenia miary LHDI jest niezwykle ważna i potrzebna, a uzyskane wyniki (zwłaszcza po naniesieniu na mapę) przekonujące. Wymaga jednak bardzo dogłębnych analiz dotyczących dwóch kwestii:

- 1) interpretacji, określenia, co właściwie jest przedmiotem pomiaru (poziom rozwoju społecznego, społeczno-gospodarczego, jakości życia itp.) oraz
- 2) refleksji metodologicznej dotyczącej użytych wskaźników, ich znaczenia dla miary syntetycznej (mocy wskaźnikowej) oraz stopnia przypadkowości (zmienności w czasie) wobec niewielkich populacji, do których się odnoszą.

Zmiana wielkości jednostek analizy w kierunku ich uszczegółowienia (a więc zmniejszenia rozmiaru, a zwiększenia ich liczby) z konieczności musi prowadzić do wzrostu stopnia złożoności użytych miar dla cech tak syntetycznych (pochodzących z języka teoretycznego, wysoce abstrakcyjnego), jak poziom rozwoju społeczno-gospodarczego, jakość życia itp. Komplikacja ta jest wynikiem faktu, że osiągnięcie takiego samego celu, jakim jest rozwój (termin z języka teorii) może dokonywać się na różne sposoby, a użyta miara nie powinna preferować sposobu osiągania celu, ale stopień zbliżenia się do niego. Oczywiście sam cel powinien być jasno określony i sformułowany w taki sposób, aby uwzględniał różne strategie zbliżania się do niego. Powoduje to poważne kłopoty z operacjonalizacją pojęcia „rozwój”, która z konieczności musi wykorzystywać całą baterię wskaźników empirycznych i relacji między nimi. Ztraca się zatem „elegancja” miary (jej prostota formalna) na rzecz wzrostu trafności.

Przykładu prób rozwiązania tego rodzaju problemów dostarcza opracowanie miary „poziom rozwoju społeczno-gospodarczego” wykorzystana dla agregacji gminami i ograniczona do obszarów wiejskich, a wykorzystana w pracach IRWiR [por. Rosner (red.) 2007; Rosner, Stanny 2014]. Autorzy przyjęli kilka założeń:

1. Głównym problemem rozwoju obszarów wiejskich na obecnym etapie zaawansowania procesów kształtowania się wydolnych struktur gospodarczych i społecznych jest dezagraryzacja rozumiana jako stopień przekształceń lokalnych struktur gospodarczych w kierunku umożliwiającym zmniejszenie udziału zatrudnienia rolniczego (zmniejszenie udziału rolnictwa w strukturach źródeł utrzymania).

2. Realizacja tego celu może odbywać się poprzez:

- rozwój pozarolniczego rynku pracy na wsi,
- migracje nadwyżek zasobów pracy do ośrodków pozarolniczych (miast),
- wykorzystanie sąsiedztwa pozarolniczych rynków pracy (migracje wahadłowe).

W zależności od lokalnych uwarunkowań może dokonywać się poprzez zmiany proporcji gospodarstw domowych rolników i gospodarstw bezrolnych lub poprzez wzrost wielodochodowości gospodarstw domowych rolników (użytkowników ziemi).

3. Destymulantą rozwoju jest dezagraryzacja poprzez ucieczkę z rynku pracy w bierność zawodową lub w kierunku systemów wspierania dochodów przez inne niezarobkowe źródła utrzymania (zasiłki zarówno dla bezrobotnych, jak i z systemu pomocy społecznej). Zjawisko tego rodzaju ucieczki z rynku pracy (statystycznie zbieżne z dezagraryzacją) sprzyja brakowi spójności w wymiarze przestrzennym.

4. Proces dezagraryzacji struktury źródeł utrzymania nie prowadzi do rozwoju społeczno-gospodarczego, jeśli odbywa się kosztem wzrostu braku równowagi na rynku pracy.

5. Społeczności lokalne wykazujące wyższy stopień rozwoju społeczno-gospodarczego wykazują się większą możliwością kreowania lokalnych finansów publicznych oraz większym kapitałem uprzedmiotowionym w postaci lokalnej infrastruktury zarówno technicznej, społecznej, jak i instytucjonalnej.

6. Osiągnięcie relatywnie wyższego poziomu rozwoju społeczno-gospodarczego kreuje nowe potrzeby w zakresie edukacji, udziału w samorządności lokalnej, stowarzyszania się w grupy o różnych interesach i polach aktywności. Łączy się to ze wzrostem potrzeby poczucia sprawstwa w zakresie działań na szczeblu lokalnym.

Wychodząc z takich założeń, przyjęto, że rozwój społeczno-gospodarczy jest pojęciem, które w zakresie konotacji zawiera komponenty gospodarcze i społeczne. Wśród komponentów gospodarczych wyabstrahowano pięć subkomponentów:

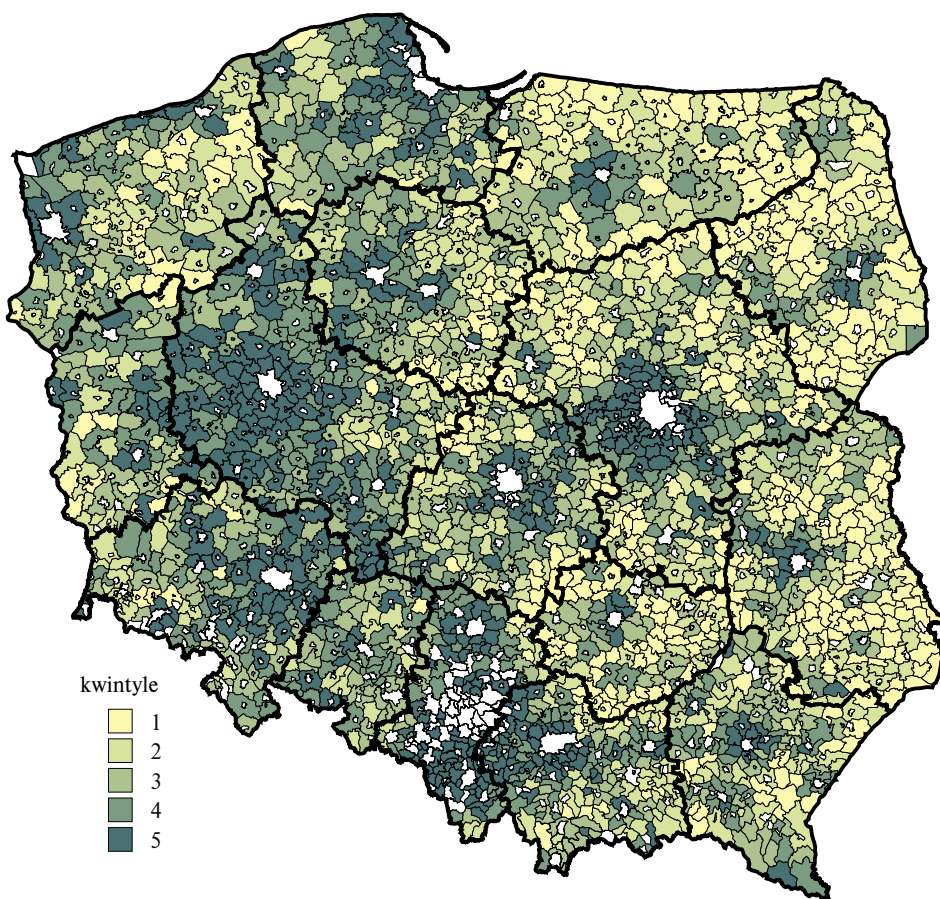
- 1) poziom dezagraryzacji,
- 2) charakter (ekonomiczny) sektora rolnego,
- 3) charakter (ekonomiczny) sektora pozarolniczego,
- 4) stopień równowagi rynku pracy oraz
- 5) zasobność zarówno „bieżąca” jak i skumulowaną z okresów wcześniejszych w majątku gminnym.

Komponent społeczny rozwoju zawiera trzy elementy:

- 1) demograficzny,
- 2) edukacyjny oraz
- 3) związany z aktywnością społeczną.

Praktyczna realizacja pomiaru poziomu rozwoju społeczno-gospodarczego tak małych jednostek, jak gminy (obszary wiejskie gmin miejsko-wiejskich) dostarczała wielu problemów, wśród których warto wymienić trzy.

1. Liczba zdarzeń (np. demograficznych) w małej populacji jest niestabilna w czasie, należy precyzyjnie określić których wskaźników empirycznych to dotyczy i dla każdego znaleźć sposób na ich „ustabilizowanie”. Jednocześnie wyklucza to możliwość wykorzystywania takich miar, jakie ze swojej natury mają sens tylko dla dużych populacji, a więc np. danych z tablic trwania życia. Taką „stabilizację” miar



RYSUNEK 2. Syntetyczna ocena poziomu rozwoju społeczno-gospodarczego obszarów wiejskich
 FIGURE 2. Synthetic evaluation of socio-economic development level of rural areas.

Źródło: Rosner A. (red.), 2007: *Zróżnicowanie poziomu rozwoju społeczno-gospodarczego obszarów wiejskich a zróżnicowanie dynamiki przemian*. IRWIR PAN, Warszawa.

Source: A. Rosner (ed.), 2007: *Zróżnicowanie poziomu rozwoju społeczno-gospodarczego obszarów wiejskich a zróżnicowanie dynamiki przemian, (Socio-economic development level of rural areas diversity, and diversity of change dynamics)*. IRWIR PAN, Warszawa.

opartych na zdarzeniach, np. demograficznych (urodzenia, zgony, migracje), można uzyskać, obliczając wskaźnik średni np. z trzech lat.

2. W wielu przypadkach, konstruując miary trzeba korzystać (w przypadku gmin miejsko-wiejskich) z danych dla całych gmin. Dotyczy to np. elementów infrastruktury punktowej, takiej jak ośrodki zdrowia, który z reguły zlokalizowany jest w mieście, ale obsługuje również teren wiejski. Podobnie w przypadku finansów publicznych nie istnieje oddzielny budżet dla miasta i obszarów wiejskich.

3. Operacjonalizacja subkomponentów prowadzi do wykorzystania dużej liczby wskaźników empirycznych. Między nimi mogą zachodzić liczne związki merytoryczne, np. charakter procesów migracyjnych ma związek zarówno z tym, jak

kształtuje się lokalna struktura demograficzna, sytuacja na rynku pracy, zatrudnienie w gospodarstwach rodzinnych itp. Bezrobocie rejestrowane jest charakterystyczną formą bezrobocia dla ludności w gospodarstwach domowych bez użytkownika gospodarstwa rolnego, a w konsekwencji stopa bezrobocia jawnego pozostaje w związku ze strukturą gospodarstw domowych. Zależności tego rodzaju jest dużo, niekiedy są to związki bardzo skomplikowane. Abstrahowanie od ich występowania może prowadzić do sytuacji niekontrolowanego wzrostu wagi pewnych zjawisk względem innych. Dobór wskaźników i ew. wprowadzenie dla nich wag musi brać pod uwagę te zależności. Nie wystarcza więc tylko statystyczna analiza zależności między wskaźnikami i eliminacja tych, między którymi zachodzą silne związki korelacyjne. Potrzebna jest również (a może zwłaszcza) refleksja nad merytorycznymi uwikłaniami występującymi między wskaźnikami.

Rezultat przeprowadzonych przez zespół IRWiR analiz prowadził do wyników przedstawionych na rysunku 2. Uzyskany rozkład potwierdza główne zależności przestrzenne ujawniane przez rozkład indeksu LHDI, pokazuje jednak, że powiat (jednostka analizy przy zastosowaniu miary LHDI) zwykle jest wewnątrznie silnie zróżnicowany.

UWAGI KOŃCOWE

Celem opracowania jest jednak nie tyle prezentacja wyników, ile wniosków natury metodologicznej dotyczących procedur pomiaru zmiennych języka teorii, takich jak poziom rozwoju społeczno-gospodarczego. Większość z nich wspomniana była wcześniej, warto jednak jeszcze raz je zebrać.

1. O ile dla dużych jednostek, takich jak kraje, można tworzyć względnie eleganckie formalnie wskaźniki dla pojęć bardzo uogólnionych, to w miarę przechodzenia do agregacji według mniejszych jednostek konieczne jest odchodzenie od elegancji na rzecz uwzględniania zróżnicowania sposobów realizacji celu, jakim jest np. rozwój. Eleganckie formalnie wskaźniki, takie jak PKB, PPP, a nawet HDI, są podatne na krytykę podnoszącą, że *implicite* zakładają skalę wartości świata zachodniego. Obrona przed tym zarzutem możliwa jest wyłącznie kosztem zmniejszenia elegancji formalnej i rozbudowy liczby wskaźników oraz relacji między nimi.

2. Skale służące do pomiaru zmiennych w wysokim stopniu uogólniających zjawiska społeczne i gospodarcze mają niemal zawsze charakter skal porządkowych, to znaczy odzwierciedlają wyłącznie relacje poprzedzania (ew. lokowania się za) oraz relację równości położenia na skali. Nawet jeśli jednostką pomiaru jest wartość mierzalna na skali silniejszej logicznie (np. dolar w przypadku miary PKB), to interpretacja miary o mianie dolara będzie porządkowa, co jest związane z charakterem cechy mierzonej. Warto zauważyć, że jeśli PKB *per capita* w jednostce „a” wynosi 100, a w jednostce „b” 200, zaś jednocześnie w jednostce „c” 10 000, natomiast w „d” 10 100, to dystans rozwojowy między „a” i „b” wydaje się większy niż między „c” i „d”, mimo że w wymiarze dolarów w PKB *per capita* jest identyczny. Najogólniej można powiedzieć, że pomiar tego rodzaju cech (pojęć języka teorii), jak poziom rozwoju, prestiż społeczny, kapitał społeczny itp., z natury samej cechy mierzalnej wyznacza typ logiczny skali pomiarowej.

3. Konsekwencją porządkowego charakteru skali pomiarowej jest to, że nie można określić, czy rozkład wykazuje duże czy też małe zróżnicowanie, oraz to, że pozostaje kwestią nierozstrzygniętą, czy jest on zbliżony do tzw. rozkładu normalnego czy też ma charakter biegunowy. Doświadczenia podpowiadają, że poza wyjątkowymi sytuacjami bardziej wiarygodne wydaje się założenie rozkładu normalnego i posługiwanie się miarami opartymi na liczebności podzbiorów (mediana, kwartał itp.) niż na interwałach skali (w konsekwencji w prezentacji danych wskazane jest wykorzystywanie przedziałów równolicznych, a nie o równych interwałach osi skali).

4. Dokonując operacjonalizacji pojęć uogólniających zjawiska społeczne i gospodarcze, trzeba przede wszystkim dokładnie określić ich konotację. Wyznacza ona pola merytoryczne, które powinny znajdować odzwierciedlenia w zbiorze przyjętych wskaźników. Pozostaje otwarty problem, czy fakt, iż rozkłady wskaźników wykazują związki statystyczne między nimi oznacza, że wskazują one to samo *indicatum* i że należy przeprowadzić ich redukcję do indeksów niepowiązanych. Doświadczenia z analiz prowadzonych w IRWiR prowadzą do wniosku, że nie istnieje w tym względzie bezwyjątkowa reguła, jeśli zmienne te dotyczą różnych pól merytorycznych wyróżnionych w procesie analizy konotacji pojęcia będącego przedmiotem badań – redukcja taka jest niewskazana. Jeśli są to zmienne z tego samego pola merytorycznego, można rozważyć specyfikę każdego z rozwiązań: redukcji ich do jednego, pozostawienia obu skorelowanych wskaźników lub pozostawienia obu ze zmniejszoną wagą. Wybór sposobu postępowania należy do autora i zawsze zawiera element arbitralności.

BIBLIOGRAFIA

- Arak P., Ivanov A., Peleah M., Płoszaj A., Rakocy K., Rok J., Wyszkowski K., 2012: *Krajowy Raport o Rozwoju Społecznym. Polska 2012. Rozwój regionalny i lokalny*. Ministerstwo Rozwoju Regionalnego – Biuro Projektowe UNDP w Polsce, Warszawa.
- Nowak S., 1965: *Studia z metodologii nauk społecznych*. PWN, Warszawa.
- Pawłowski T., 1966: Rodzaje skal a struktura zdań stwierdzających związki między wielkościami. [w:] *Logiczna Teoria Nauki*. PWN, Warszawa.
- Rocznik Statystyki Międzynarodowej 2012*. GUS Warszawa 2012.
- Rosner A. (red.), 2007: *Zróżnicowanie poziomu rozwoju społeczno-gospodarczego obszarów wiejskich a zróżnicowanie dynamiki przemian*. IRWiR PAN, Warszawa.
- Rosner A., Stanny M., 2014: *Monitoring rozwoju obszarów wiejskich. Etap I. Przestrzenne zróżnicowanie poziomu rozwoju społeczno-gospodarczego obszarów wiejskich w 2010 roku*. Fundacja Europejski Fundusz Rozwoju Wsi Polskiej, IRWiR PAN, Warszawa.
- Stiglitz J.E., Sen A., Fitoussi J.P., 2013: *Błąd pomiaru. Dlaczego PKB nie wystarcza. Raport Komisji ds. Pomiaru Wydajności Ekonomicznej i Postępu Społecznego*. PTE, Warszawa.
- Suppes P., Zinnes J., 1962: *Basic measurement theory*. Institute for Mathematical Studies in the Social Science, Stanford University, Technical Report No. 45.

THE ISSUE OF SOCIO-ECONOMIC DEVELOPMENT LEVEL. MEASUREMENT SCALE AND ITS PROPERTIES

Abstract: The subject of the analysis are methodological restrictions of variable measurements defined in theoretical terms (on a high level of abstraction), such as the level of socio-economic development. It defends the thesis relating to the fact, that even if all components of the measurement (empirical indicators taking part in creation of measurement scale of such concept) share the character of measurement on at least interval scales, still the synthetic measurement scale shows a weaker, ordinal character. The sentences using results of such measurement have a character of historical and spatial relativization.

The article also talks about a relation between the simplicity (formal elegance) of measurement algorithm and spatial aggregation of analyzed units. This relation leads to the point where the smaller spatial units, the more complex algorithm is, which is shown as a result of multiply possible ways leading to the realization of the same goal (for example the progress in reached level of development).

Key words: socio-economic development, development's spatial diversity, measurement scale, operationalization of theoretical terms