

MAREK KWIEK*, WOJCIECH ROSZKA**

Produktywność publikacyjna polskich naukowców w ujęciu podłużnym: przykład doktorów i doktorów habilitowanych

Wprowadzenie

W prezentowanym artykule szukamy odpowiedzi na proste pytanie badawcze dotyczące wpływu wcześniejszej indywidualnej produktywności publikacyjnej na późniejszą produktywność publikacyjną. Zakładając, że naukowcy mogą zmieniać klasy produktywności w trakcie swoich karier, poszukujemy wzorców mobilności między klasami produktywności w pięciu dziedzinach nauki (obejmujących łącznie 12 dyscyplin naukowych).

Analizujemy zmieniającą się produktywność polskich naukowców ($N = 4165$) reprezentujących nauki ścisłe, techniczne, inżynieryjne, matematyczne oraz medyczne (STEMM) w miarę ich awansu: przed i po habilitacji. Wszyscy analizowani naukowcy są doktorami habilitowanymi, zatrudnionymi na pełen etat w sektorze szkolnictwa wyższego.

Łącząc dane demograficzne i biograficzne pochodzące z krajowego rejestru naukowców ($N = 99\,935$) z własnymi obliczeniami wykonanymi na podstawie metadanych Scopus dotyczących wszystkich polskich artykułów naukowych indeksowanych w ciągu ostatniego półwiecza (1973–2021, $N = 935\,167$), analizujemy poszczególnych naukowców, którzy zmieniają klasy produktywności w czasie, w okresie obejmującym maksymalnie 40 lat (zakres wieku biologicznego w próbie: 30–70 lat). Skupiamy się tylko na dwóch etapach kariery: przed i po habilitacji.

Metodologicznym punktem wyjścia jest przypisanie wszystkich doktorów habilitowanych widocznych międzynarodowo w bazie Scopus do 10 aktualnych klas produktywności w okresie 2018–2021 (czyli według 10 decyli produktywności). Następnie analizujemy ich wcześniejszą produktywność z czasów, gdy byli doktorami, porównujemy ich z rówieśnikami w obrębie ich dziedzin nauki i retrospektywnie przypisujemy ich do 10

* 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 Wojciech Roszka (wojciech.roszka@ue.poznan.pl), Uniwersytet Ekonomiczny w Poznaniu, Centrum Studiów nad Polityką Publiczną UAM w Poznaniu ORCID: orcid.org/0000-0003-4383-3259

wcześniejszych klas produktywności (ponownie od najwyższych decyli do najniższych). Szczegółowo analizujemy indywidualne trajektorie karier naukowców, łącząc obecną i wcześniejszą produktywność dla każdego naukowca, a następnie badając mobilność między klasami produktywności w zależności od dziedziny nauki i typu produktywności.

W szczególności interesuje nas porównanie wzorców mobilności między klasami produktywności pod względem czterech typów produktywności – pełnego i częściowego zliczania – zarówno w wersji znormalizowanej do prestiżu czasopism, jak i w wersji nieznormalizowanej.

Nasze podejście jest zatem podłużne (śledzimy produktywność tych samych naukowców w czasie) oraz klasyfikacyjne (analizujemy zmiany produktywności w ramach 10 klas produktywności opartych na decylach, a nie zmiany dotyczące liczby publikacji; górnych 10% to klasyczna miara w analizach nierówności produktywności, zob. np. prace na temat tak definiowanej klasy *top scientists* w Abramo i in. 2017). Analizujemy naukowców z najwyższych i najniższych klas produktywności, którzy zmieniają klasy w czasie, w ujęciu relatywnym: identyfikacja klasy jest możliwa przez badanie produktywności naukowców w stosunku do produktywności innych naukowców (podobnie jak w badaniach „gwiazd nauki”; zob. np. Aguinis i O’Boyle, 2014: 313–315; DiPrete i Eirich, 2006: 282).

W oparciu o wcześniejsze badania dotyczące wysokiej produktywności badawczej (Abramo i in. 2009; Fox i Nikivincze 2021; Kwiek 2016; Yin i Zhi 2017), skupiamy się na trwałości najwyższej i najniższej produktywności, w czasie gdy naukowcy awansują w hierarchii akademickiej. Nasze intuicje bazują na wcześniejszych teoriach z zakresu socjologii i ekonomii nauki, zgodnie z którymi naukowcy o najwyższej produktywności mają tendencję do utrzymywania wysokiej produktywności, a naukowcy o najniższej produktywności pozostają na tym samym niskim poziomie; natomiast osoby o marginalnej produktywności często opuszczają sektor nauki akademickiej (Allison i Stewart 1974: 596; Allison i in. 1982: 615; Cole i Cole 1973: 114; Turner i Mairesse 2005: 3).

Silne zróżnicowanie osiągnięć naukowców oraz utrwalone nierówności w produkcji wiedzy były przedmiotem badań od dawna, a podstawowe analizy zostały przedstawione przez Alfreda Lotkę (1926), Dereka J. de Solla Price’a (1963), Roberta K. Mertona (1968), Jonathana R. Cole’a i Stephena Cole’a (1973) oraz wielu innych badaczy, inspirowując kolejne pokolenia teoretyków. Kluczowym tematem tych teorii indywidualnej produktywności naukowej jest stwierdzenie, że „większość pracy naukowej wykonywana jest przez stosunkowo niewielką liczbę naukowców” (Crane 1965: 714).

Mechanizmy stojące za kumulatywną przewagą (i kumulatywną stratą) badano przez dekady (Allison i in. 1982; Allison i Stewart 1974; Cole i Cole 1973; DiPrete i Eirich 2006; Merton 1968), podobnie jak inne ważne teorie produktywności naukowej, takie jak teoria „iskry bożej” (Allison i Stewart 1974; Cole i Cole 1973; Fox 1983; Zuckerman

1970) i teoria maksymalizacji użyteczności (Kyvik 1990; Stephan i Levin, 1992). „Niedemokratyczność” nauki jest jej integralną cechą, a „nierówność była i zawsze będzie jej nieodłącznym elementem” (Xie 2014: 809). W Polsce, podobnie jak gdzie indziej, naukowcy o niskiej produktywności badawczej pracują w laboratoriach obok naukowców o bardzo wysokiej produktywności (Abramo i in. 2013; Piro i in. 2016). Co więcej, w Polsce – podobnie jak w obszarze OECD – 10% najbardziej produktywnych naukowców odpowiada za powstawanie aż 50% wszystkich publikacji (Kwiek 2018). „Reguła 10/50” obowiązuje w polskim systemie nauki od przynajmniej 30 lat (Kwiek i Roszka 2024a). Rola „gwiazd nauki” skoncentrowanych po prawej stronie długiego ogona rozkładu produktywności w każdym krajowym systemie nauki utrzymuje się w czasie (Agrawal i in. 2017: 1).

Prawostronna asymetria produktywności w nauce była tematem licznych publikacji bibliometrycznych (np. Albarrán i in. 2011; Carrasco i Ruiz-Castillo 2014; Ruiz-Castillo i Costas 2014). Najnowsze badania obejmują analizę naukowców o różnie określonej wysokiej produktywności: gwiazd i supergwiazd (Abramo i in. 2009; Aguinis i O’Boyle 2014; Sidiropoulos i in. 2016; Yair i in. 2017; Agrawal i in. 2017), najlepszych naukowców (O’Boyle i Aguinis 2012), „płodnych profesorów” (Piro i in. 2016), czołowych badaczy (Abramo i in. 2013; Cortés i in. 2016) czy też elity akademickiej (Yin i Zhi 2016, Kwiek 2016).

W większości systemów nauki produktywność jest jedną z najważniejszych determinant – choć nie jedyną – trajektorii rozwoju kariery akademickiej (Leisyte i Dee 2012; Stephan 2015). Produktywność była szeroko badana zarówno z perspektywy jednego kraju, jak i w porównawczym ujęciu międzynarodowym (zob. np. Allison i in. 1982; Fox 1983; Lee i Bozeman 2005; Shin i Cummings 2010; Ramsden 1994; Stephan i Levin 1992; Teodorescu 2000; Wanner i in. 1981; Kwiek 2018). Oprócz liczby publikacji, na sukces w karierze akademickiej wpływają takie czynniki jak uzyskane zewnętrzne finansowanie badań, wzorce międzynarodowej współpracy, nagrody i wyróżnienia, pochodzące z wyboru członkostwo w stowarzyszeniach naukowych i akademiach, międzynarodowa mobilność fizyczna i doświadczenie międzynarodowe, sieci zawodowe, afiliacja instytucjonalna oraz szczęście (Carvalho 2017; Hermanowicz 2012).

Sukces w nauce determinowany jest również przez umiędzynarodowienie badań, liczbę otrzymywanych cytowań, sposób dystrybucji czasu pracy, podział ról akademickich i inne czynniki. Główne czynniki wpływające na produktywność dzielą się na dwie grupy: indywidualne i środowiskowe (obejmujące zarówno instytucje, np. klimat w miejscu pracy, jak pokazały Fox i Mohapatra 2007, jak i całe krajowe systemy nauki, np. systemy awansu i uznania akademickiego, jak pokazali Leisyte i Dee 2012).

W prezentowanej pracy zastosowaliśmy analizę regresji logistycznej do identyfikacji głównych determinant przynależności do najwyższych klas produktywności. Literatura

przedmiotu wskazuje na kilka istotnych luk, które zamierzamy wypełnić: po pierwsze, zdecydowana większość badań nad produktywnością opiera się na danych przekrojowych (głównie bazujących na ankietach), a nie na danych longitudinalnych.

Po drugie, nieliczne badania podłużne koncentrują się na mobilności w obrębie najwyższych klas produktywności, co nie odzwierciedla wzorców kariery wielu naukowców, którzy poruszają się w górę lub w dół w swojej produktywności (znormalizowanej w odniesieniu do dyscypliny). Z perspektywy indywidualnej, niezwykle rzadkie przypadki mobilności z najniższych do najwyższych klas są tak samo istotne, jak znacznie częstsze przypadki mobilności w obrębie najwyższych klas z perspektywy instytucjonalnej.

Po trzecie, literatura nie przedstawia trwałości produktywności w czasie w bardziej szczegółowy sposób (np. można badać wszystkie decyle produktywności, gdzie większość mobilności do najwyższych decyli pochodzi z sąsiednich decyli, a żadna mobilność nie pochodzi z najniższych decyli); ogólne wzorce mobilności – takie jak przejścia naukowców między poszczególnymi kwantylami produktywności – ukrywają bardziej szczegółowe wzorce, które wymagają bardziej precyzyjnego podejścia.

Po czwarte, dotychczasowe badania całkowicie pomijały wpływ różnych metod zliczania publikacji w ramach obliczania produktywności na skalę zaobserwowanej mobilności; w szczególności, w ogóle nie uwzględniano roli pionowej struktury hierarchicznej czasopism naukowych (oraz roli prestiżu czasopism mierzonego na podstawie liczby cytowań w ujęciu czasowym).

Różne metody zliczania odgrywają niezwykle istotną rolę w takich systemach naukowych, jak polski, w których przyznawanie grantów, awanse i decyzje dotyczące zatrudnienia są ściśle związane z publikacjami. Standardowa produktywność (bez normalizacji do prestiżu czasopism) wydaje się nieprzydatna w realiach, w których publikacje w niektórych czasopismach – oficjalnie zdefiniowanych przez Ministerstwo Nauki i nieoficjalnie uznawanych przez społeczność naukową – mają znaczenie dla jednostek i instytucji, a publikacje w innych czasopismach nie mają znaczenia. Polska nadal stoi w szerokim rozkroku i nie definiuje czasopism drapieżnych (typu MDPI) jako takich, co pozwala niemal wszystkim polskim instytucjom szkolnictwa wyższego na utrzymywanie w ostatnich latach udziału artykułów w czasopismach z grupy MDPI na poziomie co najmniej kilkunastu procent wszystkich artykułów indeksowanych w bazie Scopus.

Ponadto w dotychczasowych badaniach modele ekonometryczne stosowane w badaniach longitudinalnych nie wykorzystywały danych biograficznych pochodzących z rejestrów naukowców typu dane z OPI PIB, co oznacza, że klasy szybkości awansu (okres od stopnia do stopnia liczony w latach) i klasy wieku uzyskania awansu (wiek biologiczny) nie były używane do wyjaśniania wysokiej produktywności w taki sposób, w jaki są one wykorzystywane w naszym badaniu (zob. pełną metodologię zastosowaną w Kwiek i Roszka 2023).

Zbiór danych, próba badawcza i metodologia

Zbiór danych

W pracy korzystamy z dwóch źródeł danych: krajowego i międzynarodowego. Krajowy zestaw danych to utrzymywana i aktualizowana przez nas baza „Obserwatorium Polskiej Nauki”. Międzynarodowym zestawem danych jest natomiast jednostkowa surowa baza Scopus zawierająca informacje o publikacjach i cytowaniach z lat 1973–2021 wszystkich polskich naukowców aktywnych w tym okresie.

Baza „Obserwatorium” została utworzona poprzez integrację krajowego biograficznego i administracyjnego rejestru polskich naukowców ($N = 99\,935$) z bibliometryczną bazą Scopus z lat 2009–2018 (metadane dotyczące $N = 380\,000$ publikacji autorów afiliowanych w Polsce). Baza „Obserwatorium” zawiera m.in. informacje na temat płci, roku urodzenia, dat kolejnych awansów naukowych (stopień doktora, stopień doktora habilitowanego, tytuł profesora tytularnego – o ile dotyczy), aktualnych afiliacji instytucjonalnych oraz dyscyplin, w których uzyskano stopnie naukowe.

Oficjalny rejestr krajowy oraz publikacyjna i cytowaniowa baza Scopus z lat 2009–2018 zostały połączone z wykorzystaniem metod probabilistycznych i deterministycznych (zob. Kwiek i Roszka 2021: 4–6). Baza „Obserwatorium” została następnie wzbogacona o metadane publikacji wszystkich naukowców afiliowanych w Polsce z ostatniego półwiecza, zebrane z bazy Scopus i pozyskane dzięki wieloletniej umowie o współpracy z ICSR Lab (International Center for the Study of Research) – platformą chmurową firmy Elsevier utrzymywaną na potrzeby badań naukowych ($N = 935\,167$ artykułów z lat 1973–2021).

Próba badawcza

Nasza próba badawcza ($N = 4165$ naukowców z łączną liczbą $N_{\text{art}} = 71\,841$ artykułów) obejmuje osoby zatrudnione w pełnym wymiarze czasu pracy w instytucjach szkolnictwa wyższego, posiadające stopień doktora habilitowanego i pracujące w jednej z pięciu dziedzin nauk STEMM (obejmujących łącznie 12 dyscyplin STEMM; ich lista została zamieszczona pod tabelą 1).

W naszej próbie jedną trzecią stanowią kobiety, a dwie trzecie – mężczyźni (odpowiednio 37,3% i 62,7%), co ogólnie odzwierciedla strukturę polskiego środowiska akademickiego w obszarach STEMM na etapach pracy ze stopniem doktora i ze stopniem doktora habilitowanego. Podobnie jak w innych krajach, również w Polsce odsetek kobiet jest najwyższy na niższych stanowiskach, a najniższy na wyższych stanowiskach, sięgając 28,34% w gronie profesorów tytularnych we wszystkich dyscyplinach łącznie (GUS, 2023: tabela 1/42).

Niemal dwie trzecie doktorów habilitowanych w naszej próbie mieści się w przedziale wiekowym 40–54 lata (61,8%), a wśród nich zarówno mężczyźni, jak i kobiety są

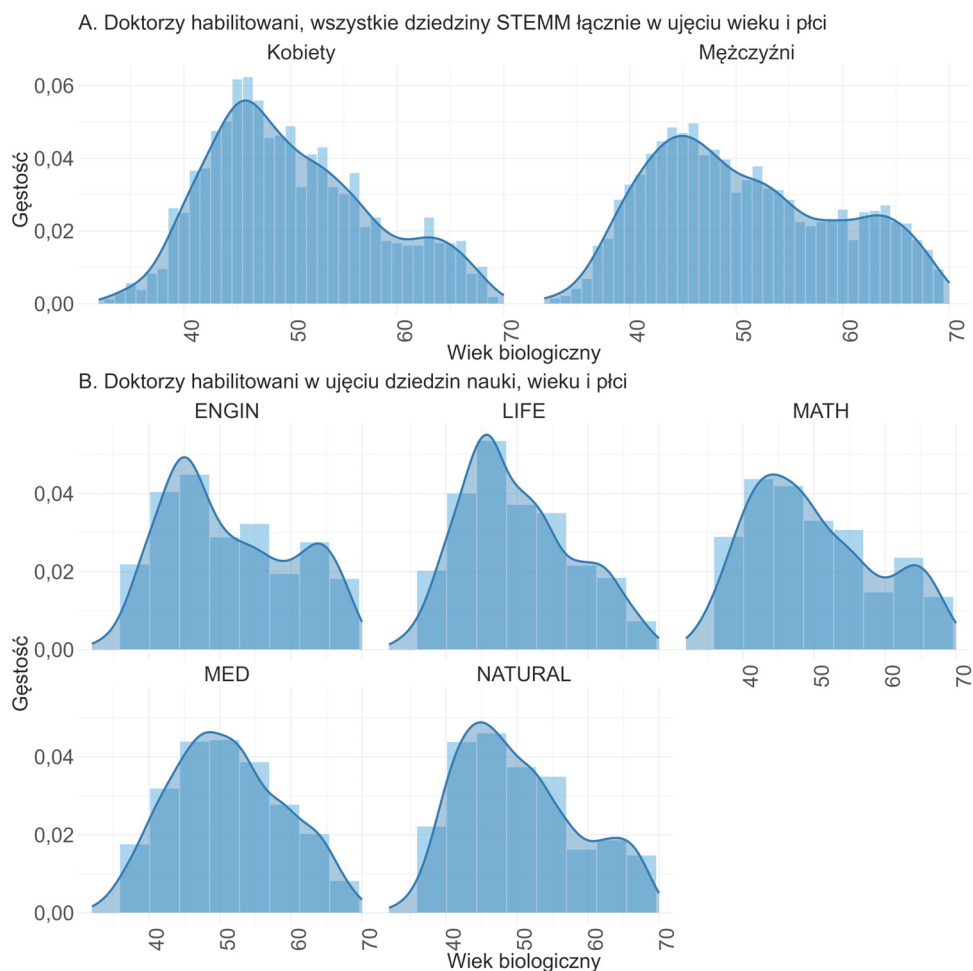
rozproszeni w trzech grupach wiekowych. Jedynie jedna piętnasta ma mniej niż 40 lat (7,2%). Wykres gęstości jądrowej przedstawiony na ryc. 1 pokazuje, że obecny rozkład wiekowy doktorów habilitowanych w podziale na płeć różni się zwłaszcza w starszych grupach wiekowych. W szczególności udział starszych doktorów habilitowanych jest wyższy wśród mężczyzn niż wśród kobiet, co może odzwierciedlać rosnący napływ kobiet do dyscyplin STEMM 30 lat temu i wcześniej.

Niemal co trzeci naukowiec (29,8%) wywodzi się z 10 uczelni o wysokiej intensywności badawczej zdefiniowanych jako instytucje uczestniczące w programie IDUB, a więc spośród uczestników pierwszej polskiej inicjatywy doskonałości naukowej. Rozkład wieku naukowców w poszczególnych dziedzinach nauki jest nieznacznie zróżnicowany, co potwierdzają wspomniane wykresy gęstości jądrowej: podczas gdy w dwóch dziedzinach (NATURAL i LIFE) przeważają młodsze grupy wiekowe, w MATH rozkład jest bardziej płaski, a w ENGI występuje większy odsetek starszych naukowców.

Tabela 1. Struktura próby wszystkich polskich doktorów habilitowanych widocznych na arenie międzynarodowej według płci, grupy wiekowej i dziedziny nauk STEMM ($N=4165$); liczebność i odsetki

		Łącznie		Kobiety			Mężczyźni		
		<i>n</i>	% kolum.	<i>n</i>	% kolum.	% wiersz.	<i>n</i>	% kolum.	% wiersz.
Grupa wiekowa	Łącznie	4165	100,0	1553	100,0	37,3	2612	100,0	62,7
	Poniżej 40 r.ż.	301	7,2	92	5,9	30,6	209	8,0	69,4
	40–54 lata	2575	61,8	1036	66,7	40,2	1539	58,9	59,8
	55 lat i więcej	1289	30,9	425	27,4	33,0	864	33,1	67,0
Typ uczelni	IDUB	1240	29,8	354	22,8	28,5	886	33,9	71,5
	Pozostałe	2925	70,2	1199	77,2	41,0	1726	66,1	59,0
Dziedzina nauki	ENGI	959	23,0	184	11,8	19,2	775	29,7	80,8
	LIFE	897	21,5	485	31,2	54,1	412	15,8	45,9
	MATH	400	9,6	76	4,9	19,0	324	12,4	81,0
	MED	630	15,1	335	21,6	53,2	295	11,3	46,8
	NATURAL	1279	30,7	473	30,5	37,0	806	30,9	63,0

Dwanaście dyscyplin z bazy Scopus (według ASJC: All Science Journal Classification) należących do obszaru STEMM zostało pogrupowanych w pięć dziedzin nauki. W skład analizowanych dziedzin wchodzi: ENGI (inżynieria, obejmująca inżynierię i materiałoznawstwo), LIFE (nauki o życiu, obejmujące nauki rolnicze i biologiczne oraz biochemię, genetykę i biologię molekularną), MATH (matematyka, obejmująca matematykę i informatykę), MED (medycyna, obejmująca nauki medyczne) oraz NATURAL (nauki przyrodnicze, obejmujące inżynierię chemiczną, chemię, fizykę i astronomię, nauki o Ziemi i planetach oraz nauki o środowisku).



Ryc. 1. Rozkład wieku biologicznego: wykresy gęstości jądrowej. Panel A. doktorzy habilitowani w pięciu dziedzinach STEMM łącznie, według płci. Panel B. doktorzy habilitowani w podziale na dziedziny nauk STEMM ($N = 4165$)

Metodologia

Jednostka analizy: poszczególny naukowiec, a nie poszczególna publikacja

Jednostką analizy są indywidualni naukowcy o jednoznacznie zdefiniowanych cechach biograficznych i cechach portfeli publikacyjnych. Każdy naukowiec otrzymał indywidualny numer identyfikacyjny (ID), dzięki czemu możliwe stało się pozyskanie jego danych biograficznych związanych z karierą naukową z baz OPI PIB.

Nasze badanie koncentruje się na indywidualnych ścieżkach kariery akademickiej autorów publikacji naukowych rozwijanych w czasie – w tym przypadku na zmianach klas produktywności w trakcie awansów akademickich – a nie na samych publikacjach.

Konstrukcja indywidualnych portfeli publikacyjnych

Dla każdego naukowca opracowaliśmy unikalny „portfel publikacyjny” obejmujący wszystkie metadane wszystkich publikacji zgromadzonych w bazie Scopus. W szczególności metadane te obejmują informacje o czasopiśmie (np. percentylowy ranking Scopus CiteScore) oraz o samej publikacji (np. rok publikacji, liczba współautorów wraz z ich afiliacjami, liczba cytowań). Każda publikacja jest powiązana z konkretnymi datami z biografii naukowca, czyli z dwoma etapami kariery: pracą z doktoratem oraz pracą z habilitacją, które są jasno zdefiniowane jako okres między uzyskaniem stopnia doktora a uzyskaniem stopnia doktora habilitowanego oraz okres po uzyskaniu stopnia doktora habilitowanego. Data pierwszej publikacji (dowolnego typu) w bazie Scopus pozwala dodatkowo wyznaczyć tzw. wiek akademicki oparty na publikacjach, będący przybliżeniem doświadczenia naukowego i doświadczenia publikacyjnego. Wiek ten jest wykorzystywany w naszych modelach regresji logistycznej.

Konstrukcja indywidualnych ścieżek biografii naukowej

Dla każdego naukowca, oprócz unikalnego portfela publikacyjnego, skonstruowaliśmy również indywidualną ścieżkę biograficzną, zawierającą kluczowe daty: rok urodzenia (pozwalający na określenie wieku biologicznego na początku obu etapów kariery: etapu pracy z doktoratem i z habilitacją), rok uzyskania stopnia doktora oraz rok uzyskania stopnia doktora habilitowanego. Wszyscy naukowcy z naszej próby są obecnie doktorami habilitowanymi, a wcześniej, co oczywiste, pracowali jako doktorzy.

Pierwszy etap ich kariery akademickiej rozpoczął się więc wraz z uzyskaniem stopnia doktora, a drugi – wraz z uzyskaniem stopnia doktora habilitowanego. W przypadku obu stopni posiadamy pełne dane administracyjne, takie jak data nadania stopnia, tytuł rozprawy czy instytucja i miasto zatrudnienia, dyscyplina naukowa oraz dziedzina nauki. Dane te, pochodzące z oficjalnego krajowego rejestru naukowców, można uznać za w pełni wiarygodne.

Podejście podłużne w badaniach karier akademickich

Na potrzeby analiz wybraliśmy wszystkich obecnych, międzynarodowo widocznych doktorów habilitowanych (tj. autorów posiadających przynajmniej jeden artykuł opublikowany w czasopiśmie indeksowanym w bazie Scopus), a następnie przyjrzelśmy się retrospektywnie ich karierom zawodowym. Przeanalizowaliśmy ich aktualną aktywność publikacyjną w czteroletnim okresie 2018–2021 oraz ich wcześniejszą aktywność publikacyjną w czasach, gdy byli doktorami, również ujętą w czteroletnie okresy. Zintegrowana baza danych (OPI-Scopus) została pierwotnie skonstruowana dla lat 2009–2018, natomiast na potrzeby niniejszego badania została rozszerzona o metadane publikacji z kolejnych lat (2019–2021). Analizy produktywności oparte są zatem na czteroletnim

okresie referencyjnym 2018–2021, co pozwala na porównywalność w obrębie aktualnej aktywności publikacyjnej.

W tradycyjnych projektach badawczych opartych na założeniach longitudinalnych te same osoby są obserwowane w kolejnych momentach, co umożliwia prowadzenie badań porównawczych (Menard 2002; Singer i Willett 2003). W kontekście badania karier akademickich podejście longitudinalne dotychczas nie było stosowane z powodów technicznych i finansowych, jednak w ostatnim czasie pojawiły się podejścia o charakterze kohortowym z wykorzystaniem danych bibliometrycznych (zob. np. Milojevic i in. 2018; Huang i in. 2020; Wang i Barabási 2021).

Połączenie indywidualnych ścieżek biograficznych (dane o przebiegu kariery) z indywidualnymi portfelami publikacyjnymi (dane o publikacjach i cytowaniach) pozwala na uzyskanie retrospektywnego spojrzenia, w którym możemy śledzić przez kilka dekad aktywność publikacyjną licznej grupy naukowców. Podejście longitudinalne otwiera nowe możliwości badania ścieżek kariery akademickiej w dłuższej perspektywie czasowej.

Tym samym możliwe staje się zastosowanie czterech głównych wymiarów do biograficznych i bibliometrycznych analiz naukowców: płeć, wiek, dziedzina nauki (dyscyplina naukowa) oraz – co najważniejsze – czas. Zamiast ograniczania się do kilkuletnich ujęć „migawkowych” z kolejnych badań przekrojowych, zastosowanie podejścia podłużnego umożliwia analizę zmian zachodzących w populacji naukowców w różnych wymiarach (np. produktywności naukowej) na przestrzeni lat.

Definiowanie płci, wieku biologicznego, wieku akademickiego i dziedzin nauki

W naszej próbie wszyscy naukowcy mają jednoznacznie określoną płeć (zgodnie z danymi z krajowego rejestru OPI PIB) i rok urodzenia. Dzięki temu łatwo jest ustalić ich wiek biologiczny w dowolnym momencie kariery zawodowej. Wiek akademicki, czyli liczba lat, jakie upłynęły od pierwszej publikacji indeksowanej w bazie Scopus (dowolnego typu), potrzebny do modeli regresji logistycznej, został ustalony na podstawie daty pierwszej publikacji pobranej z bazy Scopus za pomocą protokołu API Elseviera (Scopus API). API zostało użyte do automatycznego pozyskania metadanych publikacyjnych dla każdego naukowca, w tym roku pierwszej publikacji, który posłużył jako punkt początkowy do obliczeń wieku akademickiego.

Posługujemy się indywidualnymi portfelami publikacyjnymi (obejmującymi całość publikacji zindeksowanych w bazie Scopus), aby określić dominującą dyscyplinę na podstawie najczęściej występującego kodu ASJC (All Science Journal Classification) przypisywanego każdemu naukowcowi. Łączymy wszystkie publikacje (artykuły w czasopiśmie oraz prace zawarte w materiałach konferencyjnych) w portfelach z dyscypli-

nami ASJC. Jeśli w portfelu występują dwie lub więcej dyscyplin z jednakowo wysoką częstością, losowo wybieramy jedną z nich. Następnie grupujemy 12 dyscyplin w pięć obszarów nauk, aby zwiększyć reprezentatywność zarówno kobiet, jak i mężczyzn, oraz aby uniknąć zbyt małej liczby obserwacji w niektórych dyscyplinach.

Pomiar indywidualnej produktywności publikacyjnej

Produktywność publikacyjna mierzona jest w czteroletnim okresie referencyjnym 2018–2021 (określanym jako „aktualna produktywność doktorów habilitowanych”) oraz we wcześniejszym okresie kariery, gdy naukowcy pracowali jako doktorzy (określanym jako „przeszła produktywność doktorów habilitowanych w okresie, gdy byli doktorami”), korzystając z danych o publikacjach (artykułach w czasopismach) pochodzących z indywidualnych portfeli publikacyjnych. Potrzebowaliśmy dokładnych dat z danych OPI PIB, aby ustalić, kiedy obecni doktorzy habilitowani pracowali jako doktorzy, i przyporządkować ich publikacje do okresu 2018–2021 oraz do okresu pracy z doktoratem (o różnej długości trwania dla różnych osób). W obu przypadkach wykorzystywaliśmy czteroletnie przedziały.

Podejście znormalizowane do prestiżu czasopism w pomiarze produktywności publikacyjnej

W niniejszym badaniu przyjmujemy, że uzasadnione jest analizowanie zmian produktywności w czasie przy wykorzystaniu klas produktywności wewnątrz dyscyplin (zob. Costas i Bordons, 2005; Costas i Bordons, 2007) – zamiast analizy wyłącznie liczby publikacji. Nasze podejście znormalizowane do prestiżu czasopism uwzględnia silnie zróżnicowaną strukturę czasopism w skali globalnej, opierając się na założeniu, że artykuły publikowane w wysoko cenionych czasopismach wymagają średnio większego nakładu pracy naukowej niż artykuły publikowane w czasopismach o niskim prestiżu.

Prestiż czasopisma (w naszym ujęciu wyrażony jako ranga percentylowa w bazie Scopus w przedziale 0–99) stanowi istotny element indywidualnej produktywności, szczególnie w systemach takich jak polski (Antonowicz i in., 2021), w których zarówno liczba publikacji, jak i ich jakość mierzona za pomocą wskaźników prestiżu czasopisma wpływają na awanse akademickie. Wartości percentylowe p zostały przeskalowane do przedziału 0–1 poprzez podstawienie $x = p/100$. Następnie zastosowaliśmy transformację potęgową $w = x^{2.5}$. Aby uniknąć zerowych lub bliskich zera wag dla artykułów z najniższymi rangami, przyjęliśmy dolne ograniczenie $p = 10$. Podstawowa zasada ewaluacji osiągnięć naukowych w ostatnich dwóch dekadach głosi, że „czasopisma nie są sobie równe”: publikacje w różnych czasopismach otrzymują różną liczbę punktów (z zakresu 20–200).

Dysponując indywidualnymi „portfelami publikacyjnymi” każdego naukowca w naszej próbie, zastosowaliśmy cztery sposoby pomiaru produktywności: dwa znormalizowane do prestiżu czasopism i dwa niezależne od prestiżu czasopism. Szczególnie interesujące jest podejście znormalizowane do prestiżu czasopism, nad którym pracowaliśmy w ostatnich latach (zob. Kwiek i Roszka 2023). Polega ono na tym, że artykuły zostają powiązane z czasopismami w bazie Scopus, w których zostały opublikowane. Wszystkie czasopisma Scopus ($N = 46\,702$ w 2024 r.) mają swoje miejsca w ramach rang percentylowych Scopus CiteScore (od 0 do 99), przy czym najbardziej prestiżowe czasopisma zazwyczaj plasują się w przedziale 90–99.

W podejściu do produktywności bez normalizacji do prestiżu czasopisma każdy artykuł ma w obliczeniach produktywności wagę 1 (stosując metodę zliczania całkowitego, zob. Waltman i van Eck, 2019). Natomiast w podejściu znormalizowanym do prestiżu wartość ta jest przekształcana proporcjonalnie do rangi percentylowej czasopisma w bazie Scopus. W praktyce w naszym szczegółowym ujęciu przyjęliśmy, że różnice w prestiżu czasopism mierzonym w Scopus lepiej odzwierciedla funkcja potęgowa niż funkcja liniowa. Użyliśmy funkcji $y = x^{2.5}$ (w przeciwieństwie do funkcji liniowej $y = x$, gdzie x oznacza rangę percentylową CiteScore). Na podstawie testów różnych funkcji (z wykładnikami 1,5; 2; 2,5; 3) uznaliśmy, że wykładnik 2,5 najlepiej oddaje nasze założenia: zwiększa on wartość artykułów w najwyższej ocenianych czasopismach (szczególnie w 95–99 percentylu CiteScore) kosztem wartości artykułów z czasopism z dolnych przedziałów (poniżej 50 percentyla CiteScore; szczegóły znajdują się w Materiale uzupełniającym).

Powiązanie artykułów z miejscem, jakie czasopisma zajmują w silnie zhierarchizowanym systemie czasopism naukowych, przy użyciu funkcji potęgowej zamiast liniowej, dodatkowo podkreśla fakt, że publikacje w prestiżowych czasopismach (z uwagi na bardziej rygorystyczne procedury recenzyjne i wyższe wymagania stawiane przez recenzentów oraz wobec recenzentów) przeciętnie wymagają znacznie większego nakładu pracy i czasu na przygotowanie, poprawki i ponowne zgłoszenie niż publikacje w czasopismach niskoprestiżowych. Dotyczy to zwłaszcza czasopism o bardzo niskich wskaźnikach akceptacji, często poniżej 10%.

Reguły są stosunkowo proste: najlepsi recenzenci pracują niemal wyłącznie dla najlepszych czasopism, ponieważ sami w nich publikują. Jednocześnie co do zasady nie recenzują artykułów w czasopismach o niskim prestiżu, ponieważ w nich nie publikują, a pula indywidualnego czasu przeznaczanego na recenzje jest zawsze ograniczona.

W najlepszych czasopismach jest ogromny tłok i najlepsi recenzenci, w słabych czasopismach teksty jest prosto wydać, a recenzenci są radykalnie mniej kompetentni. Dlatego artykuły w kiepskich czasopismach, choć wydane – nie są szeroko czytane, a tym bardziej szeroko cytowane. Nie stoi za nimi renoma czasopisma i renoma jego suro-

wych i kompetentnych recenzentów. Oczywiście tak wygląda uśredniony i wyidealizowany obraz – w praktyce najbardziej cenieni i znani naukowcy mogą publikować w dowolnych czasopismach, choć najczęściej tego nie robią.

Metoda zliczania pełnego i metoda zliczania ułamkowego w pomiarze produktywności publikacyjnej

Nasza próba badawcza obejmuje wyłącznie naukowców reprezentujących dziedzinę STEMM, gdzie normą są publikacje wieloautorskie, a publikacje indywidualne stanowią rzadkość (Wagner 2018; Olechnicka i in. 2019). W metodzie zliczania pełnego (*full counting*) każdy współautor otrzymuje pełen udział w publikacji, natomiast w metodzie zliczania ułamkowego (*fractional counting*) udział dzieli się przez liczbę współautorów (nasza próba nie obejmuje artykułów mających ponad 100 współautorów, które najczęściej występują w subdyscyplinach fizyki i astronomii; zob. Waltman i van Eck 2019). Uwzględnienie obu metod zliczania (pełnego i ułamkowego) w połączeniu z podejściem znormalizowanym i niezależnym od prestiżu czasopism dało nam w sumie cztery typy produktywności w niniejszym badaniu:

1. Produktywność 1 (znormalizowana do prestiżu, zliczanie pełne),
2. Produktywność 2 (znormalizowana do prestiżu, zliczanie ułamkowe),
3. Produktywność 3 (bez normalizacji, zliczanie pełne),
4. Produktywność 4 (bez normalizacji, zliczanie ułamkowe).

Przypisanie naukowców do klas produktywności

Kluczowym elementem metodologicznym niniejszego badania jest przyporządkowanie naukowców do 10 (opartych na decylach) klas produktywności. Najpierw, w obrębie każdej dziedziny nauki, uszeregowaliśmy wszystkich obecnych doktorów habilitowanych w kolejności malejącej według ich czteroletniej produktywności w okresie referencyjnym 2018–2021.

Ponieważ zdefiniowaliśmy cztery typy produktywności, przeprowadziliśmy cztery procedury rankingowe. Górnych 10% naukowców w każdym obszarze nauk (tj. 10. decyl produktywności według któregoś z czterech typów) zakwalifikowano jako grupę najwyższej produktywności, a dolnych 10% (1. decyl produktywności) – jako grupę najniższej produktywności, oczywiście z uwzględnieniem odpowiednich punktów odcięcia. W efekcie w klasie najwyższej produktywności znalazło się 419 naukowców, natomiast w klasie najniższej – 412 (413 w przypadku Produktywności 4).

Jednym ze sposobów wyróżniania grup o najwyższej i najniższej produktywności jest oparcie się na danych (np. stosując przedziały produktywności – liczby publikacji) i grupowanie osób, które przekraczają wyznaczone progi. W naszym podejściu, polegającym

na wyznaczaniu decyli, łatwiej jest porównywać te grupy przy użyciu modeli statystycznych, w których analizujemy predyktory zwiększające prawdopodobieństwo przynależności do najwyższej lub najniższej klasy.

Klasy produktywności oparte na decylach traktujemy jako wygodną procedurę klasyfikacyjną, która umożliwia pracę z silnie skośnym, ciągłym rozkładem. Nie odzwierciedlają one w sposób bezpośredni rzeczywistych doświadczeń życiowych poszczególnych naukowców; istnieje wiele różnych dróg prowadzących do wysokiej produktywności i wiele powodów niskiej produktywności (zob. Wang i Barabási 2021: 13–15 – model Shockleya; Bornmann 2024 – zasada „Anny Kareniny”). Produktywność publikacyjna to tylko jeden z wymiarów aktywności naukowej, która obejmuje także m.in. wpływ cytowań, liczbę publikacji w wiodących czasopismach (górny 10% rankingu czasopism), uzyskane granty badawcze i ich prestiż, wystąpienia plenarne na konferencjach czy opiekę nad doktorantami. Jednocześnie aktywność naukowa jest jedynie częścią szerszej działalności akademickiej, tradycyjnie obejmującej również dydaktykę i prace organizacyjne.

Szczegółowej analizie poddano dane, na których opiera się bieżąca klasyfikacja decylowa: w każdym obszarze nauk prześledziliśmy punkty odcięcia między poszczególnymi decylami produktywności. Oddzielna analiza dla wszystkich obszarów nauk oraz dla dwóch etapów kariery (doktoratu i habilitacji) wykazała znaczne różnice w punktach odcięcia między naukowcami na tych etapach – dotyczy to zarówno najniższych (1. decyl), jak i najwyższych (10. decyl) wartości produktywności. Na przykład w przypadku granicy między 9. a 10. decylem produktywności różnice mieszczą się w przedziale 4–6-krotnym (od średnio 4,73 w MATH do średnio 5,98 w ENGI, przy wartości 5,33 dla wszystkich obszarów łącznie). Pokazuje to, że doktorzy habilitowani publikują znacznie więcej (i częściej w bardziej prestiżowych czasopismach) niż na etapie pracy z doktoratem.

Następnie, ponownie w obrębie każdego obszaru nauk i dla każdego z czterech typów produktywności, uszeregowaliśmy wszystkich obecnych doktorów habilitowanych według średniej czteroletniej produktywności z czasów, gdy pracowali jako doktorzy. Analizowany okres jest różny w zależności od długości trwania tej fazy kariery, jednak zawsze bazuje na ujednoliconym czteroletnim ekwiwalencie produktywności.

Wyniki

Wzorce mobilności między klasami produktywności

Skupiamy się na mobilności między klasami produktywności, w szczególności między klasami najwyższymi i najniższymi oraz klasami im najbliższymi: decylami 8., 9. i 10. w górnej strefie oraz decylami 1., 2. i 3. w dolnej strefie produktywności. Doktorzy z najwyższych i najniższych klas produktywności mogą zmieniać swoje klasy produk-

tywności na etapie pracy jako doktorzy habilitowani, przechodząc do klasy wyższej, niższej bądź do jakiegokolwiek innego decyla produktywności. W szczególności analizujemy następujące typy mobilności, zróżnicowane w zależności od obszaru nauk i typu produktywności:

Mobilność z najwyższej do najwyższej klasy (doktorzy należący do najwyższej klasy produktywności nadal należą do najwyższej klasy jako doktorzy habilitowani: przejście z 10. decyla produktywności do 10. decyla produktywności).

Mobilność z najniższej do najniższej klasy (doktorzy należący do najniższej klasy produktywności nadal należą do najniższej klasy jako doktorzy habilitowani: przejście z 1. decyla produktywności do 1. decyla produktywności).

Ekstremalna mobilność pionowa: przejście z najwyższej do najniższej klasy oraz z najniższej do najwyższej klasy (doktorzy należący do najwyższej klasy produktywności przechodzą do najniższej klasy jako doktorzy habilitowani; analogicznie doktorzy należący do najniższej klasy przechodzą do najwyższej klasy; odpowiednio: przejście z 10. do 1. decyla oraz z 1. do 10. decyla produktywności).

Poza powyższymi typami mobilności, które odnoszą się wyłącznie do najwyższej (10.) i najniższej (1.) klasy produktywności, omawiamy również szerszą mobilność między górnymi decylami produktywności (8.–10.) a dolnymi decylami (1.–3.), uwzględniając przy tym znaczenie wartości granicznych (*cut-off points*) liczby publikacji.

W badanych danych występują przypadki bliskie odcięciu zarówno przy dolnej, jak i przy górnej granicy: naukowcy minimalnie powyżej granicy 1. decyla i minimalnie poniżej granicy 10. decyla. Z tego powodu szersze ujęcie, obejmujące sąsiednie decyle, wydaje się zasadne.

Zasadnicze pytanie dotyczy tego, jak obecni doktorzy habilitowani – zaliczający się do grupy o najwyższej produktywności (10. decyl, $N = 419$) – byli rozmieszczeni w rozkładzie rankingu percentylowego produktywności (zakres 0–99) w przeszłości, gdy pracowali jako doktorzy. Analogicznie interesuje nas również, jak obecni doktorzy habilitowani o najniższej produktywności (1. decyl, $N = 412$) byli rozmieszczeni w rankingach percentylowych w przeszłości, na etapie pracy ze stopniem doktora.

W celu odpowiedzi na te pytania przeanalizowaliśmy rozkłady z uwzględnieniem obszaru nauk, płci, intensywności badawczej instytucji zatrudniających (IDUB) oraz dwóch zmiennych związanych z wiekiem (wiek akademicki i wiek biologiczny). Zgodnie z przewidywaniami mediana pierwotnej rangi percentylowej (z etapu pracy z doktorem) dla obecnych naukowców w 10. decylu produktywności jest bardzo zbliżona do późniejszej pozycji (na etapie pracy z habilitacją), zaś dla 1. decyla – jest analogicznie. Mediana wynosi 87,9 percentyla w przypadku naukowców z obecnej grupy o najwyższej produktywności oraz 18,3 percentyla w przypadku naukowców z grupy o najniższej produktywności (tabela 2). Zróżnicowanie między obszarami nauk jest niewielkie dla

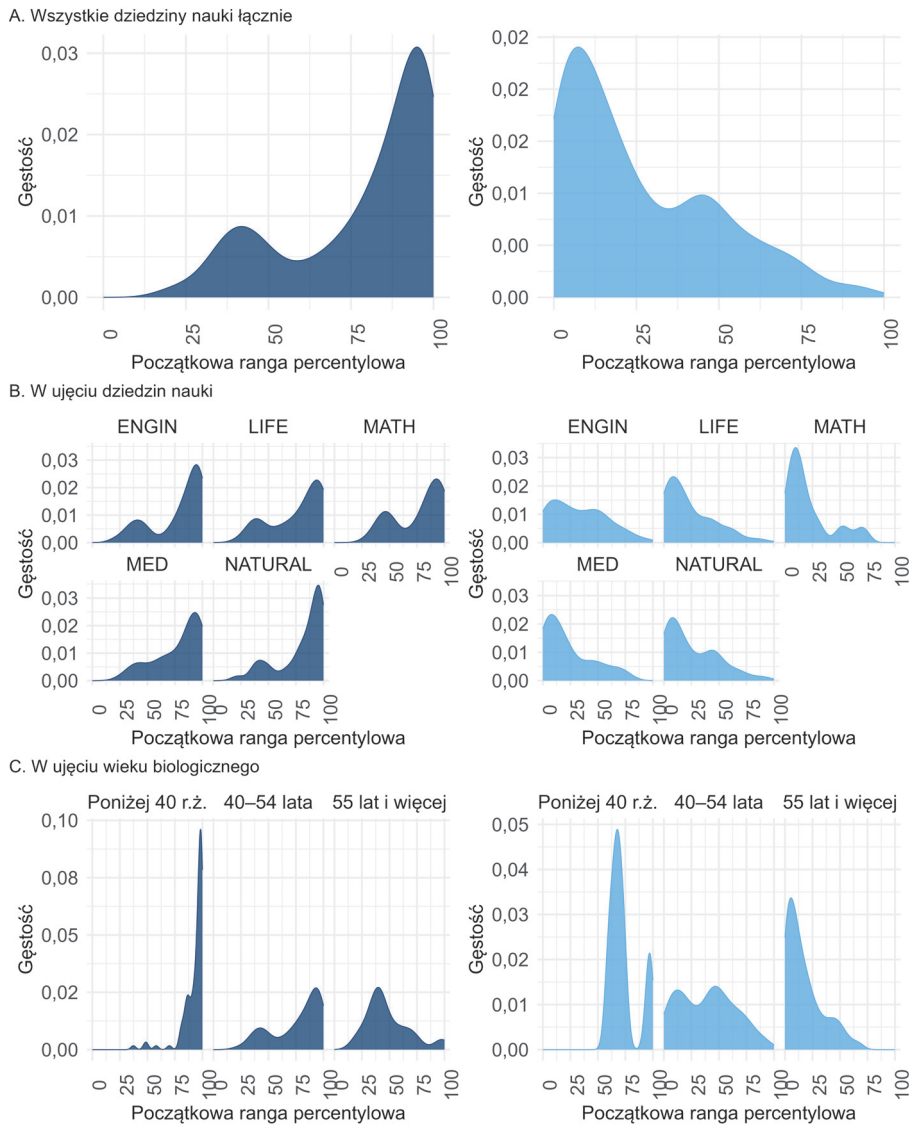
decyla 10. (od 84,7 w LIFE do 89,8 w NATURAL) oraz zauważalne dla decyla 1. (od 13,6 w MATH do 29,6 w ENGI). Różnice między mężczyznami i kobietami są marginalne: zarówno mężczyźni, jak i kobiety z grup obecnie najwyższej lub najniższej produktywności pochodzą przeciętnie z podobnych rang percentylowych na etapie pracy z doktoratem.

Tabela 2. Jak doktorzy habilitowani z obecnie najwyższą (decyl 10.) i najniższą (decyl 1.) produktywnością (odpowiednio: panel lewy, $N=419$) i panel prawy, $N=412$) byli ulokowani w percentylowym rozkładzie produktywności (zakres: 0–100) na etapie pracy z doktoratem? Doktorzy habilitowani: początkowy (w okresie pracy z doktoratem) rozkład rang percentylowych, Produktywność 1 (zliczanie pełne, znormalizowane do prestiżu czasopism), w podziale na dziedzinę, płeć, intensywność badawczą instytucji, grupę wieku akademickiego oraz grupę wieku biologicznego

		Aktualnie najbardziej produktywni (górných 10%) doktorzy habilitowani – rozkład według mediany produktywności				Aktualnie najmniej produktywni (dolnych 10%) doktorzy habilitowani			
		<i>N</i>	Średnia	SD	Mediana	<i>N</i>	Średnia	SD	Mediana
Łącznie		419	78,6	22,2	87,9	412	26,1	23,2	18,3
Dziedzina nauki	ENGI	96	78,5	23,4	88,2	95	33	24,8	29,6
	LIFE	90	76,3	23,3	84,7	89	23,3	22	17,2
	MATH	41	77,3	21,9	86,3	39	21,2	20,7	13,6
	MED	64	78,4	21	86,6	62	22,7	22,1	15,5
	NATURAL	128	80,8	21,2	89,8	127	25,9	23,2	18
Płeć	Kobiety	140	78,5	21,8	88,5	133	28	24,8	18,7
	Mężczyźni	279	78,6	22,4	87,6	279	25,1	22,4	18,1
Typ uczelni	IDUB	161	81,2	21,7	91,3	91	29,3	24,7	21,2
	Pozostałe	258	76,9	22,4	85,2	321	25,1	22,7	18
Grupa wieku akademickiego	Początkujący	6	78,5	16,2	80,7	7	40,1	21,5	44,6
	Wczesny etap	201	89	14,2	93,7	128	31,8	26,5	30
	Średni etap	198	69,2	24	74,5	195	25,1	23,1	16,5
	Późny etap	14	60,9	23,2	55,4	82	18	13,3	16,2
Grupa wieku biologicznego	Poniżej 40 r.ż.	104	92,7	10,9	96,9	5	72,9	14,3	68,6
	50–54 lata	289	76,2	21,9	84,3	148	39,3	25,1	42,8
	55 lat i więcej	26	48,7	20,3	42,2	259	17,6	16,7	12,0

Dobrym sposobem na pokazanie, jak obecni doktorzy habilitowani z najwyższą (i najniższą) produktywnością byli ulokowani w poszczególnych decylach produktywności w okresie, gdy pracowali jako doktorzy, jest zastosowanie wykresów gęstości jądrowej (ryc. 2). Wykresy te wykorzystują estymację gęstości jądrowej do utworzenia wygładzonej, ciągłej krzywej przybliżającej rozkład danych. Ta metoda lepiej odzwierciedla

kształt rozkładu niż histogramy, ponieważ jej wynik nie zależy ani od liczby przedziałów w histogramie, ani od dużych różnic między nimi; umożliwia też elastyczne porównywanie rozkładów kilku zbiorów danych.



Ryc. 2. Rozmieszczenie obecnych doktorów habilitowanych o najwyższej produktywności (po lewej, $N = 419$, 10. decyl produktywności) oraz o najniższej produktywności (po prawej, $N = 412$, 1. decyl produktywności) w rozkładzie percentylowym produktywności (zakres: 0–100) z okresu, gdy byli doktorami. (Panel A) wszystkie dziedziny nauki łącznie, (Panel B) w podziale na dziedziny, (Panel C) w podziale na grupy wieku biologicznego. Wykresy gęstości jądrowej, początkowy rozkład rang percentylowych, Produktywność 1: zliczanie pełne, znormalizowane do prestiżu czasopism

Tabela 3. Mobilność naukowców o najwyższej produktywności między decylami produktywności na dwóch etapach kariery – przed habilitacją (etap początkowy) i po habilitacji (etap docelowy). Z których początkowych decyli produktywności (z okresu przed habilitacją) pochodzą obecni doktorzy habilitowani z najwyższą (panel górny) i najniższą (panel dolny) produktywnością? Doktorzy habilitowani z najwyższą ($N = 419$) i najniższą produktywnością ($N = 412$) według dziedzin nauki i początkowego decyla produktywności. Produktywność 1: zliczanie pełne, znormalizowane do prestiżu czasopism (liczebności i odsetki)

		Łącznie	Dolnych 10 %	Decyl 2.	Decyl 3.	Decyl 4.	Decyl 5.	Decyl 6.	Decyl 7.	Decyl 8.	Decyl 9.	Górnych 10%
Najbardziej produktywni doktorzy habilitowani – według wyjściowego decyla produktywności przed habilitacją												
Łącznie	<i>N</i>	419	0	1	7	34	37	15	20	36	74	195
	%	100	0	0,2	1,7	8,1	8,8	3,6	4,8	8,6	17,7	46,5
ENGI	<i>N</i>	96	0	0	2	10	9	1	3	7	18	46
	%	100	0	0	2,1	10,4	9,4	1	3,1	7,3	18,8	47,9
LIFE	<i>N</i>	90	0	0	2	10	7	4	6	7	16	38
	%	100	0	0	2,2	11,1	7,8	4,4	6,7	7,8	17,8	42,2
MATH	<i>N</i>	41	0	0	0	2	9	1	0	4	6	19
	%	100	0	0	0	4,9	22	2,4	0	9,8	14,6	46,3
MED	<i>N</i>	64	0	0	1	4	4	4	5	7	11	28
	%	100	0	0	1,6	6,3	6,3	6,3	7,8	10,9	17,2	43,8
NATURAL	<i>N</i>	128	0	1	2	8	8	5	6	11	23	64
	%	100	0	0,8	1,6	6,3	6,3	3,9	4,7	8,6	18	50
Najmniej produktywni doktorzy habilitowani – według wyjściowego decyla produktywności przed habilitacją												
Łącznie	<i>N</i>	412	137	81	45	30	49	24	23	15	3	5
	%	100	33,3	19,7	10,9	7,3	11,9	5,8	5,6	3,6	0,7	1,2
ENGI	<i>N</i>	95	22	14	12	10	14	9	5	6	1	2
	%	100	23,2	14,7	12,6	10,5	14,7	9,5	5,3	6,3	1,1	2,1
LIFE	<i>N</i>	89	32	19	11	6	10	2	6	1	1	1
	%	100	36	21,3	12,4	6,7	11,2	2,2	6,7	1,1	1,1	1,1
MATH	<i>N</i>	39	14	12	4	2	0	4	1	2	0	0
	%	100	35,9	30,8	10,3	5,1	0	10,3	2,6	5,1	0	0
MED	<i>N</i>	62	25	13	6	2	6	4	3	3	0	0
	%	100	40,3	21	9,7	3,2	9,7	6,5	4,8	4,8	0	0
NATURAL	<i>N</i>	127	44	23	12	10	19	5	8	3	1	2
	%	100	34,6	18,1	9,4	7,9	15	3,9	6,3	2,4	0,8	1,6

Analizując wszystkie dziedziny nauki łącznie, okazuje się, że zdecydowana większość obecnych doktorów habilitowanych o najwyższej produktywności należała wcześniej do 8.–10. decyla produktywności, podczas gdy zdecydowana większość doktorów habilitowanych o najniższej produktywności wywodzi się z 1.–3. decyla.

Najwyższa koncentracja najbardziej produktywnych doktorów habilitowanych występuje w obszarze NATURAL, natomiast największy odsetek najmniej produktywnych naukowców obserwujemy w obszarze MATH. Z kolei w podziale na grupy wiekowe najbardziej usytuowani w czołówce są najmłodszy naukowcy (do 39 lat), co ilustruje ryc. 2.

Dzięki naszemu zestawowi danych możemy szczegółowo analizować mobilność między decylami produktywności poszczególnych osób. Tabela 3 (panel górny) ukazuje, z których decyli produktywności (z okresu, gdy byli doktorami) pochodzą obecni doktorzy habilitowani zaliczani obecnie do 10. decyla. Prawie połowa (46,5%) obecnych doktorów habilitowanych o najwyższej produktywności wywodzi się z 10. decyla produktywności w czasach, gdy byli doktorami – a więc pozostaje w tej samej klasie produktywności po zrobieniu habilitacji (17,7% pochodzi z 9. decyla, a 8,6% z 8. decyla). W sumie trzy czwarte z nich (72,8%) należało w przeszłości do 8.–10. decyla produktywności. Niemal żaden z nich nie znajdował się w najniższych trzech decylach; nie zanotowano przypadku ekstremalnej mobilności w górę (z 1. do 10. decyla), a tylko jeden naukowiec (z obszaru NATURAL) awansował z 2. do 10. decyla. W bazie posiadamy pełne dane biograficzne i publikacyjne o wszystkich naukowcach, w tym również o tym jednym wyjątku.

Rozpatrując poszczególne dziedziny nauki: połowa doktorów habilitowanych o wysokiej produktywności (50,0%) była również wysoko produktywna w okresie pracy jako doktorzy w obszarze NATURAL, zaś w obszarze ENGI odsetek ten wynosi 47,9%. W ENGI 74,0% najbardziej produktywnych doktorów habilitowanych pochodzi z trzech najwyższych decyli, a żaden nie wywodzi się z trzech najniższych decyli produktywności. W LIFE i MATH wartości te wynoszą odpowiednio 67,8% i 70,7%, a z trzech najniższych decyli pochodzi w tych obszarach zaledwie 2,2% oraz 0% analizowanych naukowców.

W dolnym panelu tabeli 3 zaprezentowano, z których decyli produktywności pochodzą obecni doktorzy habilitowani należący dziś do 1. decyla (najniższa produktywność). Obserwowane wzorce są podobne do tych dla grupy o najwyższej produktywności, lecz mniej wyraźne. Niemal dwie trzecie (63,9%) doktorów habilitowanych o najniższej produktywności wywodzi się z trzech najniższych decyli (1.–3.), w tym jedna trzecia z najniższego (decyla 1, 33,3%). Tylko 5,5% (23 naukowców) pochodzi z trzech najwyższych decyli.

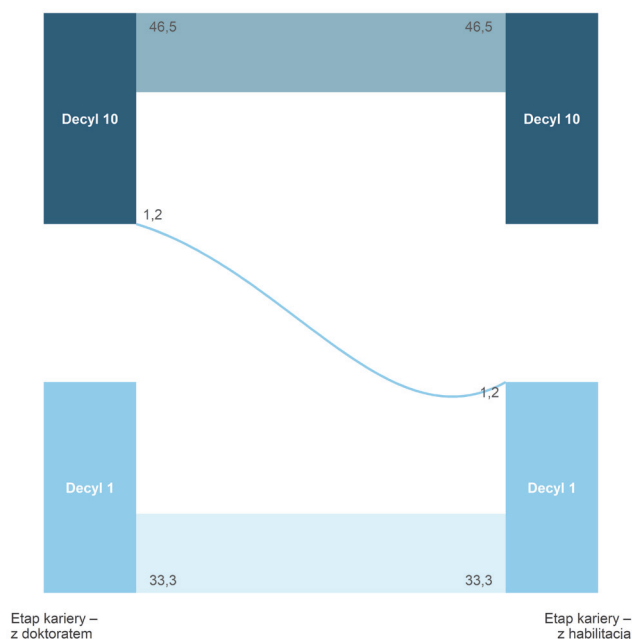
Mobilność między klasami produktywności według typu produktywności: wszystkie dziedziny nauki łącznie

Wykres strumieniowy (nazywany również diagramem Sankeya, ryc. 3) pomaga pokazać to, co rozumiemy przez mobilność naukowców pomiędzy różnymi klasami produk-

tywności. Ukazuje on przepływy naukowców między decylami produktywności na etapie pracy z doktoratem (lewa strona: najwyższa i najniższa klasa) oraz na etapie pracy z habilitacją (prawa strona: najwyższa i najniższa klasa). W szczególności interesuje nas mobilność pozioma i skrajna mobilność pionowa.

Przykład na rycinie 3 obrazuje mobilność naukowców ze wszystkich dziedzin nauki łącznie, stosując Produktywność 1 (zliczanie pełne, znormalizowane do prestiżu czasopism). Lewa kolumna pokazuje rozkład doktorów w najwyższych i najniższych decylach produktywności (oba łącznie 100%), prawa kolumna – odpowiedni rozkład doktorów habilitowanych w tych samych klasach produktywności.

Grube strumienie poziome odpowiadają mobilnościom poziomym. Skrajna, pionowa mobilność od najwyższej do najniższej klasy zdarza się rzadko i jest przedstawiona jako cienki strumień wiodący w dół; skrajny strumień do góry nie występuje w ogóle: jedynie 1,2% doktorów z najwyższej klasy produktywności (dokładnie pięciu naukowców) trafia do klasy najniższej jako doktorzy habilitowani, natomiast żaden doktor z najniższej klasy produktywności (dokładnie 0 przypadków) nie osiąga klasy najwyższej produktywności.



Ryc. 3. Mobilność naukowców między klasami produktywności na dwóch etapach kariery naukowej (typ produktywności: pełne zliczanie z normalizacją do prestiżu czasopisma). Wszystkie dziedziny STEM łącznie, obecni doktorzy habilitowani. 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)

Patrząc całościowo na wszystkie dziedziny nauki (tabela 4 oraz tabela uzupełniająca 1): zaobserwowane wzorce mobilności są jednoznaczne – niemal połowa (46,5%) naukowców przypisanych do najwyższej klasy produktywności (decyl 10.) utrzymuje się w tej samej klasie w dalszej karierze akademickiej, a jedna trzecia (33,3%) naukowców z najniższej klasy produktywności (decyl 1.) pozostaje w tej najniższej klasie (w ramach Produktywności 1). Zjawisko to można określić mianem „mechanizmu zakotwiczenia” w karierze akademickiej, co zasługuje na dalsze pogłębione badania. Jest to o tyle ciekawe, że w specyficznych polskich warunkach awans w karierze akademickiej zależy niemal wyłącznie od publikacji i produktywności publikacyjnej, przy znikomej roli dydaktyki i pracy organizacyjnej.

Analizując aktualny rozkład wieku biologicznego doktorów habilitowanych (ryc. 1) i ich podział na grupy wiekowe (tabela 1), można wnioskować, że naukowcy w obszarze STEMM „zastygają” w danej klasie produktywności na wiele lat, a niekiedy nawet dekad: niemal jedna trzecia (30,9%) doktorów habilitowanych ma 55 lat lub więcej; stanowisko doktora uzyskali średnio w wieku około 28–32 lat. W systemie, w którym tytuł profesora tytularnego jest ukoronowaniem kariery naukowej, osiąganym przez stosunkowo nieliczne grono naukowców, doktorzy habilitowani są obecni we wszystkich grupach wiekowych. Dlatego też nasza analiza obejmuje kilka dekad karier akademickich obecnych doktorów habilitowanych.

Tabela 4. 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 (etap wyjściowy)	Etap pracy (docelowy)	Decyl produktywności po habilitacji (etap 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

Co istotne, w naszym podejściu nie odwołujemy się do bezwzględnej liczby publikacji, ponieważ produktywność w Polsce w ostatnich latach rosła w skali całego systemu. Wszystkich obecnych doktorów habilitowanych uszeregowaliśmy pod względem

bieżącej produktywności („docelowe miejsce w mobilności produktywności” w tabeli 4), przydzielając ich do klas produktywności oddzielnie dla każdej dziedziny nauki i dla każdego typu produktywności. Następnie odtworzyliśmy retrospektywnie, w którym miejscu w klasyfikacji znajdowali się ci sami naukowcy w okresie, gdy byli doktorami („wyjściowe miejsce w mobilności produktywności”), opierając się na danych o ich produktywności z wybranego czteroletniego okresu. W sumie przeanalizowaliśmy 4165 indywidualnych ścieżek kariery akademickiej w pięciu dziedzinach nauki STEM, korzystając z pełnych danych biograficznych, administracyjnych i bibliometrycznych na poziomie poszczególnych naukowców.

Jak kształtują się szanse na radykalny awans (z 1. do 10. decyla) bądź radykalny spadek (z 10. do 1. decyla) w klasyfikacji produktywności? Czy naukowcy mogą w sposób zasadniczy zmieniać swoje zachowania publikacyjne (w porównaniu z rówieśnikami pod względem stopnia naukowego w poszczególnych dziedzinach nauki)?

Z naszych danych jasno wynika, że w podejściach uwzględniających prestiż czasopism nie występuje w ogóle przypadek skrajnego awansu w górę: żaden spośród 412 obecnych doktorów habilitowanych zaliczanych do najwyższej klasy produktywności w drugim etapie kariery nie należał wcześniej do najniższej klasy (0%). Z kolei w tradycyjnym podejściu bez normalizacji (zliczanie całkowite) znaleziono tylko jednego naukowca (0,24%) z taką ścieżką mobilności. Szanse na skrajny spadek są nieco wyższe, lecz w dalszym ciągu znikome – około 1% (5 naukowców spośród 419, czyli 1,19% w przypadku trzech typów produktywności, oraz 4 naukowców, czyli 0,95% dla Produktywności 4).

Różnice między dziedzinami nauki

Zbiorcze ujęcie wszystkich dziedzin nauki STEM maskuje bardziej zniuansowany obraz poszczególnych dziedzin, które charakteryzują się odmiennymi wzorcami mobilności między klasami produktywności. Jeśli uwzględnimy Produktywność 1 i skupimy się na mobilności poziomej (przejście z 10. decyla do 10. decyla), to w skali wszystkich dziedzin nauki 40–50% doktorów utrzymuje najwyższą klasę produktywności po habilitacji (tabela 5). Najwyższy odsetek zaobserwowano w naukach przyrodniczych (NATURAL) – 50,0%, natomiast w inżynierii (ENGI) jest to 47,9%. Najniższy odsetek zanotowano w naukach biologicznych (LIFE), gdzie 42,2% doktorów pozostaje po habilitacji w tej samej najwyższej klasie produktywności. Znaczne zróżnicowanie między dziedzinami nauki widać również przy analizie mobilności według pozostałych trzech typów produktywności (zob. tabela uzupełniająca 1).

Podobnie przeanalizowaliśmy poziomą mobilność z decyla 1. do decyla 1. w podziale na dziedziny nauki i typy produktywności. W skali wszystkich dziedzin nauki około jedna trzecia (33,3%) doktorów habilitowanych charakteryzuje się obecnie najniższą produk-

tywnością i utrzymuje ją przez oba etapy kariery. Najwyższe wartości obserwujemy przy zastosowaniu Produktywności 1, a najniższe przy Produktywności 4 (odpowiednio 33,3% i 26,9%).

Tabela 5. Cztery typy mobilności w ujęciu dziedzin nauki (Produktywność 1, zliczanie pełne, znormalizowane do prestiżu czasopism), wartości procentowe ($N = 4165$)

Dziedzina nauki	Doktorzy: Decyl 10. do decyla 10. (%)	Doktorzy: Decyl 1. do decyla 1. (%)	Doktorzy: Decyl 10. do decyla 1. (%)	Doktorzy: Decyl 1. do decyla 10. (%)
ENGI	47,9	22,4	2,1	0
LIFE	42,2	36	1,1	0
MATH	46,3	36,8	0	0
MED	43,8	40,3	0	0
NATURAL	50	35,2	1,6	0
Łącznie	46,5	33,3	1,2	0

Jednocześnie zróżnicowanie między dziedzinami nauki dla poszczególnych typów produktywności jest wyraźnie większe niż w przypadku mobilności z decyla 10. do decyla 10. (szczegółowe dane: tabela uzupełniająca 2). W dziedzinie ENGI (przy Produktywności 1) odsetek naukowców rozpoczynających karierę jako doktor w najniższej klasie produktywności i utrzymujących ją po habilitacji wynosi 22,4%, natomiast w dziedzinie MED 40,3%. Ogólnie rzecz biorąc, od 20% do 40% osób o najniższej produktywności na pierwszym etapie kariery utrzymuje najniższą produktywność na jej drugim etapie.

Przejścia z 10. do 1. decyla praktycznie się nie zdarzają; najwyższe empiryczne prawdopodobieństwo (2,1%) występuje w dziedzinie ENGI, a następnie 1,6% w dziedzinie NATURAL. Jedynie 1,2% doktorów zanotowało skrajny spadek z najwyższej do najniższej klasy produktywności; w dziedzinach MATH i MED w ogóle nie występują takie przypadki (0%).

Naukowcy przejawiający mobilność w górę (z decyla 1. do decyla 10.), co szczególnie interesuje badaczy produktywności, nie występują w żadnej dziedzinie nauki w typach produktywności uwzględniających prestiż czasopism (0%, tabela 6). Wyjątek stanowi zaledwie jeden naukowiec w naukach biologicznych (1,1% w LIFE) w każdym z dwóch typów bez normalizacji do prestiżu, łącznie więc jest dwoje takich naukowców.

Przykład ten dobrze pokazuje wartość danych mikro na poziomie jednostkowym, jakie gromadzimy w naszym Laboratorium Polskiej Nauki. Owe dwie osoby (mężczyzna i kobieta), pracujące w dyscyplinie AGRI w instytucjach o niższej intensywności badawczej (czyli nie typu IDUB) mają odpowiednio 44 i 48 lat, uzyskały stopień doktora w wieku 30 lat, a stopień doktora habilitowanego w wieku 41 i 39 lat. Publikują głównie w średnio- lub niskoprestiżowych czasopismach Scopus (mediana percentyla prestiżu: 32 i 25); ich zespoły badawcze są dość liczne jak na AGRI (średnio 5,21 i 4,77 autorów

na artykuł). Łączna liczba publikacji w dorobku tych naukowców to odpowiednio 14 i 40 artykułów, a w rozkładzie produktywności przeskoczyli z 6. i 4. percentyla do 90. percentyla. Z łatwością możemy obliczyć kilkanaście innych wskaźników dotyczących ich karier akademickich (tak jak w przypadku wszystkich pozostałych naukowców), w tym m.in. ich znormalizowany do dyscypliny wpływ cytowań, wiek i szybkość awansu w porównaniu z rówieśnikami w AGRI, zmieniające się w czasie wzorce publikacyjne czy wzorce współpracy.

Tabela 6. Mobilność pionowa w górę (z decyla 1. do decyla 10.) w podziale na dziedziny nauki i typ produktywności (wartości procentowe) ($N = 4165$)

Dziedzina nauki	Produktywność 1. znormalizowana do prestiżu, pełne zliczanie	Produktywność 2. znormalizowana do prestiżu, zliczanie cząstkowe	Produktywność 3. nieznormalizowana do prestiżu, pełne zliczanie	Produktywność 4. nieznormalizowana do prestiżu, zliczanie cząstkowe
	Doktorzy: z decyla 1. do decyla 10. (%)	Doktorzy: z decyla 1. do decyla 10. (%)	Doktorzy: z decyla 1. do decyla 10. (%)	Doktorzy: z decyla 1. do decyla 10. (%)
ENGI	0	0	0	0
LIFE	0	0	1,1	1,1
MATH	0	0	0	0
MED	0	0	0	0
NATURAL	0	0	0	0
Łącznie	0	0	0,2	0,2

Modelowanie statystyczne: regresja logistyczna

Na początku przedstawiamy cztery modele (po jednym dla każdego typu produktywności), w których oszacowano ilorazy szans przynależności do grupy doktorów habilitowanych o najwyższej produktywności (górnym 10%). Analogiczną analizę przeprowadzono dla grupy doktorów habilitowanych o najniższej produktywności (dolnym 10%), jednak z powodu ograniczonego miejsca nie pokazujemy jej tutaj. Każdy model ocenia wpływ wielu zmiennych na prawdopodobieństwo zaliczenia się do grupy bardzo wysoko produktywnych doktorów habilitowanych.

W każdym modelu sukcesem jest znalezienie się w grupie 10% najbardziej produktywnych doktorów habilitowanych. Wybór zmiennych był oparty na literaturze dotyczącej produktywności (np. Lee i Bozeman 2005; Ramsden 1994; Shin i Cummings 2010; Teodorescu 2000) oraz wysokiej produktywności (np. Abramo i in. 2009; Fox i Nikivincze 2021; Kwiek 2016) i dostępności danych. Przeprowadzono również analizę współliniowości między zmiennymi niezależnymi i w tym celu przebadano odwrócone macierze korelacji, wykorzystując wartości z głównej przekątnej, nie znajdując istotnego stopnia skorelowania wektora zmiennych niezależnych (zob. tabela uzupełniająca 4).

Najważniejszym predyktorem w czterech wykorzystanych modelach okazała się wcześniejsza przynależność doktorów habilitowanych do klasy doktorów o najwyższej

produktywności (śledzimy dokładnie tych samych naukowców zmieniających klasy produktywności w czasie; zmienna: najbardziej produktywni doktorzy, tabela 7). W Modelu 1 zwiększa to szanse niemal sześciokrotnie ($\text{Exp}(B) = 5,978$, 95% CI: 4,493–7,954). Podobnie silne efekty występują w pozostałych modelach z ilorazami szans równymi 6,735 (Model 3), 3,677 (Model 2) oraz 6,305 (Model 4). W każdym przypadku p -wartość w teście na istotność parametru jest mniejsza niż 0,001.

Tym samym wielowymiarowa analiza zdecydowanie potwierdza wyniki dwuwymiarowych analiz zaprezentowanych w poprzednich sekcjach pracy: przy uwzględnieniu łącznego wpływu wszystkich zmiennych, wcześniejsza przynależność do najwyższej klasy produktywności przed habilitacją (przy pozostałych czynnikach niezmiennych) jest w modelu zdecydowanie najsilniejszym predyktorem przynależności do tej klasy po habilitacji.

Uzyskane modele potwierdzają intuicje dotyczące zmian klas produktywności na dwóch etapach kariery w perspektywie longitudinalnej: jeśli naukowiec jest obecnie wysoce produktywny po habilitacji, zwykle był wysoce produktywny również wcześniej, przed habilitacją. Niezależnie od przyjętej metody pomiaru produktywności (zliczanie pełne, zliczanie cząstkowe, podejście znormalizowane do prestiżu czasopism bądź nieznormalizowane), wzorce wyłaniające się z analizy regresji pozostają podobne.

Czynnikiem statystycznie istotnym we wszystkich modelach jest również wcześniejsza przynależność do szybkiej ścieżki awansu wśród doktorów habilitowanych definiowanej jako 20% naukowców, którzy uzyskali stopień doktora habilitowanego w najkrótszym czasie od uzyskania doktoratu. Innymi słowy, chodzi o zmienną *habilitacja uzyskana szybko*, wskazującą na naukowców z relatywnie krótkim okresem przejścia od etapu doktoratu do etapu habilitacji. W Modelu 1 przynależność do tej klasy zwiększa szanse na sukces o 47,1% ($\text{Exp}(B) = 1,471$, 95% CI: 1,015–2,13), a w Modelu 4 efekt jest jeszcze silniejszy: iloraz szans wzrasta ponad dwukrotnie ($\text{Exp}(B) = 2,128$, 95% CI: 1,479–3,062; zob. materiały uzupełniające odnoszące się do szybkości i wieku uzyskania awansu).

Wydaje się zatem, że szybka ścieżka kariery akademickiej stanowi istotny czynnik sprzyjający wysokiej produktywności. Z kolei przynależność do klasy naukowców, którzy uzyskali doktorat w młodym wieku (*habilitacja uzyskana młodo*, górne 20% rozkładu), nie jest istotna statystycznie. Podobnie intensywność badawcza instytucji (typ IDUB) nie odgrywa tu znaczącej roli.

Interesująco wypada kwestia płci, która pojawia się jedynie w dwóch modelach – za każdym razem w tych modelach, gdzie zastosowano liczenie ułamkowe. W Modelu 2 bycie mężczyzną zwiększa szanse na dołączenie do grupy wysoko produktywnych doktorów habilitowanych o 49% ($\text{Exp}(B) = 1,49$, 95% CI: 1,165–1,905), a w Modelu 4 o 33% ($\text{Exp}(B) = 1,33$, 95% CI: 1,042–1,697).

W trzech z czterech modeli (Modele 1–3) wiek biologiczny wpływa istotnie i negatywnie na prawdopodobieństwo znalezienia się w grupie najbardziej produktywnych naukowców. W Modelu 1 każdy dodatkowy rok zmniejsza iloraz szans o ok. 11% ($\text{Exp}(B) = 0,888$, 95% CI: 0,85–0,929), zaś w Modelach 2 i 3 spadek wynosi 8–12% na rok. Wynika stąd, że młodszy doktorzy habilitowani częściej osiągają wysoką produktywność. Z kolei wiek akademicki, definiowany jako liczba lat od pierwszej publikacji indeksowanej w bazie Scopus, wpływa pozytywnie na szanse w Modelach 1 i 2 (oba podejścia są znormalizowane do prestiżu czasopism). W Modelu 1 każdy dodatkowy rok zwiększa szansę o 8% ($\text{Exp}(B) = 1,079$, 95% CI: 1,049–1,11), a w Modelu 2 o 5,8% ($\text{Exp}(B) = 1,058$, 95% CI: 1,029–1,088).

Warto zaznaczyć, że zarówno kierunek, jak i w dużej mierze siła działania predyktorów, nie zależą od konkretnego modelu (czyli od przyjętego typu produktywności). Niezależnie od sposobu pomiaru produktywności, statystycznie istotne pozostają te same predyktory. Najsilniejszym z nich okazuje się wcześniejsza przynależność do klasy o najwyższej produktywności przed habilitacją, co potwierdzają też omówione wcześniej analizy dwuwymiarowe mobilności typu poziomego z decyla 10. do decyla 10.

Modele regresji logistycznej wskazują na kilka istotnych predyktorów przynależności do grupy najmniej produktywnych doktorów habilitowanych. Najsilniejszym predyktorem w modelu jest wcześniejsza przynależność do klasy najmniej produktywnych doktorów, a w dalszej kolejności – wiek biologiczny. Wiek akademicki oraz wiek uzyskania habilitacji również odgrywają istotne role: dłuższa kariera akademicka oraz niższy wiek uzyskania habilitacji zmniejszają prawdopodobieństwo niskiej produktywności. W przeciwieństwie do modeli dotyczących doktorów habilitowanych o najwyższej produktywności, płeć jako zmienna nie okazała się istotna statystycznie. Natomiast intensywność badawcza instytucji (uczelnia typu IDUB) jest istotna w dwóch modelach, zwiększając szanse na niską produktywność o 31–38%.

W modelach konstruowanych dla najniższych klas produktywności predyktory mają zazwyczaj słabsze działanie niż w modelach dla klas najwyższych. Kierunki wpływu są zazwyczaj przeciwne – widać to szczególnie w przypadku zmiennych związanych z wiekiem. W obu przypadkach, niezależnie od przyjętego typu produktywności, najważniejszym predyktorem przynależności do klasy najwyższej (bądź najniższej) produktywności jest wcześniejsza przynależność do analogicznej klasy produktywności na etapie pracy z doktoratem.

Dyskusja i wnioski

Po pierwsze, nasza analiza ma mocne fundamenty empiryczne: zbadaliśmy produktywność publikacyjną wszystkich polskich doktorów habilitowanych widocznych międzynarodowo w 12 dyscyplinach STEMM (zgrupowanych w pięć dziedzin nauki, $N = 4165$) i przeanalizowaliśmy ich kompletne portfele publikacyjne (Scopus) oraz pełne dane

biograficzne (OPI PIB). Dane biograficzne i demograficzne zostały połączone z metadanyami o publikacjach, obejmującymi wszystkie polskie artykuły opublikowane w ostatnim półwieczu i indeksowane w bazie Scopus (1973–2021, $N = 935\,167$).

Po drugie, w ramach szerszego podejścia metodologicznego przyjęliśmy za jednostkę analizy indywidualnego naukowca, a nie pojedynczą publikację. Zastosowaliśmy cztery główne typy pomiaru produktywności, aby ocenić, w jakim stopniu różne metody zliczania produktywności mogą wpływać na uzyskiwane wzorce (aby nie komplikować badań, pominęliśmy kilka kolejnych sposobów: zliczanie inaczej publikacji w charakterze pierwszego i ostatniego autora, autora korespondującego etc. w różnych wariantach).

Ponadto posłużyliśmy się klasyfikacjami opartymi na decylach produktywności (a nie na bezwzględnej liczbie publikacji). Szczególnie istotne okazało się podejście znormalizowane do prestiżu czasopisma (w którym liczy się lokalizacja opublikowanego artykułu w zhierarchizowanym globalnym systemie czasopism, czyli w systemie Scopus CiteScore) zestawione z podejściem nieznormalizowanym (ze zliczaniem pełnym lub ułamkowym). I wreszcie zastosowaliśmy longitudinalny schemat badawczy, w którym możemy śledzić produktywność publikacyjną na przestrzeni lat i dekad, uwzględniając dwa kolejne etapy kariery: etap pracy przed habilitacją i po niej. Udało się to dzięki dysponowaniu pełnymi danymi o awansach, cechach demograficznych oraz o publikacjach (indeksowanych w Scopus) na poziomie indywidualnych naukowców.

W naszym zbiorze danych produktywność doktorów habilitowanych rozciąga się na wiele dekad, ponieważ w polskim systemie nie ma zewnętrznych wymogów czasowych związanych z awansami. Część analizowanych naukowców (30,9%) jest w wieku 55 lat lub więcej i pracuje w systemie szkolnictwa wyższego od ponad 20 lat.

W analizach longitudinalnych skupiliśmy się na wzorcach mobilności pomiędzy klasami produktywności na etapie doktoratu i etapie habilitacji. W naszym głównym typie produktywności (zliczanie pełne, znormalizowane do prestiżu czasopism) wyniki pokazują, że niespełna połowa (46,5%) naukowców z najwyższej klasy (górne 10%, decyl 10.) utrzymuje tę pozycję, a około jedna trzecia (33,3%) naukowców z najniższej klasy (dolne 10%, decyl 1.) pozostaje w tej samej klasie.

Ekstremalna mobilność okazała się w Polsce zjawiskiem marginalnym: żaden (0%) z obecnych doktorów habilitowanych nie doświadczył takiej ścieżki w trakcie swojej kariery. W badanej próbie 4165 naukowców nie odnotowaliśmy przypadków awansu z najniższej do najwyższej klasy w typach znormalizowanych do prestiżu (oraz zaledwie dwie osoby w podejściach bez normalizacji) oraz jedynie pięcioro naukowców przeszło z najwyższej do najniższej klasy. Obecni doktorzy habilitowani z najwyższej klasy produktywności w przeszłości prawie zawsze należeli do grupy bardzo produktywnych już na etapie pracy z doktoratem (mediana: 87,9 percentyla), a doktorzy habilitowani z klasy najniższej – do grona najniższych klas produktywności (mediana: 18,3 percentyla).

Tabela 7. Statystyki regresji logistycznej: oszacowania ilorazów szans na przynależność do klasy najbardziej produktywnych doktorów habilitowanych

Model	Model 1: Produktynność 1. Znormalizowana do prestiżu, pełne zliczanie			Model 2: Produktynność 2. Znormalizowana do prestiżu, zliczanie częściowe			Model 3: Produktynność 3. nieznormalizowana do prestiżu, pełne zliczanie			Model 4: Produktynność 4. nieznormalizowana do prestiżu, zliczanie częściowe		
	pseudo $R^2 = 0,22, N = 4165$	95% przedział ufności		pseudo $R^2 = 0,21, N = 4165$	95% przedział ufności		pseudo $R^2 = 0,20, N = 4165$	95% przedział ufności		pseudo $R^2 = 0,18, N = 4165$	95% przedział ufności	
	Exp(B)	Dolny	Górny	Exp(B)	Dolny	Górny	Exp(B)	Dolny	Górny	Exp(B)	Dolny	Górny
Mężczyzna	1,147	0,899	1,464	1,49**	1,165	1,905	1,255	0,983	1,602	1,33*	1,042	1,697
Intensywność badawcza: nie IDUB	0,923	0,725	1,174	1,079	0,847	1,374	0,856	0,675	1,086	0,912	0,72	1,157
Wiek biologiczny	0,888***	0,85	0,929	0,878***	0,839	0,918	0,922***	0,884	0,961	0,981	0,945	1,018
Wiek akademicki	1,079***	1,049	1,11	1,058***	1,029	1,088	1,025	0,998	1,052	0,988	0,964	1,012
Wiek: doktorat	1,066	0,993	1,145	1,126**	1,048	1,21	1,054	0,987	1,126	1,013	0,949	1,081
Wiek: habilitacja	0,959	0,905	1,016	0,949	0,891	1,011	0,97	0,917	1,026	0,936	0,887	0,987
Klasa: najbardziej pro- duktywni doktorzy (10%)	5,978***	4,493	7,954	3,677***	2,776	4,871	6,735***	5,156	8,797	6,305***	4,85	8,197
Klasa: doktorat uzyskany młodo (10%)	1,301	0,939	1,804	1,54	1,119	2,12	1,154	0,837	1,592	1,057	0,768	1,455
Klasa: habilitacja uzyskana młodo (10%)	0,736	0,502	1,08	0,781	0,537	1,137	0,661*	0,45	0,972	0,568**	0,387	0,834
Klasa: habilitacja uzyskana szybko (10%)	1,471*	1,015	2,13	1,63*	1,129	2,354	1,825**	1,265	2,633	2,128***	1,479	3,062
Constant	4,01	0,62	25,93	2,117	0,354	12,671	1,453	0,269	7,84	2,274	0,423	12,241

*** - $p < 0,001$; ** - $p < 0,01$; * - $p < 0,05$

Zarówno dla samych naukowców, jak i dla decydentów w obszarze polityki naukowej na różnych poziomach nasze ustalenia stanowią poważne wyzwanie: wyniki analiz pokazują, że radykalne zmiany poziomu produktywności publikacyjnej (skok w górę bądź w dół, tutaj rozumiane jako przejście między najwyższymi 10% a najniższymi 10% w rozkładzie produktywności) praktycznie nie zdarzają się w obszarze STEMM.

Innymi słowy, niektórzy naukowcy przez lata i dekady utrzymują bardzo wysoki poziom produktywności, podczas gdy inni – pracujący w tych samych instytucjach i w tych samych dziedzinach nauki – latami utrzymują bardzo niski poziom produktywności. Prawdopodobieństwo, że naukowiec przechodząc kolejne szczeble kariery akademickiej, stanie się nagle względnie znacznie bardziej produktywny, jest zerowe – a prawdopodobieństwo, że stanie się względnie znacznie mniej produktywny, jest marginalne.

Istotne jest też zróżnicowanie między poszczególnymi dyscyplinami: w naukach przyrodniczych (NATURAL) odsetek naukowców doświadczających mobilności z 10. decyla do 10. decyla sięga 50%. Ponadto ogólny obraz dla wszystkich dziedzin STEMM, w którym 33,3% doktorów pracujących wcześniej jako doktorzy w najniższej klasie produktywności pozostaje na tym samym poziomie jako doktorzy habilitowani, kryje w sobie dużo większe zróżnicowanie. W zależności od dziedziny, między 30% a 60% doktorów o najniższej produktywności kontynuuje karierę w tej samej najniższej klasie także na późniejszym etapie. Porównując te wyniki z trwałością „gwiazdorstwa” w nauce zbadaną przez Abramo i in. (2017) dla Włoch, widać, że w Polsce odsetek naukowców utrzymujących wysoką produktywność na obu etapach kariery (46,5% we wszystkich dziedzinach STEMM łącznie) jest wyższy niż udział naukowców utrzymujących we Włoszech swój status gwiazdorski w ciągu 12 lat (35%). Stratyfikacja produktywności w Polsce okazuje się głębsza i trwalsza niż we Włoszech, prawdopodobnie z uwagi na wieloletnie, poważne niedofinansowanie badań naukowych.

Analiza regresji logistycznej wyraźnie potwierdza wyniki analizy dwuwymiarowej. W przypadku oszacowań ilorazów szans przynależności do klasy doktorów habilitowanych o najwyższej produktywności najważniejszym predyktorem (we wszystkich czterech badanych modelach) okazuje się wcześniejsza przynależność do klasy najbardziej produktywnych doktorów. Efekt ten jest istotny statystycznie we wszystkich modelach i zwiększa szanse na znalezienie się w najwyższej klasie nawet 4–6-krotnie (w zależności od typu produktywności). Również wcześniejsza przynależność do klasy szybko awansujących habilitowanych jest istotna statystycznie w każdym modelu i zwiększa szanse na sukces od 50% do 130%. Z kolei w przypadku oszacowań ilorazów szans przynależności do klasy najmniej produktywnych doktorów habilitowanych najsilniejszym predyktorem w modelach jest wcześniejsza przynależność do najmniej produktywnej klasy doktorów; w takim przypadku szanse wzrastają o 150–300% (również w zależności od typu produktywności).

Z szerszej perspektywy nasze badanie pokazuje, że tradycyjne pojęcie silnego i słabego indywidualnego dorobku naukowego – tutaj rozumianego wyłącznie jako produktywność publikacyjna (bez uwzględniania np. cytawalności znormalizowanej do dyscypliny czy pozyskiwania funduszy badawczych) – jest przydatne do oceny poszczególnych naukowców. Osoby z bardzo słabym dorobkiem w przeszłości mają, zgodnie z naszymi wynikami, znikome szanse, aby w przyszłości osiągnąć wyjątkowo wysoki poziom produktywności (w obszarze STEM).

Wniosek płynący z naszych badań dla rozwoju karier akademickich jest taki, że wczesne osiągnięcia naukowe, mierzone tutaj wysoką publikacyjną produktywnością w początkowej fazie kariery, istotnie wpływają na osiągnięcia na późniejszym etapie (wczesne „zakotwiczenie” w klasach produktywności). Innymi słowy, jeśli okres pracy na etapie doktoratu wyróżniał się wysoką liczbą i jakością publikacji, to analogicznie okres pracy na etapie habilitacji również będzie najpewniej charakteryzował się wysoką produktywnością. Jeśli zaś ten pierwszy etap cechował się słabym dorobkiem – to i kolejne lata zwykle pozostaną na niskim poziomie produktywności.

Zgromadzone dane na poziomie mikro wyraźnie wskazują, że naukowcy przez lata, a nawet dekady, tkwią w tych samych klasach produktywności publikacyjnej we własnych dziedzinach. Radykalne przejścia między skrajnymi klasami są wyjątkową rzadkością, w praktyce tacy skoczki w nauce nie istnieją, podobnie jak w całym obszarze OECD, zob Kwiek i Szymula 2025). W efekcie wczesne osiągnięcie górnych granic indywidualnej produktywności przekłada się na dalsze etapy kariery: późniejsza produktywność jest silnie powiązana z wcześniejszą.

Należy jednak podkreślić, że w tym badaniu nasz pomiar produktywności ma charakter względny, a nie bezwzględny. Produktywność poszczególnych naukowców, pogrupowanych w klasy, jest porównywana do produktywności innych naukowców, również pogrupowanych w klasy. Choć interesuje nas zmiana, to chodzi o zmianę między klasami (czyli względną), a nie zmianę dotyczącą samej liczby publikacji (nominalną).

Nie zajmujemy się więc tym, czy produktywność mierzona liczbą publikacji rośnie bądź spada wraz z wiekiem naukowca (lub wraz z postępami w jego karierze). W tym sensie nie możemy stwierdzić, czy najbardziej produktywni naukowcy faktycznie zwiększają swą produktywność w czasie – zgodnie z hipotezą intensyfikacji publikacyjnej (Hermanowicz i Scheitle 2023). Skupiamy się wyłącznie na tym, czy dochodzi do zmiany klasy produktywności, co jest zagadnieniem niezależnym od klasycznej dyskusji nad relacjami między wiekiem a osiągnięciami naukowymi (np. Wang i Barabási 2021).

Zaobserwowane wzorce zmian można zrozumieć, odwołując się do tradycyjnych teorii produktywności naukowej (zwłaszcza teorii „iskry Bożej,” np. Allison i Stewart 1974; Cole i Cole 1973; Fox 1983 czy teorii akumulacji przewag i efektu św. Mateusza,

np. David 1994; DiPrete i Eirich 2006; Merton 1968). Nieco zaskakująca jest jednak siła tych wzorców w tak różnych dziedzinach nauki i dla różnych miar produktywności: uderzająca okazuje się wysoka trwałość przynależności do klasy najwyższej produktywności przez całą karierę akademicką, a także zerowe prawdopodobieństwo mobilności między klasą najniższą a najwyższą na kolejnych szczeblach awansu akademickiego.

Nasze badanie podkreśla zatem długofalowy charakter kariery naukowej, w którym już na etapie pracy z doktoratem kształtują się wzorce produktywności (i zapewne inne wzorce aktywności zawodowej), wpływające na produktywność w okresie pracy z habilitacją. Można założyć, że pewne wzorce pracy, często przejmowane od promotorów i opiekunów naukowych – nastawienie na dydaktykę lub na badania, podział czasu pracy w tygodniu (w tym czas poświęcony na badania), nawyki publikacyjne i nawyki związane ze współpracą – kształtują się latami i utrzymują na dalszych etapach kariery. Nauka wymaga znacznych nakładów czasowych (tygodniowych, miesięcznych, rocznych); a kariery naukowe o bardzo wysokiej produktywności zwykle formują się przez dekady – i nasze dane potwierdzają, że potrzebne są do tego lata konsekwentnie wysokiej produktywności.

Oczywiście nasze badanie ma ograniczenia. Nie analizujemy klas produktywności wszystkich doktorów, ponieważ skupiamy się na doktorach habilitowanych publikujących w ostatnim dostępnym czteroletnim okresie (2018–2021) i retrospektywnie na ich produktywności z czasów pracy z doktoratem. Tym samym nie obejmujemy naszym badaniem doktorów, którzy opuścili polski system nauki. Ponadto badanie nie dotyczy tych naukowców, którzy nigdy nie uzyskali stopnia doktora habilitowanego (ponieważ nie znaleźli się w naszej próbie).

W naszym zbiorze danych nie ujęto też wielu czynników środowiskowych, które mogą silnie wpływać na indywidualną produktywność – np. klimatu w miejscu pracy (Fox i Mohapatra 2007), który ma duże znaczenie szczególnie dla kobiet w STEMM (Branch 2016). Dysponujemy bowiem tylko jedną zmienną opisującą intensywność badawczą uczelni, tj. przynależność do programu doskonałości naukowej (IDUB). Nie mamy też dostępu do danych na temat nastawień i zachowań akademickich (typowych dla ankiet), *work-life balance* czy sytuacji rodzinnej, co jest istotne w badaniach ankietowych (zob. Kwiek 2019).

Należy również pamiętać, że na poziomie jednostkowym długotrwała zerowa lub bardzo niska produktywność jest w zasadzie niemożliwa, zwłaszcza w górnych warstwach systemu nauki: tacy naukowcy zwykle opuszczają polski system szkolnictwa wyższego i nauki. Zatem nasza grupa doktorów habilitowanych i tak stanowi grono osób, które odniosły sukces – w sensie pozostania w systemie (co można określić mianem *success bias*). To tylko nasz statystyczny podział decylowy sprawia, że pewna część pozostających w systemie zalicza się do dolnych klas produktywności. Bez względu na

ogólny poziom produktywności w systemie zawsze można go podzielić na dziesięć decyli, z których jeden będzie tym najniższym, zachowując wartości progowe.

Nasze badanie obejmuje jeden system nauki (Polskę). Możliwość uogólniania wniosków zależy więc od podobieństw i różnic między systemami nauki na świecie – choćby w zakresie zróżnicowania ścieżek kariery akademickiej, wewnętrznej konkurencyjności, struktury zachęt, proporcji czasu poświęconego na dydaktykę i badania, dostępu do finansowania czy atrakcyjności zawodu naukowca. Nasze badanie jest jednak elementem szerszego projektu porównawczego obejmującego 38 krajów OECD i analizującego mobilność pomiędzy klasami produktywności na podstawie ścieżek zawodowych ponad 320 tys. naukowców znajdujących się na zaawansowanych etapach kariery (co najmniej 25 lat kariery publikacyjnej, zob. Kwiek i Szymula 2025).

Podziękowania

Autorzy składają serdeczne podziękowania organizatorom i uczestnikom seminariów, na których Marek Kwiek prezentował koncepcje badań longitudinalnych, w szczególności w University of Oxford (Simon Marginson, CGHE, Center for Global Higher Education, czerwiec 2022), Stanford University (John Ioannidis, METRICS, Meta-Research Innovation Center, czerwiec 2022), DZHW Berlin (Torger Moeller, German Center for Higher Education Research and Science Studies, czerwiec 2023) oraz Leiden University (Ludo Waltman, CWTS, Centre for Science and Technology Studies, czerwiec 2023). Szczególne podziękowania kierujemy do dr. Łukasza Szymuli z Zakładu Sztucznej Inteligencji UAM w Poznaniu za wsparcie w zakresie zbierania i analizowania danych z obszaru OECD. Wyrażamy też wdzięczność International Center for the Study of Research (ICSR) Lab, zwłaszcza dla Kristy James i Alicka Birda, oraz dziękujemy za wsparcie w postaci grantu MNISW NDS nr NdS/529032/2021/2021.

Bibliografia

- Abramo G., Cicero T., D'Angelo C.A. (2013). The impact of unproductive and top researchers on overall university research performance. *Journal of Informetrics*. 7(1), 166–175.
- 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., Di Costa F. (2009). Research collaboration and productivity: Is there correlation? *Higher Education*. 57(2), 155–171.
- Abramo G., D'Angelo C.A., Soldatenkova A. (2017). How long do top scientists maintain their stardom? An analysis by region, gender and discipline: Evidence from Italy. *Scientometrics*. 110(2), 867–877.
- Adams J. (2013). The fourth age of research. *Nature*, 497(30 May), 557–560.
- Agrawal A., McHale J., Oettl A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*. 46(4), 853–867.

- 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.
- Allison P.D., Long J.S., Krauze T.K. (1982). Cumulative advantage and inequality in science. *American Sociological Review*. 47(5), 615–625.
- Allison P.D., Stewart J.A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review*. 39(4), 596–606.
- Antonowicz D., Kulczycki E., Budzanowska A. (2021). Breaking the deadlock of mistrust? A participative model of the structural reforms in higher education in Poland. *Higher Education Quarterly*. 74(4), 391–409.
- 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.
- Bornmann L. (2024). Skewed distributions of scientists' productivity: a research program for the empirical analysis. *Scientometrics*. 129, 2455–2468.
- Branch E.H., red. (2016). *Pathways, Potholes, and the Persistence of Women in Science: Reconsidering the Pipeline*. Lanham, MD: Lexington Books.
- Carrasco R., Ruiz-Castillo J. (2014). The evolution of the scientific productivity of highly productive economists. *Economic Inquiry*. 52(1), 1–16.
- Carvalho T. (2017). The study of the academic profession – Contributions from and to the sociology of professions. [W:] J. Huisman and M. Tight (red.), *Theory and Method in Higher Education Research*. Bingley, UK: Emerald Group Publishing Limited, First edition, 59–76.
- Chatterjee S., Hadi A.S. (1986). Influential observations, high leverage points, and outliers in linear regression. *Statistical Science*. 1(3), 379–393.
- Cole J.R., Cole S. (1973). *Social Stratification in Science*. Chicago, IL: University of Chicago Press.
- Cortés L.M., Mora-Valencia A., Perote J. (2016). The productivity of top researchers: A semi-nonparametric approach. *Scientometrics*. 109(2), 891–915.
- Costas R., Bordons M. (2007). A classificatory scheme for the analysis of bibliometric profiles at the micro level. *Proceedings of ISSI 2007: 11th international conference of the ISSI*, t. I i II, 226–230.
- Costas R., Bordons M. (2005). Bibliometric indicators at the micro-level: Some results in the area of natural resources at the Spanish CSIC. *Research Evaluation*. 14(2), 110–120.
- Crane D. (1965). Scientists at major and minor universities: A study of productivity and recognition. *American Sociological Review*. 30(5), 699–714.
- Croux C., Dhaene G., Hoorelbeke D. (2004): *Robust Standard Errors for Robust Estimators*. Center for Economic Studies Discussions Paper Series (DPS) 03.16. Available from: <http://www.econ.kuleuven.be/ces/discussionpapers/default.htm>
- David P.A. (1994). Positive feedbacks and research productivity in science: Reopening another black box. [W:] O. Granstrand, *Economics of Technology*. Amsterdam: Elsevier, 65–89.
- DiPrete T.A., Eirich G.M. (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology*. 32(1), 271–297.
- Fender B.F., Taylor S.W., Burke K.G. (2005). Making the big leagues: Factors contributing to publication in elite economics journals. *Atlantic Economic Journal*. 33(1), 93–103.

- Fox M.F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*. 13(2), 285–305.
- Fox M.F., Nikivincze I. (2021). Being highly prolific in academic science: Characteristics of individuals and their departments. *Higher Education*. 81, 1237–1255.
- Fox M.F., Mohapatra S. (2007). Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *The Journal of Higher Education*. 78(5), 542–571.
- GUS (2023). *Higher Education and its Finances in 2022*. Main Statistical Office of Poland. Dostęp: 12.VI.2024 <https://stat.gov.pl/obszary-tematyczne/edukacja/edukacja/szkolnictwo-wyzsze-i-jego-finanse-w-2022-roku,2,19.html>
- Hermanowicz J.C. (2012). The sociology of academic careers: Problems and prospects. [W:] J. C. Smart and M. B. Paulsen (red.), *Higher Education: Handbook of Theory and Research*. T. 27. Dordrecht: Springer.
- Hermanowicz J.C., Scheitle C.P. (2023). Predicting future publishing success among sociologists in the US Higher education system, *Research Evaluation*. 32(2), 384–394.
- 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.
- 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. <https://doi.org/10.1007/s10734-015-9910-x>
- Kwiek M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers? *Scientometrics*. 115(1), 415–462. <https://doi.org/10.1007/s11192-018-2644-7>
- Kwiek M. (2019). *Changing European Academics. A Comparative Study of Social Stratification, Work Patterns and Research Productivity*. London and New York: Routledge.
- Kwiek M., Roszka W. (2021). Gender disparities in international research collaboration: A large-scale bibliometric study of 25,000 university professors. *Journal of Economic Surveys*. 35(5), 1344–1388. <https://doi.org/10.1111/joes.12395>
- Kwiek M., Roszka W. (2022). Academic vs. biological age in research on academic careers: A large-scale study with implications for scientifically developing systems. *Scientometrics*. 127, 3543–3575. <https://doi.org/10.1007/s11192-022-04363-0>
- Kwiek M., Roszka W. (2023). The young and the old, the fast and the slow: a large-scale study of productivity classes and rank advancement. *Studies in Higher Education*, 1–16. <https://doi.org/10.1080/03075079.2023.2288172>
- Kwiek M., Roszka W. (2024a). Top research performance in Poland over three decades: A multi-dimensional micro-data approach. *Journal of Informetrics*, 18(4). November 2024. 101595. 1–16. <https://doi.org/10.1016/j.joi.2024.101595>
- Kwiek M., Roszka W. (2024b). Once highly productive, forever highly productive? Full professors' research productivity from a longitudinal perspective. *Higher Education*. 87, 519–549. <https://doi.org/10.1007/s10734-023-01022-y>
- Kwiek M., Szymula L. (2025). Quantifying lifetime productivity changes: A longitudinal study of 320,000 late-career scientists. *Quantitative Science Studies*. <https://doi.org/10.1162/qss.a.16.1-38>.

- Kyvik S. (1990). Age and scientific productivity: Differences between fields of learning. *Higher Education*. 19, 37–55.
- Lee S. Bozeman B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*. 35(5), 673–702.
- Leišyte L., Dee J.R. (2012). Understanding academic work in a changing institutional environment. [W:] J.C. Smart, M.B. Paulsen (red.), *Higher Education: Handbook of Theory and Research*. Dordrecht: Springer Netherlands, 27, 123–206.
- Lotka A.J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*. 16(12), 317–323.
- Menard S. (2002). *Longitudinal Research*. Thousand Oaks, CA: Sage.
- Merton R.K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*. 159(3810), 56–63.
- Milojevic S., Radicchi F., Walsh J.P. (2018). Changing demographics of scientific careers: The rise of the temporary workforce. *Proceedings of the National Academy of Sciences*. 115, 12616–12623.
- O’Boyle Jr. E., Aguinis H. (2012). The best and the rest: Revisiting the norm of normality of individual performance: Personnel psychology. *Personnel Psychology*, 65(1), 79–119.
- Olechnicka A., Ploszaj A., Celinska-Janowicz D. (2019). *The Geography of Scientific Collaboration*. London and New York: Routledge.
- Piro F.N., Rørstad K., Aksnes D.W. (2016). How does prolific professors influence on the citation impact of their university departments? *Scientometrics*. 107(3), 941–961.
- Price D.J. de S. (1963). *Little Science, Big Science*. New York: Columbia University Press.
- Ramsden P. (1994). Describing and explaining research productivity. *Higher Education*. 28(2), 207–226.
- Ruiz-Castillo J., Costas R. (2014). The skewness of scientific productivity. *Journal of Informetrics*. 8(4), 917–934.
- Shin J.C., Cummings W.K. (2010). Multilevel analysis of academic publishing across disciplines: Research preference, collaboration, and time on research. *Scientometrics*. 85(2), 581–594.
- Sidiropoulos A., Gogoglou A., Katsaros D., Manolopoulos Y. (2016). Gazing at the skyline for star scientists. *Journal of Informetrics*. 10(3), 789–813.
- Singer J.D., Willett J.B. (2003). *Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence*. Oxford: Oxford University Press.
- Stephan P.E. (2015). *How Economics Shapes Science*. Boston, MA: Harvard University Press.
- Stephan P.E., Levin S.G. (1992). *Striking the Mother Lode in Science: The Importance of Age, Place, and Time*. New York: Oxford University Press.
- Sugimoto C., Larivière V. (2018). *Measuring Research: What Everyone Needs to Know*. Oxford: Oxford University Press.
- Teodorescu D. (2000). Correlates of faculty publication productivity: A cross-national analysis. *Higher Education*. 39(2), 201–222.
- 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)*. Paris: CNRS. Retrieved from <http://piketty.pse.ens.fr/files/Turner2005.pdf>
- Wagner C.S. (2018). *The Collaborative Era in Science. Governing the Network*. Cham: Palgrave Macmillan.

- Waltman L., van Eck N.J. (2019). Field normalization of scientometric indicators. [W:] W. Glänzel H.F. Moed U. Schmoch M. Thelwall (red.), *Springer Handbook of Science and Technology Indicators*. Cham: Springer, 281–300.
- Wang D., Barabási A. (2021). *The Science of Science*. Cambridge: Cambridge University Press.
- Wanner R.A., Lewis L.S., Gregorio D.I. (1981). Research productivity in academia: A comparative study of the sciences, social sciences and humanities. *Sociology of Education*. 54(4), 238.
- Wuchty S., Jones B.F., Uzzi B. (2007) The increasing dominance of teams in production of knowledge. *Science*. 316(5827), 1036–1039.
- Xie Y. (2014). ‘Undemocracy’: Inequalities in science. *Science*, 344(6186), 809–810.
- Yair G., Gueta N., Davidovitch N. (2017). The law of limited excellence: Publication productivity of Israel Prize laureates in the life and exact sciences. *Scientometrics*. 113(1), 299–311.
- Yin Z., Zhi Q. (2017). Dancing with the academic elite: A promotion or hindrance of research production? *Scientometrics*. 110(1). 17–41.
- Zuckerman H. (1970). Stratification in American science. *Sociological Inquiry*. 40(2), 235–257.

Material uzupełniający

Podejście znormalizowane do prestiżu czasopism (funkcja potęgowa)

W przypadku zastosowania funkcji liniowej wartość artykułu opublikowanego w czasopiśmie z 99. percentyla czasopism w bazie Scopus (np. *Nature*, *Science*, *Cell*, *Lancet* itp.) wynosiłaby 0,99, co jest tylko nieznacznie niższą wartością niż dwukrotność wagi artykułu opublikowanego w czasopiśmie z 50. percentyla. Taka sytuacja nie odzwierciedlałaby w sposób odpowiedni nakładu pracy naukowej.

Dlatego w naszej metodzie (wykorzystywanej w dwóch wariantach produktywności znormalizowanych do prestiżu czasopism) wartość artykułu w czasopiśmie o niskiej randze percentylowej (np. z 1.–50. percentyla) rośnie powoli, natomiast w czasopismach wysoko notowanych (90.–99. percentyl) – rośnie bardzo stromo, tak by różnica między artykułem opublikowanym w 99. a 50. percentylu wynosiła pięciokrotność (dokładnie 4,67).

Takie podejście pozwala uchwycić różnorodność indywidualnych wzorców publikacyjnych i zróżnicowane drogi do wysokiej produktywności: jedni badacze wybierają kilka publikacji w bardzo prestiżowych czasopismach, inni skupiają się na licznych publikacjach w czasopismach o niższym prestiżu (niższych rangach percentylowych). (Analogicznym ujęciem byłaby praca na danych z bazy Web of Science w oparciu o współczynnik wpływu; uważamy jednak, że praca w ramach CiteScore z bazy Scopus prowadzi do bardziej czytelnych wyników).

W analizowanym czteroletnim okresie (2018–2021) nie rozróżniamy zmian w rangach percentylowych czasopism w bazie Scopus w kolejnych latach czy dekadach (histo-

ryczne dane o rangach percentylowych nie są dostępne). Wykorzystujemy najnowsze (z 2023 roku) rangi percentylowe Scopus jako przybliżenia. Dla zdecydowanej większości czasopism z obszaru STEMM zmiany w rangach percentylowych w kolejnych latach są raczej umiarkowane.

W Scopus system rangowania czasopism opiera się na cytowaniach, jakie w ciągu poprzednich czterech lat uzyskały wszystkie publikacje z danego czasopisma. Choć rangi percentylowe stanowią jedynie przybliżenie jakości (odzwierciedlają wpływ całego czasopisma na środowisko naukowe, a nie wpływ konkretnego artykułu), to jednak artykuły w czasopismach o wysokim prestiżu są na ogół przeciętnie wyżej cytowane. Zastosowane w niniejszej pracy miary mogą być uznane za przybliżone, jednak obecnie brak bardziej wiarygodnych, zintegrowanych danych o publikacjach polskich naukowców z ostatniego półwiecza.

W podejściu do produktywności bez normalizacji do prestiżu czasopisma (zliczanie pełne) każdy artykuł, niezależnie od czasopisma, otrzymuje wartość 1. Z kolei w naszym podejściu znormalizowanym do prestiżu (wersja potęgowa) w przypadku zliczania pełnego artykułu w czasopiśmie z 90. percentylem ma wartość 0,77, zaś artykuł w czasopiśmie z 50. percentylem – 0,18.

Wzór przyjmuje postać:

$$p_{adj}(exp) = (perc100)^{2,5} p_{\mathit{adj}(exp)} = \left(\frac{\textit{perc}}{100} \right)^{2,5} p_{adj}(exp) = (100perc)^{2,5}$$

gdzie $p_{\mathit{adj}(exp)}$ to potęgowo znormalizowana do prestiżu wartość artykułu, a $perc$ oznacza percentyl czasopisma, w którym opublikowano artykuł (według bazy Scopus).

Tabela uzupełniająca 1. Mobilność pozioma z decyla 10. do decyla 10. w podziale na dziedziny nauki i typ produktywności (wartości procentowe) ($N = 4165$)

Dziedzina nauki	Produktywność 1. Zliczanie pełne, znormalizowane do prestiżu czasopism	Produktywność 2. Zliczanie ułamkowe, znormalizowane do prestiżu czasopism	Produktywność 3. nieznormalizowana do prestiżu, pełne zliczanie	Produktywność 4. nieznormalizowana do prestiżu, zliczanie cząstkowe
	Doktorzy: decyl 10. do decyl 10. (%)	Doktorzy: decyl 10. do decyl 10. (%)	Doktorzy: decyl 10. do decyl 10. (%)	Doktorzy: decyl 10. do decyl 10. (%)
ENGI	47,9	46,9	44,8	41,7
LIFE	42,2	37,8	45,6	32,2
MATH	46,3	34,1	46,3	48,8
MED	43,8	34,4	39,1	37,5
NATURAL	50	41,4	49,2	48,4
Łącznie	46,5	40,1	45,6	41,8

Tym samym cztery artykuły w 50. percentylu (wydane w okresie czterech lat) dają łącznie wartość $4 \times 0,177 = 0,7084$, a następnie średnią dzieloną przez cztery lata, co daje średnią roczną wartość produktywności 0,177 w oknie 2018–2021. Natomiast pojedynczy artykuł w 99. percentylu będzie miał wartość 0,975 podzielone przez 4 lata (0,248). Funkcja potęgowa wprowadza karę za publikacje w czasopismach z niską rangą percentylową – o niskim prestiżu – która maleje wraz ze wzrostem prestiżu czasopisma.

Tabela uzupełniająca 2. Mobilność pozioma z decyla 1. do decyla 1. w podziale na dziedziny nauk i typ produktywności (wartości procentowe) ($N = 4165$)

Dziedzina nauki	Model 1: Produktywność 1. Znormalizowana do prestiżu, pełne zliczanie	Model 2: Produktywność 2. Znormalizowana do prestiżu, zliczanie cząstkowe	Model 3: Produktywność 3. nieznormalizowana do prestiżu, pełne zliczanie	Model 4: Produktywność 4. nieznormalizowana do prestiżu, zliczanie cząstkowe
	Doktorzy: decyl 1. do decyl 1. (%)	Doktorzy: decyl 1. do decyl 1. (%)	Doktorzy: decyl 1. do decyl 1. (%)	Doktorzy: decyl 1. do decyl 1. (%)
ENGI	22,4	22,1	29,8	25,3
LIFE	36	33,7	36,7	25,8
MATH	36,8	30,8	35	30
MED	40,3	29	32,3	24,2
NATURAL	35,2	32,3	31,7	29,1
Łącznie	33,3	29,6	32,8	26,9

Tabela uzupełniająca 3. Mobilność pionowa z decyla 10. do decyla 1. w podziale na dziedziny nauk i typ produktywności (wartości procentowe) ($N = 4165$)

Dziedzina nauki	Model 1: Produktywność 1. Znormalizowana do prestiżu, pełne zliczanie	Model 2: Produktywność 2. Znormalizowana do prestiżu, zliczanie cząstkowe	Model 3: Produktywność 3. nieznormalizowana do prestiżu, pełne zliczanie	Model 4: Produktywność 4. nieznormalizowana do prestiżu, zliczanie cząstkowe
	Doktorzy: decyl 10. do decyl 1. (%)	Doktorzy: decyl 10. do decyl 1. (%)	Doktorzy: decyl 10. do decyl 1. (%)	Doktorzy: decyl 10. do decyl 1. (%)
ENGI	2,1	2,1	2,1	1
LIFE	1,1	0	1,1	2,2
MATH	0	0	0	0
MED	0	0	1,6	0
NATURAL	1,6	2,3	0,8	0,8
Łącznie	1,2	1,2	1,2	1

Diagnostyka współliniowości wektora zmiennych niezależnych w modelach regresji

Zaprezentowane wartości obrazują stopień korelacji danej zmiennej z pozostałymi zmiennymi. Zmienne, które cechują się istotnie wyższymi wartościami w porównaniu

z innymi, uznaje się za istotnie skorelowane. W naszym przypadku do zmiennych o względnie wysokim stopniu korelacji wielowymiarowej należą: wiek biologiczny oraz wiek uzyskania habilitacji. Ze względu na kluczowe znaczenie obu zmiennych dla rozwoju kariery akademickiej (oraz dla niniejszej analizy) i fakt, że poziom tej korelacji nie jest bardzo wysoki, pozostają one w modelu. W poniższej części omówiono rozkład statystyk reszt.

Tabela uzupełniająca 4. Odwrócona macierz korelacji, główna przekątna (modele dotyczące najwyższych klas produktywności)

Zmienna	Produktywność 1. znormalizowana do prestiżu, pełne zliczanie	Produktywność 2. znormalizowana do prestiżu, zliczanie cząstkowe	Produktywność 3. nieznormalizowana do prestiżu, pełne zliczanie	Produktywność 4. nieznormalizowana do prestiżu, zliczanie cząstkowe
Mężczyzna	1,046	1,049	1,046	1,05
IDUB: Pozostałe	1,032	1,028	1,028	1,027
Wiek biol.	5,911	5,775	5,85	5,674
Wiek akad.	2,008	2,026	2	2,01
Wiek: prof.tyt.	2,624	2,627	2,625	2,629
Wiek: hab.	5,595	5,535	5,585	5,509
Klasa najb. produkt. doktorów	1,234	1,163	1,203	1,135
Klasa dr. w młodym wieku	1,864	1,866	1,865	1,869
Klasa hab. w młodym wieku	2,513	2,508	2,512	2,502
Klasa szyb- ko uzyska- nych hab.	2,151	2,148	2,147	2,154

Tabela uzupełniająca 5. Statystyki reszt standaryzowanych (modele dotyczące najwyższych klas)

Statystyka	Produktywność 1. Znormalizowana do prestiżu, pełne zliczanie	Produktywność 2. Znormalizowana do prestiżu, zliczanie cząstkowe	Produktywność 3. nieznormalizowana do prestiżu, pełne zliczanie	Produktywność 4. nieznormalizowana do prestiżu zliczanie cząstkowe
Średnia	0	0	0	0
Mediana	-0,19	-22	-0,17	-0,23
Odchyl. st.	1,00	1,00	1,00	1,00

IQR	0,01	0,01	0,01	0,01
Max	18,26	18,66	31,55	14,16
Stat. K-S	0,264	0,234	0,295	0,272
<i>p</i> -wartość	< 0,001	< 0,001	< 0,001	< 0,001

Rozkład statystyk reszt w naszych modelach nie jest normalny (w teście normalności Kołmogorowa-Smirnowa statystyka *D* osiąga wartości 0,268–0,292, przy $p < 0,001$, co oznacza odrzucenie hipotezy zerowej o zgodności rozkładu z rozkładem normalnym). Choć w metodzie największej wiarygodności kluczowa jest asymptotyczna normalność, nie zaś normalność *per se*, ze względu na możliwe naruszenia założeń wariancji stosujemy odporne błędy standardowe (Croux i in., 2004). Kolejny etap analizy rozkładu reszt wykazał, że nie ma obserwacji „wpływowych” (*influential observations*), ponieważ zakres reszt standaryzowanych mieści się w granicach ± 3 odchyłeń standardowych. W konsekwencji wnioski z naszych modeli pozostają ważne.

Zmienne o stosunkowo wysokiej korelacji wielowymiarowej to ponownie wiek biologiczny oraz wiek uzyskania habilitacji. Z uwagi na centralne znaczenie tych zmiennych dla przeprowadzanej analizy, a także fakt, że ich korelacja nie jest bardzo wysoka, pozostają one w modelach. Rozkład reszt w naszych modelach nie jest normalny. Aby skorygować niezgodność z założeniami, stosujemy odporne błędy standardowe, na podstawie których przeprowadzamy testy istotności współczynników (Croux i in., 2004), analogicznie do modeli dotyczących najwyższych klas opisanych powyżej.

Wiek uzyskania awansu i szybkość awansu

Dodatkowo, tylko w modelach regresji logistycznej, zastosowaliśmy dwie zmienne dotyczące wieku uzyskania awansu i szybkości awansu. Dla wszystkich naukowców przyjęliśmy schemat 20/60/20, oparty na indywidualnych ścieżkach biograficznych. Wszystkich naukowców podzieliliśmy na trzy grupy w zależności od wieku, w jakim uzyskali kolejno stopień doktora oraz doktora habilitowanego („młodzi”, „średni”, „starsi”), a także w zależności od liczby lat, jaka upłynęła między uzyskaniem doktoratu i uzyskaniem habilitacji („szybko”, „typowo”, „wolno”) (zob. np. Costas i in., 2010 w kontekście Hiszpanii). Wszyscy naukowcy, w swoich dziedzinach nauki, zostali przydzieleni do odpowiednich klas wieku awansu i szybkości awansu.

Produktywność publikacyjna polskich naukowców w ujęciu podłużnym: przykład doktorów i doktorów habilitowanych

W naszym badaniu podchodzimy do zagadnienia produktywności naukowej w ujęciu longitudinalnym (podłużnym), śledząc kariery naukowców na przestrzeni czasu

(maksymalnie przez 40 lat). Najpierw klasyfikujemy naukowców w ramach decylnych klas produktywności publikacyjnej – od najniższych 10% do najwyższych 10%. Następnie analizujemy wzorce mobilności między tymi klasami na dwóch etapach kariery: przed i po habilitacji (czyli tylko z doktoratem i tylko z habilitacją). Nasze wyniki potwierdzają, że radykalne zmiany poziomu produktywności publikacyjnej (zarówno wzrost, jak i spadek) w zasadzie w polskim systemie nauki nie występują. Podobnie jak w badanych gdzie indziej zbiorczo państwach wysoko rozwiniętych (zob. Kwiek i Szymula 2025). Naukowcy z bardzo słabym dotychczasowym dorobkiem publikacyjnym mają znikome szanse na osiągnięcie bardzo wysokiego poziomu produktywności w przyszłości, niezależnie od dziedziny nauki (w ramach dziedzin STEMM: nauki ścisłe, techniczne, inżynieryjne, matematyczne i medyczne). Nasze badania wskazują na długoterminowy charakter kariery naukowej: produktywność publikacyjna na etapie przed habilitacją silnie wpływa na produktywność w bardziej niezależnym okresie pracy naukowej na etapie po habilitacji. Wykorzystujemy mikrodane dotyczące karier naukowych (pochodzące z baz danych OPI PIB) oraz metadane dotyczące publikacji (z surowej bazy danych Scopus). Polscy doktorzy habilitowani często pozostają w przypisanych im klasach produktywności przez wiele lat: naukowcy o wysokiej produktywności publikacyjnej utrzymują swój status, podobnie jak naukowcy o niskiej produktywności. Wyniki analizy regresji logistycznej silnie wspierają nasze dwuwymiarowe ustalenia. Badanie obejmuje wszystkich polskich doktorów habilitowanych widocznych na arenie międzynarodowej – w postaci artykułów naukowych indeksowanych w bazie Scopus w pięciu obszarach STEMM ($N = 4165$; $N_{\text{art}} = 71\,841$ artykułów).

Słowa kluczowe: polscy naukowcy, kariera naukowa, badanie longitudinalne, podejście klasyfikacyjne, kariera akademicka, mikrodane, normalizacja do prestiżu czasopism, mobilność, regresja logistyczna

Publication productivity of Polish scientists in a longitudinal perspective: doctoral and postdoctoral career stages

In our study, we approach scientific productivity in a longitudinal perspective, tracking the careers of scientists over time (for up to 40 years). First, we classify scientists into decile-based classes of publication productivity – from the lowest 10% to the highest 10%. We then analyze mobility patterns between these classes at two stages of academic career: before and after habilitation degree (i.e., with only a PhD and with only habilitation degree). Our results confirm that radical changes in the level of publication productivity (both increases and decreases) do not occur in the Polish science system, similarly to other highly developed countries. Researchers with a very weak publication record to date have little chance of achieving very high productivity in the future, regardless of their field of science (within STEMM fields: science, technology, engineering, mathematics and medicine). Our research indicates a long-term nature of academic careers: publication productivity at the pre-habilitation stage strongly influences productivity in the more independent period of academic work at the post-habilitation stage. We use microdata on academic careers (from the OPI PIB database) and metadata on publications (from the raw Scopus database). Polish scientists with

habilitation degree often remain in their assigned productivity classes for many years: researchers with very high publication productivity maintain their status, as do researchers with very low productivity. The results of logistic regression analysis strongly support our two-dimensional findings. The study covers all Polish scientists with habilitation degree visible on the international stage through scientific articles indexed in the Scopus database in five STEM fields ($N = 4165$; $N_{\text{art}} = 71,841$ articles).

Key words: Polish scientists, scientific career, longitudinal study, classificatory approach, academic career, microdata, journal-prestige normalization, mobility, logistic regression

