

Sugerowane cytowanie:

Michalik, A. (2017). Statystyka a wnioskowanie logiczne w pracy naukowej: kilka uwag na temat najczęstszych błędów i pułapek w procesie badawczym. W: M. Maciejewski, K. Wach (red.), *Handel zagraniczny i biznes międzynarodowy we współczesnej gospodarce*. Kraków: Uniwersytet Ekonomiczny w Krakowie, s. 481-492.

Statystyka a wnioskowanie logiczne w pracy naukowej: kilka uwag na temat najczęstszych błędów i pułapek w procesie badawczym*

Adam Michalik

Uniwersytet Ekonomiczny w Krakowie
Wydział Ekonomii i Stosunków Międzynarodowych
Katedra Handlu Zagranicznego
ul. Rakowicka 27, 31-510 Kraków
e-mail: michalia@uek.krakow.pl

Streszczenie:

Liczne badania w dziedzinie handlu międzynarodowego bazują na danych statystycznych pozyskiwanych z instytucji krajowych i międzynarodowych. Analiza tych danych wskazuje na bardzo dużą rozbieżność pomiędzy nimi, która może prowadzić w efekcie badacza danego zjawiska do całkowicie fałszywych wniosków. Podobny efekt może przynieść zgodna z poprawnością obliczeniową ale błędna pod względem logicznym interpretacja wyników prowadzonych badań. Niniejszy artykuł podkreśla znaczenie wiarygodności i rzetelności analiz ekonomicznych prowadzonych w oparciu o dane statystyczne, a także identyfikuje w tej sferze wybrane błędy i pułapki z jakimi może zetknąć się badacz zjawiska.

Słowa kluczowe: badania naukowe, statystyka, metodyka badań

Klasyfikacja JEL: C82, F14

* Publikacja została sfinansowana ze środków przyznanych Wydziałowi Ekonomii i Stosunków Międzynarodowych Uniwersytetu Ekonomicznego w Krakowie, w ramach dotacji na utrzymanie potencjału badawczego.

1. WSTĘP

„Są trzy rodzaje kłamstw: kłamstwa, bezczelne kłamstwa i statystyka” – ten cytat przypisywany Markowi Twainowi wydaje się dezawuować znaczenie tej nauki oraz danych statystycznych w wyjaśnianiu praw otaczającego nas świata. Tak naprawdę wskazuje on jednak tylko na pułapkę wynikającą z automatycznego przypisywania danym statystycznym mocy prawdy objawionej i bezkrytycznego spojrzenia prezentowanego w wielu analizach i pracach o charakterze naukowym czy publicystycznym. W niniejszym artykule zostaną przedstawione wybrane błędy i pułapki, związane z niewłaściwym wykorzystaniem danych statystycznych, z jakimi spotykamy się zarówno w pracy naukowej, publikacjach różnych autorów,

a także pracach dyplomowych studentów. Celem artykułu, co należy podkreślić jednak z całą mocą, nie jest przekreślanie sensu badań statystycznych, lecz zwrócenie uwagi na dość istotną rangę problemu wiarygodności wyników i wniosków obciążonych poważnymi wadami popełnionymi na etapie metodyki pracy badawczej, choć często będącymi poza kontrolą badacza.

2. DANE, INFORMACJA A WIEDZA

Jak słusznie podkreśla B. Nadolna (2009, s.135) celem nauki jest gromadzenie wiedzy w celu zrozumienia świata. Tymczasem zasadniczym problemem w badaniach naukowych może być opisanie za pomocą liczb i faktów otaczającej nas rzeczywistości w taki sposób, aby proces badawczy prowadził do powiększania zasobu wiedzy na jej temat. W świecie fizyki czy chemii możemy posłużyć się niejednokrotnie eksperymentem przeprowadzonym w warunkach laboratoryjnych, przez nas odpowiednio przygotowanych, od początku do końca eksperymentu śledząc całość procesu. Tymczasem badacz zjawisk ekonomicznych (w ujęciu makro), w tym handlu międzynarodowego w procesie analizy przepływów towarów, usług czy kapitału jest zmuszony do korzystania ze źródeł wtórnych zawierających potrzebne dane makroekonomiczne. Wobec powyższego deterministycznego znaczenia nabiera problem dostępności, kompletności oraz przede wszystkim wiarygodności danych, w oparciu o które przeprowadzany jest proces badawczy.

B. Sundgren (Stefanowicz, 2013, s. 17) określa dane jako „wycinek rzeczywistości przystosowany do reprezentowania innego wycinka rzeczywistości”. Ujęcie to zwraca uwagę na funkcję danych jako swoistą podstawę wnioskowania statystycznego na temat rzeczywistości.

Drugim czynnikiem niezbędnym w procesie gromadzenia wiedzy jest informacja. Może ona wynikać bezpośrednio z danych, bądź stanowić ich uzupełnienie, poprzez umiejscowienie danych w szerszym kontekście. Jak zauważa B. Stefanowicz (2013, s.8) w najogólniejszym ujęciu informację określa się jako „odbicie (odwzorowanie) różnorodności cechującej otaczającą rzeczywistość (obiekt, zdarzenie, proces, zjawisko)”. Natomiast zdaniem E. Niedzielskiej (1988, s. 20), informacja „jest specyficznym dobrem niematerialnym, które w miarę postępu gospodarczego oraz rozwoju środków i form komunikowania się społecznego nabiera coraz większego znaczenia, przeobrażając oblicze wielu tradycyjnie zorganizowanych gospodarek świata”. Ten pogląd znajduje w pełni uzasadnienie w dobie cyfrowego świata XXI wieku, kiedy informacja (bądź dezinformacja) stała się orężem w walce politycznej i ekonomicznej nieraz o większej skuteczności niż działania o charakterze militarnym. Wraz z rozwojem świata Internetu zwiększył się nie tylko dostęp do informacji, ale także wzrosła możliwość popularyzacji informacji fałszywych, czy to poprzez świadome działania manipulacyjne, czy też w wyniku rozprzestrzeniania się błędnych przekazów dzięki nieświadomym „posłańcom”, którzy nie poddali kolportowanej treści procedurze weryfikacji.

Rozpatrując problem wiarygodności i prawdziwości przekazu informacyjnego, należy pamiętać, iż jedną z właściwości informacji jest poddanie jej procesowi przetworzenia, odbywającemu się na trzech poziomach (Stefanowicz, 2013, s.55):

- na poziomie danych wchodzących w skład odebranego komunikatu. Zjawisko jest to typowe dla osób przywiązujących wagę do wybranych danych statystycznych i faktów bez podjęcia próby ich usytuowania w określonym kontekście,
- na poziomie wzajemnego połączenia danych – a więc na poziomie sformułowanego całościowo komunikatu. Zestawienie przekazywanych treści może prowadzić do poddania odbiorcy procesowi zamierzonej manipulacji, bądź w sposób niezamierzony wywołać u odbiorcy błędną interpretację,
- na poziomie poszerzonego spojrzenia na odebraną informację jako elementu większego zasobu opisującego rozpatrywany wycinek rzeczywistości. Mnogość informacji i komunikatów składających się na dany zasób informacyjny może utrudniać prowadzony wywód o charakterze analitycznym.

Wspomniane właściwości powinny prowadzić badacza do kilku istotnych wniosków o charakterze ogólnym:

1. Konieczna w procesie badań jest weryfikacja posiadanych danych oraz wszelkich informacji istotnych przy powstawaniu wyводу naukowego;
2. Niezbędne w procesie badawczym staje się poszukiwanie potwierdzenia zgromadzonych danych i informacji z innych źródeł, co prowadzi do konieczności pozyskiwania informacji z pozorów nadmiarowych, które należy traktować jako kontrolne;
3. Badacza naukowego powinna cechować dociekliwość, a także zdolność krytycznego spojrzenia wobec dostępnych zasobów informacyjnych, tak aby umiał ocenić wiarygodność danych wykorzystywanych w procesie badawczym;
4. Wobec nieuniknionych ograniczeń w procesie poznawczym prezentowane na jego podstawie tezy powinny charakteryzować się pewną „ostrożnością wnioskowania”.

Jak zauważa J. Brdulak (2005, s.15) to właśnie zdolność interpretacji i prawidłowego przetworzenia informacji na podstawie procesu wnioskowania prowadzi do powstania wiedzy. Chociaż, jak zaznacza T. Kaczmarek, „w znaczeniu ogólnym wiedzę stanowi zbiór informacji, jakie człowiek przyswaja sobie poprzez narządy zmysłów” (...), to „wiedza jest tym, co przekształca informację w nową informację lub w nową wiedzę, czy nowe preferencje, a preferencje są tym, co wskazuje jakiej wiedzy należy użyć”. (2010, s.14-23). Nieco szerzej natomiast pojęcie wiedzy ujmuje A. Zaliwski (2000, s.23) wskazując jej funkcjonalny charakter i stwierdzając, iż “wiedza wykracza poza informacje, gdyż implikuje zdolność do rozwiązywania problemów, do inteligentnego zachowania się i działania. (...) Tak pojmowaną wiedzę utożsamia się ze zbiorem reguł (bazą wiedzy), podczas gdy informacje utożsamia się z bazą faktów”.

Ponieważ niniejszy artykuł został przygotowany na Konferencję Handlu Międzynarodowego, będącą jednocześnie ukoronowaniem Jubileuszu Pracy Prof. Stanisława Wydymusa, którego dorobek naukowy związany jest głównie z procesem

analiz i wnioskowania statystycznego, dalsza jego część skupi się na problematyce wykorzystania danych statystycznych w pracy badawczej, oraz wskazaniu trudności oraz błędów niejednokrotnie ten proces badawczy determinujących.

Analizując obszary pracy badawczej prowadzonej w oparciu o dane statystyczne, można wyróżnić zasadniczo dwie główne płaszczyzny, w ramach których podjęty wywód naukowy może prowadzić do błędnych wniosków:

- Zewnętrzną – będącą poza kontrolą i wpływem badacza, związaną z ułomnością danych statystycznych publikowanych przez instytucje zewnętrzne;
- Wewnętrzną – związaną z błędnie przyjętą przez badacza metodyką działań, bądź błędnym wnioskowaniem w oparciu o poprawnie dobraną metodykę.

3. WNISKOWANIE A RZETELNOŚĆ DANYCH

W literaturze przedmiotu można znaleźć sformułowanie, iż wnioskowanie statystyczne należy do wnioskowań uprawdopodobniających albo subiektywnie niepewnych (Daszkowski, 2014; Stanosz, 1988; Ziemiński, 1993) Tymczasem należy uświadomić sobie istnienie jednego dość istotnego z punktu widzenia badań naukowych problemu. Otóż dane statystyczne prezentowane przez poszczególne uprawnione do tego instytucje bardzo często odbiegają od rzeczywistości. I ta różnica może być na tyle znacząca, że w efekcie użyteczność powstałych na tej podstawie wniosków będzie znikoma. Bardzo celnie problem ten określa E. Mączyńska (2013, s.44), pisząc o swoistym „teatrze liczb, gdzie dane i statystyki liczbowe tworzą kurtynę, która odcina nas od rzeczywistości”.

Powodów tego zjawiska można upatrywać w różnorodnych przyczynach, począwszy chociażby od niedokładnych metodologii badań instytucji gromadzących dane (np. pomijających zdarzenia o mniejszych wielkościach, które choć indywidualnie nie mają wpływu na prezentowane dane, to jednak po zsumowaniu wszystkich takich przypadków może się okazać, iż ich pominięcie było z perspektywy prezentacji dokładnego obrazu rzeczywistości błędem), przez mnogość informacji podlegających raportowaniu (co niejednokrotnie utrudnia ich precyzyjną prezentację), kończąc na błędzie ludzkim.

Z perspektywy analiz ekonomicznych niewątpliwie najistotniejszym problemem jest narastający dynamizm procesów zachodzących w gospodarce światowej oraz wszechobecny wpływ internacjonalizacji działalności przedsiębiorstw. W efekcie statystyka międzynarodowa nie jest w stanie zidentyfikować precyzyjnie efektów dynamicznych działań podejmowanych na różnych kontynentach, w postaci np. przepływów kapitałowych pomiędzy poszczególnymi podmiotami. W ciągu danego dnia mnogość zawieranych przez największe firmy i korporacje międzynarodowe transakcji, często o przeciwnych kierunkach przepływów finansowych powoduje, iż w raportach statystycznych znajdziemy jedynie pewną przybliżoną wartość końcową tych działań, natomiast nie będziemy w stanie odtworzyć ich charakteru. Dla ilustracji tego problemu warto posłużyć się przykładem amerykańskiej firmy X, posiadającej szereg powiązanych ze sobą kapitałowo podmiotów w różnych częściach

świata, w tym także w Polsce. Otóż firma ta wśród szeregu swoich transakcji postanowiła dokapitalizować własną inwestycję w Rosji, jednak czyniąc to za pośrednictwem spółki córki zlokalizowanej w Polsce. W najprostszym przypadku działanie takie może spowodować zmianę odnotowanej przez Narodowy Bank Polski (NBP) wartości inwestycji amerykańskich w Polsce (w wyniku transferu kapitału z firmy amerykańskiej do polskiej spółki) oraz wzrost wartości polskich inwestycji w Rosji (transfer tej samej wielkości kapitału do spółki zlokalizowanej w Rosji). Na podstawie tych danych badacz analizujący przykładowo atrakcyjność inwestycyjną Polski w oczach firm amerykańskich może wysnuć błędny wniosek o wzroście napływu inwestycji amerykańskich do naszego kraju.

Lecz należy pamiętać, że zaprezentowany przykład charakteryzuje się uproszczeniem w stosunku do obserwowanej rzeczywistości gospodarczej. Bardzo często ten napływ kapitału do Polski może być realizowany przez amerykański koncern z wykorzystaniem innych podmiotów zlokalizowanych w różnych krajach. Także dalszy transfer kapitału „na trasie Polska – Rosja” może odbywać się (i zwykle odbywa się) z wykorzystaniem podmiotów zlokalizowanych w krajach o liberalnej jurysdykcji podatkowej (np. w Luksemburgu, bądź na Cyprze). W efekcie badacz wykorzystujący w swojej analizie dane Narodowego Banku Polski nie będzie w stanie wychwycić wzrostu zaangażowania kapitałowego polskiej spółki w Rosji, tylko odnotuje przykładowo wzrost polskich inwestycji w Luksemburgu. Wobec powyższego, gdyby bezrefleksyjnie przyjąć jako podstawę analizy dane dotyczące polskich inwestycji zagranicznych prezentowane przez Narodowy Bank Polski, można by postawić fałszywą z gruntu naukowego tezę, iż głównymi krajami napływu polskich bezpośrednich inwestycji zagranicznych w XXI wieku były Cypr i Luksemburg (NBP, 2017). Problem przydatności danych prezentowanych przez NBP do analizy naukowej pogłębia fakt, iż w przypadku kilku państw wartości polskiego zaangażowania inwestycyjnego są ujemne, co wynika m.in. z faktu zobowiązań polskich firm względem swoich zagranicznych spółek „córek”, za pośrednictwem których pozyskują kapitał z emisji obligacji na rynku europejskim, przewyższających kwotę zainwestowanego w te podmioty kapitału.

Kolejnym problemem dla badacza jest aktualizacja danych prezentowanych przez poszczególne instytucje wraz z kolejnymi corocznymi zestawieniami. Porównajmy w tym celu dwa raporty NBP o polskich inwestycjach zagranicznych z 2015 i 2016 roku. Tabela 1. prezentuje „Transakcje z tytułu polskich inwestycji bezpośrednich...” według danych zawartych w raporcie NBP z 2015 roku, natomiast Tabela 2. według raportu NBP z roku kolejnego.

Jak widać z przytoczonego zestawienia danych publikowanych przez NBP dla tego samego okresu, ale w dwóch kolejnych latach, różnice w prezentowanych danych są w niektórych przypadkach olbrzymie. Przykładowo suma transakcji z tytułu polskich inwestycji bezpośrednich w 2012 roku w raporcie z 2015 roku była ujemna, co miałyby świadczyć o wycofaniu kapitału do Polski na kwotę ponad 8,6 mld zł, natomiast według aktualizacji rok później wartość tej pozycji była dodatnia i to na kwotę ponad 9,4 mld zł!. W efekcie w zależności od tego, z którego raportu

zostaną pozyskane dane do analizy naukowej, prezentowane liczby mogą diametralnie się różnić doprowadzając do falsyfikacji wniosków.

Tabela 1. Dane dotyczące wielkości transakcji z tytułu polskich BIZ w raporcie NBP z 2015 roku (w mln PLN)

Rok	Akcje i inne formy udziałów kapitałowych	Reinwestycje zysków	Instrumenty dłużne	Ogółem
2011	9 776	636	465	10 877
2012	12 523	4 215	-342	-8 650
2013	1 202	-672	-4 784	-4 255
2014	12 938	1 536	-9 441	5 033

Źródło: dane NBP, www.nbp.pl (dostęp: 14.09.2017 r.).

Tabela 2. Dane dotyczące wielkości transakcji z tytułu polskich BIZ w raporcie NBP z 2016 roku (w mln PLN)

Rok	Akcje i inne formy udziałów kapitałowych	Reinwestycje zysków	Instrumenty dłużne	Ogółem
2011	1 940	636	465	3 041
2012	5 574	4 215	-342	9 447
2013	4 030	-672	-4 784	-1 426
2014	17 481	1 610	-9 948	9 142

Źródło: dane NBP, www.nbp.pl (dostęp: 14.09.2017 r.).

Powyższe przykłady, wskazujące na ograniczoną użyteczność danych prezentowanych przez Narodowy Bank Polski w zakresie bezpośrednich inwestycji zagranicznych, zostały przytoczone jedynie jako przykładowa ilustracja omawianego problemu, gdyż także pozostałe źródła danych nie są wolne od istotnych wad utrudniających proces analizy naukowej. Dane o inwestycjach zagranicznych prezentowane przez Polską Agencję Inwestycji i Handlu bazują na deklaracjach firm i uwzględniają np. tylko inwestycje o wartości przekraczającej 1 mln USD, co powoduje, iż część przepływów kapitałowych nie jest uwzględniona w raportach tej instytucji. Podobne ograniczenia charakteryzuje metodyka badań GUS.

Wspomniana powyżej znacząca odmienność danych prezentowanych przez poszczególne instytucje prowadzi w konsekwencji także do braku możliwości przeprowadzenia rzetelnych analiz na poziomie zarówno globalnym, krajowym jak i lokalnym. Przykładowo analizując znaczenie Rosji jako kraju lokalizacji polskich inwestycji bezpośrednich badacz jest zmuszony odwołać się do danych polskich, prezentujących strukturę geograficzną polskich Zagranicznych Inwestycji Bezpośrednich (ZIB) i miejsce Rosji w tej strukturze, jak i rosyjskich, prezentujących miejsce Polski jako inwestora w Rosji na tle innych krajów, czy też zróżnicowanie regionalne polskich inwestycji w Rosji. Tymczasem według danych Narodowego Banku Polski wartość netto polskich ZIB w Rosji wyniosła w 2015 roku 54,8 mln USD (NBP, 2017), a według danych Centralnego Banku Rosji 20 mln USD (CBR, 2017).

Przykład ten pokazuje, iż taka pogłębiona analiza na podstawie tych danych jest znacząco utrudniona, o ile pod kątem statystycznym w ogóle możliwa¹.

Podobne, choć nieco mniejsze trudności badacz może napotkać analizując sytuację w polskim eksporcie. Tabela 3. przedstawia widoczne różnice w prezentowanej wartości polskiego eksportu w latach 2013 - 2014 według wybranych czterech źródeł danych.

Tabela 3. Dane o wartości eksportu Polski w latach 2013-2014 wg baz danych wybranych instytucji (w mld USD)

Eksport Polski	Narodowy Bank Polski	Główny Urząd Statystyczny	Konferencja Narodów Zjednoczonych ds. Handlu i Rozwoju	Bank Światowy
2013 r.	198,1	206,1	204,9	205,6
2014 r.	210,6	222,3	220,2	214,5
% zmiana 2014/2013	6,3	7,8	7,4	4,3

Źródło: opracowanie własne na podstawie (NBP, 2017), (MR, 2017), (UNCTAD, 2017) oraz (Bank Światowy, 2017).

Według przytoczonych danych badacz przeprowadzający analizę rozwoju polskiego eksportu, w przypadku skorzystania z bazy danych Banku Światowego, napisze w swojej publikacji, że wartość polskiego eksportu w 2014 roku wzrosła w stosunku do roku poprzedniego o 4,3%, a w przypadku wykorzystania danych GUS, zamieści informację, że wzrost ten odnotowano na poziomie 7,8%.

Powyższe przykłady dowodzą skali problemu doboru i oceny wiarygodności danych w procesie badań naukowych. Udostępniana na stronach internetowych poszczególnych instytucji metodyka badań nie stanowi precyzyjnego wytłumaczenia różnic w prezentowanych danych, a co istotniejsze nie przesądza o stopniu ich użyteczności w procesie badawczym.

4. BŁĘDY WNISKOWANIA POPEŁNIANE PRZEZ BADACZA

We wcześniejszym punkcie wskazano przykłady utrudnień i pułapek, jakie nieśie dla badacza zmierzenie się ze zróżnicowaną na poziomie różnych krajów i często niewyjaśnioną przez poszczególne instytucje metodyką zbierania i opracowania danych, a także jak czynnik zewnętrzny, jakim jest postępująca dynamika przepływów towarowo – kapitałowych w dzisiejszym zglobalizowanym świecie, utrudnia obserwację i analizę tych procesów.

Niestety w pracach o charakterze naukowym, czy publicystycznym można także spotkać błędy wynikające z winy autora, który *stricte* statystyczną analizę danych przedłożył nad proces o charakterze myślowo-logicznym. Przyczyn tego zjawia-

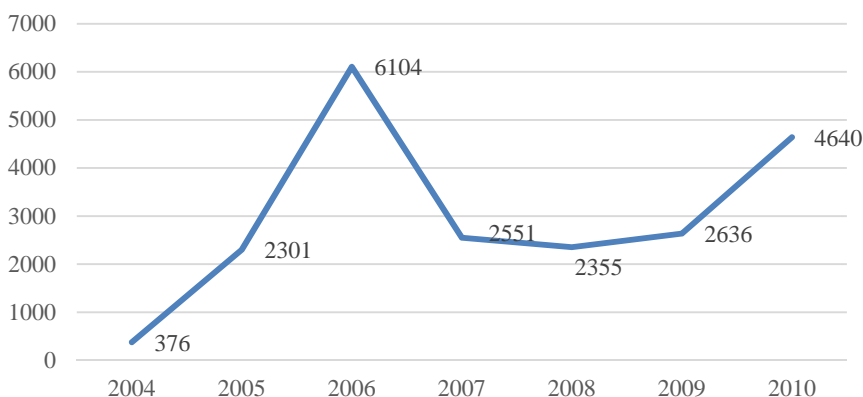
¹ Na koniec tego wątku należy dodać, iż odrębnym problemem pozostaje rozstrzygnięcie, czy w obliczu sytuacji, kiedy amerykańska firma transferuje kapitał do Rosji z wykorzystaniem własnej spółki córki zlokalizowanej w Polsce, powinno się w analizie dotyczącej aktywności inwestycyjnej polskich przedsiębiorstw na rynkach zagranicznych takie transakcje uwzględniać.

ska naukowcy upatrują w słabościach ludzkiego umysłu, podatnego na różnego rodzaju uproszczenia i złudzenia. F. Bacon w tzw. teorii idoli przypisał umysłowi ludzkiemu cztery formy złudzeń, jakiemu on podlega (Kaczmarek, 2010, s.18-19):

- złudzenia plemienne (*idola tribus*) – wynikające z natury ludzkiej, takie jak antropomorfizm i doszukiwanie się celowości w świecie,
- złudzenia jaskini (*idola specus*) – przesady jednostek, spowodowane wpływem wychowania i otoczenia,
- złudzenia rynku (*idola fori*) – powodowane przez niedoskonałość języka, niedokładność, nieadekwatność i wieloznaczność pojęć,
- złudzenia teatru (*idola theatri*) – powodowane przez błędne spekulacje filozoficzne, których wyniki są przez odbiorców przekazu automatycznie przyjmowane na mocy autorytetu.

Przykład 1. Analiza trendu a kupno rafinerii

Jednym z częstych błędów wnioskowania jest oparcie wywodu myślowego na wybranych danych cząstkowych, bez uwzględnienia innych czynników, takich przykładowo jak czasowe zaburzenia trendu, wysoka dynamika zmian zjawiska, czy też niewielka jego skala, czyniąca daną zmienną podatną na silne fluktuacje. Jako przykład posłuży analiza wspomnianych wcześniej polskich inwestycji zagranicznych w latach 2004 – 2010 według danych Narodowego Banku Polski (Rysunek 1).



Rysunek 1. Wartość polskich inwestycji zagranicznych w latach 2004-2010 w mln EUR wg danych NBP

Źródło: dane NBP, www.nbp.pl (dostęp: 20.09.2017).

Zadaniem naukowca w procesie badawczym jest pełnienie roli bezstronnego obserwatora, który przy wykorzystaniu warsztatu badawczego przetwarza zgromadzone dane oraz informacje w zasób wiedzy. Naukowiec nie powinien ulegać pokusie dopasowywania wybranych danych do z góry przyjętych tez i teorii. Jak zauważa jednak T. Kaczmarek (2010, s. 14) w praktyce zgromadzone informacje podlegają subiektywnej obróbce, a uczeni okazują nieraz stronną lojalność wobec swoich teorii.

Przykładowo, kiedy opublikowano w 2008 roku dane za rok wcześniejszy, w kilku poczytnych tytułach prasowych w dziale analiz ekonomicznych wyciągnięto wnioski o poważnym zahamowaniu tempa aktywności inwestycyjnej polskich przedsiębiorstw na rynkach zagranicznych, o czym miał świadczyć znaczący spadek wielkości zaangażowanego kapitału z ponad 6 mld euro do niewiele ponad 2,5 mld. Analiza tych danych nie uwzględniła jednakże jednego istotnego faktu (abstrahując od innych wątpliwości metodologicznych zawartych przez mnie wcześniej), całkowicie przekreślając swoją rzetelność. Mianowicie skala polskiego zaangażowania inwestycyjnego na rynkach zagranicznych w XXI wieku jest relatywnie niewielka i jej wartość może być nieraz zdeterminowana pojedynczymi większymi projektami inwestycyjnymi polskich firm. Tymczasem w 2006 roku doszło do największej w tamtym okresie inwestycji w Europie Środkowo-Wschodniej, mianowicie przejęcia przez PKN Orlen litewskiej rafinerii w Możejkach. Była to wówczas największa polska inwestycja zagraniczna w historii a wartość tej transakcji znacząco w tamtym czasie wpłynęła na sumaryczne zestawienie polskich BIZ w 2006 roku, wobec czego nie można było oczekiwać choćby zbliżonego wyniku w kolejnym roku.

Przykład 2. Magia liczb - Łatwiej zwielokrotnić eksport do Australii niż do Niemiec

Zdarza się, iż obliczane na podstawie danych statystycznych wskaźniki przesłaniają badaczowi podstawowe prawa logiczne. Prezentowane przez Ministerstwo Rozwoju dane dotyczące polskiego eksportu wykazują niezbitnie, iż największą dynamikę w 2016 roku w porównaniu do roku wcześniejszego wykazywał polski eksport na kontynent Australii i Oceanii (145%) (MR, 2017). Czy jednak na tej podstawie można zbudować następującą konstrukcję logiczną: „*W 2016 roku najlepiej rozwijał się polski eksport na kontynent Australii i Oceanii*”? Otóż absolutnie nie, gdyż pojęcie najwyższej dynamiki nie może być traktowane tożsamo z pojęciem największego rozwoju. Należy zwrócić tu uwagę, iż wartość omawianego zjawiska wyniosła wówczas zaledwie 0,9 mld euro, wobec blisko 163 mld w przypadku polskiego eksportu do Europy. Zatem czy mniejsza dynamika eksportu do Europy, lecz w wartościach bezwzględnych przynosząca znacznie większe dochody, oznacza, iż rozwój ten jest słabszy niż dla kontynentu australijskiego?

Przykład 3. Iluzja korelacji – współzależność czy współwystępowanie?

Bardzo istotnym problemem dla badacza zjawisk zachodzących w świecie gospodarczym jest trafna identyfikacja występujących zależności pomiędzy szeregiem zmiennych. Jeżeli rząd danego państwa obniży gruntownie wysokość cła importowego na sprowadzane samochody, to zapewne ich cena na tamtejszym rynku w kolejnych miesiącach ulegnie obniżeniu. Na tej podstawie można wykazać korelację pomiędzy wysokością ceł importowych a ceną importowanego towaru.

Jednak Huff zauważył, iż zarobki prezbiteriańskich pastorów w stanie Massachusetts rosną z roku na rok wprost proporcjonalnie do cen rumu w Hawanie (Francuz, Mackiewicz, 2005, s. 472). Czy na tej podstawie można stwierdzić, iż istnieje korelacja pomiędzy tymi zjawiskami, spowodowana faktem, iż wzrost zarobków

pastorów powoduje wzrost popytu na rum w Hawanie? Dlatego tak istotne dla właściwej analizy procesów zachodzących w świecie gospodarczym jest odróżnienie sytuacji współzależności pomiędzy dwoma zmiennymi od ich przypadkowego współwystępowania. Co prawda przytoczony przykład charakteryzuje dość wysoki poziom abstrakcji, jednak fakt ulegania złudzeniu korelacji pomiędzy dwoma zjawiskami ekonomicznymi jest dość często obserwowany w pracach naukowych.

Przykład 4. Opóźnienia reakcji w zależnościach przyczynowo – skutkowych

Badacze zjawisk ekonomicznych prawidłowo identyfikując związki pomiędzy zachodzącymi procesami i występującymi zmiennymi, mogą także popełnić błąd wynikający z automatycznego przypisania „natychmiastowej mocy sprawczej” pomiędzy przyczyną a konsekwencją danego zjawiska. Załóżmy, iż rząd danego kraju ogłosi w styczniu wprowadzenie od marca br. znacznych ułatwień administracyjnych dla potencjalnych inwestorów zagranicznych. Opublikowane dane statystyczne za marzec tego roku pokażą istotny wzrost napływu kapitału w formie ZIB do tego kraju. Oczywiście udowodniono, iż ułatwienia dla inwestycji mogą stanowić determinantę ich rozwoju, ale czy tak mogło być i w tym przypadku? Aby odpowiedzieć na to pytanie, należałoby uwzględnić długość procesu inwestycyjnego, związanego z uruchomieniem nowej inwestycji, czyli inaczej rzecz ujmując odpowiedzieć sobie na pytanie, ile czasu minie, od momentu podjęcia przez przedsiębiorstwo decyzji o podjęciu inwestycji do czasu rzeczywistego przepływu kapitału. We wspomnianym przykładzie jest niemożliwe, aby napływ kapitału był spowodowany zapowiedziami rządu. Badacz danego zjawiska gdyby wysnuł prosty wniosek o zależności między działaniami rządu a decyzjami inwestorów popełniłby błąd, m.in. wynikający z nieuwzględnienia efektu opóźnienia reakcji. Natomiast niezależnie od tego problemu, rzetelna analiza badanych zależności powinna zostać pogłębiona o badania mikroekonomiczne, wyjaśniające rzeczywiste determinanty decyzji inwestorskich poszczególnych podmiotów gospodarczych.

5. PODSUMOWANIE

W klasycznej tzw. korespondencyjnej koncepcji prawdy podkreśleniu ulega zgodność osądu, zdania, myśli lub przekonania z rzeczywistością. W świecie nauki często można jednak zetknąć się z tzw. socjologiczną koncepcją prawdy, według której prawdą nie jest to, co wiernie oddaje analizowaną rzeczywistość, lecz to co uznaje większość. (Kaczmarek, 2010, s.25-26). Niestety to dość znaczące oddalenie pomiędzy tym „co wiemy”, a tym „jak jest naprawdę” występuje również w pracach o charakterze naukowym, eksperckim, czy publicystycznym. Jak zauważa P. Romer (2015, s. 141-150) bardzo często w tym naukowym procesie powstawania owej socjologicznej koncepcji prawdy pomaga zjawisko tzw. *mathiness*, określone przez autora jako błędnie pojęte i niewłaściwie lub fragmentarycznie realizowane wykorzystanie matematyki w ekonomii, prowadzące do „pseudomatematyzacji” teorii ekonomicznych. Wykorzystanie różnego rodzaju

wzorów i modeli matematycznych opartych na wątpliwych metodologicznie założeniach ma ową „prawdę” ubrać w szaty nauki. Dane statystyczne stanowią pierwotne elementy „wejściowe” w procesie analiz naukowych o charakterze statystycznym. W przypadku, kiedy charakteryzują się one znacznym błędem ich dokładności bądź przekaz z nich płynący ulega deformacji spowodowanej błędnym procesem wnioskowania o charakterze logicznym, wnioski, modele i powstałe na tej bazie koncepcje automatycznie będą posiadać znikomą użyteczność i nie będą powiększały zasobu wiedzy na temat otaczającego nas świata.

Reasumując należy podkreślić, że celem artykułu nie było podważanie celowości prowadzenia badań statystycznych, lecz właśnie z uwagi na duży szacunek wobec tej dyscypliny nauki, zwrócenie uwagi na problem, z jakim badacze stykają się codziennie w pracy naukowej. Należy zaznaczyć, że autor niniejszego artykułu nie jest statystykiem, zatem w swoim wywodzie celowo pomija błędy wnioskowania statystycznego powstałe w wyniku niewłaściwego zastosowania wzorów oraz reguł i praw statystyki, skupiając się na pułapkach ulokowanych w pozostałych obszarach analizy naukowej.

Dlatego tak istotne dla procesów badawczych w świecie ekonomii, w tym handlu międzynarodowego, powinno być podejmowanie działań w zakresie poprawy jakości danych oraz ujednoczenie standardów ich gromadzenia i przetwarzania. Problem integracji badań statystycznych w ramach systemu informacyjnego statystyki publicznej oraz jakości produkowanych danych i wiarygodności instytucji statystycznych był także przedmiotem obrad 60. Światowego Kongresu Statystyki w Rio de Janeiro w 2015 roku (Bielak i inni, 2016). Nakreślone trudności natury statystycznej nie powinny jednak zwalniać badacza od rzetelnej analizy danych na poziomie wyvodu o charakterze logicznym, co pozwala w wielu przypadkach uniknąć falsyfikacji powstałych w procesie badawczym wniosków.

LITERATURA

- Bielak, R. Cierpiat-Wolan, M. Okrasa, W. Witkowski, J. (2016). Statystyka dla lepszego Świata” — oczekiwania i trendy. 60. Światowy Kongres Statystyki, Rio de Janeiro 2015. *Wiadomości Statystyczne*, 3, 1-24.
- Brdulak, J. (2005). *Zarządzanie wiedzą a proces innowacji produktu*. Warszawa: Oficyna Wydawnicza SGH.
- Daszkowski, J. (2014). O retoryczności i nielogiczności statystyki stosowanej. *Wiadomości Statystyczne*, 2, 17-35.
- Francuz, P. Mackiewicz, R. (2015). *Liczy nie wiedzą skąd pochodzą. Przewodnik po metodologii i statystyce nie tylko dla psychologów*. Lublin: Wydawnictwo KUL.
- Kaczmarek, T. (2010). Metodologia badań naukowych. O wiedzy i prawdzie w naukach ekonomicznych. *Mysł Ekonomiczna i Polityczna*, 2, 13-27.
- Mączyńska, E. (2013). Błędy pomiaru w gospodarce. W: Owsiak, S. Pollok, A. (red.), *W poszukiwaniu nowego ładu gospodarczego*, 41-52, Warszawa: Wydawnictwo PTE.
- Nadolna, B. (2009). Metodologiczne aspekty prac naukowych. *Folia Pomer. Univ. Technol. Stetin, Oeconomica*, 56, 145 -158.

- Niedzielska, E. (1998). *Informatyka ekonomiczna*. Wrocław: Wydawnictwo AE.
- Romer, P. (2015). Niewłaściwe wykorzystanie matematyki w teorii wzrostu gospodarczego. *Gospodarka Narodowa*, 6, 141-150.
- Stanosz, B. (1998). *Wprowadzenie do logiki formalnej. Podręcznik dla humanistów*. Warszawa: Wydawnictwo Naukowe PWN.
- Stefanowicz, B. (2013). *Informacja, wiedza, mądrość*. Warszawa: Główny Urząd Statystyczny.
- Zaliwski, A. (2000). *Korporacyjne bazy wiedzy*. Warszawa: Wydawnictwo PWE.
- Ziemiński, Z. (1993). *Logika praktyczna*. Warszawa: Wydawnictwo Naukowe PWN.

Źródła internetowe:

- Bank Światowy (2017). Portal Banku Światowego, <http://wits.worldbank.org> (dostęp: 23.09.2017)
- CBR (2017). Portal Centralnego Banku Rosji, www.cbr.ru (dostęp: 23.09.2017)
- MR (2017). Portal Ministerstwa Rozwoju, www.mr.gov.pl (dostęp: 23.09.2017)
- NBP (2017). Portal Narodowego Banku Polskiego, www.nbp.pl (dostęp: 23.09.2017)
- UNCTAD (2017). Portal Konferencji Narodów Zjednoczonych ds. Handlu i Rozwoju, <http://unctadstat.unctad.org> (dostęp: 23.09.2017)

Statistics and the logical inference in the scientific research: a few remarks about the most frequent mistakes and traps in the research process

Abstract: The scientific research in the area of international trade are based on statistical data obtained from national and international institutions The present article is emphasizing meaning of the credibility and diligences of conducted economic analyses based on statistical data, and is also identifying mistakes in this sphere chosen and traps on the researcher's way.

Keywords: scientific researches; statistics; methodology of researches

JEL codes: C82, F14