

Młodzi naukowcy o niskiej produktywności? Żadnych szans na wysoką produktywność w przyszłości (analiza 320 000 karier naukowych w 38 krajach OECD)

1. Wprowadzenie

Nasze badanie koncentruje się na trwałości bardzo wysokiej i bardzo niskiej indywidualnej produktywności badawczej z perspektywy całości kariery naukowej. Analizujemy produktywność naukowców znajdujących się na późnym etapie kariery, czyli publikujących od co najmniej 25 lat ($N = 320\,564$) i pracujących w 38 krajach OECD (w tym w Polsce). Badamy wzorce ich mobilności pomiędzy dziesięcioma klasami produktywności – od najniższej (dolnych 10%) do najwyższej (górnym 10%).

W pracy pokazujemy, że już na stosunkowo wczesnym etapie kariery naukowej rozkład produktywności w ramach globalnej profesji naukowej na jej dwóch krańcach jest w dużej mierze ustalony. Początkowy rozkład utrzymuje się w czasie, przez lata i dziesięciolecia. Najmniej produktywni naukowcy niemal nigdy nie stają się najbardziej produktywni, a najbardziej produktywni – najmniej produktywni. Uderzająca jest trwałość przynależności do klas najwyższej i najniższej produktywności z perspektywy cyklu życia naukowców. Nasze badanie pokazuje, że globalny system nauki pod względem rozkładu produktywności jest wyjątkowo sztywny.

W tym celu przekształcamy potężny zbiór danych bibliometrycznych dotyczących publikacji i cytowań (surowe dane z bazy Scopus) w globalne, kompleksowe, wielowymiarowe i co najważniejsze podłużne – czyli obejmujące zmiany w czasie – źródło danych na temat karier naukowych setek tysięcy naukowców. Jak dotąd nie badano indywidualnej produktywności publikacyjnej w ujęciu globalnym z perspektywy podłużnej, nie śledzono indywidualnych karier akademickich od pierwszej publikacji przez kolejnych 25–50 lat po to, aby porównywać między sobą zmiany przynależności do klas produktywności w czasie. W szczególności nie badano, czy w nauce zdarzają się sytuacje skrajne: czy najmniej produktywni naukowcy stają się naukowcami najbardziej produktywnymi (określamy ich mianem „skoczków”; i odwrotnie, najbardziej produktywni stają się naukowcami najmniej produktywnymi, nazywamy ich tu „spadkowiczami”)?

* Prof. dr hab. Marek Kwiek (marek.kwiek@amu.edu.pl), Institute for Advanced Studies in Social Sciences and Humanities (IAS), Uniwersytet im. Adama Mickiewicza w Poznaniu, ORCID: orcid.org/0000-0001-7953-1063

** Dr Łukasz Szymula (lukasz.szymula@amu.edu.pl), Wydział Matematyki i Informatyki, Uniwersytet im. Adama Mickiewicza w Poznaniu, ORCID: orcid.org/0000-0001-8714-096X

Odchodzimy w naszym badaniu po raz kolejny od indywidualnych publikacji (i ich właściwości) i zajmują nas pojedynczy naukowcy (i ich cechy) jako jednostki analizy. Nie zajmuje nas bibliometria – ale nowe naukoznawstwo (*science of science*), które w dużej skali pyta o funkcjonowanie zarazem nauki i naukowców.

W sensie metodologicznym budujemy indywidualne historie publikacyjne i cytowania dla każdego naukowca znajdującego się w naszej próbie, ograniczając zakres badań do 16 dyscyplin z dziedzin STEMM (nauki ścisłe, techniczne, inżynieryjne, matematyczne i medyczne) i z nauk społecznych.

Nasze badanie ma charakter podłużny, ponieważ wykorzystujemy globalny zbiór danych bibliometrycznych do badania zmian karier naukowych w czasie (Menard, 2002; Rowland, 2014). Obserwujemy kariery publikacyjne naukowców trwające dziesięciolecia. Testujemy przydatność metadanych dotyczących publikacji i cytowań do badania globalnej profesji akademickiej; owe metadane to cyfrowe ślady pozostawiane przez naukowców we własnych publikacjach przez całe życie zawodowe (lub pozostawiane tak długo, jak długo publikują oni w czasopiśmie akademickich).

Dostęp do baz danych ze śladami cyfrowymi (Liu i in. 2023; Salganik 2018) pozwolił na wyłonienie się nowej multidyscyplinarnej dziedziny naukoznawstwa w rozumieniu *science of science*, czyli nauk o nauce (Clauzet i in. 2017; Wang i Barabási 2021), umożliwiając tym samym radykalne wyjście badań karier naukowych poza tradycyjne ankiety i przeprowadzane na małą skalę wywiady z naukowcami. Cyfrowe ślady pozostawiane przez naukowców w globalnych zbiorach danych (typu Scopus) pozwalają badaczom karier akademickich na zmianę punktu ciężkości z krajowych systemów naukowych – na globalny system nauki.

W sensie praktycznym – prezentowane badanie pozwala na analizę mobilności naukowców między 10 indywidualnymi, decyłowymi klasami produktywności (od najwyższej do najniższej) w trakcie długiej kariery naukowej, obejmującej jej wczesny, środkowy i późny etap. Nasze wstępne hipotezy zakładały, po pierwsze, że naukowcy zazwyczaj przez lata tkwią w swoich – niskich bądź wysokich – klasach produktywności; po drugie, że najbardziej produktywni naukowcy zazwyczaj kontynuują swoją karierę właśnie jako najbardziej produktywni naukowcy; i wreszcie, że radykalne zmiany między klasami produktywności, zwłaszcza zmiany w górę, wbrew popularnym przekonaniom o zawsze możliwych, nieograniczonych karierach akademickich, są w praktyce wysoce nieprawdopodobne ze względu na kumulatywny charakter przewag (i ograniczeń) w nauce. Przewagi jednych naukowców i ograniczenia drugich budują się w systemie nauki przez wiele lat; zakładaliśmy zatem, że w zasadzie – z perspektywy 25 czy 35 lat pracy naukowej – są one nie do przeskoczenia. Stąd nasze określenie tych nieprawdopodobnych przypadków: to skoczki w nauce. Nie możemy być radykalnie gdzie indziej pod względem produktywności publikacyjnej badawczy mi 25 czy 35 lat temu; a może właś-

nie możemy? Poszukiwaliśmy odpowiedzi empirycznej na ogromnym materiale badawczym, a nie odpowiedzi zwyczajowej.

Prześledziliśmy bardzo szczegółowo kariery 320 564 naukowców będących na późnym etapie kariery i pochodzących z 38 różnych systemów naukowych, co dało nam możliwość przetestowania hipotez dotyczących profesji akademickiej w bardziej ogólnym ujęciu niż ujęcie jednego kraju (tradycyjnie badano niemal wyłącznie system amerykański: tam były dane, środki i zainteresowana publiczność naukowa).

Istnieją jak dotąd tylko trzy badania podłużne – przeprowadzone jednak na bardzo małą skalę – w jednym kraju, podobne do naszego (Turner i Mairesse 2005; Kelchtermans i Veugelers 2013; oraz Abramo i in. 2017). Nasze badanie ma jednak inny zakres i wykorzystuje zupełnie inną metodologię. Prześledziliśmy kariery naukowców ze wszystkich sektorów nauki (w tym szkolnictwa wyższego); zbadaliśmy zmiany produktywności w dłuższym okresie czasu (25–50 lat) we wszystkich dyscyplinach STEMM i w trzech wybranych dyscyplinach nauk społecznych. Nasze podejście było podłużne (zob. Menard 2002; Rowland 2014; Ruspini 1999) i klasyfikacyjne (czyli bazujące na relatywnych klasach produktywności, a nie na liczbach publikacji) (zob. Costas i Bardons 2007; Costas i in. 2010).

Szukaliśmy stabilności w produktywności badawczej (lub jej braku) na przestrzeni długich karier akademickich. Śledząc kariery setek tysięcy naukowców, poszukaliśmy niewidocznych w inny sposób, globalnych wzorców mobilności między klasami produktywności publikacyjnej (ilekroć używamy terminu „globalny”, odnosimy się do 38 krajów OECD). Nasza próba naukowców będących na późnym etapie kariery w krajach OECD obejmuje 79,42% wszystkich naukowców będących na późnym etapie kariery na całym świecie w 2023 r. (ze wszystkich krajów pochodziłoby $N = 403\,653$ naukowców); a ich dorobek naukowy (30 695 679 artykułów naukowych) obejmuje 83,03% wszystkich artykułów naukowych opublikowanych przez tę kategorię naukowców na całym świecie w 2023 r. ($N = 36\,969\,473$).

Postawiliśmy następujące dwa pytania badawcze dotyczące zmian produktywności publikacyjnej w trakcie kariery akademickiej: po pierwsze, jaka jest skala przejść poziomych (czyli między najwyższymi i między najniższymi decylami produktywności) i radykalnych przejść pionowych (najwyższe decyle – najniższe decyle) pomiędzy klasami produktywności? Po drugie, jakie są różnice między dyscyplinami we wzorcach mobilności między globalnymi klasami produktywności?

2. Dane i metodologia

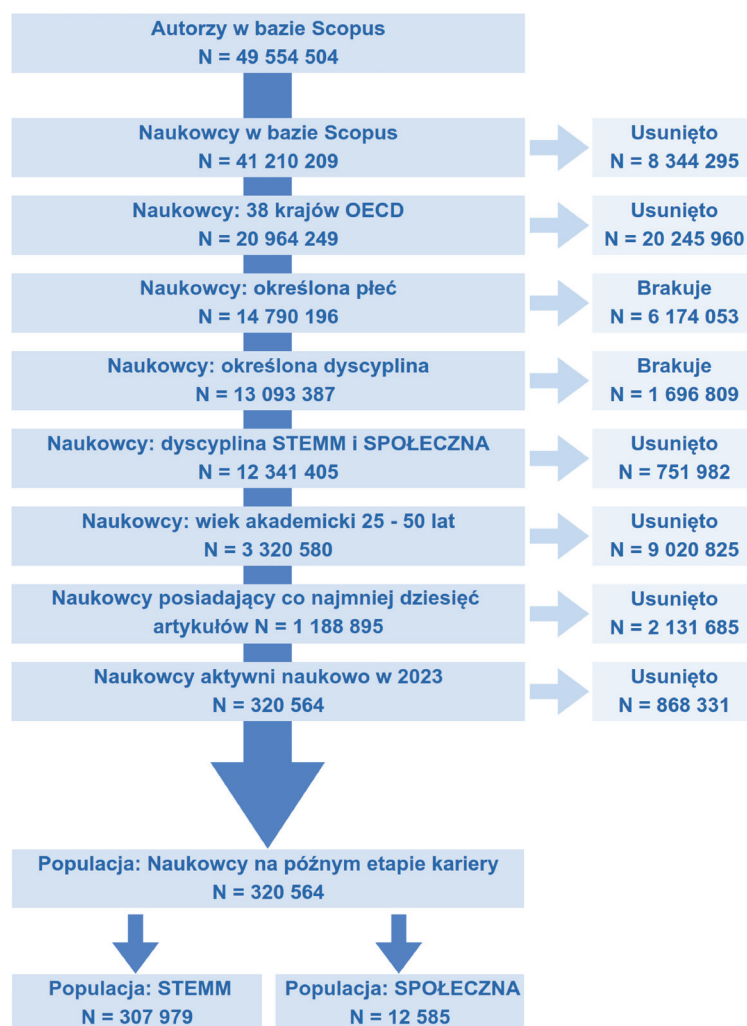
Analizowane dane zostały pozyskane z bibliometrycznej bazy danych Scopus i uzyskane w ramach wieloletniej umowy o współpracy z laboratorium International Center for the Study of Research (ICSR) Lab, które umożliwiło nam dostęp do platformy obli-

czeniowej w chmurze wykorzystywanej do celów badawczych przez firmę Elsevier. Nasza końcowa próba obejmowała wszystkich naukowców będących na późnym etapie kariery, którzy byli aktywni badawczo w 2023 r. (posiadających co najmniej 25-letnie doświadczenie publikacyjne) działających w ramach 16 dyscyplin z dziedzin STEM i dyscyplin z nauk społecznych i pochodzących z 38 krajów OECD ($N = 320\,564$ naukowców z $N = 16\,345\,891$ artykułami naukowymi; zob. najważniejsze etapy wstępnego przetwarzania danych na ryc. 1; „autorzy” w naszym zbiorze danych zostali zdefiniowani jako naukowcy posiadający publikacje dowolnego typu, a „naukowcy” jako autorzy artykułów publikowanych wyłącznie w czasopiśmie naukowych i materiałach konferencyjnych.

Do obliczeń wykorzystaliśmy surową bazę danych Scopus z 29 marca 2024 roku. Nasza próba (zob. tabele uzupełniające 1 i 2) obejmowała 12 585 naukowców społecznych (z dyscyplin BUS, ECON i PSYCH) i 307 979 naukowców z dyscyplin STEM, przy czym ci ostatni stanowili 95,76% próby. Jedna czwarta próby obejmowała kobiety naukowców (26,34%), a ich odsetek był nieco wyższy w naukach społecznych niż w dyscyplinach STEM. W najmłodszych kohortach naukowców znajdowało się około 20 000–25 000 naukowców, a w najstarszych około 2000–3000 naukowców, przy czym udział kobiet zwiększał wraz z każdą kolejną młodszą kohortą, od 13–15% w najstarszych kohortach (49–50 lat doświadczenia publikacyjnego) do jednej trzeciej w najmłodszej kohorcie (25 lat doświadczenia publikacyjnego: 32,77%).

Aby uzyskać wyniki na poziomie zagregowanym, wykorzystaliśmy środowisko Databricks, które pozwala na zarządzanie i wykonywanie obliczeń w chmurze na zasobach usługi Amazon EC2. Skrypty do generowania wyników zostały opracowane przy użyciu biblioteki PySparkSQL. Uruchamianie odbywało się przy użyciu klastra w trybie standardowym z Databricks Runtime w wersji 11.2 ML, technologią Apache Spark w wersji 3.3.0, Scala 2.12 oraz instancją i3.2xlarge z 61 GB pamięci, ośmioma rdzeniami, od jednego do sześciu procesów roboczych dla węzła roboczego oraz instancją c4.2xlarge z 15 GB pamięci i czterema rdzeniami dla węzła sterownika. Czas wykonania skryptu wyniósł sześć godzin, a operację pozyskania wyników uruchomiono 25 czerwca 2024 roku. Wyniki uzyskaliśmy w formacie CSV.

Kariery naukowców zostało retrospektywnie podzielone na trzy etapy: wczesny, środkowy i późny. Z definicji wszyscy naukowcy znajdujący się na późnym etapie byli wcześniej zarówno naukowcami na wczesnym etapie (w latach publikowania 5–14), jak i naukowcami na etapie środkowym (w latach publikowania 15–24). Przeanalizowaliśmy ich obecne, pięcioletnie wzorce publikacyjne (z lat 2019–2023) i wróciliśmy do ich wcześniejszych wzorców, aby zbadać, w jaki sposób zmieniali swoje klasy produktywności w czasie. Tylko w taki sposób mogliśmy pokazać zmiany w ujęciu czasu – potrzebowaliśmy do tego karier naukowych trwających co najmniej 25 lat.



Ryc. 1. Schemat blokowy pokazujący najważniejsze etapy wstępnego przetwarzania danych: przejście od wszystkich autorów znajdujących się w bazie danych Scopus do naukowców znajdujących się na późnym etapie kariery i spełniających pozostałe wymagania tworzące próbę

Na każdym wcześniejszym etapie naukowcy z naszej próby wykazywali pewną mierzalną roczną indywidualną produktywność. Ich produktywność została obliczona dla ostatniego pięcioletniego okresu (2019–2023) oraz dla dwóch wcześniejszych okresów: pierwszego, kiedy byli naukowcami na wczesnym etapie kariery, i drugiego, kiedy byli naukowcami na jej środkowym etapie.

Nasze analizy opierają się na koncepcji podziału naukowców na klasy: naukowcy na późnym etapie kariery zostali najpierw podzieleni według obecnych klas produktywności (oddzielnie w ramach każdej z 16 dyscyplin), a następnie, retrospektywnie, po-

dzieleni według wcześniejszych klas produktywności na dwóch wcześniejszych etapach kariery. To istota badania podłużnego – a zarazem zaleta mikrodanych z surowej bazy Scopus.

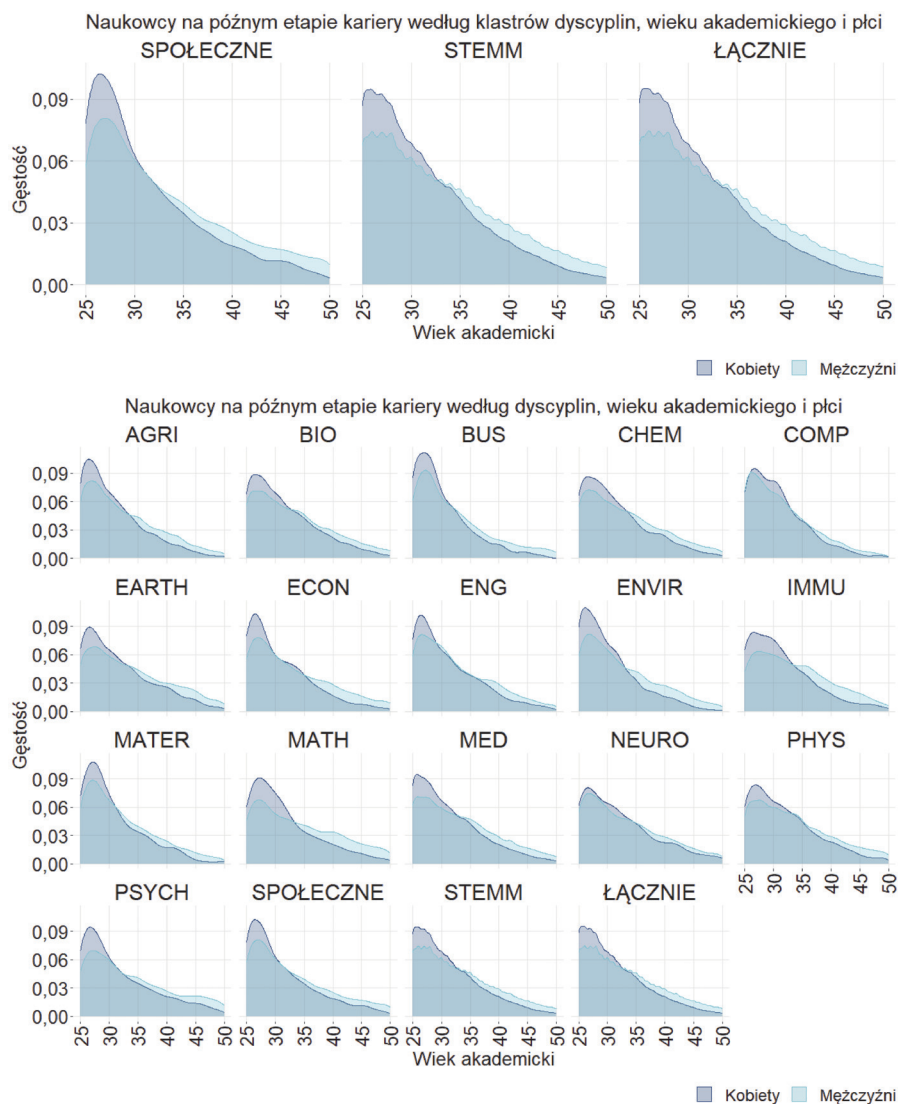
Naukowcy na wczesnym etapie kariery mogą zachować lub zmienić swoje (oparte na decylach, od decyla 1. do decyla 10.) klasy produktywności, kiedy stają się naukowcami na środkowym etapie kariery. Podobnie dzieje się z naukowcami na środkowym etapie kariery, kiedy stają się naukowcami na późnym etapie kariery. Obserwowaliśmy zatem tych samych naukowców przez dekady i porównywaliśmy ich produktywność między sobą (na tym samym etapie kariery akademickiej i w tej samej dyscyplinie).

Po pierwsze, prześledziliśmy produktywność naukowców odkąd stali się naukowcami na wczesnym etapie kariery, czyli w naszym ujęciu pięć lat po ich pierwszej globalnie indeksowanej publikacji (lata 5–14). Po drugie, nie porównywaliśmy zmian produktywności w czasie (w miarę rozwoju indywidualnej kariery naukowej) pod względem zmieniającej się liczby publikacji – porównywaliśmy produktywność pod względem stabilnej lub zmieniającej się przynależności do klas produktywności, gdy naukowcy wspinali się po szczeblach kariery zawodowej.

Naukowców zawsze można przypisać do najwyższej i najniższej decylowej (podobnie jak percentylowej) klasy produktywności, niezależnie od faktycznej liczby publikacji, jeśli tylko pozwalają na to punkty odcięcia, dlatego też terminów „najmniej produktywni” i „najbardziej produktywni” używamy, nie oceniając poziomu produktywności, a jedynie go klasyfikując na prostej skali.

Do określenia indywidualnych cech naukowców wykorzystaliśmy globalny zbiór danych bibliometrycznych. Określenie niektórych atrybutów zostało już szczegółowo opisane w naszych wcześniejszych pracach: określenie płci (binarne: mężczyzna lub kobieta, zob. metody w Karimi i in. 2016; Santamaria i Mihaljević 2018; Sebo 2021, 2023), określenie dyscypliny (przy użyciu wszystkich przytaczanych odniesień bibliograficznych we wszystkich publikacjach wydanych na przestrzeni całego życia), określenie kraju afiliacji (przy użyciu wartości modalnej wszystkich afiliacji ze wszystkich publikacji, na przestrzeni całego życia), określenie statusu naukowców w nauce światowej (epizodyczny/nieepizodyczny, przy użyciu minimalnego dorobku naukowego w postaci 10 artykułów naukowych) oraz określenie wieku akademickiego (przy użyciu dystansu czasowego wyrażonego w latach między pierwszą publikacją dowolnego typu a rokiem 2023; Kwiek i Szymula, 2023, 2024; Kwiek i Roszka 2024b). Rozkład próby według wieku akademickiego (tj. doświadczenia w publikowaniu) przedstawiono na ryc. 2, a dalsze szczegóły podano w tabeli uzupełniającej 1 i tabeli uzupełniającej 2.

Skupiliśmy się na 13 dużych dyscyplinach z obszaru STEMM i 3 dyscyplinach z obszaru nauk społecznych, zgodnie z systemem klasyfikacji czasopism stosowanym w bazie Scopus (All Science Journal Classification, ASJC): AGRI, nauki rolnicze i biologicz-



Ryc. 2. Rozkład wieku akademickiego: wykresy gęstości jądrowej. Naukowcy na późnym etapie kariery, według klastra dyscyplin naukowych i płci (górny panel) oraz dyscypliny naukowej i płci (dolny panel) ($N = 320\,564$)

ne; BIO, biochemia, genetyka i biologia molekularna; BUS biznes, zarządzanie i rachunkowość; CHEM, chemia; COMP, informatyka; EARTH, nauki o Ziemi i planetach; ECON, ekonomia, ekonometria i finanse; ENG, inżynieria; ENVIR, nauka o środowisku; IMMU, immunologia i mikrobiologia; MATER, materiałoznawstwo; MATH, matematyka; MED, medycyna, NEURO, neuronauka; PHYS, fizyka i astronomia oraz PSYCH psychologia.

3. Wyniki badań

3.1. Mobilność między klasami produktywności na przestrzeni kariery naukowej

Skupiliśmy się na analizie mobilności między klasami produktywności, a w szczególności na przejściach pomiędzy najwyższą i najniższą klasą, oraz pomiędzy klasami sąsiadującymi: czyli na decylach produktywności 8–10. na górze skali produktywności i na decylach 1–3. na jej dole. Przeanalizowaliśmy trzy etapy kariery oraz przejścia od wczesnego etapu do etapu środkowego i od etapu środkowego do późnego.

Naukowcy na wczesnym etapie kariery, należący do najwyższych i najniższych klas produktywności, mogą zmieniać swoje klasy w miarę przechodzenia do środkowego etapu kariery (przenosząc się do dowolnego decyla). Analogicznie, naukowcy znajdujący się na środkowym etapie kariery w najwyższych i najniższych klasach produktywności mogą doświadczać zmian swoich klas w miarę przechodzenia do późnego etapu kariery, przesuwając się w górę, w dół lub pozostając w tym samym decylu produktywności. Chcielibyśmy poznać dominujące wzorce ewolucji produktywności w trakcie trwania kariery naukowej i poznać zakres mobilności między różnymi klasami produktywności.

Przeanalizowaliśmy trzy podstawowe typy mobilności (we wszystkich dyscyplinach łącznie i w ramach poszczególnych dyscyplin):

1. *Mobilność pozioma między najwyższymi klasami:* naukowcy pozostają w najwyższym decylu produktywności na kolejnych etapach rozwoju kariery (mobilność między decylami 10. i 10.).
2. *Mobilność pozioma między najniższymi klasami:* naukowcy pozostają w najniższym decylu produktywności na kolejnych etapach rozwoju kariery (mobilność między decylami 1. i 1.).
3. *Radykalna mobilność pionowa:* naukowcy zmieniają klasy produktywności między najwyższym i najniższym decylem produktywności w dowolnym kierunku na kolejnych etapach rozwoju kariery (mobilność między decylami 1. i 10. oraz 10. i 1., czyli fenomen skoczków i spadkowiczów w nauce).

Oprócz zbadania podstawowej mobilności między najwyższymi (decyl 10.) i najniższymi (decyl 1.) decylami produktywności, przyjrzymy się również mobilności z szerszej perspektywy, która uwzględnia przechodzenie między górnymi (8–10.) i dolnymi decylami (1–3.). Niektórzy naukowcy znajdują się bowiem tuż powyżej progu przynależności do decyla 1., a inni tuż poniżej progu przynależności do decyla 10. Dlatego też przydatne wydaje się również bardziej kompleksowe podejście, które obejmuje mobilność między sąsiednimi decylami (1–3. i 8–10.).

Nasze szczegółowe pytania dotyczą tego, w jaki sposób naukowcy osiągający najwyższą produktywność (decyl produktywności 10., $N = 32\,063$) na środkowym etapie

kariery byli rozmieszczeni według percentyli produktywności (zakres: 0–100) na wczesnym etapie kariery w przeszłości; oraz, analogicznie, w jaki sposób naukowcy osiągający najwyższą produktywność (decyl produktywności 10., $N = 32\,063$) na późnym etapie kariery byli rozmieszczeni według percentyli produktywności na środkowym etapie kariery.

Tabela 1. Jak najbardziej produktywni naukowcy na środkowym etapie kariery (decyl produktywności 10.) (lewy panel) i najmniej produktywni naukowcy na środkowym etapie kariery (decyl produktywności 1.) (prawy panel) zostali rozmieszczeni według percentyli produktywności (zakres: 0–100), kiedy byli na wczesnym etapie kariery? Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Naukowcy najbardziej/najmniej produktywni na środkowym etapie kariery, mediana wyjściowych wartości rozkładu percentylowego według dyscypliny ($N_{\text{top}} = 32\,063$, $N_{\text{bottom}} = 32\,063$)

Dyscyplina	Najbardziej produktywni naukowcy na środkowym etapie kariery – rozkład według mediany produktywności na wczesnym etapie kariery				Najmniej produktywni naukowcy na środkowym etapie kariery – rozkład według mediany produktywności na wczesnym etapie kariery			
	N	Średnia	SD	Mediana	N	Średnia	SD	Mediana
AGRI	2373	85,03	15,84	91	2373	18,66	17,13	14
BIO	4582	82,95	18,69	90	4582	21,05	19,58	15
BUS	326	82,00	19,74	89,5	326	24,43	21,44	18
CHEM	1490	84,52	17,55	91	1490	17,57	16,68	13
COMP	765	80,80	20,73	89	765	23,12	20,07	17
EARTH	1437	85,01	16,27	91	1437	18,32	17,48	13
ECON	385	82,32	19,85	90	385	22,30	19,20	17
ENG	1282	82,40	18,84	89	1282	20,85	19,02	15
ENVIR	652	84,82	16,78	91	652	20,99	19,48	15,5
IMMU	315	82,28	19,21	89	315	21,24	20,46	14
MATER	584	83,26	17,91	90	584	17,89	16,18	14
MATH	701	85,58	16,79	92	701	19,90	18,01	15
MED	13108	83,97	17,04	90	13108	20,15	18,49	15
NEURO	587	83,72	18,50	90	587	19,90	18,88	14
PHYS	2928	83,25	20,90	92	2928	18,49	19,15	12
PSYCH	548	83,60	18,35	91	548	18,79	17,74	13
SPOŁECZNE	1259	82,79	19,17	90	1259	21,32	19,14	17
STEMM	30 804	83,78	17,77	90	30 804	19,88	18,53	15
ŁĄCZNIE	32 063	83,74	17,83	90	32 063	19,94	18,55	15

Ponadto interesuje nas, w jaki sposób naukowcy o najniższej produktywności (decyl produktywności 1, $N = 32\,063$) na środkowym etapie kariery byli rozmieszczeni według

percentyli produktywności (zakres: 0–100) na wczesnym etapie kariery, oraz, analogicznie, jak obecni naukowcy o najniższej produktywności (decyl produktywności 1., $N=32\,075$) znajdujący się na późnym etapie kariery byli rozmieszczeni według percentyli produktywności, gdy byli naukowcami na środkowym etapie kariery.

We wszystkich analizowanych przypadkach badamy retrospektywnie obecnych naukowców na późnym etapie kariery (którzy stanowią naszą próbę) na dwóch wcześniejszych etapach: na wczesnym i na środkowym etapie kariery.

W przypadku mobilności z wczesnego do środkowego etapu mediana początkowej rangi percentylowej (jako naukowca na wczesnym etapie kariery) jest bardzo zbliżona do mediany docelowej rangi percentylowej (jako naukowca na środkowym etapie kariery): mediana to 90. percentyl dla naukowców o najwyższej produktywności i 15. percentyl dla naukowców o najniższej produktywności (tabela 1), przy ograniczonej zmienności pod względem dyscyplin dla naukowców o najwyższej produktywności (od 89. w COMP, ENG i IMMU do 92. w MATH) i nieco większej zmienności pod względem dyscyplin dla naukowców o najniższej produktywności (od 12. w PHYS do 17. w COMP i ECON).

Podobnie, jeśli chodzi o mobilność od środkowego do późnego etapu kariery, mediana wyjściowego percentyla (jako naukowca na środkowym etapie kariery) dla obecnych najbardziej produktywnych naukowców na późnym etapie kariery jest bardzo zbliżona do docelowego percentyla (jako naukowca na późnym etapie kariery): mediana to 90. percentyl dla najbardziej produktywnych naukowców i 19. percentyl dla najmniej produktywnych naukowców (tabela 2), przy ograniczonej zmienności pod względem dyscyplin (od 89. w BUS do 92. w CHEM oraz od 15. w PSYCH i EARTH do 21. w BIO i ECON).

Przydatną metodą wizualizacji rozkładu najbardziej produktywnych naukowców na późnym etapie kariery według decyli produktywności na ich środkowym i wczesnym etapie kariery są wykresy gęstości jądrowej (ryc. 3 i 4). Wykresy te wykorzystują estymację gęstości jądrowej do generowania gładkiej krzywej, która przedstawia rozkład danych wyjściowych. W przeciwieństwie do histogramów, na wykresy gęstości jądrowej nie ma wpływu liczba zastosowanych przedziałów ani znaczące różnice między nimi, co sprawia, że bardziej efektywnie ilustrują one kształt rozkładu; pozwalają również na bardziej elastyczne porównywanie między sobą różnych zbiorów. Biorąc pod uwagę wszystkie dyscypliny naukowe łącznie, większość najbardziej produktywnych naukowców znajdowała się wcześniej między 8. i 10. decylem produktywności (a większość najmniej produktywnych naukowców – między 1. a 3. decylem, co nie zostało tutaj pokazane z powodu ograniczeń miejsca). Największą koncentrację odnotowano w dyscyplinach PHYS i CHEM.

Nasz zbiór danych pozwala na szczegółową analizę mobilności między decylami produktywności (na poziomie poszczególnych naukowców). Tabela 3 pokazuje wyjściowe

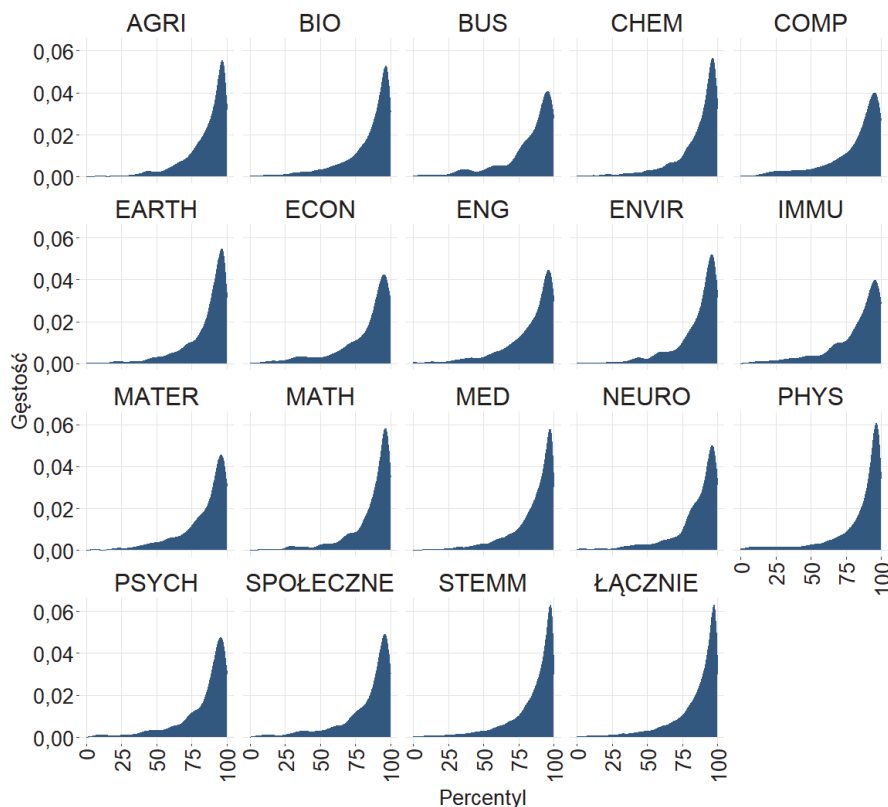
decyle produktywności (na wczesnym etapie kariery) w przeszłości najbardziej produktywnych naukowców na środkowym etapie kariery w różnych dyscyplinach.

Tabela 2. W jaki sposób najbardziej produktywni naukowcy na późnym etapie kariery (decyl produktywności 10.) (lewy panel) i najmniej produktywni naukowcy na późnym etapie kariery (decyl produktywności 1.) (prawy panel) zostali rozmieszczeni według percentyli produktywności (zakres: 0–100), kiedy znajdowali się na środkowym etapie kariery? Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Najbardziej/najmniej produktywni naukowcy na późnym etapie kariery, mediana wyjściowych wartości rozkładu percentylowego według dyscypliny ($N_{\text{top}} = 32\,063$, $N_{\text{bottom}} = 32\,075$)

Dyscyplina	Najbardziej produktywni naukowcy na późnym etapie kariery – rozkład według mediany produktywności na środkowym etapie kariery				Najmniej produktywni naukowcy na późnym etapie kariery – rozkład według mediany produktywności na środkowym etapie kariery			
	<i>N</i>	Średnia	SD	Mediana	<i>N</i>	Średnia	SD	Mediana
AGRI	2373	84,34	17,98	91	2373	21,90	19,08	17
BIO	4582	82,97	19,52	90	4582	26,37	22,31	21
BUS	326	80,33	21,99	89	326	25,52	20,82	20
CHEM	1490	85,20	17,92	92	1490	21,92	19,39	16
COMP	765	81,76	21,08	90	767	24,55	20,68	19
EARTH	1437	82,50	19,38	90	1437	21,70	19,70	15
ECON	385	80,38	22,31	90	385	25,44	20,76	21
ENG	1282	83,06	19,74	91	1282	23,38	19,26	19
ENVIR	652	83,58	19,25	91	652	24,69	20,37	19
IMMU	315	82,78	19,81	90	319	24,12	22,07	17
MATER	584	85,03	17,84	91	590	22,17	18,93	17
MATH	701	84,21	19,00	91	701	22,07	19,02	18
MED	13108	83,01	19,61	90	13 108	24,50	21,09	19
NEURO	587	84,53	18,35	91	587	24,55	21,75	18
PHYS	2928	83,28	21,94	92	2928	22,37	20,84	16
PSYCH	548	83,84	18,10	91	548	21,29	19,69	15
SPOŁECZNE	1259	81,87	20,39	90	1259	23,65	20,31	20
STEMM	30 804	83,29	19,56	90	30 816	23,98	20,78	19
ŁĄCZNIE	32 063	83,23	19,60	90	32 075	23,96	20,76	19

Ponad połowa najbardziej produktywnych naukowców na środkowym etapie kariery znajdowała się w decylnym 10. produktywności na wczesnym etapie kariery (52,39%), przy czym 20,94% zaczynało w decylnym 9., a 10,33% w decylnym 8. W sumie ponad 80% naukowców znajdowało się w decylach produktywności 8–10 na wczesnym etapie kariery (83,66%). Tylko niewielka część tych naukowców awansowała z najniższych

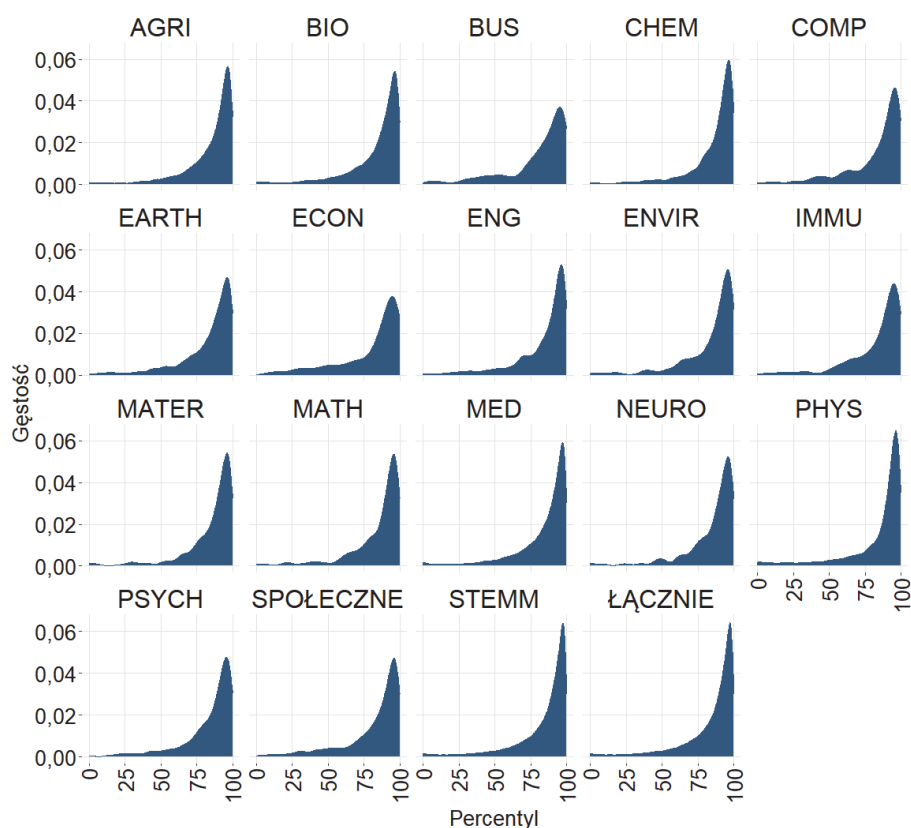
trzech decyli do decyla najwyższego, przy czym tylko 162 dokonało ekstremalnego (maksymalnego) skoku: z decyla 1. do decyla 10. (0,51%; to nasi skoczki) i 232 z decyla 2. do decyla 10. (0,72%).



Ryc. 3. Jak najbardziej produktywni naukowcy ($N = 32\,063$, decyl produktywności 10.) na środkowym etapie kariery byli rozmieszczeni według percentyli produktywności (zakres: 0–100), gdy byli na wczesnym etapie kariery? Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Wykresy gęstości jądrowej, wyjściowy rozkład percentyli produktywności według dyscypliny

Dysponujemy profilami biograficznymi i publikacyjnymi każdego naukowca, w tym również owych kilkuset przypadków odstających. Tylko 2,2% (717 naukowców z 32 063) z decyli 1–3. osiągnęło później decyl 10. W dyscyplinach nauk społecznych prawdopodobieństwo tak ekstremalnej mobilności w górę było nieco wyższe niż w dyscyplinach STEMM (3,13% vs. 2,22%), ale nadal niezwykle niskie.

Różnice w zakresie ekstremalnej mobilności pomiędzy decylem 1. i decylem 10. (skoczki) między dyscyplinami są znaczące, a wskaźniki wahają się od 0,26% w ECON do 1,33% w PHYS.



Ryc. 4. Jak najbardziej produktywni naukowcy ($N=32\,063$, decyl produktywności 10.) na późnym etapie kariery byli rozmieszczeni według percentyli produktywności (zakres: 0–100), gdy byli na środkowym etapie kariery? Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Wykresy gęstości jądrowej, wyjściowy rozkład percentyli produktywności według dyscypliny

W 38 krajach OECD tylko jeden ekonomista (0,26%) i tylko jeden immunolog (0,32%) przeskoczyli w swoim życiu naukowym z decyla 1. do decyla 10. (odpowiednio na 385 i 315 badanych naukowców). Na podstawie danych bibliometrycznych dysponujemy szczegółowymi informacjami na temat ich karier (natomiast nie znamy ich biografii naukowych opartych na danych administracyjnych pochodzących z krajowych rejestrów naukowców, jak ma to miejsce w przypadku badań prowadzonych w jednym kraju, na przykład w Polsce, przy wykorzystaniu danych z zasobów OPI PIB – dane te nie są dostępne w przypadku badań prowadzonych w wielu krajach). Mobilność między decylem 10. i decylem 10. wykazuje również pewną zmienność, z mniej niż 50% naukowców z COMP, ENG i IMMU pozostających w decylnie 10. w porównaniu z 57–58% naukowców z MATH i PHYS.

Tabela 3. Mobilność aktualnie najbardziej produktywnych naukowców pomiędzy dwoma etapami kariery: na wczesnym etapie (wyjściowym) i na środkowym etapie (docelowym): z których wyjściowych decyli produktywności (na wczesnym etapie kariery) pochodzą najbardziej produktywni naukowcy na środkowym etapie kariery? Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Najbardziej produktywni naukowcy na późnym etapie kariery ($N= 32\ 063$), którzy byli najbardziej produktywni na środkowym etapie kariery według dyscypliny naukowej i wyjściowego decyla produktywności (częstości i wartości procentowe)

		Łącznie	Dolne 10%	Decyl 2.	Decyl 3.	Decyl 4.	Decyl 5.	Decyl 6.	Decyl 7.	Decyl 8.	Decyl 9.	Górne 10%
Najbardziej produktywni naukowcy na środkowym etapie kariery – według wyjściowego decyla produktywności na wczesnym etapie kariery												
ŁĄCZ- NIE	<i>N</i>	32 063	162	232	323	540	799	1175	2007	3312	6714	16 799
	%	100	0,51	0,72	1,01	1,68	2,49	3,66	6,26	10,33	20,94	52,39
SPO- ŁĘCZNE	<i>N</i>	1259	7	14	12	35	35	48	65	149	234	660
	%	100	0,56	1,11	0,95	2,78	2,78	3,81	5,16	11,83	18,59	52,42
STEMM	<i>N</i>	30 804	155	218	311	505	764	1127	1942	3163	6480	16139
	%	100	0,50	0,71	1,01	1,64	2,48	3,66	6,30	10,27	21,04	52,39
AGRI	<i>N</i>	2373	7	11	9	25	57	72	147	273	510	1262
	%	100	0,29	0,46	0,38	1,05	2,40	3,03	6,19	11,50	21,49	53,18
BIO	<i>N</i>	4582	24	41	50	100	121	187	296	476	933	2354
	%	100	0,52	0,89	1,09	2,18	2,64	4,08	6,46	10,39	20,36	51,37
BUS	<i>N</i>	326	2	4	2	12	7	16	15	39	66	163
	%	100	0,61	1,23	0,61	3,68	2,15	4,91	4,60	11,96	20,25	50,00
CHEM	<i>N</i>	1490	7	11	14	21	34	50	85	149	306	813
	%	100	0,47	0,74	0,94	1,41	2,28	3,36	5,70	10,00	20,54	54,56
COMP	<i>N</i>	765	4	9	20	20	25	29	57	82	150	369
	%	100	0,52	1,18	2,61	2,61	3,27	3,79	7,45	10,72	19,61	48,24
EARTH	<i>N</i>	1437	5	4	14	14	32	51	81	149	297	790
	%	100	0,35	0,28	0,97	0,97	2,23	3,55	5,64	10,37	20,67	54,98
ECON	<i>N</i>	385	1	5	5	15	11	13	23	44	64	204
	%	100	0,26	1,30	1,30	3,90	2,86	3,38	5,97	11,43	16,62	52,99
ENG	<i>N</i>	1282	7	12	14	26	32	59	88	153	259	632
	%	100	0,55	0,94	1,09	2,03	2,50	4,60	6,86	11,93	20,20	49,30
ENVIR	<i>N</i>	652	4	2	4	7	17	25	36	64	137	356
	%	100	0,61	0,31	0,61	1,07	2,61	3,83	5,52	9,82	21,01	54,60
IMMU	<i>N</i>	315	1	3	5	7	11	10	26	28	69	155
	%	100	0,32	0,95	1,59	2,22	3,49	3,17	8,25	8,89	21,90	49,21
MATER	<i>N</i>	584	3	2	5	11	19	27	38	64	115	300
	%	100	0,51	0,34	0,86	1,88	3,25	4,62	6,51	10,96	19,69	51,37
MATH	<i>N</i>	701	3	3	9	11	9	22	37	60	141	406
	%	100	0,43	0,43	1,28	1,57	1,28	3,14	5,28	8,56	20,11	57,92
MED	<i>N</i>	13 108	46	63	113	205	330	491	869	1360	2902	6729
	%	100	0,35	0,48	0,86	1,56	2,52	3,75	6,63	10,38	22,14	51,34
NEURO	<i>N</i>	587	5	6	3	12	16	18	34	51	137	305
	%	100	0,85	1,02	0,51	2,04	2,73	3,07	5,79	8,69	23,34	51,96
PHYS	<i>N</i>	2928	39	51	51	46	61	86	148	254	524	1668
	%	100	1,33	1,74	1,74	1,57	2,08	2,94	5,05	8,67	17,90	56,97
PSYCH	<i>N</i>	548	4	5	5	8	17	19	27	66	104	293
	%	100	0,73	0,91	0,91	1,46	3,10	3,47	4,93	12,04	18,98	53,47

Tabela 4. Mobilność najbardziej produktywnych naukowców pomiędzy dwoma etapami kariery: środkowym (wyjściowym) i późnym (docelowym): z których wyjściowych decyli produktywności (na środkowym etapie kariery) pochodzą najbardziej produktywni naukowcy na jej późnym etapie? Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Najbardziej produktywni naukowcy na późnym etapie kariery ($N=32\,063$) według dyscypliny i wyjściowego decyla produktywności na środkowym etapie kariery (częstości i wartości procentowe)

	Łącznie	Dolne 10%	Decyl 2.	Decyl 3.	Decyl 4.	Decyl 5.	Decyl 6.	Decyl 7.	Decyl 8.	Decyl 9.	Górne 10%	
Najbardziej produktywni naukowcy na późnym etapie kariery – według wyjściowego decyla produktywności na środkowym etapie kariery												
ŁĄCZNIE	<i>N</i>	32 063	436	305	345	541	756	1125	1816	3139	6339	17 261
	%	100	1,36	0,95	1,08	1,69	2,36	3,51	5,66	9,79	19,77	53,83
SPO- LECZNE	<i>N</i>	1259	10	14	21	32	44	55	58	131	249	645
	%	100	0,79	1,11	1,67	2,54	3,49	4,37	4,61	10,41	19,78	51,23
STEMM	<i>N</i>	30 804	426	291	324	509	712	1070	1758	3008	6090	16 616
	%	100	1,38	0,94	1,05	1,65	2,31	3,47	5,71	9,76	19,77	53,94
AGRI	<i>N</i>	2373	23	16	17	34	49	81	131	245	474	1303
	%	100	0,97	0,67	0,72	1,43	2,06	3,41	5,52	10,32	19,97	54,91
BIO	<i>N</i>	4582	65	38	45	85	104	170	266	467	952	2390
	%	100	1,42	0,83	0,98	1,86	2,27	3,71	5,81	10,19	20,78	52,16
BUS	<i>N</i>	326	6	3	4	11	14	16	9	38	70	155
	%	100	1,84	0,92	1,23	3,37	4,29	4,91	2,76	11,66	21,47	47,55
CHEM	<i>N</i>	1490	14	7	17	22	32	42	60	133	291	872
	%	100	0,94	0,47	1,14	1,48	2,15	2,82	4,03	8,93	19,53	58,52
COMP	<i>N</i>	765	11	8	9	19	32	25	55	59	145	402
	%	100	1,44	1,05	1,18	2,48	4,18	3,27	7,19	7,71	18,95	52,55
EARTH	<i>N</i>	1437	12	21	12	23	42	61	83	155	308	720
	%	100	0,84	1,46	0,84	1,60	2,92	4,24	5,78	10,79	21,43	50,10
ECON	<i>N</i>	385	2	8	8	13	14	21	21	31	69	198
	%	100	0,52	2,08	2,08	3,38	3,64	5,45	5,45	8,05	17,92	51,43
ENG	<i>N</i>	1282	13	12	19	26	33	43	81	118	240	697
	%	100	1,01	0,94	1,48	2,03	2,57	3,35	6,32	9,20	18,72	54,37
ENVIR	<i>N</i>	652	8	9	3	11	12	20	47	58	128	356
	%	100	1,23	1,38	0,46	1,69	1,84	3,07	7,21	8,90	19,63	54,60
IMMU	<i>N</i>	315	3	4	5	6	3	16	22	30	61	165
	%	100	0,95	1,27	1,59	1,90	0,95	5,08	6,98	9,52	19,37	52,38
MATER	<i>N</i>	584	8	1	6	10	5	15	30	63	117	329
	%	100	1,37	0,17	1,03	1,71	0,86	2,57	5,14	10,79	20,03	56,34
MATH	<i>N</i>	701	8	3	12	11	14	14	44	71	123	401
	%	100	1,14	0,43	1,71	1,57	2,00	2,00	6,28	10,13	17,55	57,20
MED	<i>N</i>	13 108	187	131	125	200	312	480	771	1344	2671	6887
	%	100	1,43	1,00	0,95	1,53	2,38	3,66	5,88	10,25	20,38	52,54
NEURO	<i>N</i>	587	9	1	7	6	14	10	33	62	119	326
	%	100	1,53	0,17	1,19	1,02	2,39	1,70	5,62	10,56	20,27	55,54
PHYS	<i>N</i>	2928	65	40	47	56	60	93	135	203	461	1768
	%	100	2,22	1,37	1,61	1,91	2,05	3,18	4,61	6,93	15,74	60,38
PSYCH	<i>N</i>	548	2	3	9	8	16	18	28	62	110	292
	%	100	0,36	0,55	1,64	1,46	2,92	3,28	5,11	11,31	20,07	53,28

Analogicznie, badając mobilność pomiędzy środkowym i późnym etapem kariery w przypadku najbardziej produktywnych naukowców (tab. 4), stabilność najwyższej produktywności okazuje się jeszcze bardziej uderzająca: 53,83% naukowców, którzy znajdowali się w decylnym 10. – pozostało w decylnym 10. Tylko jeden na sześciu (16,61%) naukowców w decylnym 10 nie pochodzi wyjściowo z decyli 8–10., a tylko co trzydziesty pochodzi z decyli 1–3. (3,39%).

Wśród wszystkich obecnie najbardziej produktywnych naukowców społecznych na późnym etapie kariery w 38 krajach jest tylko 10 skoczków: dwóch ekonomistów (ECON: 0,52%), dwóch psychologów (PSYCH: 0,36%) i sześciu naukowców z dyscypliny biznes, zarządzanie i rachunkowość (BUS: 0,61%). Również w przypadku tych naukowców dysponujemy kompleksowymi danymi dotyczącymi ich demografii, wzorców publikowania i współpracy naukowej oraz wpływu na naukę poprzez cytowania.

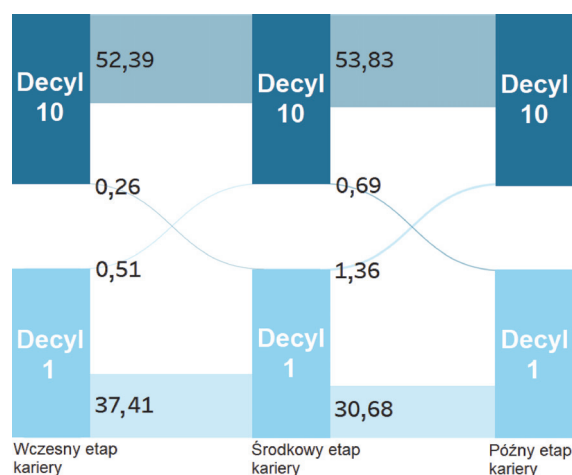
Analogiczne analizy przeprowadzono dla najniższych decyli produktywności. Tabela uzupełniająca 3 pokazuje pochodzenie decylowe aktualnie najmniej produktywnych naukowców (decyl produktywności 1.), przedstawiając ich wyjściowe decyle produktywności na wczesnych etapach kariery. Zaobserwowane wzorce są podobne, ale mniej wyraźne niż wzorce występujące w przypadku najbardziej produktywnych naukowców. Trzy czwarte (75,31%) najmniej produktywnych naukowców pochodziło z trzech najniższych decyli produktywności (decyle 1–3.), a ponad jedna trzecia z najniższego decyla (37,41%). I odwrotnie, tylko 2,28% (728 naukowców) pochodziło z trzech najwyższych decyli, w tym zaledwie 0,26% (82 naukowców) z najwyższego decyla 10. (spadkowicze). Tabela uzupełniająca 4 pokazuje pochodzenie decylowe obecnych najmniej produktywnych naukowców na środkowym etapie kariery, odsłaniając podobne wzorce mobilności.

Szczególnie interesująca dla polityki naukowej jest skrajna mobilność w górę, ponieważ to jeden z wymiarów sukcesu publikacyjnego; natomiast skrajna mobilność w dół może być przypisywana osobistym okolicznościom, takim jak względy zdrowotne czy problemy rodzinne, których nie można poddać analizie przy wykorzystaniu zbiorów danych bibliometrycznych.

3.2. Mobilność między klasami produktywności: wszystkie dyscypliny naukowe łącznie

Wykres strumieniowy (wykres Sankeya, ryc. 5) może posłużyć jako wizualny przewodnik pozwalający lepiej zrozumieć koncepcję mobilności naukowców między różnymi klasami produktywności w trakcie ich kariery. Wykres ten ilustruje przechodzenie naukowców między decylami produktywności na różnych etapach kariery: pomiędzy jej wczesnym (lewa kolumna), środkowym (środkowa kolumna) i późnym etapem kariery (prawa kolumna). Koncentrujemy się na mobilności poziomej (decyl 10. – decyl 10. i decyl 1. – decyl 1.) oraz na przejściach obejmujących ekstremalną mobilność w dół i ekstremalną mobilność w górę, charakterystycznych dla spadkowiczów i skoczków.

Rycina 5 prezentuje mobilność naukowców we wszystkich dyscyplinach naukowych łącznie ($N=320\,564$), ryc. 6 w naukach społecznych ($N=12\,585$), a ryc. 7 – w dyscyplinach STEMM ($N=307\,979$). Lewe kolumny wykresów odnoszą się do naukowców na wczesnym etapie kariery w najwyższym i najniższym decylnie produktywności (suma wartości dla każdego decylu wynosi 100%), środkowe kolumny do naukowców na środkowym etapie, a prawe – naukowców na późnym etapie kariery w tych samych dwóch klasach produktywności. Aby zwiększyć przejrzystość wykresu, decyle od 2. do 9. zostały usunięte. Dane wskazują na odsetki naukowców zaangażowanych w przejścia między klasami produktywności między trzema etapami kariery.

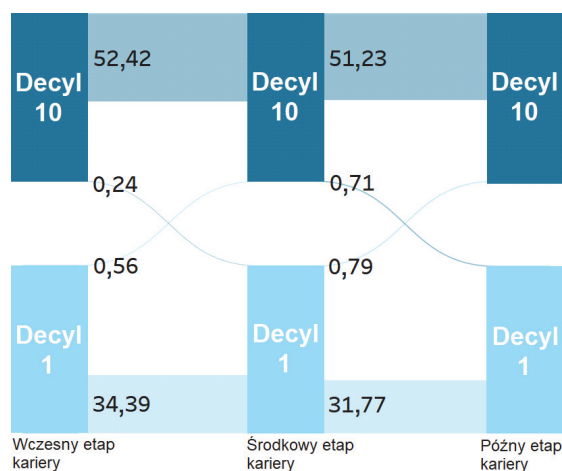


Ryc. 5. Wykres strumieniowy. Mobilność naukowców między klasami produktywności na trzech etapach kariery naukowej. Wszystkie dyscypliny naukowe łącznie, naukowcy aktualnie znajdujący się na późnym etapie kariery. Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Wszystkie obserwacje uszeregowane i pogrupowane w decyle produktywności, pokazane tylko najwyższe (górne 10%, decyl produktywności 10.) i najniższe (dolne 10%, decyl produktywności 1.) klasy produktywności ($N=320\,564$) (wartości procentowe, górna klasa i dolna klasa, każda klasa łącznie obejmuje 100% naukowców)

Pozioma mobilność między wczesnym i środkowym etapem kariery jest na rysunkach reprezentowana przez grube strumienie: jest tak dlatego, że ponad połowa globalnie najbardziej produktywnych naukowców kontynuuje karierę jako najbardziej produktywni naukowcy (52,39%), a ponad jedna trzecia globalnie najmniej produktywnych naukowców – jako najmniej produktywni naukowcy (37,41%).

Ekstremalna mobilność pionowa jest rzadka i pokazana jako niewielkie strumienie prowadzące w dół i w górę: tylko 0,51% najmniej produktywnych naukowców na wczesnym etapie kariery (162 osoby) przenosi się do klasy najbardziej produktywnych naukowców na środkowym etapie kariery (i tylko 0,26%, czyli 82 osoby, przenosi się w przeciwną stronę).

Wzorce mobilności są bardzo podobne między naukami społecznymi (ryc. 6) i dyscyplinami STEMM (ryc. 7): zaskakujące jest to, że mimo odmiennych wzorców publikowania i wzorców współpracy w nauce (zob. zwłaszcza Kwiek 2021 dla krajów europejskich), mobilność między najwyższym i najniższym decylem produktywności jest zbliżona. W naukach społecznych ekstremalna mobilność w górę – od decyla 1. do decyla 10. – jest równie rzadka, jak w dyscyplinach STEMM.

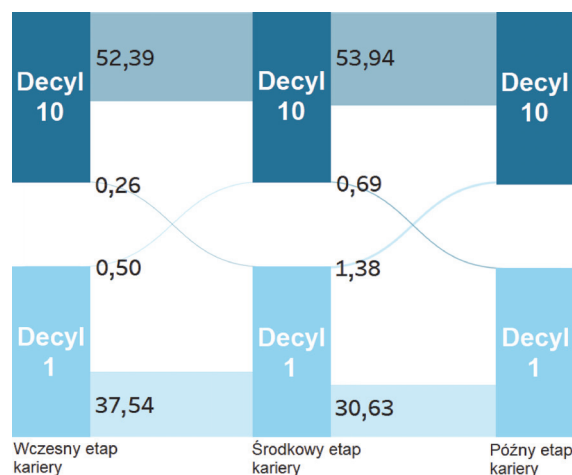


Ryc. 6. Wykres strumieniowy. Mobilność naukowców między klasami produktywności na trzech etapach kariery naukowej. Wszystkie dyscypliny nauk społecznych, naukowcy aktualnie będący na późnym etapie kariery. Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Wszystkie obserwacje uszeregowane i pogrupowane w decyle produktywności, tylko najwyższe (górne 10%, decyl produktywności 10.) i najniższe (dolne 10%, decyl produktywności 1.) klasy produktywności ($N = 12\,585$) (wartości procentowe, górna i dolna klasa, każda klasa łącznie obejmuje 100% naukowców)

W naszej próbie naukowców należących do nauk społecznych ($N = 12\,585$) znalazło się tylko 7 osób zaangażowanych w pierwsze przejście i tylko 10 osób zaangażowanych w drugie przejście (na 1259 naukowców, tabela 5). Zarówno tutaj, jak i w innych miejscach pracy nie pokazujemy istotności statystycznej różnic, ponieważ pracujemy na populacji naukowców (wszyscy naukowcy spełniający kryteria włączenia), a nie na ich próbie (wyborze spośród wszystkich naukowców).

Z perspektywy zagregowanej, obejmującej wszystkie dyscypliny łącznie (tabela 5), wzorce mobilności są jednoznaczne: ponad połowa (52,39% i 53,83%) najbardziej produktywnych naukowców (decyl 10.) pozostaje w tym samym, najwyższym decylnie. Jednocześnie jedna trzecia (37,41% i 30,68%) naukowców z najniższego decylu produktywności nadal w nim pozostaje. Sugeruje to istnienie intrygującego mechanizmu stabilności (czy blokowania) w karierze akademickiej, który wymaga dalszych analiz, zarówno ilościowych, jak i jakościowych.

Warto podkreślić, że w ramach naszego podejścia metodologicznego klasyfikujemy wszystkich obecnych naukowców znajdujących się na późnym etapie kariery pod względem produktywności, przypisując ich do określonych klas produktywności w ramach odpowiednich dyscyplin. Następnie retrospektywnie klasyfikujemy tych naukowców na podstawie ich produktywności na wczesnym i środkowym etapie kariery, wykorzystując czteroletnie okresy do pomiaru ich produktywności w tym czasie („decyle wyjściowe” w mobilności na każdym etapie pod względem produktywności w tabeli 5).



Ryc. 7. Wykres strumieniowy. Mobilność naukowców między klasami produktywności na trzech etapach kariery naukowej. Wszystkie dyscypliny akademickie STEM łącznie, naukowcy aktualnie znajdujący się na późnym etapie kariery. Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Wszystkie obserwacje uszeregowane i pogrupowane w decyle produktywności, tylko najwyższe (górne 10%, decyl produktywności 10) i najniższe (dolne 10%, decyl produktywności 1.) klasy produktywności ($N=307\,979$) (wartości procentowe, górna i dolna klasa, każda klasa łącznie obejmuje 100% naukowców)

Prawdopodobieństwo wystąpienia skrajnej mobilności w górę (przejście od decyla 1. do decyla 10.) i skrajnej mobilności w dół (przejście od decyla 10. do decyla 1.) między klasami produktywności jest niezwykle niskie. Z naszych danych (tabela 5) wynika, że w kontekście zastosowanego przez nas znormalizowanego pod względem prestiżu podejścia do zliczania publikacji (*full counting, journal prestige normalized approach*), szanse na radykalną zmianę zachowań publikacyjnych między etapami kariery naukowej są minimalne (pokażemy dalej, że w przypadku Polski, szanse te są zerowe, 0%).

Tylko 162 najbardziej produktywnych naukowców (0,51%), którzy stali się najbardziej produktywni na środkowym etapie kariery, znajdowało się wcześniej w najniższej klasie produktywności na wczesnym etapie kariery. Jednocześnie tylko 436 najbardziej produktywnych naukowców (1,36%), którzy zostali najbardziej produktywnymi naukowcami na późnym etapie kariery, znajdowało się początkowo w najniższej klasie produk-

Tabela 5. Mobilność między najwyższymi (decyl 10.) i najniższymi (decyl 1.) klasami produktywności podczas przechodzenia z wczesnego do środkowego etapu kariery oraz ze środkowego do późnego etapu kariery dla wszystkich dyscyplin naukowych łącznie ($N = 320\ 564$), dyscyplin nauk społecznych ($N = 12\ 585$) i dyscyplin STEM ($N = 307\ 979$). Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma (częstości i wartości procentowe)

Przejście z etapu...	Wyjściowy decyl produktywności	Przejście do etapu...	Docelowy decyl produktywności	Liczba naukowców w danej mobilności	Liczba naukowców w danej klasie produktywności	%
ŁĄCZNIE						
Wczesny etap	Najniższy	Środkowy etap	Najniższy	11 996	32 063	37,41
Wczesny etap	Najniższy	Środkowy etap	Najwyższy	162	32 063	0,51
Wczesny etap	Najwyższy	Środkowy etap	Najniższy	82	32 063	0,26
Wczesny etap	Najwyższy	Środkowy etap	Najwyższy	16 799	32 063	52,39
Środkowy etap	Najniższy	Późny etap	Najniższy	9836	32 063	30,68
Środkowy etap	Najniższy	Późny etap	Najwyższy	436	32 063	1,36
Środkowy etap	Najwyższy	Późny etap	Najniższy	222	32 063	0,69
Środkowy etap	Najwyższy	Późny etap	Najwyższy	17 261	32 063	53,83
DYSCYPLINY SPOŁECZNE						
Wczesny etap	Najniższy	Środkowy etap	Najniższy	433	1259	34,39
Wczesny etap	Najniższy	Środkowy etap	Najwyższy	7	1259	0,56
Wczesny etap	Najwyższy	Środkowy etap	Najniższy	3	1259	0,24
Wczesny etap	Najwyższy	Środkowy etap	Najwyższy	660	1259	52,42
Środkowy etap	Najniższy	Późny etap	Najniższy	400	1259	31,77
Środkowy etap	Najniższy	Późny etap	Najwyższy	10	1259	0,79
Środkowy etap	Najniższy	Późny etap	Najniższy	9	1259	0,71
Środkowy etap	Najwyższy	Późny etap	Najwyższy	645	1259	51,23
DYSCYPLINY STEM						
Wczesny etap	Najniższy	Środkowy etap	Najniższy	11 563	30 804	37,54
Wczesny etap	Najniższy	Środkowy etap	Najwyższy	155	30 804	0,5
Wczesny etap	Najwyższy	Środkowy etap	Najniższy	79	30 804	0,26
Wczesny etap	Najwyższy	Środkowy etap	Najwyższy	16 139	30 804	52,39
Środkowy etap	Najniższy	Późny etap	Najniższy	9436	30 804	30,63
Środkowy etap	Najniższy	Późny etap	Najwyższy	426	30 804	1,38
Środkowy etap	Najwyższy	Późny etap	Najniższy	213	30 804	0,69
Środkowy etap	Najwyższy	Późny etap	Najwyższy	16 616	30 804	53,94

tywności na środkowym etapie kariery. Szanse na ekstremalną mobilność w dół są również rzadkie.

Dane te wskazują, że radykalne zmiany w zachowaniach publikacyjnych, które prowadzą do tak znaczących przemieszczeń w obrębie klas produktywności, są zjawiskiem nadzwyczajnym. I nie można na nie liczyć przy planowaniu przyszłości przez instytucje naukowe.

3.3. Mobilność między klasami produktywności: różnicowanie dyscyplinarne

Zagregowane ujęcia dyscyplin łącznie przesłaniają znacznie bardziej zniuansowany obraz poszczególnych dyscyplin, z odmiennymi wzorcami mobilności między klasami produktywności. Dla prawie wszystkich dyscyplin, ponad 50% najbardziej produktywnych naukowców pozostaje naukowcami o najwyższej produktywności (tabela 6 – podsumowanie wyników). Najwyższy odsetek zaobserwowano dla MATH i PHYS w obu przejściach (a także CHEM w drugim przejściu), osiągający poziom 60% dla PHYS w drugim przejściu. Naukowcy reprezentujący skrajną mobilność pionową (skoczkiwie), pojawiają się niezwykle rzadko we wszystkich dyscyplinach: ich odsetek waha się od 0,29% w przypadku AGRI do 1,33% w przypadku PHYS na pierwszym etapie kariery i od 0,36% dla PSYCH i 1,84% dla BUS w jej drugim etapie.

Tabela 6. Podsumowanie czterech typów mobilności pod względem produktywności między wczesnym i środkowym oraz środkowym i późnym etapem kariery. Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma, według dyscypliny, w procentach ($N=320\ 564$)

Dyscyplina	Mobilność decyl 10. – decyl 10.		Mobilność decyl 1. – decyl 1.		Mobilność decyl 10. – decyl 1. (spadkowicze)		Mobilność decyl 1. – decyl 10. (skoczkiwie)	
	Wczesny – środkowy etap kariery (%)	Środkowy – późny etap kariery (%)	Wczesny – środkowy etap kariery (%)	Środkowy – późny etap kariery (%)	Wczesny – środkowy etap kariery (%)	Środkowy – późny etap kariery (%)	Wczesny – środkowy etap kariery (%)	Środkowy – późny etap kariery (%)
AGRI	53,18	54,91	38,6	32,2	0,17	0,25	0,29	0,97
BIO	51,37	52,16	36,56	28,26	0,33	0,87	0,52	1,42
BUS	50	47,55	30,37	26,38	0,31	0,61	0,61	1,84
CHEM	54,56	58,52	41,01	33,42	0,07	0,4	0,47	0,94
COMP	48,24	52,55	32,42	28,37	0,13	0,26	0,52	1,44
EARTH	54,98	50,1	40,36	35	0,21	0,28	0,35	0,84
ECON	52,99	51,43	29,87	29,09	0,26	0,78	0,26	0,52
ENG	49,3	54,37	36,97	29,88	0,23	0,47	0,55	1,01
ENVIR	54,6	54,6	35,58	28,68	0,46	0,46	0,61	1,23
IMMU	49,21	52,38	35,87	35,87	0,63	–	0,32	0,95
MATER	51,37	56,34	38,7	32,53	–	0,34	0,51	1,37
MATH	57,92	57,2	36,23	32,38	0,29	0,29	0,43	1,14
MED	51,34	52,54	36,37	29,48	0,16	0,82	0,35	1,43
NEURO	51,96	55,54	39,18	31,35	0,34	0,34	0,85	1,53
PHYS	56,97	60,38	42,25	34,53	0,75	1,13	1,33	2,22
PSYCH	53,47	53,28	39,96	36,86	0,18	0,73	0,73	0,36
SPOŁECZ.	52,42	51,23	34,39	31,77	0,24	0,71	0,56	0,79
STEMM	52,39	53,94	37,54	30,63	0,26	0,69	0,5	1,38
ŁĄCZNIE	52,39	53,83	37,41	30,68	0,26	0,69	0,51	1,36

Uwaga: „–” = brak naukowców biorących udział w mobilności

3.4. Przypadek Polski: doktorzy i doktorzy habilitowani

Przedstawiamy tu również krótko polski przypadek przeanalizowany przez nas szczegółowo w innym miejscu (Kwiek i Roszka 2024b). Polskie badania były prowadzone równoległe z badaniami międzynarodowymi, przy czym mikrodane pozyskane z OPI PIB były o wiele bardziej szczegółowe niż mikrodane pozyskane z bazy Scopus. Zbadaliśmy zmieniającą się produktywność 4165 polskich naukowców specjalizujących się w dziedzinach STEMM.

Wszyscy przeanalizowani polscy naukowcy byli doktorami habilitowanymi zatrudnionymi w pełnym wymiarze czasu pracy w sektorze szkolnictwa wyższego. Łącząc dane demograficzne i biograficzne pochodzące z krajowego rejestru naukowców (OPI PIB, $N = 99\,935$) z danymi pochodzącymi z własnych obliczeń opartych na metadanych pochodzących z bazy Scopus i dotyczących wszystkich polskich artykułów naukowych indeksowanych w ciągu ostatniego półwiecza (1973–2021, $N = 935\,167$), przeanalizowaliśmy, jak indywidualni naukowcy zmieniali klasy produktywności w czasie. W polskim badaniu był to okres obejmujący maksymalnie 40 lat (zakres wieku biologicznego w próbie: 30–70 lat). Skoncentrowaliśmy się na dwóch etapach kariery: na etapie kariery po doktoracie i po habilitacji (od roku uzyskania danego stopnia).

Ponieważ w polskim badaniu liczba naukowców w poszczególnych dyscyplinach była niska, pogrupowaliśmy 12 dyscyplin STEMM (w ujęciu AJSC z bazy Scopus) w 5 dziedzin STEMM. Uwzględniono następujące dziedziny nauki: ENG (inżynieria, obejmująca inżynierię i materiałoznawstwo); LIFE (nauki przyrodnicze, obejmujące nauki rolnicze i biologiczne oraz biochemię, genetykę i biologię molekularną); MATH (matematyka, obejmująca matematykę i informatykę); MED (nauki medyczne); oraz NATURAL (nauki przyrodnicze, obejmujące inżynierię chemiczną, chemię, fizykę i astronomię, nauki o Ziemi i nauki o środowisku).

Polskie wyniki odzwierciedlają wyniki badań dla 38 krajów OECD (tabela 7). Prawie połowa aktualnie najbardziej produktywnych doktorów habilitowanych (decyl 10.) wywodzi się z 10. decyla produktywności w okresie, gdy byli oni doktorami (46,5%): nadal znajdują się w tym samym decylu produktywności (ponadto 17,7% pochodzi z decyla 9., a 8,6% z decyla 8.). W sumie trzy czwarte z nich należało w przeszłości do decyli produktywności 8–10. (72,8%). Prawie żaden z najbardziej produktywnych doktorów habilitowanych nie należał wcześniej do najniższych trzech decyli produktywności.

Co najważniejsze, nie znalazł się w Polsce ani jeden skoczek: żaden naukowiec nie doświadczył ekstremalnej mobilności w górę z decyla 1. do decyla 10. i tylko jeden naukowiec (znajdujący się w NATURAL) doświadczył mobilności w górę z decyla 2. do decyla 10. (mamy pełne profile biograficzne i publikacyjne każdego naukowca, w tym tego jednego, wyjątkowego przypadku).

Biorąc pod uwagę dziedziny naukowe, połowa najbardziej produktywnych doktorów habilitowanych (50,0%) była również najbardziej produktywnymi doktorami w dziedzinie nauk przyrodniczych, a 47,9% w dziedzinie inżynierii. W przypadku inżynierii 74,0% najbardziej produktywnych doktorów habilitowanych pochodzi z trzech najwyższych

decyli, a żaden z nich nie pochodzi z trzech najniższych decyli. W dziedzinach LIFE i MATH odsetki te wynoszą odpowiednio 67,8% i 70,7% oraz 2,2% i 0%.

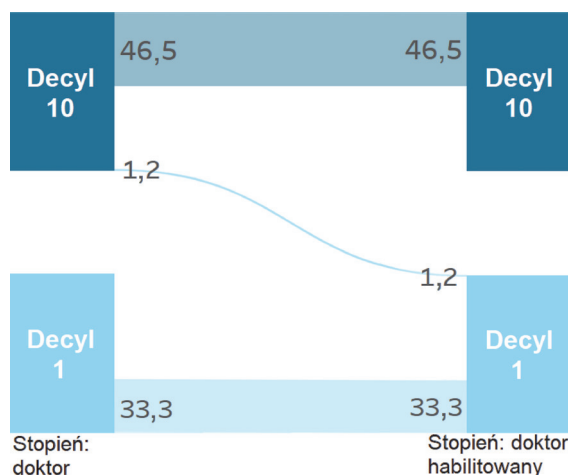
Tabela 7. Polscy naukowcy: mobilność najbardziej produktywnych naukowców między decylami produktywności pomiędzy dwoma etapami kariery – etapem pracy po doktoracie (wyjściowym) i etapem pracy po habilitacji (docelowym). Z których wyjściowych decyli produktywności (jako doktorzy) pochodzą obecni najbardziej produktywni doktorzy habilitowani? Najbardziej produktywni (N = 419) doktorzy habilitowani według dziedziny i wyjściowego decyla produktywności (typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma) (częstości i wartości procentowe)

		Łącznie	Dolne 10%	Decyl 2.	Decyl 3.	Decyl 4.	Decyl 5.	Decyl 6.	Decyl 7.	Decyl 8.	Decyl 9.	Górne 10%
Najbardziej produktywni doktorzy habilitowani – według wyjściowego decyla produktywności na etapie pracy po doktoracie												
ŁĄCZ- NIE	N	419	0	1	7	34	37	15	20	36	74	195
	%	100	0,0	0,2	1,7	8,1	8,8	3,6	4,8	8,6	17,7	46,5
ENG	N	96	0	0	2	10	9	1	3	7	18	46
	%	100,0	0,0	0,0	2,1	10,4	9,4	1,0	3,1	7,3	18,8	47,9
LIFE	N	90	0	0	2	10	7	4	6	7	16	38
	%	100,0	0,0	0,0	2,2	11,1	7,8	4,4	6,7	7,8	17,8	42,2
MATH	N	41	0	0	0	2	9	1	0	4	6	19
	%	100,0	0,0	0,0	0,0	4,9	22,0	2,4	0,0	9,8	14,6	46,3
MED	N	64	0	0	1	4	4	4	5	7	11	28
	%	100,0	0,0	0,0	1,6	6,3	6,3	6,3	7,8	10,9	17,2	43,8
NATU- RAL	N	128	0	1	2	8	8	5	6	11	23	64
	%	100,0	0,0	0,8	1,6	6,3	6,3	3,9	4,7	8,6	18,0	50

Rycina 8 przedstawia mobilność polskich naukowców we wszystkich dziedzinach łącznie. Lewa kolumna odnosi się do doktorów w najwyższej i najniższej klasie produktywności (łącznie 100% w każdej klasie), a prawa kolumna do doktorów habilitowanych w tych samych dwóch klasach produktywności. Strumienie pokazują odsetek naukowców doświadczających cztery typy mobilności (w tym: skoczków i spadkowiczów).

Podobnie jak w przypadku wizualizacji w postaci wykresów strumieniowych sporządzonych dla krajów OECD, pozioma mobilność jest przedstawiona za pomocą grubych strumieni. Ekstremalna pionowa mobilność w dół między decylem 10. i decylem 1. (spadkowicze) jest rzadka i jest przedstawiona jako cienki przepływ w dół między dwoma kolumnami); natomiast – i to jest chyba najciekawsze – przepływ w górę między decylem 1. i decylem 10. w ogóle nie jest pokazany, ponieważ po prostu w Polsce nie występuje (0%).

Tylko 1,2% najbardziej produktywnych doktorów (dokładnie pięciu naukowców) spadło do klasy najmniej produktywnych doktorów habilitowanych i żaden najmniej produktywny doktor (0% naukowców) nie dostał się do klasy najbardziej produktywnych doktorów habilitowanych (tabela 8). Skoczkiw w Polsce nie występują (a w krajach OECD pojawiają się w nieznacznych ilościach).



Ryc. 8. Polscy naukowcy: mobilność naukowców między klasami produktywności na dwóch etapach kariery naukowej (typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma). Wszystkie dziedziny STEMM łącznie, obecni doktorzy habilitowani. Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma. Wszystkie obserwacje uszeregowane i pogrupowane w decyle produktywności, najwyższe klasy produktywności (górnych 10%, decyl produktywności 10, $N=419$) i najniższe klasy produktywności (dolnych 10%, decyl produktywności 1, $N=412$) (wartości procentowe, tylko najwyższa klasa i najniższa klasa, każda klasa łącznie obejmuje 100% naukowców)

Jakie są szanse na skrajną mobilność w górę pomiędzy klasami produktywności (od decyla 1. do decyla 10.) lub skrajną mobilność w dół (od decyla 10. do decyla 1.)? Nasze dane (tabela 8) wyraźnie pokazują, że w Polsce nie ma żadnych szans na maksymalną mobilność w górę: żaden z obecnych 412 najbardziej produktywnych naukowców z habilitacją nie był najmniej produktywnym naukowcem z doktoratem (0%). Szanse na skrajną mobilność w dół są nieco wyższe, ale nadal znikome i wynoszą około 1% (5 naukowców na 419, czyli 1,19%).

Podsumowując: polscy doktorzy habilitowani zazwyczaj są osadzeni w swoich klasach produktywności przez wiele lat, a często – dekad: najbardziej produktywni pozostają najbardziej produktywni, a najmniej produktywni – pozostają najmniej produktywni. Najbardziej produktywni doktorzy habilitowani byli w przeszłości przede wszystkim najbardziej produktywnymi doktorami (mediana percentyli produktywności na etapie przed habilitacją: 87,9), a najmniej produktywni doktorzy habilitowani byli w przeszłości

przede wszystkim najmniej produktywnymi doktorami (mediana percentyli produktywności na etapie przed habilitacją: 18,3).

Tabela 8. Polscy naukowcy: mobilność między najwyższymi (decyl 10.) i najniższymi (decyl 1.) klasami produktywności podczas przechodzenia od etapu pracy po doktoracie do etapu pracy po habilitacji. Typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma, wszystkie dziedziny nauki łącznie (częstości i wartości procentowe) ($N = 4165$)

Etap pracy (wyjściowy)	Decyl produktywności po doktoracie (e. wyjściowy)	Etap pracy (docelowy)	Decyl produktywności po habilitacji (e. docelowy)	Liczba naukowców w danej mobilności	Liczba naukowców w klasie produktywności	%
Po doktoracie	Najniższy (1)	Po habilitacji	Najniższy (1)	137	412	33,3
Po doktoracie	Najniższy (1)	Po habilitacji	Najwyższy (10)	0	412	0
Po doktoracie	Najwyższy (10)	Po habilitacji	Najniższy (1)	5	419	1,2
Po doktoracie	Najwyższy (10)	Po habilitacji	Najwyższy (10)	195	419	46,5

Innymi słowy, niektórzy naukowcy są wysoce produktywni przez lata i dziesięciolecia (jak pokazujemy gdzie indziej w przypadku profesorów tytularnych, Kwiek i Roszka 2024a), a inni – ich koledzy w ramach tych samych dziedzin nauki – są z kolei mało produktywni przez lata. Istnieje zerowe prawdopodobieństwo, że naukowcy będą radykalnie bardziej (i znikome prawdopodobieństwo, że będą radykalnie mniej) produktywni, wspinając się po szczeblach kariery akademickiej.

Wyniki analizy regresji logistycznej potwierdziły nasze dwuwymiarowe wyniki. W przypadku szacowania ilorazu szans przynależności do klasy najbardziej produktywnych doktorów habilitowanych, jeden predyktor okazał się najważniejszy we wszystkich czterech modelach: przynależność do klasy najbardziej produktywnych doktorów na wcześniejszym etapie kariery. Ta wcześniejsza przynależność była statystycznie istotna we wszystkich modelach w podobnie wysokim stopniu, zwiększając szanse na sukces 4–6 razy, w zależności od wybranego typu produktywności (zastosowaliśmy cztery typy). Uprzednia przynależność do klasy szybko awansujących doktorów habilitowanych była również statystycznie istotna we wszystkich modelach. Przynależność ta zwiększa szanse na sukces o 50–130%, w zależności od modelu. W przypadku szacowania ilorazu szans przynależności do klasy najmniej produktywnych doktorów habilitowanych, najsilniejszym czynnikiem predykcyjnym okazała się wcześniejsza przynależność do klasy najmniej produktywnych doktorów. Prawdopodobieństwo sukcesu rosło o 150–300% (Kwiek i Roszka 2024b).

Z naszych polskich badań wynika, że jeśli okres pracy przed habilitacją jest silny pod względem osiągnięć publikacyjnych, to okres pracy po habilitacji jest również silny; i analogicznie, jeśli pierwszy okres jest słaby pod względem publikacji, to okres drugi jest również słaby. Skoczki w badanym polskim kontekście – praca po doktoracie/praca po habilitacji – nie istnieją.

3.5. Ograniczenia: metodologiczne, bazodanowe, praktyczne

Nasze analizy z konieczności opierają się na kilku zmiennych przybliżonych i wiążą się z kilkoma ograniczeniami, co może wpływać na ostateczny kształt wyników i ich interpretację. Analizy opierają się na następujących przybliżeniach:

- (1) komercyjna, globalna klasyfikacja czasopism i dyscyplin naukowych (system ASJC), a nie bogactwo 38 klasyfikacji krajowych;
- (2) dane dotyczące indywidualnych identyfikatorów autora w bazie Scopus, a nie dane dotyczące „realnych naukowców” z identyfikatorami osadzonymi w krajowych rejestrach (co prowadzi do fundamentalnej różnicy ontologicznej między tradycyjnymi badaniami karier akademickich – i badaniami karier opartymi na danych bibliometrycznych);
- (3) płeć domyślna, a nie zadeklarowana lub potwierdzona administracyjnie (w tym przypadku na podstawie algorytmów określających płeć – zastosowany próg prawdopodobieństwa: 0,85; zob. prace z bazą OPI PIB: Kwiek i Roszka 2021a; 2021b);
- (4) afiliacja do jednego, dominującego kraju i jednej, dominującej instytucji, a nie zmieniające się afiliacje w trakcie kariery akademickiej (przynajmniej dla pewnego odsetka naukowców);
- (5) kariera naukowa rozpoczynająca się od pierwszej indeksowanej publikacji i dwudziestopięcioletnie doświadczenia badawcze liczone na podstawie daty pierwszej publikacji (tj. kariera publikacyjna) – a nie pierwsze zatrudnienie w sektorze akademickim lub poza nim (tj. kariera akademicka czy naukowa).

W naszych badaniach w zaproponowanej definicji produktywności pominęliśmy szeroki zakres działalności naukowców w środowisku akademickim (m.in. kształcenie, opiekę nad studentami, recenzowanie artykułów naukowych, recenzowanie wniosków grantowych czy pracę przy redagowaniu czasopism). Dlatego szeroka kategoria „produktywność badawcza” oznacza w praktyce węższą kategorię „produktywność publikacyjna”.

Mówiąc bardziej ogólnie: prezentowane badanie odzwierciedla kompromis między tym, co jest teoretycznie wskazane, a tym, co jest praktycznie możliwe w badaniu globalnych wzorców indywidualnej produktywności (tutaj: w oparciu o obecnie dostępne dane pochodzące z globalnych zbiorów danych bibliometrycznych).

Istnieją również kompromisy i ograniczenia związane z danymi i przyjętą przez nas metodologią.

Po pierwsze, nie ma innych globalnie dostępnych i niezawodnych, podłużnych zbiorów danych (dla 38 krajów OECD) niż Scopus (lub Web of Science), które można by w sensowny sposób wykorzystać do analizy zmieniającej się produktywności znormalizowanej do dyscypliny i do prestiżu czasopisma na przestrzeni całego życia naukowców. Pod względem ujednocznienia autorów publikacji baza Scopus jest dokładniejsza niż baza Web of Science (Sugimoto i Larivière 2018: 36), ale z pewnością nie jest doskonała. Bazy otwarte (typu OpenAlex) na razie do podobnych badań się nie nadają.

Krajowe zbiory danych są dostępne tylko dla wybranych krajów i tylko dla określonych parametrów. Dlatego nie jest obecnie możliwe podłużne, globalne i oparte na dyscyplinach (w przeciwieństwie do badań wybranych krajów) podejście do produktywności publikacyjnej bez dostępu do globalnych zbiorów danych bibliometrycznych, które zapewniają metadane dotyczące wszystkich indeksowanych publikacji w ujęciu czasu. Jednak globalne zbiory danych bibliometrycznych mają swoje własne ograniczenia, szeroko dyskutowane od co najmniej dwóch dekad (np. język angielski i dyscypliny STEM).

Po drugie, charakter naszego zbioru danych wymusił zawężone rozumienie indywidualnej produktywności, w którym liczone są tylko publikacje indeksowane w bazie Scopus, z pominięciem nieindeksowanych publikacji w języku angielskim i większości publikacji w językach lokalnych (na przykład po polsku). Jednak wybór dyscyplin z obszaru STEM i trzech dyscyplin z nauk społecznych (szeroko wykorzystujących język angielski do globalnej komunikacji naukowej), sprawia, że nasze badanie jest mniej stronnicze, niż gdyby obejmowało wszystkie dyscypliny (dyscypliny z obszaru STEM są uwzględnione w globalnych zbiorach danych w znacznie większej mierze niż tradycyjne dyscypliny humanistyczne).

Po trzecie, podłużny charakter naszego badania sprawia, że koncentrujemy się tylko na tych naukowcach, którzy przetrwali w nauce (*survivors in science*): pomijamy wszystkich, którzy nie są aktywni badawczo od co najmniej 25 lat, zgodnie z naszą definicją naukowców znajdujących się na późnym etapie kariery. Z tego względu, zdając sobie sprawę z istnienia wysokich wskaźników rezygnacji z nauki w dyscyplinach STEM (w krajach OECD, jak ostatnio przeanalizowaliśmy w Kwiek i Szymula 2024; oraz w USA, jak ostatnio pokazano w Spoon i in. 2023), uznajemy pewne „skrzywienie związane z sukcesem” w naszych badaniach. Przyjmujemy perspektywę długoterminową, w której z konieczności, ze względu na wskaźnik rezygnacji z pracy naukowej sięgający 70–80% po 20 latach, większość obecnie aktywnych naukowców nie jest reprezentowana. Nie pracuje od co najmniej 25 lat i nie da się przypisać ich retrospektywnie do analogicznych klas produktywności na wcześniejszych etapach kariery.

Po czwarte, zastosowana metodologia ma ograniczenia, które są szczególnie widoczne, jeśli porównamy prezentowane badanie z naszymi badaniami podłużnymi przeprowadzonymi dla Polski (Kwiek i Roszka 2024b). W badaniach krajowych wykorzystujemy wiele typów danych krajowych – jak choćby dane dotyczące roku urodzenia, roku uzyskania stopni naukowych i tytułu profesorskiego – które nie są dostępne dla 38 krajów OECD. Ponadto nasze globalne badanie obejmuje naukowców pochodzących z systemów o różnych poziomach finansowania badań i różnej przeciętnej indywidualnej produktywności publikacyjnej.

I po piąte, nasze podejście do produktywności publikacyjnej wykorzystuje wskaźniki wpływu mierzone na poziomie czasopisma, a nie bardziej szczegółowe wskaźniki wpływu mierzone na poziomie artykułu (np. 4-letni wskaźnik cytowań dla każdej publikacji), ze wszystkimi tego ograniczeniami. Drugi wskaźnik jest wykorzystywany wyłącznie w analizach regresji logistycznej, których tu nie przywołujemy.

4. Dyskusja i wnioski

Nasze badanie dotyczące krajów OECD obejmuje 79,42% wszystkich naukowców znajdujących się na późnym etapie kariery naukowej na świecie w 2023 r. (i 83,03% wszystkich artykułów badawczych opublikowanych przez naukowców znajdujących się w 2023 r. na tym etapie). Na podstawie podłużnego badania opartego na mikrodanych dotyczących setek tysięcy naukowców wskazujemy, że już na stosunkowo wczesnym etapie kariery naukowej rozkład produktywności w ramach globalnej profesji naukowej na jej dwóch krańcach (górnym 10% i dolnym 10%) jest w dużej mierze ustalony. Ten początkowy globalny rozkład utrzymuje się w czasie, to znaczy przez lata i dziesięciolecia.

Wyjątki są bardzo rzadkie: skoczkowie i spadkowicze w nauce światowej – tak jak ich tu definiujemy – niemal nie występują. Najmniej produktywni naukowcy w ujęciu globalnym niemal nigdy nie stają się globalnymi najbardziej produktywnymi naukowcami. I analogicznie, globalnie najbardziej produktywni naukowcy niemal nigdy nie stają się globalnymi najmniej produktywnymi naukowcami (a pokazuje to analiza próby obejmującej ponad 320 000 naukowców).

Uderzająca jest pokazywana przez nas trwałość przynależności do globalnych klas najwyższej i najniższej produktywności z perspektywy cyklu życia naukowców. Na późniejszym etapie kariery większość globalnych najbardziej produktywnych naukowców (decyl 10. w rozkładzie produktywności) nadal jest najbardziej produktywna, a jedna trzecia globalnych najmniej produktywnych naukowców (decyl 1. w rozkładzie produktywności) nadal jest najmniej produktywna. W ich przypadku prawdopodobieństwo pozostania w najwyższych i najniższych globalnych klasach produktywności – pozioma mo-

bilność pod względem produktywności – przez dziesięciolecia kariery naukowej jest bardzo wysokie i przekracza najczęściej 50%.

Nasze badanie pokazuje, że globalny system nauki jest wyjątkowo sztywny pod względem przynależności do klas produktywności: przypadki skoczków i spadkowiczów są niezwykle rzadkie (tylko 0,51% naukowców przechodzi z najniższej klasy na wczesnym etapie kariery do najwyższej klasy na jej środkowym etapie; i tylko 0,26% naukowców przechodzi z najwyższej klasy na wczesnym etapie kariery do najniższej klasy na środkowym etapie kariery; istnieją odpowiednio tylko 162 i 82 takie przypadki odstające na 32 063 naukowców z 38 krajów we wszystkich dziedzinach nauki łącznie; zob. ryc. 5 i tabela 5).

Przypomnijmy empiryczną skalę zjawiska: wśród wszystkich aktualnie (2020–2023) najbardziej produktywnych ekonomistów i psychologów na późnym etapie kariery na świecie udało nam się zidentyfikować tylko czterech skoczków: dwóch ekonomistów (0,52%) i dwóch psychologów (0,36%), którzy doświadczyli w swoim życiu ekstremalnej mobilności z decyla 1. produktywności do decyla 10. (tabela 4). Siła naszych mikrodatywnych związanych z indywidualnymi identyfikatorami autora polega na wglądzie w unikalne trajektorie rozwoju naukowego.

Obliczenia na dużą skalę przeprowadzone na surowych danych Scopus (np. 1,8 mld cytowanych odniesień bibliograficznych użytych przez nas do zdefiniowania unikalnej dominującej dyscypliny naukowej dla każdego naukowca w naszym zbiorze danych) pozwalają nam badać nie tylko te wyjątkowe przypadki – ale także wszelkie wybrane grupy naukowców. Dla każdego naukowca w naszej próbie dysponujemy kompleksowym zestawem danych demograficznych, publikacyjnych, cytowaniowych i instytucjonalnych na mikropoziomie.

Moglibyśmy na przykład zbadać historię publikacyjną i historię współpracy naukowej naszych czterech imponujących globalnych skoczków w dziedzinie ekonomii i psychologii, w tym ich ewoluujący w czasie wpływ na naukę akademicką, ich krajowe afiliacje na różnych etapach kariery, intensywność badań prowadzonych przez ich instytucje, rok rozpoczęcia kariery naukowej, wskaźnik współpracy międzynarodowej (zarówno przez całe życie, jak i w określonych okresach), ogólny wskaźnik współpracy, medianę wielkości ich zespołów badawczych, znormalizowany do dyscypliny wpływ cytowań ich każdej publikacji w wybranych ramach czasowych (np. w czteroletnim przedziale czasowym), średni percentyl prestiżu ich wszystkich publikacji w czasopiśmie czy też całokształt ich dorobku naukowego według typu czasopisma (np. topowe czasopisma, czasopisma z otwartym dostępem itp.).

Zasadniczo moglibyśmy uzyskać kompleksową wiedzę – w ramach ograniczeń bazy danych i naszych metod obliczeniowych – na temat tego, kim są wspomniani nietypowi

naukowcy, w jaki sposób współpracują, publikują, pracują oraz w jaki sposób ich dorobek był i jest odbierany przez globalną społeczność naukową.

Aż 8 na 10 globalnych najbardziej produktywnych naukowców (sklasyfikowanych w decylnym rozkładzie produktywności) pochodzi z decyli 8–10. na wcześniejszych etapach kariery (83,66% naukowców na środkowym etapie kariery i 83,39% naukowców na późnym etapie kariery); i analogicznie, globalni najmniej produktywni naukowcy (sklasyfikowani w decylnym rozkładzie produktywności) pochodzą głównie z decyli 1–3. (od 75,31% w pierwszym etapie do 68,40% w drugim etapie), z pewnymi różnicami między dyscyplinami. Jako naukowcy, w dużej skali, nie zaskakujemy publikacyjnie – raczej podążamy utartymi szlakami przez dekady.

Dlaczego wcześniejsza przynależność do skrajnej klasy produktywności (najwyższej, najniższej) w dużym stopniu determinuje późniejszą przynależność do klasy produktywności (najwyższej, najniższej)? Istnieją co najmniej dwa wyjaśnienia.

Po pierwsze, wcześniejsze badania pokazały, że rozkład produktywności wśród naukowców jest zawsze silnie skośny (Abramo i in. 2017; David 1994; Kwiek 2016; Kwiek 2018) i że mniejszość naukowców zawsze jest odpowiedzialna za zdecydowaną większość publikacji (Ruiz-Castillo i Costas, 2014; Xie, 2014; również w Polsce od 30 lat obowiązuje „reguła 10/50”: 10% polskich naukowców odpowiada za połowę wszystkich publikacji, zob. Kwiek i Roszka 2024c). Znany temat badań i polityki naukowej, który można podsumować stwierdzeniem, że „większość pracy naukowej jest wykonywana przez stosunkowo niewielką liczbę naukowców” (Crane 1965: 714) leży u podstaw teorii dotyczących indywidualnej produktywności badawczej.

Uznanie w nauce pochodzi niemal wyłącznie od wspólnoty naukowców, a system nagród w nauce opiera się niemal wyłącznie na publikacjach. Ponadto awanse akademickie i perspektywy zatrudnienia, poziom wynagrodzeń, czas wolny przeznaczony na badania i dostęp do grantów badawczych – są mniej lub bardziej bezpośrednio związane z produktywnością publikacyjną (w ujęciu ilości i jakości). Wcześniejszy sukces rodzi sukces aktualny i przyszły, a w analizowanym w tej pracy przypadku sukcesem jest przynależność do niewielkiej klasy najbardziej produktywnych naukowców na świecie.

Zdajemy sobie sprawę z funkcjonowania w globalnym systemie nauki, liczącym kilka milionów naukowców, przypuszczalnie kilku tysięcy naukowców, którzy starają się z różnych powodów nie stosować do tradycyjnych norm akademickich, pracując w ramach farm cytowań i papierni publikacyjnych. Dzieje się tak w każdym kraju, również w Polsce. Jak się jednak wydaje, odstępstwa od reguł są ciągle rzadkie i nie podmywają fundamentów systemu nauki tworzonych przez dekady, chociaż są głośne medialnie i wymagają spokojnych analiz i zdecydowanych kroków zaradczych.

Po drugie, wyższa produktywność publikacyjna na poziomie indywidualnym zasadniczo przyczynia się do lepszego finansowania badań, jak pokazuje model cyklu wia-

rygodności w karierze akademickiej (Latour i Woolgar 1986). W ramach tego modelu badania publikowane w prestiżowych czasopismach (ilość, jakość) są przekuwane w uznanie wspólnoty naukowców; pomyślnie rozpatrzone wnioski grantowe są przekuwane w nowy sprzęt badawczy, młodszych współpracowników w badaniach, nowe argumenty i wreszcie nowe artykuły.

Cykl wiarygodności może być jednak bardziej brzemienne w skutki, jeżeli określa możliwości kariery już na jej wczesnych etapach. Po uzyskaniu finansowania na podstawie uznanych artykułów w prestiżowych czasopismach, prawdopodobieństwo ponownego finansowania wcześniej nagrodzonych naukowców jest wyższe niż w przypadku ich mniej produktywnych kolegów, przynajmniej w bardziej merytokratycznych krajowych systemach finansowania badań, ze znacznym udziałem indywidualnego finansowania w oparciu o granty. Po uzyskaniu finansowania i doskonałych publikacjach, naukowcy mają większe szanse na ponowne finansowanie i szybszy awans na wyższe stanowiska, co odzwierciedla ideę tego modelu, głoszącą, że każdy element cyklu wiarygodności w karierze akademickiej „jest tylko jedną częścią niekończącego się cyklu inwestycji i konwersji” (Latour i Woolgar 1986: 200). Konwersji środków finansowych – na badania i publikacje.

Jeśli chodzi o zmiany klas produktywności z perspektywy całego cyklu życia, to naukowcy, którzy odnoszą mniejsze sukcesy na początku swojej kariery (a sukces w tym przypadku wymaga połączenia m.in. produktywności, motywacji, determinacji, aspiracji, mentoringu, zasobów, jakości wykształcenia, wrodzonych zdolności i szczęścia), mają trudności z udowodnieniem, że są tak dobrzy, jak ich odnoszący większe sukcesy, bardziej produktywni, bardziej zdeterminowani, bardziej zdolni, lepiej wykształceni, mający więcej szczęścia i prawdopodobnie lepiej finansowani koledzy.

Początkowy sukces publikacyjny jest bowiem silnie skorelowany z późniejszym sukcesem publikacyjnym, co można wyjaśnić na dwa sposoby: naukowcy od początku kariery są różni, a niektórzy od czasów doktoranckich są znacznie bardziej produktywni. Ponadto zdarza się, że odnoszą początkowy sukces publikacyjny z powodów niezwiązanych ze swoją wyjątkowością, trochę przypadkowo. W obu przypadkach otoczenie może postrzegać ich bardziej pozytywnie, co z kolei może prowadzić do kolejnych sukcesów w pozyskiwaniu grantów, przyjmowaniu artykułów do druku etc. Wyjaśnienia te odnoszą się do poszczególnych naukowców, wzmacniając ich indywidualny cykl wiarygodności w karierze akademickiej.

Nasze badania po raz kolejny potwierdzają znaczenie dla rozwoju kariery naukowej bardzo mocnego dorobku publikacyjnego (za każdym razem, gdy naukowcy są oceniani przez panele finansujące badania, komisje awansowe czy zespoły redakcyjne): z różnych powodów – których nie jesteśmy w stanie zbadać, wykorzystując nasz zbiór danych – prawdopodobieństwo, że globalni mistrzowie produktywności w przeszłości będą nadal

globalnymi mistrzami w przyszłości jest bardzo wysokie. Zarazem prawdopodobieństwo, że staną się oni globalnymi pariasami produktywności jest marginalne. Doganianie naukowej czołówki w analizowanym aspekcie uprawiania nauki – produktywności publikacyjnej – po prostu się nie zdarza, poza nielicznymi wyjątkami (a w niektórych systemach, jak w Polsce, fenomen skoczków nie zdarza się w ogóle): obliczone przez nas prawdopodobieństwo pojawienia się skoczka dla przejścia z etapu przed habilitacją do etapu po habilitacji wynosi 0%; Kwiek i Roszka, 2024b).

W naszym badaniu sprawdzaliśmy słuszność tradycyjnego założenia, zgodnie z którym produktywni naukowcy stają się w przyszłości „jeszcze bardziej produktywni, a naukowcy, którzy nie tworzą zbyt wielu oryginalnych prac – najprawdopodobniej jeszcze bardziej obniżają swoją produktywność” (Allison i Stewart 1974: 596). Zasadnicze i z góry określone różnice między naukowcami mają ogromny wpływ na przebieg ich karier (Cole i Cole, 1973; Fox, 1983). Początkowy sukces może zwiększać produktywność; natomiast słaby start w nauce może prowadzić do stopniowego porzucania kariery (Turner i Mairesse 2005). Niektórzy naukowcy są zawsze bardzo produktywni, a zróżnicowany rozkład zdolności wpływa na nierówności w produktywności publikacyjnej znacznie bardziej niż system uznania obowiązujący w nauce akademickiej (Stephan i Levin 1992).

Być może jest tak, możemy spekulować, że najbardziej produktywni naukowcy starają się nie zawieść swoich kolegów i samych siebie, a naukowcy osiągający słabe wyniki stopniowo tracą wiarę w swoje możliwości. Wcześniejsza wysoka produktywność zawsze znacząco i pozytywnie wpływa na obecną wysoką produktywność (Kelchtermans i Veugelers 2013), i to pokazują nasze modele regresji logistycznej.

Nasze wyniki mogą przekładać się na politykę instytucjonalną, zwłaszcza w odniesieniu do zatrudniania i awansowania. Pokazujemy w bardzo dużej skali, że naukowcy są mocno osadzeni w klasach produktywności już na pierwszym etapie rozwoju kariery (czyli po pierwszych 5–15 latach pracy). Z tego względu decyzje zatrudnieniowe i awansowe podejmowane na poziomie poszczególnych wydziałów wywierają długofalowy wpływ na produktywność całych instytucji – i to przez wiele lat.

Szanse instytucji, które zatrudniają i promują głównie wysoce produktywnych naukowców, na posiadanie produktywniej kadry są zatem duże; natomiast zatrudnianie i promowanie naukowców o niskiej produktywności w praktyce oznacza milczącą zgodę instytucji na ich utrzymywanie przez wiele lat, co niesie z sobą konsekwencje instytucjonalne odczuwane przez dekady (zwłaszcza w takich systemach jak polski czy włoski, w których nawet na najlepszych uczelniach nie działają systemy zabezpieczeń przed niską produktywnością). Wiara w istnienie fenomenu skoczków w nauce nie znajduje potwierdzenia w bardzo rozległym, wielopłaszczyznowym i podłużnym materiał empirycz-

nym. Kumulatywny charakter sukcesu w nauce powoduje, że naukowcy swoją wysoką produktywność wypracowują przez dziesięciolecia.

Wniosek dla instytucjonalnych strategii rozwoju jest prosty: identyfikacja, zatrudnianie i promowanie naukowców o dużym potencjale publikacyjnym przed 40. rokiem życia (i równoległa rezygnacja z młodych naukowców o niskim potencjale publikacyjnym) jest rozsądnym kierunkiem rozwoju dla instytucji, które mają aspiracje badawcze.

Podziękowania

Jesteśmy wdzięczni gospodarzom i słuchaczom pięciu zaproszonych seminariów: na Uniwersytecie Oksfordzkim (Simon Marginson, CGHE, czerwiec 2022), Uniwersytecie Stanforda (John Ioannidis, METRICS, czerwiec 2022), w DZHW w Berlinie (Torger Möller, czerwiec 2023), na Uniwersytecie w Lejdzie (Ludo Waltman, CWTS, czerwiec 2023) oraz w Sciences Po (Christine Musselin, czerwiec 2023), na których Marek Kwiek omawiał mocne strony i ograniczenia podłużnych badań karier akademickich z wykorzystaniem ustrukturyzowanych Big Data. Dziękujemy za współpracę ICSR Lab, szczególnie Kristy James i Aliczowi Birdowi. Chcemy również serdecznie podziękować dr. Wojciechowi Roszce z UEP w Poznaniu za cenne dyskusje. Dziękujemy również za wsparcie udzielone przez MNISW w ramach grantu nr NdS/529032/2021/2021.

Bibliografia

- Abramo G., D'Angelo C.A., Caprasecca, A. (2009). The contribution of star scientists to overall sex differences in research productivity. *Scientometrics*, 81(1), 137–156.
- Abramo G., D'Angelo C.A., Soldatenkova A. (2017). An investigation on the skewness patterns and fractal nature of research productivity distributions at field and discipline level. *Journal of Informetrics*, 11(1), 324–335.
- Aguinis H., O'Boyle E. (2014). Star performers in twenty-first century organizations. *Personnel Psychology*, 67(2), 313–350.
- Albarrán P., Crespo J.A., Ortuño I., Ruiz-Castillo J. (2011). The skewness of science in 219 sub-fields and a number of aggregates. *Scientometrics*, 88(2), 385–397.
- Baas J., Schotten M., Plume A., Côté G., Karimi R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377–386. 10.1162/qss_a_00019
- Carrasco R., Ruiz-Castillo J. (2014). The evolution of the scientific productivity of highly productive economists. *Economic Inquiry*, 52(1), 1–16.
- Clauset A., Larremore D.B., Sinatra R. (2017). Data-driven predictions in the science of science. *Science*, 355, 477–480.
- Cole J.R., Cole S. (1973). *Social stratification in science*. The University of Chicago Press.
- Crane D. (1965). Scientists at major and minor universities: A study of productivity and recognition. *American Sociological Review*, 30(5), 699–714.

- David P.A. (1994). Positive feedbacks and research productivity in science: Reopening another black box. [w:] O. Granstrand (red.), *Economics of technology* (s. 65–89). Elsevier.
- Hermanowicz J. (2012). The sociology of academic careers: Problems and prospects. [w:] J.C. Smart, M.B. Paulsen (red.), *Higher education: Handbook of theory and research* (s. 207–248). Springer.
- Horta H., Santos J.M. (2016). The impact of publishing during PhD studies on career research publication, visibility, and collaborations. *Research in Higher Education*, 57(1), 28–50.
- Huang J., Gates A.J., Sinatra R., Barabási A.-L. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, 117(9), 4609–4616.
- Ioannidis J.P.A., Boyack K.W., Klavans R. (2014). Estimates of the continuously publishing core in the scientific workforce. *PLOS One*, 9(7), e101698.
- Karimi F., Wagner C., Lemmerich F., Jadidi M., Strohmaier M. (2016). Inferring gender from names on the web: A comparative evaluation of gender detection methods. [w:] *Proceedings of the 25th International Conference Companion on World Wide Web* (s. 53–54). Montreal, Canada.
- Kelchtermans S., Veugelers R. (2013). Top research productivity and its persistence: Gender as a double-edged sword. *Review of Economics and Statistics*, 95(1), 273–285.
- Kwiek M. (2016). The European research elite: A cross-national study of highly productive academics across 11 European systems. *Higher Education*, 71(3), 379–397
- Kwiek M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers? *Scientometrics*, 115(1), 415–462.
- Kwiek M. (2021). What large-scale publication and citation data tell us about international research collaboration in Europe: Changing national patterns in global contexts. *Studies in Higher Education*, 46(12), 2629–2649.
- Kwiek M., Roszka W. (2021a). Gender disparities in international research collaboration: A large-scale bibliometric study of 25,000 university professors. *Journal of Economic Surveys*, 35(5), 1344–1388.
- Kwiek M., Roszka W. (2021b). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, 15(3), 1–38.
- Kwiek M., Roszka W. (2024a). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*, 87, 519–549.
- Kwiek M., Roszka W. (2024b). Are scientists changing their research productivity classes when they move up the academic ladder? *Innovative Higher Education*, Online first, 1–40. <https://doi.org/10.1007/s10755-024-09735-3>.
- Kwiek M., Roszka W. (2024c). Top research performance in Poland over three decades: A multi-dimensional micro-data approach. *Journal of Informetrics*, 18(4). November 2024. 101595. 1–16.
- Kwiek M., Szymula Ł. (2024). Quantifying attrition in science: A cohort-based, longitudinal study of scientists in 38 OECD countries. *Higher Education* (accepted August 1, 2024), Online first, 1–29. <https://doi.org/10.1007/s10734-024-01284-0>.
- Kwiek M., Szymula L. (2023). Young male and female scientists: A quantitative exploratory study of the changing demographics of the global scientific workforce. *Quantitative Science Studies*, 4(4), 902–937.
- Larivière V., Ni C., Gingras Y., Cronin B., Sugimoto C.R. (2013). Global gender disparities in science. *Nature*, 504, 211–213.

- Latour B., Woolgar S. (1986) *Laboratory life. The construction of scientific facts*. Princeton University Press.
- Leišytė L., Dee J.R. (2012). Understanding academic work in changing institutional environment. *Higher Education: Handbook of Theory and Research*, 27, 123–206.
- Li W., Aste T., Caccioli F., Livan G. (2019). Early coauthorship with top scientists predicts success in academic careers. *Nature Communications*, 10, 5170.
- Liu L., Jones B.F., Uzzi B., Wang D. (2023). Data, measurement and empirical methods in the science of science. *Nature Human Behaviour*, 7, 1046–1058.
- Menard, S. (2002). *Longitudinal research*. Sage.
- Merton R.K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Ni C., Smith E., Yuan H., Larivière V., Sugimoto C.R. (2021). The gendered nature of authorship. *Science Advances*, 7, eabe4639.
- Nielsen, M. W., Andersen J.P. (2021). Global citation inequality is on the rise. *Proceedings of the National Academy of Sciences*, 118(7), e2012208118.
- Rowland D.T. (2014). *Demographic methods and concepts*. Oxford University Press.
- Ruiz-Castillo J., Costas R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, 8(4), 917–934.
- Ruspini E. (1999). Longitudinal research and the analysis of social change. *Quality and Quantity*, 33(3), 219–227.
- Salganik M.J. (2018). *Bit by bit. Social research in a digital age*. Princeton University Press.
- Santamaría L., Mihaljević H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science*, 4, e156. <https://doi.org/10.7717/peerj-cs.156>
- Savage W.E., Olejniczak A.J. (2021). Do senior faculty members produce fewer research publications than their younger colleagues? Evidence from Ph.D. granting institutions in the United States. *Scientometrics*, 126, 4659–4686.
- Sebo P. (2021). Performance of gender detection tools: a comparative study of name-to-gender inference services. *Journal of the Medical Library Association*, 109(3), 414–421.
- Sebo P. (2023). How well does NamSor perform in predicting the country of origin and ethnicity of individuals based on their first and last names? *PLOS One*, November 16, 2023, <https://doi.org/10.1371/journal.pone.0294562>
- Spoon K., LaBerge N., Wapman K.H., Zhang S., Morgan A.C., Galesic M., Fosdick B.K., Larremore D.B., Clauset A. (2023). Gender and retention patterns among U.S. faculty. *Science Advances*, 9, eadi2205. 10.1126/sciadv.adi2205
- Stephan P. (2012). *How economics shapes science*. Harvard University Press.
- Sugimoto C., Larivière V. (2018). *Measuring research: What everyone needs to know*. Oxford University Press.
- Sugimoto C., Larivière V. (2023). *Equity for women in science. Dismantling systemic barriers to advancement*. Harvard University Press.
- Turner L., Mairesse J. (2005). *Individual productivity differences in public research: How important are non-individual determinants? An econometric study of French physicists' publications and citations (1986–1997)*. CNRS.
- Wang D., Barabási A.-L. (2021). *The science of science*. Cambridge University Press.
- Wang Y., Jones B.F., Wang D. (2019). Early career setback and future career impact. *Nature Communications*, 10, 4331.

- Way S.F., Morgan A.C., Clauset A., Larremore D.B. (2017). The misleading narrative of the canonical faculty productivity trajectory. *Proceedings of the National Academy of Sciences*, *114*(44), E9216–E9223. 10.1073/pnas.1702121114
- Xie Y. (2014). ‘Undemocracy’: Inequalities in science. *Science*, *344*(6186), 809–810.
- Zhang S., Wapman K.H., Larremore D.B., Clauset A. (2022). Labor advantages drive the greater productivity of faculty at elite universities. *Science Advances*, *8*, eabq7056.

Tabele uzupełniające

Tabela uzupełniająca 1. Struktura próby, wszyscy naukowcy z 38 krajów OECD znajdujący się na późnym etapie kariery naukowej (min. 25 lat doświadczenia publikacyjnego), z co najmniej 10 opublikowanymi artykułami naukowymi lub artykułami w materiałach konferencyjnych, w podziale na płeć, dyscyplinę naukową i kraj ($N = 320\,564$) (częstości i wartości procentowe)

		Łącznie		Kobiety			Mężczyźni		
		<i>N</i>	% kolum.	<i>N</i>	% kolum.	% wiersz.	<i>N</i>	% kolum.	% wiersz.
Agre- gaty	ŁĄCZNIE	320 564	100	84 422	100	26,34	236 142	100	73,66
	SPOŁECZNE	12585	393	3582	4,24	28,46	9003	3,81	71,54
	STEMM	307 979	9607	80 840	95,76	26,25	227 139	96,19	73,75
Dyscy- pliny	AGRI	23 724	7,4	6269	7,43	26,42	17 455	7,39	73,58
	BIO	45 813	14,29	14526	17,21	31,71	31 287	13,25	68,29
	BUS	3259	1,02	813	0,96	24,95	2446	1,04	75,05
	CHEM	14 898	4,65	3251	3,85	21,82	11 647	4,93	78,18
	COMP	7644	2,38	1187	1,41	15,53	6457	2,73	84,47
	EARTH	14 370	4,48	2536	3	17,65	11 834	5,01	82,35
	ECON	3846	1,2	498	0,59	12,95	3348	1,42	87,05
	ENG	12 814	4	1166	1,38	9,1	11 648	4,93	90,9
	ENVIR	6519	2,03	1636	1,94	25,1	4883	2,07	74,9
	IMMU	3142	0,98	1055	1,25	33,58	2087	0,88	66,42
	MATER	5839	1,82	1139	1,35	19,51	4700	1,99	80,49
	MATH	7003	2,18	1139	1,35	16,26	5864	2,48	83,74
	MED	131 075	40,89	41636	49,32	31,77	89 439	37,88	68,23
	NEURO	5863	1,83	1677	1,99	28,6	4186	1,77	71,4
	PHYS	29 275	9,13	3623	4,29	12,38	25 652	10,86	87,62
PSYCH	5480	1,71	2271	2,69	41,44	3209	1,36	58,56	
Kraje	USA	95 718	29,86	26 583	31,49	27,77	69 135	29,28	72,23
	Japonia	29 358	9,16	2953	3,5	10,06	26 405	11,18	89,94
	Włochy	28 354	8,85	10 606	12,56	37,41	17 748	7,52	62,59
	W. Brytania	21 822	6,81	5512	6,53	25,26	16 310	6,91	74,74
	Francja	21 129	6,59	6313	7,48	29,88	14 816	6,27	70,12
	Niemcy	20 551	6,41	3437	4,07	16,72	17 114	7,25	83,28
	Hiszpania	12 978	4,05	4436	5,25	34,18	8542	3,62	65,82
	Kanada	12 605	3,93	3665	4,34	29,08	8940	3,79	70,92
	Australia	10 374	3,24	3118	3,69	30,06	7256	3,07	69,94
	Holandia	8055	2,51	1995	2,36	24,77	6060	2,57	75,23
	Polska	5619	1,75	1901	2,25	33,83	3718	1,57	66,17
	Szwecja	4894	1,53	1372	1,63	28,03	3522	1,49	71,97
	Korea Płd.	4847	1,51	627	0,74	12,94	4220	1,79	87,06
	Szwajcaria	4126	1,29	746	0,88	18,08	3380	1,43	81,92
	Belgia	3582	1,12	920	1,09	25,68	2662	1,13	74,32
	Turcja	3413	1,06	850	1,01	24,9	2563	1,09	75,1
	Grecja	3412	1,06	873	1,03	25,59	2539	1,08	74,41
	Izrael	3352	1,05	935	1,11	27,89	2417	1,02	72,11
	Dania	2871	0,9	759	0,9	26,44	2112	0,89	73,56
	Austria	2808	0,88	561	0,66	19,98	2247	0,95	80,02
Pozostałe	20 696	6,45	6260	7,42	30,25	14 436	6,12	69,75	

Tabela uzupełniająca 2. Struktura próby, wszyscy naukowcy z 38 krajów OECD znajdujący się na późnym etapie kariery naukowej (min. 25 lat doświadczenia publikacyjnego), z co najmniej 10 opublikowanymi artykułami naukowymi lub artykułami w materiałach konferencyjnych, według wieku akademickiego i płci ($N = 320\,564$)

Wiek akademicki (lata od pierwszej publikacji)	Kobiety	Mężczyźni	% kobiet	% mężczyzn	Łącznie
25	8692	17 836	32,77	67,23	26 528
26	8072	17 640	31,39	68,61	25 712
27	7964	17 513	31,26	68,74	25 477
28	7591	17 780	29,92	70,08	25 371
29	6289	15 160	29,32	70,68	21 449
30	5857	14 623	28,6	71,4	20 480
31	5431	13 617	28,51	71,49	19 048
32	4792	12 507	27,7	72,3	17 299
33	4152	12 017	25,68	74,32	16 169
34	4052	11 580	25,92	74,08	15 632
35	3540	10 971	24,4	75,6	14 511
36	2994	9816	23,37	76,63	12 810
37	2559	8805	22,52	77,48	11 364
38	2349	7927	22,86	77,14	10 276
39	1942	7483	20,6	79,4	9425
40	1812	6905	20,79	79,21	8717
41	1485	5970	19,92	80,08	7455
42	1340	5764	18,86	81,14	7104
43	1159	4817	19,39	80,61	5976
44	957	4267	18,32	81,68	5224
45	805	3941	16,96	83,04	4746
46	617	3460	15,13	84,87	4077
47	524	2981	14,95	85,05	3505
48	447	2581	14,76	85,24	3028
49	365	2373	13,33	86,67	2738
50	321	2166	12,91	87,09	2487

Tabela uzupełniająca 3. Mobilność najmniej produktywnych naukowców między dwoma etapami kariery: wczesnym (wyjściowym) i środkowym (docelowym): z których decyli wyjściowej produktywności (na wczesnym etapie kariery) pochodzą najmniej produktywni naukowcy na środkowym etapie kariery? Najmniej produktywni naukowcy na środkowym etapie kariery (N= 32 063) według obszaru dyscyplin naukowych i wyjściowego decyla produktywności (częstości i wartości procentowe)

	Łącznie	Dolne 10%	Decyl 2.	Decyl 3.	Decyl 4.	Decyl 5.	Decyl 6.	Decyl 7.	Decyl 8.	Decyl 9.	Górne 10%
Najmniej produktywni naukowcy na środkowym etapie kariery – według wyjściowego decyla produktywności na wczesnym etapie kariery											
ŁĄCZNIIE	N	11 996	7325	4825	3091	2036	1277	785	425	221	82
	%	37,41	22,85	15,05	9,64	6,35	3,98	2,45	1,33	0,69	0,26
SPOŁECZNE	N	1259	294	186	125	95	60	30	18	15	3
	%	34,39	23,35	14,77	9,93	7,55	4,77	2,38	1,43	1,19	0,24
STEMM	N	30 804	7031	4639	2966	1941	1217	755	407	206	79
	%	37,54	22,82	15,06	9,63	6,3	3,95	2,45	1,32	0,67	0,26

Tabela uzupełniająca 4. Mobilność najmniej produktywnych naukowców między dwoma etapami kariery: środkowym (wyjściowym) i późnym (docelowym): z których decyli wyjściowej produktywności (na środkowym etapie kariery) pochodzą najmniej produktywni naukowcy na późnym etapie kariery? Najmniej produktywni naukowcy na późnym etapie kariery (N= 32 075) według obszaru dyscyplin naukowych i wyjściowego decyla produktywności (częstości i wartości procentowe)

	Łącznie	Dolne 10%	Decyl 2.	Decyl 3.	Decyl 4.	Decyl 5.	Decyl 6.	Decyl 7.	Decyl 8.	Decyl 9.	Górne 10%
Najmniej produktywni naukowcy na późnym etapie kariery – według wyjściowego decyla produktywności na środkowym etapie kariery											
ŁĄCZNIIE	N	9836	7051	5052	3473	2450	1691	1127	726	447	222
	%	30,67	21,98	15,75	10,83	7,64	5,27	3,51	2,26	1,39	0,69
SPOŁECZNE	N	400	254	201	154	89	70	47	21	14	9
	%	31,77	20,17	15,97	12,23	7,07	5,56	3,73	1,67	1,11	0,71
STEMM	N	9436	6797	4851	3319	2361	1621	1080	705	433	213
	%	30,62	22,06	15,74	10,77	7,66	5,26	3,5	2,29	1,41	0,69

**Młodzi naukowcy o niskiej produktywności?
Żadnych szans na wysoką produktywność w przyszłości
(analiza 320 000 karier naukowych w 38 krajach OECD)**

W pracy pokazujemy, że już na stosunkowo wczesnym etapie kariery naukowej rozkład produktywności w ramach globalnej profesji naukowej na jej dwóch krańcach (górnym 10% i dolnym 10%) jest w dużej mierze ustalony. Ten początkowy rozkład utrzymuje się w czasie, przez lata i dziesięciolecia. Najmniej produktywni naukowcy niemal nigdy nie stają się najbardziej produktywni, a najbardziej produktywni – najmniej produktywni. Uderzająca jest trwałość przynależności do klas najwyższej i najniższej produktywności z perspektywy cyklu życia naukowców. Nasze badanie pokazuje, że globalny system nauki jest wyjątkowo sztywny pod względem rozkładu produktywności.

Słowa kluczowe: kariera akademicka, produktywność publikacyjna, skoczkowie i spadkowicze, Big Data, nauka o nauce, naukoznawstwo, kraje OECD

**Unproductive Young Scientists?
No Chance to Become Highly Productive in the Future
(Analysis of 320 000 Scientific Careers in 38 OECD Countries)**

In this paper, we show that already at a relatively early stage of a scientific career, the distribution of productivity within the global scientific profession at its two extremes (the top 10% and the bottom 10%) is largely fixed. This initial distribution persists over time, over years and decades. The least productive scientists almost never become the most productive, and the most productive scientists almost never become the least productive. The persistence of membership in the highest and lowest productivity classes from the perspective of the life cycle of scientists is striking. Our study shows that the global science system is extremely rigid in terms of productivity distribution.

Key words: academic career, publication productivity, jumpers and fallers, Big Data, science of science, meta-research, OECD countries