

MAREK KWIEK*, WOJCIECH ROSZKA**

Zawsze wysoce produktywni? Zawsze nisko produktywni?

Dynamiczne ujęcie dorobku naukowego profesorów tytułarnych wykorzystujące klasy produktywności

1. Wstęp

W prezentowanym studium badamy trwałość przynależności do klas produktywności badawczej na poziomie indywidualnym w ciągu całej kariery akademickiej. Analizujemy przebieg kariery 2326 polskich profesorów tytułarnych, uwzględniając ich biografie naukowe i historie publikacyjne. Badamy daty kolejnych awansów naukowych i liczbę publikacji (79 027 artykułów) pomiędzy awansami w ciągu 20–40 lat pracy naukowej w 14 dyscyplinach (nauki ścisłe, techniczne, inżynieryjne, matematyczne i medyczne, czyli w obszarze STEMM). Interesuje nas przemieszczanie się między trzema klasami produktywności – najwyższą, przeciętną i najniższą – w trakcie kariery profesorów od etapu doktoratu do etapu profesury tytułarnej.

Zastosowaliśmy tutaj unikalne podejście do produktywności: produktywność znormalizowaną do prestiżu czasopisma, w ramach której większą wagę przypisuje się artykułom publikowanym w czasopismach o dużym wpływie na rozwój nauki niż w czasopismach o niskim wpływie, uznając tym samym wysoki stopień stratyfikacji nauki akademickiej pod kątem miejsca publikacji (zasada „publikacja nierówna publikacji”).

Nasze wyniki pokazują, że połowa najbardziej produktywnych doktorów kontynuowała pracę jako najbardziej produktywni doktorzy habilitowani, a z kolei połowa najbardziej produktywnych doktorów habilitowanych kontynuowała pracę jako najbardziej produktywni profesorowie tytułarni (52,6% i 50,8%). Przechodzenie naukowców od najwyższej do najniższej i od najniższej do najwyższej klasy produktywności występowało w niewielkim stopniu: dotyczyło tylko 100 (4,3%) naukowców. W modelach regresji logistycznej dwoma silnymi czynnikami warunkującymi przynależność do najwyższej klasy produktywności wśród profesorów tytułarnych okazały się wysoka produktywność

*Prof. dr hab. Marek Kwiek (kwiekm@amu.edu.pl), Institute for Advanced Studies in Social Sciences and Humanities (IAS) UAM 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, ORCID: orcid.org/0000-0003-4383-3259

w okresie bycia doktorem i wysoka produktywność w okresie bycia doktorem habilitowanym (zwiększając szanse średnio o 179% i 361%). Ani płeć, ani wiek (biologiczny lub akademicki) nie okazały się statystycznie istotne. Nasze wyniki mają znaczące konsekwencje dla polityki zatrudnieniowej i awansowej: zatrudnianie naukowców o wysokiej i niskiej produktywności może mieć długotrwałe konsekwencje dla instytucji i krajowego systemu nauki, ponieważ naukowcy akademicy zwykle pozostają w systemie przez dziesiątki lat. Do analiz wykorzystaliśmy dane pochodzące z prowadzonego przez nas Obserwatorium Polskiej Nauki (100 000 naukowców, 380 000 publikacji z ostatniej dekady) oraz metadane z bazy Scopus dotyczące niemal miliona (935 167) polskich artykułów z ostatnich 50 lat, uzyskane w ramach umowy o współpracy z laboratorium ICSR Lab prowadzonym przez firmę Elsevier.

W pracy wysunęliśmy hipotezę, że aktualne usytuowanie profesorów tytularnych w najwyższej, środkowej i najniższej klasie produktywności (tj. w grupie 20%, 60% i 20% naukowców w ujęciu produktywności znormalizowanej do prestiżu czasopisma w każdej dyscyplinie) odpowiada w pewnym, nieznanym nam stopniu ich usytuowaniu w klasach produktywności na wcześniejszych etapach kariery naukowej. Spodziewaliśmy się, że obecni wysoce produktywni profesorowie mogli również być wysoce produktywnymi doktorami i wysoce produktywnymi doktorami habilitowanymi.

Punktem wyjścia naszych badań był aktualny rozkład profesorów tytularnych według klas produktywności w czteroletnim okresie 2014–2017. Zostali oni sklasyfikowani jako wysoce produktywni, przeciętnie produktywni lub nisko produktywni. Następnie przeanalizowaliśmy klasy produktywności, do których można ich było retrospektywnie przypisać na wcześniejszych etapach kariery, a więc w poprzednich 20–40 latach.

Motyw przewodni artykułu jest zgodny z głównymi ustaleniami badań nad wysoce produktywnymi naukowcami i ich cechami (np. Fox i Nikivincze 2021; Yin i Zhi 2017; Agrawal et al. 2017; Cortés et al. 2016; Abramo et al. 2017; Kwiek 2016). Nasze badania dotyczą trzech równoległych pytań: (1) w jakim stopniu produktywność badawcza naukowców zmienia się w ciągu ich całego życia akademickiego? (2) Czy obecnie wysoce produktywni naukowcy zawsze byli wysoce produktywni, a obecnie nisko produktywni naukowcy – zawsze byli nisko produktywni? (3) Czy często zdarzają się radykalne zmiany klas produktywności (mobilność w górę lub w dół między skrajnymi klasami) w trakcie kariery akademickiej? Większość badań nad produktywnością skupia się na indywidualnych cechach wysoce produktywnych naukowców, a niektóre łączą cechy indywidualne i organizacyjne (środowiskowe) (Fox i Nikivincze 2021; Fox i Mohapatra 2007). Nasze podejście do analizy produktywności jest zarazem dynamiczne, względne (oparte na klasach) i znormalizowane do prestiżu czasopism:

(1) *Dynamiczne (wzdłużne)*: analizujemy produktywność aktualnych profesorów tytular-

nych przez kilka dekad wstecz (od momentu ich wejścia do systemu szkolnictwa wyższego);

- (2) *Względne*: nie badamy liczby publikacji, ale koncentrujemy się na klasach produktywności, retrospektywnie przypisując poszczególne osoby do klas i porównując naukowców z ich kolegami z danej dyscypliny i na danym etapie kariery naukowej (okres pracy z doktoratem, habilitacją i profesurą: podejście wykorzystujące system stopni i tytułu oraz daty ich uzyskania okazało się lepsze od testowanego podejścia opierającego się na skomplikowanym systemie stanowisk akademickich);
- (3) *Znormalizowane do prestiżu czasopism*: większą wartość przypisuje się artykułom opublikowanym w czasopismach o dużym wpływie na naukę mierzonym średnią liczbą cytowań (system rang percentylowych czasopism używany w bazie Scopus) niż w czasopismach o niskim wpływie.

Jednostką analizy jest dla nas pojedynczy naukowiec, a nie pojedyncza publikacja. Chociaż korzystaliśmy z danych administracyjnych, biograficznych i bibliometrycznych, nasze studium nie ma charakteru bibliometrycznego i należy do obszaru badań profesji akademickiej. Nie udałooby się przeprowadzić retrospektywnych analiz całego życia akademickiego poszczególnych naukowców bez pełnego dostępu do surowych metadanych bibliometrycznych wszystkich publikacji wszystkich polskich naukowców w ciągu ostatnich 50 lat. Nie byłoby możliwe skonstruowanie retrospektywnych klas produktywności dla wszystkich naukowców według dyscypliny, etapu kariery i wybranych okresów między awansami bez dostępu do globalnych metadanych publikacji każdego naukowca, czyli bez możliwości wykorzystania ustrukturyzowanych Big Data z bazy Scopus, komercyjnej bazy danych bibliometrycznych. Nasze badanie stanowi przykład połączenia ustrukturyzowanych Big Data i danych pochodzących z krajowych rejestrów naukowców w celu przeprowadzenia szczegółowych analiz karier akademickich.

2. Podstawy teoretyczne

2.1. Wysoka produktywność badawcza

Od co najmniej pół wieku socjologia nauki i socjologia karier akademickich zajmują się zagadnieniem nierówności w akademickiej produkcji wiedzy (Hermanowicz 2012; Kwiek 2019), ponieważ, jak powszechnie wiadomo, niewielki odsetek naukowców „wnosi nieproporcjonalnie duży wkład do rozwoju nauki i otrzymuje nieproporcjonalnie dużą część nagród i zasobów potrzebnych do prowadzenia badań” (Zuckerman 1988: 526). Jak pokazaliśmy, 10% naukowców odpowiada za połowę wszystkich publikacji w 11 krajach europejskich (Kwiek 2016). W Mertonowskiej tradycji socjologii nauki istotne jest „pierwszeństwo odkrycia” (Merton 1973: 293), ponieważ jedną z bardziej znamiennych motywacji naukowców jest „pragnienie zdobycia uznania wśród innych naukowców”

(Cole i Cole 1973: 10) w oparciu o dokonywane odkrycia. Społeczność naukowa nie jest więc „towarzystwem równych sobie”, a uznanie dla pracy naukowców jest „jedynym jednoznacznym dowodem na to, że to, co robili, ma znaczenie dla nauki” (Zuckerman 1988: 526). Uznanie, jakie dają publikacje i ich cytowania, przekłada się na finansowanie dalszych badań, a rozkład osiągnięć, cytowań, nagród i finansowania badań nie jest sprawiedliwy – jest silnie rozwarstwiony (Kwiek 2022).

W każdym systemie nauki akademickiej niewielka liczba naukowców publikuje większość prac i przyciąga większość cytowań (Ruiz-Castillo i Costas 2014; Stephan 2012; Abramo et al. 2009). W każdej społeczności naukowej wysoce produktywni naukowcy zajmują prestiżowe stanowiska akademickie i są odpowiedzialni za kształtowanie tożsamości dyscyplin naukowych (Cortés et al. 2016). Produktowność badawcza wynika z (1) cech indywidualnych, (2) cech organizacyjnych (środowisko akademickie) oraz (3) cech krajowego systemu nauki akademickiej, w którym ważną rolę odgrywa przyznawanie nagród i uznanie za osiągnięcia naukowe. Nauka jest skomplikowaną instytucją społeczną, a naukowcy muszą być wspierani systemowo w ramach krajowego systemu nauki, aby utrzymywać wysoką produktywność przez dłuższy czas. Efektywne funkcjonowanie nauki zależy od tego, w jaki sposób „dzieli się nagrody i wyróżnienia za wybitne osiągnięcia oraz stwarza możliwości tym, którzy odznaczają się wyjątkowym talentem” (Cole i Cole 1973: 15).

Dostęp do zasobów niezbędnych do prowadzenia badań mają osoby cieszące się dużym szacunkiem w społeczności naukowej, które dużo publikują i są silnie zmotywowane do publikowania, ponieważ szacunek w nauce „płynie ku tym, którzy są wysoce produktywni” (Allison i Stewart 1974: 604). Wysoce produktywni naukowcy to ci, których wysoka produktywność utrzymuje się w czasie (Abramo et al. 2017); to niewielka grupa, która utrzymuje wysoką produktywność w swojej pracy, wspierana lub nie przez strukturalne cechy systemu nauki, między innymi przez mechanizmy kumulacji przewag w czasie. Kumulacja przewag to szerszy proces, w którym „niewielkie różnice początkowe kumulują się, prowadząc do dużych różnic” (Aguinis i O'Boyle 2014: 5). W nauce kumulacja przewag to tak zwany efekt Mateusza, który prowadzi do nierówności w dostępie do nagród finansowych i niefinansowych (Xie 2014): ci, którzy mają dużo, będą mieli więcej, a ci, którzy mają mało, będą mieli jeszcze mniej, zgodnie z biblijnym przesłaniem.

Z perspektywy historycznej socjologia nauki pokazuje, że uznanie w nauce jest zakorzenione prawie wyłącznie w badaniach naukowych (Cole i Cole 1967), a system nagród jest skonstruowany w taki sposób, aby przynosił korzyści naukowcom, którzy najlepiej wykonują swoją pracę naukową. Według Mertona (1973: 297) „instytucja nauki opracowała skomplikowany system przyznawania nagród tym, którzy w różnym stopniu spełniają jej normy”. W opartym na renomie i zasobach modelu kariery naukowej Mertona

nowe zasoby nie są prostą nagrodą za wysoką produktywność w przeszłości, ale pełnią podstawową funkcję stymulowania wysokiej produktywności w przyszłości. W ostatniej dekadzie intensywnie analizowano wysoką produktywność badawczą (np. Yair et al. 2017; Aguinis i O'Boyle 2014; Agrawal et al. 2017; Abramo et al. 2017; Yin i Zhi 2017; Piro et al. 2016; Kwiek 2016; Kwiek 2018). Ostatnio Fox i Nikivincze (2021) badały płodnych naukowców z perspektywy społeczno-organizacyjnej, analizując zarówno cechy indywidualne, jak i cechy poszczególnych wydziałów. Zidentyfikowały one trzy predyktory wysokiej produktywności: stanowisko, zakres współpracy i korzystny klimat w pracy (postrzegana atmosfera panująca na wydziale, która stymuluje lub hamuje produktywność) (Fox i Mohapatra 2007). Abramo, D'Angelo i Soldatenkova (2017), których badania są najbardziej zbliżone do naszych, przeanalizowali osiągnięcia badawcze włoskich profesorów z dziedziny nauk ścisłych w trzech kolejnych czteroletnich okresach (2001–2012). Ich analizy wykazały, że 35% naukowców zachowuje wysoką produktywność przez trzy kolejne okresy, a 55% przez dwa okresy. Wyższy odsetek mężczyzn niż kobiet utrzymuje swoją pozycję, przy czym istnieją różnice między dyscyplinami (Abramo et al. 2017: 793–794). Nasze badania różnią się od powyższych pod względem zakresu czasu (całe życie vs. 12 lat), doboru próby (profesorowie tytularni vs. wszyscy naukowcy akademicki) oraz metodologii (trzy klasy produktywności vs. naukowcy o największej produktywności i naukowcy nieproduktywni; produktywność znormalizowana do prestiżu czasopism vs. definiowana przez autorów „frakcjonowana siła naukowa”).

W socjologii i ekonomii nauki pojawiło się kilka fundamentalnych teorii wyjaśniających drastyczne różnice dotyczące indywidualnej produktywności badawczej, które mogą być przydatne w badaniu stratyfikacji polskich naukowców. Teoria „iskry bożej” (Cole i Cole 1973) mówi, że „istnieją znaczące, z góry określone różnice między naukowcami dotyczące ich zdolności i motywacji do prowadzenia twórczych badań naukowych” (Allison i Stewart 1974: 596). Wysoce produktywni naukowcy „są zmotywowani wewnętrznym dążeniem do uprawiania nauki i czystą miłością do tej pracy” (Cole i Cole 1973: 62). Produktywni naukowcy są silnie zmotywowaną grupą badaczy i posiadają niezbędną „zdolność do ciężkiej pracy i upór w dążeniu do realizacji długoterminowych celów” (Fox 1983: 287). Stephan i Levin (1992: 13) są podobnego zdania, twierdząc, że „istnieje powszechna zgoda co do tego, że pewni ludzie są szczególnie dobrzy w uprawianiu nauki, a niektórzy są nie tyle dobrzy, co wręcz znakomici”. Teoria kumulacji przewag (Merton 1973) zakłada, że produktywni naukowcy stają jeszcze bardziej produktywni w przyszłości, podczas gdy naukowcy o niskiej produktywności z czasem stają się jeszcze mniej produktywni. „Nagradzani naukowcy są produktywni, natomiast naukowcy, którzy nie są nagradzani, stają się mniej produktywni” (Cole i Cole 1973: 114). I wreszcie teoria maksymalizacji użyteczności, która wyrosła z ekonomii nauki,

głosi, że naukowcy z czasem zmniejszają swoje wysiłki badawcze, ponieważ uważają, że inne zadania mogą być dla nich bardziej korzystne. Omawiając kwestię starzenia się i produktywności, Stephan i Levin (1992: 35) twierdzą, że „na późniejszym etapie kariery naukowcy mają mniejszą motywację finansową do prowadzenia badań” (zob. Kyvik 1990). Te trzy główne teorie produktywności badawczej uzupełniają się wzajemnie i w różnym stopniu odnoszą się do profesji akademickiej w Polsce (Kwiek 2019: 27–32). Teorie iskry bożej i kumulacji przewag pozwalają wyjaśnić wysoką produktywność badawczą, podczas gdy niska produktywność w Polsce może być interpretowana za pomocą teorii kumulacji przewag (i kumulacji strat) oraz teorii maksymalizacji użyteczności.

Wejście do klasy najbardziej produktywnych naukowców wymaga silnej orientacji na badania i długich godzin pracy poświęcanych na badania (zob. Kwiek 2016 i Kwiek 2018), oprócz wrodzonych zdolności podkreślanych przez teorię iskry bożej i wcześniejszych osiągnięć podkreślanych przez teorię kumulacji przewag. Duża część najbardziej produktywnych naukowców zawsze będzie należała do grona najbardziej produktywnych – niezależnie od okoliczności, miejsca zajmowanego w systemie, wieku i etapu kariery – podczas gdy jedynie marginalna część naukowców o niskiej produktywności kiedykolwiek stanie się wysoce produktywna, jak pokazujemy w tym studium. W procesie kumulowania przewag wyjątkowa produktywność badawcza na początku kariery przekłada się na nowe zasoby i nagrody, które ułatwiają utrzymanie wysokiej produktywności badawczej w kolejnych latach i dekadach. Zasoby przeznaczone na badania naukowe nie stanowią nagród za przeszłą produktywność, ale mają za zadanie stymulowanie produktywności najbardziej produktywnych w przyszłości: „Społeczność naukowa faworyzuje tych, którzy osiągnęli najwięcej w przeszłości, ze względu na dodatkowe zasoby i uwagę, jaką im poświęcono” (DiPrete i Eirich 2006: 281–282).

2.2. Pytania badawcze i hipotezy

Nasze hipotezy (zob. tabela 1) dotyczą trwałości wysokiej (H1) i niskiej (H2) klasy produktywności w czasie; trwałości wysokiej klasy produktywności na początku i pod koniec kariery akademickiej (H4); zróżnicowania pod względem dyscyplin (H3) i płci (H5) w mobilności między klasami produktywności; oraz (H6) roli przynależności do klas produktywności w przeszłości w szacowaniu (za pomocą analizy regresji logistycznej) ilorazu szans aktualnej przynależności do najwyższej klasy produktywności. Nadrzędne pytanie badawcze dotyczy zmian klas produktywności z perspektywy całego życia naukowego: czy profesorowie tytularni osiągający obecnie najwyższą produktywność zawsze, na przestrzeni swojej kariery naukowej, osiągalili najwyższą produktywność, a profesorowie tytularni osiągający niską produktywność – zawsze osiągalili niską produktywność?

Tabela 1. Pytania badawcze, hipotezy i podsumowanie wyników

Pytania badawcze	Hipotezy	Wyniki
Pytania badawcze 1. Jaki jest związek między obecną wysoką produktywnością a wysoką produktywnością na dwóch wcześniejszych etapach kariery akademickiej?	Trwałość wysokiej produktywności w czasie Hipoteza 1: Obecnie wysoce produktywni profesorowie tytularni byli w znacznej części wysoce produktywnymi doktorami, a wysoce produktywni doktorzy habilitowani byli w znacznej części wysoce produktywnymi doktorami.	Potwierdzona
Pytania badawcze 2. Jaki jest związek między obecną niską produktywnością a niską produktywnością na dwóch wcześniejszych etapach kariery akademickiej?	Trwałość niskiej produktywności w czasie Hipoteza 2: Obecnie nisko produktywni profesorowie tytularni byli w znacznej części nisko produktywnymi doktorami habilitowanymi, a nisko produktywni doktorzy habilitowani byli w znacznej części nisko produktywnymi doktorami.	Potwierdzona
Pytania badawcze 3. Jaki jest związek między zmianami produktywności w ramach cyklu życia a dyscyplinami akademickimi?	Zróżnicowanie dyscyplinarne Hipoteza 3: Mobilność między klasami produktywności różni się w zależności od dyscypliny.	Potwierdzona
Pytania badawcze 4. Jaki jest związek między obecną produktywnością a produktywnością na początku kariery akademickiej?	Trwałość produktywności w trakcie kariery akademickiej Hipoteza 4: Obecni profesorowie tytularni należą w znacznej części do tej samej klasy produktywności na początku i na końcu swojej kariery akademickiej.	Potwierdzona
Pytania badawcze 5. Jaki jest związek między zmianami produktywności w ramach cyklu życia a płcią?	Zróżnicowanie ze względu na płeć Hipoteza 5: Mobilność między klasami produktywności różni się w zależności od płci.	Potwierdzona
Pytania badawcze 6. Jaki jest wpływ wcześniejszej produktywności na obecną przynależność do klasy najwyższej produktywności w oparciu o łączny wpływ innych zmiennych?	Podjęcie modelowe do obecnej klasy najwyższej produktywności, analiza regresji logistycznej. Hipoteza 6: Przynależność w przeszłości do klasy najwyższej produktywności (wysoce produktywni doktorzy i wysoce produktywni doktorzy habilitowani) znacząco zwiększa iloraz szans oszacowany dla przynależności do obecnej klasy najwyższej produktywności w oparciu o łączny wpływ innych zmiennych.	Potwierdzona

3. Dane, próba i założenia metodologiczne

3.1. Dane i próba

Dane wykorzystane w tym studium pochodzą z utworzonej i utrzymywanej przez nas bazy danych Obserwatorium Polskiej Nauki (zob. konstrukcję bazy w: Kwiek i Roszka 2021a: 4–6), składającej się z rejestru administracyjnego i biograficznego wszystkich polskich naukowców ($N = 99\,935$, *Nauka Polska*) oraz z bibliometrycznej bazy danych Scopus (2009–2018, $N = 380\,000$ publikacji). W wyniku procesu integracji baz danych, stwierdzono połączenia dla 25 463 unikalnych autorów z polskimi afiliacjami, którzy napisali 158 743 artykułów w czasopismach. Baza danych Obserwatorium została następnie wzbogacona o metadane publikacji zebrane w bazie Scopus, które uzyskaliśmy dzięki umowie o współpracy z laboratorium ICSR Lab, które jest platformą chmurową udostępnianą do celów badawczych przez firmę Elsevier ($N = 935\,167$ artykułów z lat 1973–2021 autorstwa naukowców z polską afiliacją). Wykorzystaliśmy informacje o całym dorobku naukowym poszczególnych naukowców na podstawie ich identyfikatorów Scopus Author ID. Nasza ostateczna próba obejmowała wyłącznie profesorów tytularnych pracujących w 14 dyscyplinach STEM ($N = 2326$), będących autorami 79 027 artykułów.

3.2. Definiowanie dyscyplin akademickich i wieku akademickiego

Zdefiniowaliśmy indywidualne właściwości 23 543 naukowców co najmniej z doktoratem zatrudnionych na pełnym etacie w szkolnictwie wyższym na wszystkich stanowiskach akademickich we wszystkich dyscyplinach oraz właściwości wszystkich profesorów tytularnych w 14 dyscyplinach STEM w naszej ostatecznej próbie. W systemie klasyfikacji dyscyplin All Science Journal Classification (ASJC), stosowanym w bazie Scopus, publikacja w czasopiśmie może mieć jedną lub kilka klasyfikacji dyscyplinarnych. Dominująca dyscyplina każdego profesora tytularnego została określona na podstawie wszystkich publikacji (typ: artykuł naukowy) zawartych w jego lub jej indywidualnym portfolio publikacyjnym z lat 2009–2018 (wybraliśmy wartość modalną, czyli występującą najczęściej). Jeśli nie występowała jedna wartość modalna, dyscyplina dominująca została losowo wybrana spośród najczęściej występujących. Przetestowane przez nas podejście oparte na zmieniających się w czasie polskich klasyfikacjach dyscyplin przypisywanych przy okazji otrzymywania kolejnych stopni i tytułów naukowych wymagałoby zbyt dużych przybliżeń.

Nasza baza danych zawierała rok urodzenia profesorów oraz lata, w których uzyskiwali doktorat, habilitację i profesurę. Rok pierwszej publikacji indeksowanej w bazie Scopus uzyskaliśmy za pomocą protokołu API (*application programming interface*), czyli zestawu kodów programistycznych umożliwiających przesyłanie danych między różnymi oprogramowaniami, udostępnionego w ramach bazy Scopus. Płec wszystkich

naukowców ze stopniem co najmniej doktora jest zawarta w danych pochodzących z krajowego rejestru naukowców (*Nauka Polska*, OPI PIB) i w tym badaniu została potraktowana jako zmienna binarna.

3.3. Profesorowie tytularni: rozkład

Rozkłady wybranych charakterystyk w naszej próbie były następujące: trzy czwarte profesorów tytularnych to mężczyźni (tabela 2); jedna trzecia pracowała w 10 instytucjach funkcjonujących w programie IDUB (użytym w badaniu jako zamiennik instytucji o dużej intensywności prowadzenia badań naukowych); dwie trzecie profesorów było w wieku powyżej 60 lat, a połowa w wieku 65–70 lat. W naszej próbie 16% z nich stanowili młodzi (poniżej 55 lat) profesorowie, w tym 2% znajdowało się w wieku 40–44 lat. Rozkład profesorów tytularnych w naszej próbie według płci był zbliżony do rozkładu według płci w ich populacji na polskich uczelniach w ciągu ostatnich pięciu lat (GUS 2022, tablice elektroniczne).

3.4. Podejście metodologiczne

3.4.1. Konstruowanie historii biograficznych i historii publikacyjnych w ciągu całego życia

Baza danych Laboratorium Polskiej Nauki stworzona na potrzeby naszych badań zawiera pełne historie publikacyjne wszystkich polskich naukowców pracujących w sektorze szkolnictwa wyższego w listopadzie 2017 roku, posiadających co najmniej stopień doktora i co najmniej jedną publikację w bazie Scopus. Baza danych zawiera metadane dotyczące wszystkich publikacji każdego naukowca na każdym etapie jego kariery naukowej. Baza zawiera dane dotyczące 14 271 doktorów, 7418 doktorów habilitowanych i 3774 profesorów tytularnych w dyscyplinach STEMM i spoza STEMM.

Skoncentrowaliśmy się na podpróbie profesorów tytularnych, co pozwoliło nam na prześledzenie ich indywidualnych historii biograficznych i indywidualnych historii publikacyjnych na wcześniejszych etapach kariery (tylko profesorowie tytularni mogli być porównywani na trzech wcześniejszych etapach). Analogiczną analizę przeprowadziliśmy również na podpróbie wszystkich obecnych doktorów habilitowanych, ale w tym przypadku ich dorobek naukowy był porównywany tylko na dwóch wcześniejszych etapach (a wyników nie analizujemy tutaj z powodu ograniczonego miejsca). Analiza profesorów tytularnych obejmowała długi okres działalności naukowej, trwający kilka dekad: retrospektywnie przeanalizowaliśmy przebieg kariery akademickiej profesorów, którzy pracowali przez 20–40 lat. Zebranie pełnych historii biograficznych (tj. roku urodzenia i lat kolejnych awansów akademickich) oraz pełnych historii publikacyjnych (tj. szczegółowych danych o publikacjach, współpracy, mobilności i cytowaniach), obejmujących całe kariery akademickie, pozwoliło nam retrospektywnie przeanalizować przejścia między klasami produktywności na poszczególnych etapach, czyli w czasie.

Tabela 2. Struktura próby wszystkich polskich profesorów tytularnych z podziałem na płeć, grupę wiekową i dyscyplinę STEM

		Kobiety			Mężczyźni			Razem		
		<i>n</i>	% wier.	% kol.	<i>n</i>	% wier.	% kol.	<i>n</i>	% wier.	% kol.
Grupy wiekowe	Razem	551	23,7	100,0	1775	76,3	100,0	2326	100,0	100,0
	do 50 lat	48	24,9	8,7	145	75,1	8,2	193	100,0	8,3
	51–60	164	27,2	29,8	438	72,8	24,7	602	100,0	25,9
	61–65	145	22,3	26,3	505	77,7	28,5	650	100,0	27,9
	65–70	194	22,0	35,2	687	78,0	38,7	881	100,0	37,9
IDUB	IDUB	130	16,7	23,6	650	83,3	36,6	780	100,0	33,5
	Pozostałe	421	27,2	76,4	1125	72,8	63,4	1546	100,0	66,5
Dyscypliny akademickie	AGRI	119	33,9	21,6	232	66,1	13,1	351	100,0	15,1
	BIO	66	37,9	12,0	108	62,1	6,1	174	100,0	7,5
	CHEM	41	25,2	7,4	122	74,8	6,9	163	100,0	7,0
	CHEMENG	9	21,4	1,6	33	78,6	1,9	42	100,0	1,8
	COMP	14	14,4	2,5	83	85,6	4,7	97	100,0	4,2
	EARTH	13	11,3	2,4	102	88,7	5,7	115	100,0	4,9
	ENER	6	19,4	1,1	25	80,6	1,4	31	100,0	1,3
	ENG	18	5,8	3,3	292	94,2	16,5	310	100,0	13,3
	ENVIR	57	35,6	10,3	103	64,4	5,8	160	100,0	6,9
	MATER	37	23,1	6,7	123	76,9	6,9	160	100,0	6,9
	MATH	9	6,3	1,6	133	93,7	7,5	142	100,0	6,1
	MED	138	36,4	25,0	241	63,6	13,6	379	100,0	16,3
	PHARM	14	66,7	2,5	7	33,3	0,4	21	100,0	0,9
	PHYS	10	5,5	1,8	171	94,5	9,6	181	100,0	7,8

Uwaga: Dyscypliny STEM uwzględnione w badaniu: AGRI, nauki rolnicze i biologiczne; BIO, biochemia, genetyka i biologia molekularna; CHEMENG, inżynieria chemiczna; CHEM, chemia; COMP, informatyka; EARTH, nauki o Ziemi i planetach; ENER, energia; ENG, inżynieria; ENVIR, nauka o środowisku; MATER, materiałoznawstwo; MATH, matematyka; MED, medycyna; PHARM, farmakologia, toksykologia i farmacja; oraz PHYS, fizyka i astronomia.

Do analizy przechodzenia między klasami produktywności profesorów tytularnych w trakcie ich kariery – od roku, w którym uzyskali tytuł doktora do roku 2017 – zastosowaliśmy podejście wzdlużne (longitudinalne). Analizowaliśmy produktywność poszczególnych naukowców w miarę ich przechodzenia na wyższy szczebel kariery akademickiej. Każdy publikujący naukowiec w ramach swojej unikalnej historii biograficznej (opartej na danych) i unikalnej historii publikacyjnej (opartej na metadanych publikacji) został przypisany do klas produktywności w porównaniu z kolegami z tej samej dyscypliny i na tym samym etapie kariery.

3.4.2. Konstruowanie produktywności badawczej znormalizowanej do prestiżu czasopisma

Produktywność badawcza naukowca na danym etapie kariery akademickiej została określona jako liczba wszystkich publikacji (typ publikacji: artykuł naukowy) wydanych na tym etapie, podzielona przez liczbę lat spędzonych na tym etapie. Takie podejście redukuje potencjalne różnice między pierwszymi latami po każdym awansie, kiedy produktywność może spadać, a latami tuż przed kolejnym awansem, kiedy produktywność może rosnąć. Podobnie jak w innych krajach, również w Polsce produktywność niektórych naukowców może się zmieniać w trakcie ich kariery, z okresami szczytowymi przed awansem i przerwami w publikowaniu po awansie (Katz 1973). Podzieliliśmy karierę naukową profesorów w naszej próbie na trzy etapy ze względu na daty ich rozpoczęcia i zakończenia (doktorat, habilitacja, profesura) i skonstruowaliśmy zarówno profile produktywności w całym okresie życia, jak i profile produktywności na trzech odrębnych etapach kariery. Zastosowaliśmy metodę pełnego liczenia publikacji zamiast metody liczenia frakcjonowanego: publikacje jedno- i wieloautorskie były liczone jednakowo. Zastosowaliśmy liczbę publikacji znormalizowaną do prestiżu czasopisma, a nie surową liczbę publikacji.

Nasze podejście jest nowatorskie: wykorzystaliśmy produktywność znormalizowaną do prestiżu czasopisma, która łączy produktywność badawczą z wpływem na naukę (bazującym na cytowaniach). Wskaźniki wyników mierzą wytworzoną wiedzę, a wskaźniki wpływu mierzą sposób, w jaki praca naukowa wpływa na społeczność badawczą (Sugimoto i Larivière 2018: 1). Waga artykułu zależy od jego pozycji w globalnej hierarchii czasopism akademickich. W naszym podejściu artykuły opublikowane w czasopismach, które mają średnio duży wpływ na społeczność akademicką, co można uchwycić za pomocą średniej liczby cytowań, miały większą wagę w obliczaniu produktywności niż artykuły w czasopismach o niskim wpływie, ponieważ ich napisanie i opublikowanie wymagało przeciętnie większego wysiłku. Nasze podejście do produktywności – niestosowane dotąd w świecie – uwzględnia niezwykle rozwarstwiony charakter nauki akademickiej, w której ważna jest zarówno liczba publikacji, jak i ich standaryzowana jakość (w naszym przypadku jakość ujmowana przez zamiennik rang percentylowych czasopism w bazie Scopus w zakresie 1–99).

Pomiar prestiżu czasopisma jest ściśle związany z polskim systemem oceny naukowców i jednostek naukowych oraz ze wskaźnikami stosowanymi w ramach programu „Inicjatywa Doskonałości – Uczelnie Badawcze” (IDUB). Artykuły publikowane w czasopismach o wysokim prestiżu wymagają przeciętnie większego nakładu pracy i znajdują przeciętnie większy oddźwięk w świecie nauki, co odzwierciedlają średnio otrzymywane cytowania w badanym czteroletnim okresie. W bazie Scopus prestiżowa ranga czasopisma jest określana co roku na podstawie jego miejsca w systemie CiteScore, który

jest przygotowywany dla wszystkich indeksowanych czasopism (40 562 w 2022 roku). Rangi percentylowe są oparte na wartościach z przedziału 1–99, w którym najwyższy prestiż jest związany z 99. percentylem. Wysoce prestiżowe czasopisma w każdej dziedzinie, o niskim wskaźniku akceptacji przysyłanych maszynopisów, znajdują się zwykle w 90–99. percentylu (w naszej dziedzinie *Higher Education* i *Studies in Higher Education* znajdują się w 96. percentylu, a *Quatitative Science Studies* w 98. percentylu czasopism należących do bazy Scopus). Publikacje w bardziej prestiżowych czasopismach liczą się bardziej w obliczaniu produktywności w porównaniu z publikacjami w mniej prestiżowych czasopismach – w ramach każdej dyscypliny.

Przy standardowym podejściu do produktywności artykuł opublikowany w dowolnym czasopiśmie otrzymałby wartość 1. Natomiast przy zastosowaniu produktywności znormalizowanej do prestiżu czasopisma artykuł opublikowany w czasopiśmie o randze percentylowej 90 otrzymuje wartość 0,90, a artykuł opublikowany w czasopiśmie o randze percentylowej 40 otrzymuje wartość 0,40. Artykuły opublikowane w czasopismach o randze percentylowej równej 10 lub niższej otrzymują wartość 0,1. Podejście znormalizowane do prestiżu czasopisma do indywidualnej produktywności badawczej pozwala na rzetelny pomiar wysiłku naukowego w dyscyplinach STEMM, w których pionowa stratyfikacja czasopism jest uznawana za oczywistą. Traktowanie wszystkich publikacji w ten sam sposób, niezależnie od miejsca publikacji, w ramach obliczania produktywności nie uwzględniałoby zróżnicowanego indywidualnego wysiłku naukowego włożonego w badania. Każda dyscyplina ma swoje specyficzne, wysoce konkurencyjne czasopisma z najwyższej półki, a „tyrania pierwszej piątki” wśród czasopism (Heckman & Moktan 2018) ma zastosowanie nie tylko w ekonomii.

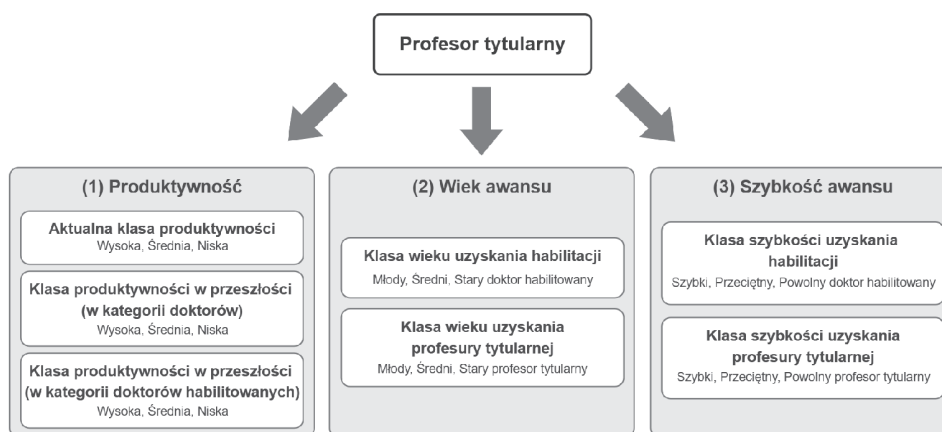
3.4.3. Konstruowanie klas kariery akademickiej: klasy produktywności, wieku awansu i szybkości awansu

W naszych badaniach zastosowaliśmy koncepcję wspinania się po drabinie akademickiej, która definiuje karierę profesorów tytularnych rozciągającą się na przestrzeni kilkudziesięciu lat. Obecni profesorowie (data awansu: przyznanie tytułu naukowego) byli wcześniej najpierw doktorami (data awansu: przyznanie stopnia doktora), a następnie doktorami habilitowanymi (data awansu: przyznanie stopnia doktora habilitowanego). Wszyscy oni pozostawali przez określoną liczbę lat na kolejnych etapach kariery akademickiej. Na każdym etapie wykazywali się określoną produktywnością związaną z liczbą publikacji i okresem czasu.

Uszeregowaliśmy wszystkich naukowców, oddzielnie w ramach poszczególnych dyscyplin, według ich czteroletniej produktywności badawczej znormalizowanej do prestiżu czasopisma na poszczególnych etapach kariery. Dla każdego profesora tytularnego policzyliśmy wszystkie artykuły opublikowane w ramach etapów określonych przez daty awansu: pierwszy etap to okres między uzyskaniem stopnia doktora a uzyskaniem stop-

nia doktora habilitowanego, drugi etap to okres między uzyskaniem stopnia doktora habilitowanego a uzyskaniem tytułu profesora, a trzeci etap to okres między uzyskaniem tytułu profesora a rokiem 2017. Na przykład, jeśli biografia profesora X wskazuje, że uzyskał on stopień doktora w 1995 roku, stopień doktora habilitowanego w 2002 roku, a tytuł profesora w 2012 roku, to w jego przypadku etap pierwszy przypada na lata 1995–2001, drugi na lata 2002–2011, a trzeci na lata 2012–2017.

Każdy profesor tytularny został przypisany do siedmiu klas kariery akademickiej (ryc. 1): trzech klas produktywności, dwóch klas wieku awansu i dwóch klas szybkości awansu. Klasy aktualnej i przeszłej produktywności to klasy najwyższej, średniej lub najniższej produktywności – czyli odpowiednio górne 20%, środkowe 60% lub dolne 20% – w podejściu znormalizowanym do prestiżu czasopisma i dyscypliny, oddzielnie w każdej z 14 dyscyplin STEMM. Klasy wieku awansu to: młodzi, średni lub starzy doktorzy habilitowani oraz młodzi, średni lub starzy profesorowie tytularni. Oznacza to odpowiednio górnych 20%, środkowych 60% lub dolnych 20% pod względem wieku awansu wyrażonego w pełnych latach. Klasy szybkości awansu obejmowały szybkich, przeciętnych i powolnych doktorów habilitowanych oraz szybkich, przeciętnych i powolnych profesorów tytularnych, czyli odpowiednio górnych 20%, środkowych 60% i dolnych 20%, jeśli chodzi o czas przejścia między kolejnymi awansami, również wyrażony w pełnych latach.



Ryc. 1. Schemat klasyfikacyjny zastosowany do profesorów tytularnych: klasy produktywności, wieku awansu i szybkości awansu

Na każdym etapie kariery profesorowie tytularni byli bardziej lub mniej produktywni i z tego powodu zmieniali klasy produktywności w stosunku do swoich kolegów z tej samej dyscypliny. Naukowcy byli konsekwentnie porównywani na tym samym etapie rozwoju kariery i w ramach tej samej dyscypliny.

3.7. Ograniczenia

Badanie ma kilka ograniczeń związanych z danymi i z metodologią. Po pierwsze, nasza próba obejmowała wszystkich naukowców (zatrudnionych w szkolnictwie wyższym w listopadzie 2017 roku) widocznych na arenie międzynarodowej dzięki swoim publikacjom indeksowanym w bazie Scopus w latach 2009–2018; w związku z tym nie uwzględniono naukowców niepublikujących (i niepublikujących na arenie międzynarodowej). Jednakże odsetek naukowców z dyscyplin STEM, którzy publikowali na arenie międzynarodowej, był wysoki; co więcej, zwiększał się on z czasem i był znacznie wyższy niż w przypadku dyscyplin spoza STEM.

Po drugie, w tym badaniu połączono (niemal doskonałe) dane administracyjne i biograficzne zbierane przez krajowy rejestr naukowców ze (znacznie mniej doskonałymi) danymi bibliometrycznymi na poziomie indywidualnym. Dlatego połączyliśmy dane o „realnych osobach” z krajowymi numerami identyfikacyjnymi z metadanymi o publikacjach według indywidualnych identyfikatorów Scopus Author ID, a nie „realnych naukowców”. Jednak związek między polskimi naukowcami i ich identyfikatorami w bazie Scopus w obszarze STEM jest niezwykle wysoki i rośnie w czasie w związku z komercyjnym charakterem tej bazy i jej wykorzystywaniem do analiz szkolnictwa wyższego na poziomie ministerialnym w ramach ostatniej fali reform.

Nasze Obserwatorium Polskiej Nauki zostało skonstruowane z wykorzystaniem deterministycznego i probabilistycznego łączenia rekordów między dwoma pierwotnymi zbiorami danych, które różniły się charakterem. Przez ostatnie dwie dekady szeroko dyskutowano, w jakim stopniu dane bibliometryczne pokazują skrzywienie pod względem językowym, geograficznym i dyscyplinarnym (Boekhout et al. 2021). Jednak źródła inne niż surowe zbiory danych Scopus (lub Web of Science Core Collection) nie mogły być wykorzystane do skonstruowania pełnej historii publikacyjnej wszystkich naukowców w ramach całego krajowego systemu nauki. Nie istnieją inne źródła metadanych dotyczące wszystkich publikacji polskich naukowców z ostatnich 50 lat. Ponadto nasze badanie wykazuje „skrzywienie pod kątem sukcesu” (*success bias*): w próbie znajdują się tylko profesorowie tytularni, czyli naukowcy, którzy dotarli na szczyt akademickiej hierarchii.

4. Wyniki

4.1. Mobilność między klasami produktywności z perspektywy całości kariery zawodowej

Pytania badawcze dotyczyły trwałości przynależności do klas produktywności profesorów tytularnych z perspektywy ich całej kariery: czy profesorowie o stosunkowo najwyższej produktywności zawsze wykazywali najwyższą produktywność? I analogicz-

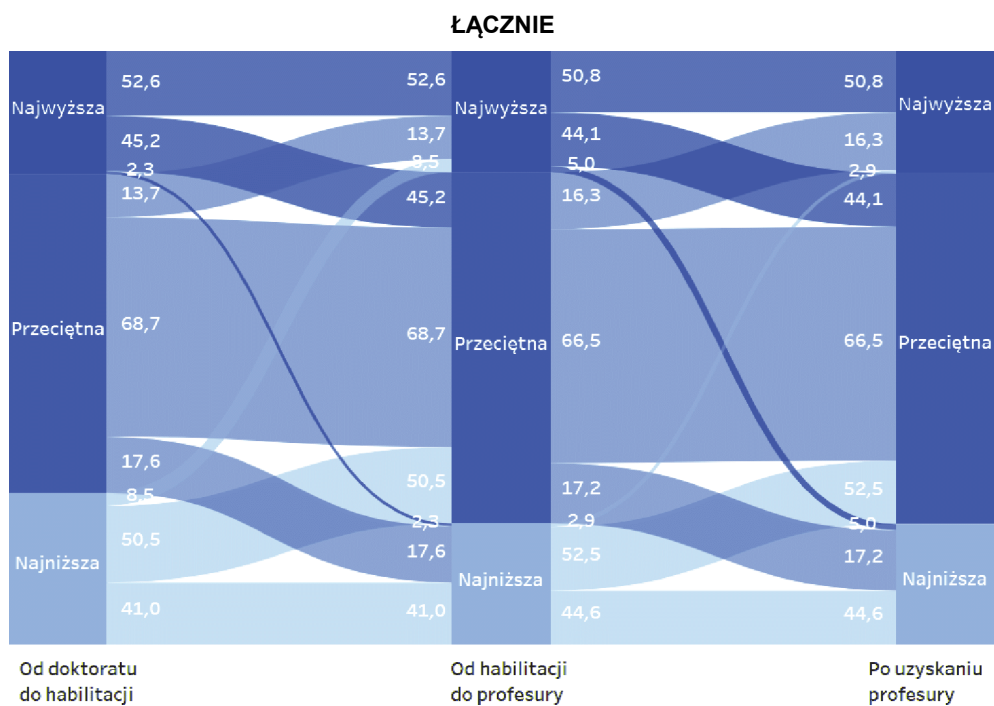
nie: czy profesorowie o stosunkowo najniższej produktywności zawsze wykazywali niską produktywność?

Rycina 2 przedstawia przebieg kariery zawodowej 2326 profesorów tytularnych w 14 dyscyplinach STEM łącznie. Ich produktywność została sklasyfikowana jako najwyższa, średnia lub najniższa (odpowiednio 20%, 60% lub 20%) w trzech okresach: pomiędzy doktoratem a habilitacją (lewa kolumna); pomiędzy habilitacją a profesurą (środkowa kolumna); oraz po uzyskaniu profesury (prawa kolumna) do 2017 roku. Skupiliśmy się na mobilności między najwyższymi i najniższymi klasami produktywności na trzech etapach kariery akademickiej. Wyniki zostały przedstawione na wykresach Sankeya (strumieniowych).

Większość wysoce produktywnych naukowców pozostawała wysoce produktywna w porównaniu ze swoimi kolegami w tej samej dyscyplinie i na tym samym etapie kariery, co widać w postaci szerokich poziomych strumieni przechodzących od strony lewej do prawej na rycinie 2. Ponad połowa wysoce produktywnych naukowców przeszła z klasy najwyższej do klasy najwyższej na pierwszym (52,6%) i drugim etapie swojej kariery akademickiej (50,8%). Tylko około 2,3% naukowców z klasy o najwyższej produktywności przeszło do klasy o niskiej produktywności w pierwszym okresie i tylko około 5% w drugim okresie. Te wyjątkowe przypadki mobilności w dół między klasami produktywności są przedstawione jako cienkie, opadające strumienie. Mobilność od najniższej do najwyższej klasy produktywności w pierwszym i drugim okresie była również bardzo ograniczona. Mobilność w górę jest przedstawiona jako cienkie, wznoszące się strumienie z klas dolnych do górnych: odpowiednio 8,5% i 2,9%. Skrajna mobilność pionowa pomiędzy klasami produktywności była charakterystyczna tylko dla 100 naukowców z grupy 2326.

Wykresy Sankeya pokazują również stałą mobilność między klasami o przeciętnej i najwyższej produktywności. Chociaż większość profesorów przypisanych do klasy o przeciętnej produktywności pozostawała w tej samej klasie, niektórzy przesuwali się w górę, a niektórzy w dół. Mobilność naukowców między klasami produktywności różniła się znacząco w zależności od dyscypliny. Szczegółowo przeanalizowaliśmy dyscyplinę o największej liczbie profesorów tytularnych (MED medycyna) oraz dyscyplinę, w której wzorce mobilności poziomej były najbardziej stabilne z porównawczej perspektywy dyscyplinarnej (MATH matematyka). Matematyka jest często badana ze względu na swoje wyjątkowe cechy, takie jak niski wskaźnik współpracy i niewielki udział kobiet wśród naukowców (np. Mihaljević-Brandt et al. 2016).

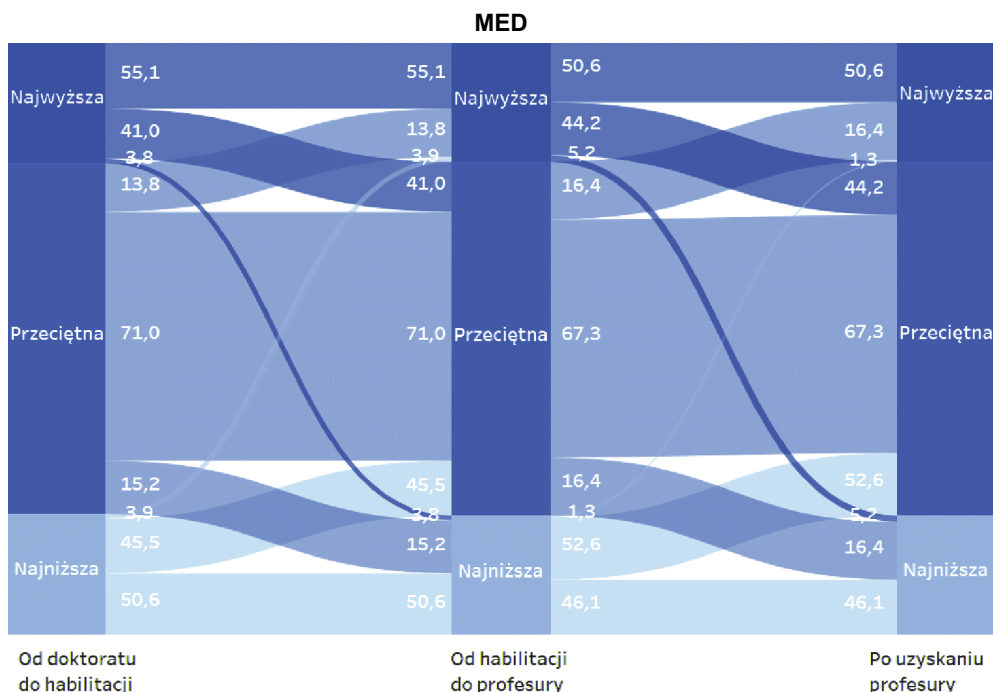
Przypadek medycyny (ryc. 3) przedstawia wyraźny wzorec mobilności między klasami produktywności: mobilność między górnymi klasami oraz mobilność między dolnymi klasami jest wysoka, a mobilność między górnymi i dolnymi oraz dolnymi i górnymi klasami jest ograniczona w ciągu całej kariery akademickiej. Ponad połowa wysoce



Ryc. 2. Wykres Sankeya retrospektywnie skonstruowanej mobilności między klasami produktywności na trzech etapach kariery akademickiej. Wszystkie dyscypliny STEM pokazane łącznie, tylko obecni profesorowie tytularni. Klasy o najwyższej (górne 20%), przeciętnej (środkowe 60%) i najniższej (dolne 20%) produktywności są przedstawione jako 100% (lub w zaokrągleniu) w każdej z klas. Klasa najniższa w lewej kolumnie jest większa niż 20%, a klasa przeciętna jest mniejsza niż 60% (punkty odcięcia nie pozwalały na inny podział na klasy). $N = 2326$

produktywnych doktorów została wysoce produktywnymi doktorami habilitowanymi; a ponad połowa nisko produktywnych doktorów została nisko produktywnymi doktorami habilitowanymi (odpowiednio 55,1% i 50,6%; zob. szerokie strumienie na ryc. 3).

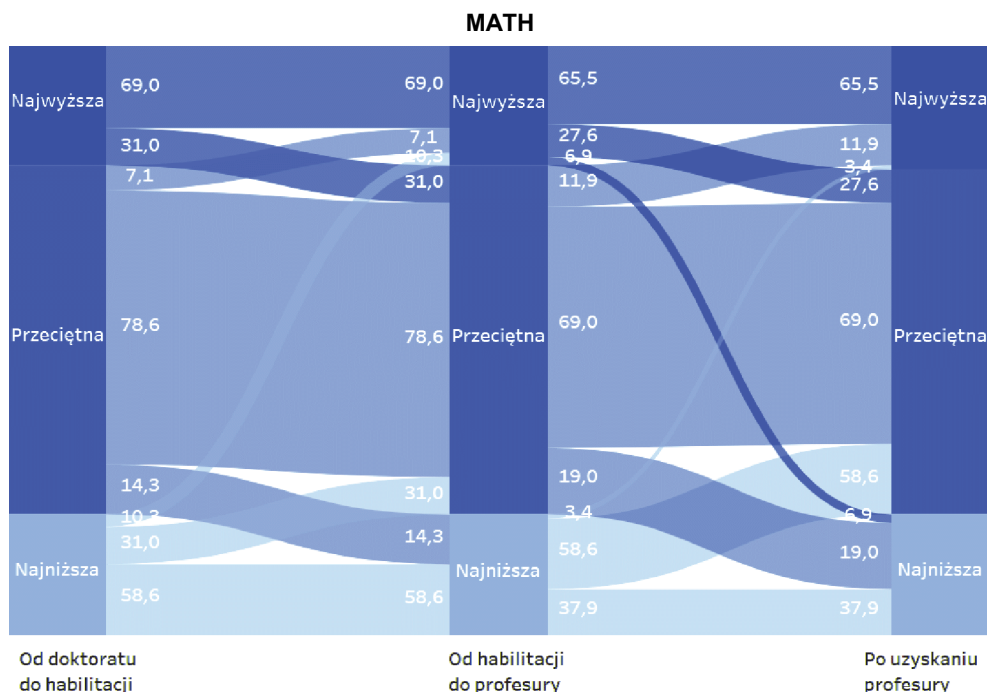
Wzorzec mobilności był podobny na dwóch etapach kariery akademickiej. Większość wysoce produktywnych doktorów habilitowanych została wysoce produktywnymi profesorami tytularnymi, a prawie połowa nisko produktywnych doktorów habilitowanych została nisko produktywnymi profesorami tytularnymi (odpowiednio 50,6% i 46,1%). Skrajne przejścia między klasami produktywności (z góry do dołu i z dołu do góry) zdarzały się rzadko, o czym świadczą bardzo cienkie strumienie łączące najwyższe i najniższe klasy produktywności w obu okresach kariery akademickiej. Skrajne przejścia wystąpiły u 3,8% (spadek) i 3,9% (wzrost) doktorów oraz u 5,2% (spadek) i 1,3% (wzrost) doktorów habilitowanych.



Ryc. 3. Wykres Sankeya przedstawiający retrospektywnie skonstruowaną mobilność między klasami produktywności na trzech etapach kariery akademickiej. Tylko profesorowie tytularni w MED medycynie, $N = 379$

Również w matematyce (ryc. 4) stabilność (czy brak mobilności) wysoce produktywnych doktorów i doktorów habilitowanych była bardzo wysoka. Dwie trzecie naukowców z klas o najwyższej produktywności pozostało w tych klasach: 69% wysoce produktywnych doktorów nadal było wysoce produktywnymi doktorami habilitowanymi, a 65,5% wysoce produktywnych doktorów habilitowanych nadal było wysoce produktywnymi profesorami tytularnymi. Prawdopodobieństwo, że nisko produktywni doktorzy habilitowani wejdą do klasy wysoce produktywnych profesorów tytularnych jest znikome (3,4%).

Pozostałe dyscypliny charakteryzowały się różnym natężeniem mobilności w górę i w dół w ramach klas produktywności (ryc. 5). W niektórych dyscyplinach żaden wysoce produktywny doktor w swojej karierze nie spadł do najniższej klasy produktywności. Mobilność w górę z klasy najniższej do najwyższej była rzadka lub nie występowała wcale (np. CHEM chemia). W innych żaden wysoce produktywny doktor ani żaden wysoce produktywny doktor habilitowany nie spadł do najniższej klasy produktywności, a awans z klasy najniższej do najwyższej nie wystąpił w przypadku doktorów habilitowanych (np. COMP informatyka oraz EARTH nauki o Ziemi i planetach).

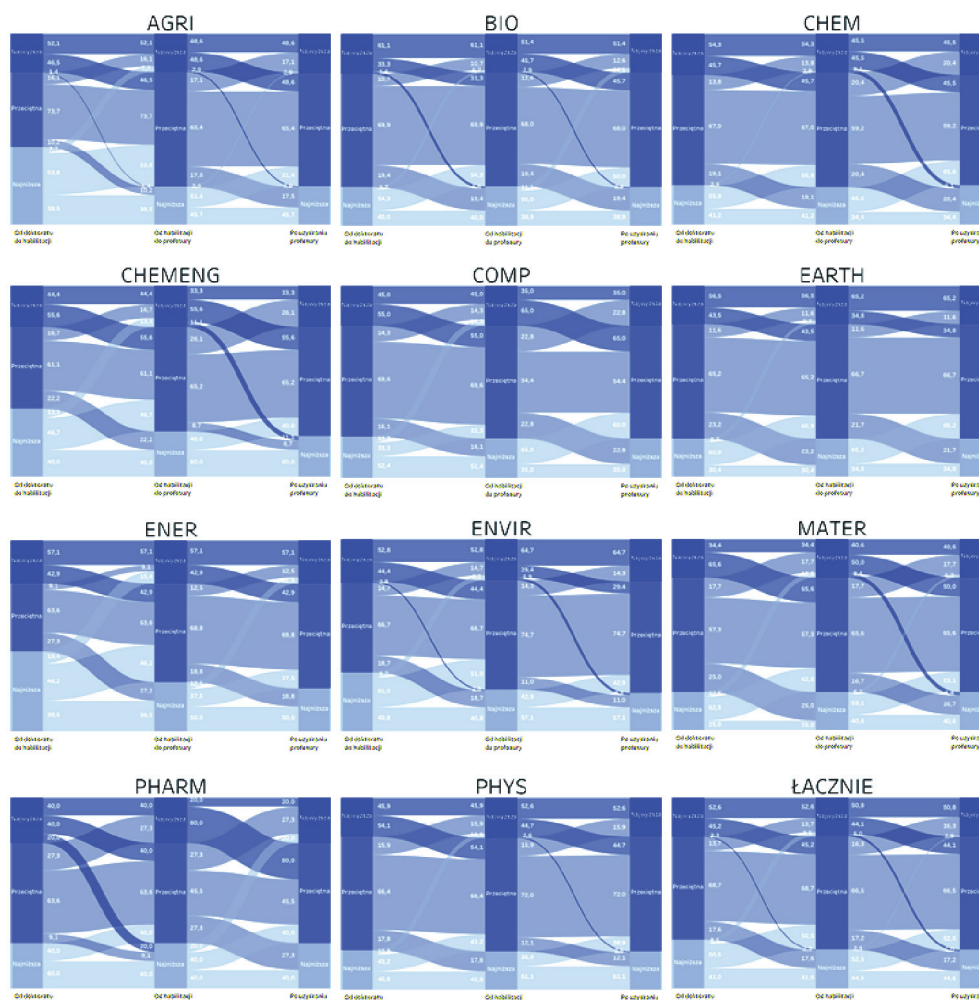


Ryc. 4. Wykres Sankeya przedstawiający retrospektywnie skonstruowaną mobilność między klasami produktywności na trzech etapach kariery akademickiej. Tylko profesorowie tytularni w MATH matematyce, $N = 142$

W jeszcze innych dyscyplinach, podczas gdy nie zaobserwowano mobilności z góry na dół w klasach produktywności, mobilność z dołu do góry była zauważalna (np. ENER energia oraz PHYS fizyka i astronomia). Co więcej, wyniki pokazały różnice w zależności od płci w obrębie dyscyplin, w których większy odsetek kobiet niż mężczyzn pozostał w najwyższych klasach produktywności, czego jednak tutaj nie analizujemy z powodu ograniczonego miejsca.

Rycina 5 pokazuje zmiany w pozostałych dyscyplinach. Stabilność klas o najwyższej produktywności była wysoka i wynosiła od 34,4% do 69,0% dla doktorów, którzy zostali doktorami habilitowanymi, oraz od 20% do 65,5% dla doktorów habilitowanych, którzy zostali profesorami tytularnymi. W pierwszym przypadku udział ten przekraczał 50% w większości dyscyplin, a w drugim – przekraczał 50% w połowie dyscyplin.

Przeprowadziliśmy również porównanie klas produktywności na pierwszym i ostatnim etapie kariery akademickiej: w ramach przejścia od doktora do profesora tytularnego. Prawie połowa obecnych wysoce produktywnych profesorów tytularnych była wysoce produktywnymi doktorami 20–40 lat wcześniej (46,8%).



Ryc. 5. Przegląd: wykresy Sankeya przedstawiające retrospektywnie skonstruowaną mobilność między klasami produktywności na trzech etapach kariery akademickiej. Tylko profesorowie tytularni w jedenastu dyscyplinach STEM oraz we wszystkich dyscyplinach STEM łącznie

Wyniki pokazały jednak interesującą różnicę między mężczyznami i kobietami: odsetek kobiet naukowców, które były wysoce produktywne przez całą swoją karierę, był znacznie wyższy niż odsetek mężczyzn naukowców (48,1% vs. 42,5%). Różnice między dyscyplinami i między mężczyznami i kobietami były znaczne: na przykład wszyscy (100%) wysoce produktywni mężczyźni profesorowie tytularni byli wysoce produktywnymi doktorami 20–40 lat wcześniej w dwóch najbardziej zdominowanych przez mężczyzn dyscyplinach, matematyce oraz fizyce i astronomii (w porównaniu z kobietami

– odpowiednio 46,4% i 44,4%). Zasada „raz wysoka produktywność, zawsze wysoka produktywność” obowiązywała zatem we wszystkich przypadkach polskich matematyków, fizyków i astronomów (obecnych profesorów tytularnych).

4.2. Modele regresji logistycznej

W tej części pracy przedstawiamy oszacowania ilorazu szans przynależności do najwyższej klasy produktywności dla obecnych profesorów tytularnych oraz, retrospektywnie, dla obecnych profesorów tytularnych na wcześniejszych etapach ich kariery akademickiej (w tych samych dyscyplinach) ($N = 2326$). Zmienne na poziomie indywidualnym obejmowały płeć, wiek biologiczny, wiek akademicki (liczba lat, jakie upłynęły od pierwszej publikacji, zob. Kwiek i Roszka 2022b) oraz wiek biologiczny, w którym nadano doktorat, habilitację i profesurę tytularną. Co najważniejsze w kontekście analiz dwuwymiarowych przedstawionych w części 4.1, zmienne na poziomie indywidualnym obejmowały również klasyfikacje z naszego ogólnego schematu klasyfikacyjnego (ryc. 1: przynależność do aktualnych i przeszłych klas produktywności, klas wieku awansu i klas szybkości awansu, z podziałem na 20%, 60% i 20% w każdym przypadku). Jedyną zmienną instytucjonalną była intensywność badawcza instytucji zatrudniającej (IDUB vs. inne instytucje).

Wyniki regresji logistycznej wzmacniają wyniki uzyskane dzięki statystykom opisowym: przynależność do grona najbardziej produktywnych profesorów tytularnych jest silnie uzależniona od tego, czy na wcześniejszych etapach kariery akademickiej naukowcy należeli do analogicznego grona wysoce produktywnych. Tak więc przynależność do klasy wysoce produktywnych doktorów zwiększa prawdopodobieństwo znalezienia się w grupie wysoce produktywnych profesorów tytularnych średnio od dwóch do czterech razy ($\text{Exp}(B) = 2,8$; 95% przedział ufności 2,1–3,6, wszędzie przy zachowaniu reguły *ceteris paribus*), natomiast przynależność do klasy wysoce produktywnych doktorów habilitowanych zwiększa prawdopodobieństwo sukcesu średnio od czterech do sześciu razy ($\text{Exp}(B) = 4,61$; 95% przedział ufności 3,6–6). Jedynym istotnym predyktorem pośrednio związanym z wiekiem jest przynależność do najmłodszej grupy 20% profesorów tytularnych pod względem wieku awansu. Przynależność do tej klasy zwiększa prawdopodobieństwo sukcesu średnio dwukrotnie ($\text{Exp}(B) = 1,942$; zob. zmienne Klasa_Wysoka_produktywnosc_doktor, Klasa_Wysoka_produktywnosc_dr_hab i Klasa_szybka_profesura_tytularna).

Podobnie wśród obecnych profesorów tytularnych pracujących wtedy, gdy byli doktorami habilitowanym (Model 2), przynależność do klasy wysoce produktywnych doktorów zwiększa prawdopodobieństwo stania się wysoce produktywnym doktorem habilitowanym średnio od pięciu do dziewięciu razy ($\text{Exp}(B) = 6,667$; 95% przedział ufności 4,7–9,4). Istotne czynniki determinujące przynależność do grupy 20% najbardziej

produktywnych naukowców są związane z wiekiem zarówno biologicznym, jak i akademickim. Wiek biologiczny ma negatywny wpływ, przy czym jest on znacząco silniejszy w przypadku doktorów habilitowanych niż doktorów. Podwyższenie wieku biologicznego o jeden rok zmniejsza prawdopodobieństwo wejścia do klasy wysoce produktywnych doktorów o 20–25%. Wśród doktorów habilitowanych wzrost o jeden rok zmniejsza to prawdopodobieństwo nawet o jedną trzecią/jedną czwartą. Wśród doktorów wzrost wieku akademickiego o jeden rok (a więc wzrost doświadczenia w publikowaniu lub liczba lat, jaka upłynęła od wydania pierwszej publikacji) powoduje średni wzrost prawdopodobieństwa sukcesu o 10–15%, natomiast wśród doktorów habilitowanych wzrost ten wynosi zaledwie 0,2–4,1%.

Inną zmienną związaną z wiekiem, która znacząco wpływa na prawdopodobieństwo sukcesu, jest wiek uzyskania doktoratu. Wśród doktorów habilitowanych wzrost wieku uzyskania doktoratu ma negatywny wpływ, zmniejszając prawdopodobieństwo sukcesu średnio o 5,8% (przy 95% przedziale ufności 0,5–10,8%), natomiast wśród doktorów kierunku zmian jest pozytywny i wysoki; wzrost wieku uzyskania doktoratu o rok zwiększa prawdopodobieństwo sukcesu średnio o 20,7% (14–27%). Wiek uzyskania habilitacji znacząco i silnie wpływa na prawdopodobieństwo sukcesu; wzrost wieku uzyskania habilitacji o rok zwiększa prawdopodobieństwo wejścia do grupy 20% najbardziej produktywnych doktorów habilitowanych o połowę (średnio 47,5%; 40–55%). Zmienna ta nie mogła być uwzględniona w modelu dla doktorów, ponieważ nie otrzymali oni jeszcze habilitacji. Zmienną (pośrednio) związaną z wiekiem, która ma znaczenie dla prawdopodobieństwa znalezienia się wśród 20% najbardziej produktywnych doktorów, jest obecność wśród 20% najmłodszych naukowców mających doktorat. Przynależność do tej grupy zwiększa prawdopodobieństwo sukcesu średnio o 73,9% (choć przedział ufności jest w tym przypadku dość szeroki: 23,2–145,5%). Płeć ma znaczący wpływ tylko wśród doktorów habilitowanych. Bycie mężczyzną zwiększa prawdopodobieństwo sukcesu średnio o 42,6%, ale zakres przedziału ufności (3–97%) sugeruje, że znaczenie tego predyktora należy interpretować z dużą ostrożnością (o różnicach między mężczyznami i kobietami w nauce, zob. Kwiek i Roszka 2021a, 2021b, 2022a).

Podsumowując, w przypadku obecnych profesorów tytularnych najsilniejszym predyktorem przynależności do klasy wysoce produktywnych naukowców jest przynależność do tej klasy podczas pracy w charakterze doktorów i doktorów habilitowanych; trzecim silnym predyktorem jest przynależność do klasy wcześniej awansowanych profesorów tytularnych. Retrospektywnie, dla obecnych profesorów tytularnych, gdy byli doktorami habilitowanymi, najsilniejszym czynnikiem predykcyjnym jest przynależność do klasy wysoce produktywnych doktorów; inne czynniki predykcyjne to przynależność do klasy profesorów tytularnych, którzy otrzymali wcześniej awans ($\text{Exp(B)} = 1,475$) i ewentualnie bycie mężczyzną ($\text{Exp(B)} = 1,426$).

Tabela 3. Statystyki regresji logistycznej: oszacowanie ilorazu szans przynależności do grupy wysoce produktywnych profesorów tytularnych oraz profesorów tytularnych retrospektywnie – gdy byli doktorami habilitowanymi i doktorami (górne 20%, osobno dla każdej dyscypliny; tylko obecni profesorowie tytularni, $N = 2326$)

Model	Model 1: Obecni profesorowie tytularni $R^2 = 0,254$ $N = 2326$				Model 2: Obecni profesorowie tytularni retrospektywnie – gdy byli doktorami habilitowanymi $R^2 = 0,582$ $N = 2326$				Model 3: Obecni profesorowie tytularni retrospektywnie – gdy byli doktorami $R^2 = 0,355$ $N = 2326$			
	Exp(B)	95% przedział ufności		Istotność	Exp(B) Dolny	95% przedział ufności		Istotność	Exp(B)	95% przedział ufności		Istotność
		Dolny	Górny			Dolny	Górny			Dolny	Górny	
Mężczyzna					1,426	1,03	1,974	0,033				
Instytucja_badawcza_IDUB												
Wiek_biologiczny					0,694	0,665	0,724	< 0,001	0,774	0,753	0,796	< 0,001
Wiek_akademicki					1,021	1,002	1,041	0,028	1,122	1,098	1,148	< 0,001
Wiek_doktoratu					0,942	0,892	0,995	0,032	1,207	1,143	1,273	< 0,001
Wiek_habilitacji					1,475	1,404	1,549	< 0,001	-	-	-	-
Wiek_profesury_tytularnej					-	-	-	-	-	-	-	-

kontynuacja tabeli 3.

Klasa_Wysoka _produktywnosc_doktor	2,793	2,14	3,646	< 0,001	6,667	4,72	9,416	< 0,001	-	-	-	-
Klasa_Wysoka _produktywnosc_dr_hab	4,61	3,558	5,974	< 0,001	-	-	-	-	-	-	-	-
Klasa_mlody_doktor									1,739	1,232	2,455	0,002
Klasa_mlody_habilitowany									-	-	-	-
Klasa_mlody_profesor _tytularny	1,942	1,503	2,509	< 0,001	-	-	-	-	-	-	-	-
Klasa_szybka_habilitacja									-	-	-	-
Klasa_szybka_profesura _tytularna					-	-	-	-	-	-	-	-
Stała	0,1			< 0,001	46,17			< 0,001	128,62			< 0,001

„-” brak obserwacji z powodów strukturalnych

Wreszcie, również retrospektywnie dla profesorów tytularnych gdy byli doktorami, najsilniejszym predyktorem przynależności do klasy najbardziej produktywnych doktorów jest przynależność do klasy wcześniej wypromowanych doktorów ($\text{Exp(B)} = 1,739$).

5. Podsumowanie i wnioski

Wysoce produktywni naukowcy są często analizowani jako szczególna grupa akademicka: jako „wybitni” i „wysoce płodni” naukowcy, jako „gwiazdy nauki” i jej *top performers* (Fox i Nikivincze 2021; Kwiek 2018; Agrawal et al. 2017; Kwiek 2016; Cortés et al. 2016; Abramo et al. 2009). Są oni „zmotywowani wewnętrznym dążeniem do uprawiania nauki i czystą miłością do tej pracy” (Cole i Cole 1973: 62) i choć pewni naukowcy są szczególnie dobrzy w uprawianiu nauki, „niektórzy są nie tyle dobrzy, co wręcz znakomici” (Stephan i Levin 1992: 13). Zgodnie z tym stwierdzeniem, niektórzy profesorowie tytularni z naszej próby byli po prostu znakomici w uprawianiu nauki od momentu rozpoczęcia kariery akademickiej aż do jej późnych etapów. Połowa wysoce produktywnych profesorów tytularnych była zawsze wysoce produktywna, niezależnie od zmian w życiu osobistym czy od okoliczności zewnętrznych (np. pracy w okresie postkomunistycznej transformacji w polskiej gospodarce, który poważnie wpłynął na sektor akademicki, zob. Kwiek 2015b i Kwiek 2015c). Wysoce produktywni profesorowie w wieku 60 lat byli również wysoce produktywni, kiedy byli doktorami i doktorami habilitowanymi w wieku 30, 40 czy 50 lat.

Wnioski płynące z naszej analizy regresji logistycznej są proste: przynależność do klas produktywności w przeszłości (tzn. indywidualne portfolio publikacyjne) w znacznym stopniu warunkuje przynależność do klas produktywności w przyszłości, przy czym pozostałe predyktory odgrywają znacznie mniejszą rolę. Nasze wielowymiarowe modele regresji silnie wspierają wyniki naszych analiz dwuwymiarowych, zgodnie z którymi naukowcy, którzy wcześniej byli wysoce produktywni, z reguły pozostają nadal wysoce produktywni, a ci, którzy wcześniej charakteryzowali się niską produktywnością, mają niewielkie szanse na przejście do klasy wysokiej produktywności (co widać jako cienkie strumienie na wykresach Sankeya, biegnące w górę między najniższą i najwyższą klasą produktywności we wszystkich dyscyplinach na ryc. 2).

Występują tylko dwa silne predyktory wysokiej produktywności wśród profesorów tytularnych: przynależność do klasy wysoce produktywnych doktorów i przynależność do klasy wysoce produktywnych doktorów habilitowanych, które zwiększają szanse średnio odpowiednio prawie trzy- i pięciokrotnie (o 179% i 361%). Najsilniejszym predyktorem zostania wysoce produktywnym doktorem habilitowanym (w próbie obecnych profesorów tytularnych) jest bycie wcześniej wysoce produktywnym doktorem, o czym świadczy imponujący wzrost szans: prawie siedmiokrotny (o 570%). W przypadku wyso-

ce produktywnych doktorów najsilniejszym czynnikiem predykcyjnym jest uzyskanie stopnia doktora w młodym wieku. Ponadto nasze wyniki potwierdzają wcześniejsze ustalenia, że profesorowie mianowani wcześniej są bardziej produktywni niż profesorowie mianowani w późniejszym okresie kariery (Abramo et al. 2016). Przynależność do klasy młodych profesorów tytularnych zwiększa szanse na przynależność do klasy wysoce produktywnych profesorów tytularnych średnio dwukrotnie (o 94,2%). Ani płeć, ani wiek (biologiczny czy akademicki) nie okazują się predyktorem przynależności do klasy wysoce produktywnych profesorów. Uzyskane wyniki nie potwierdzają bezpośrednio tezy, że produktywność naukowców o najwyższej i przeciętnej produktywności rośnie lub pozostaje stabilna wraz z wiekiem (Costas et al. 2010: 1578), ponieważ nasze badanie koncentrowało się na zmianach klas produktywności, a nie na ewolucji produktywności w czasie.

Wyniki naszego badania pokazują zaskakująco wysoki poziom immobilności (czy stabilności) w polskim systemie: przynależność do klasy produktywności w okresie posiadania doktoratu i habilitacji w dużym stopniu określa przynależność do klasy produktywności w okresie profesury tytularnej (przypominamy, że z powodów charakterystyki naszej bazy nie posługujemy się określeniami asystent, adiunkt, profesor nadzwyczajny i zwyczajny). Czy odkryta przez nas zasada „raz wysoce produktywny, zawsze wysoce produktywny” obowiązuje we wszystkich dyscyplinach STEM? Wyniki naszych badań wskazują na to, że tak. Połowa obecnych profesorów tytularnych należała przez całą swoją karierę akademicką do tej samej klasy produktywności. Przez dziesięciolecia pozostawali oni w dolnej lub górnej klasie produktywności w stosunku do swoich kolegów i w ramach swoich dyscyplin. Połowa obecnych profesorów tytularnych zmieniła swoją klasę produktywności tylko o jedną klasę w trójstopniowym podziale na klasy, przy pewnym zróżnicowaniu pod względem dyscyplin i płci. Różnice między dyscyplinami oraz między kobietami i mężczyznami były znaczne: na przykład wszyscy wysoce produktywni profesorowie tytularni byli wysoce produktywnymi doktorami 20–40 lat wcześniej w dwóch najbardziej zdominowanych przez mężczyzn dyscyplinach: matematyce i fizyce/astronomii. Zasada ta obowiązywała więc we wszystkich badanych przypadkach polskich matematyków, fizyków i astronomów.

Średnio ponad połowa wysoce produktywnych doktorów została wysoce produktywnymi doktorami habilitowanymi (w porównaniu z kolegami na tym samym etapie rozwoju zawodowego i w tej samej dyscyplinie). Średnio ponad połowa wysoce produktywnych doktorów habilitowanych została wysoce produktywnymi profesorami tytularnymi (odpowiednio 52,6 i 50,8%). Co więcej, analiza bezpośredniej mobilności od początku do końca kariery – od doktoratu do profesury – pokazuje, że średnio prawie połowa wysoce produktywnych doktorów została wysoce produktywnymi profesorami tytularnymi. Nie zmienili oni przynależności do niższej klasy produktywności w trakcie swojej

kariery akademickiej (46,8%), przy czym istnieje duże zróżnicowanie między dyscyplinami. Podobne procesy zmiany przynależności do klasy produktywności dotyczyły naukowców o niskiej produktywności: czyli analogicznie, „raz nisko produktywny, zawsze nisko produktywny”.

Najbardziej radykalne zmiany w przynależności do klasy produktywności, czyli przejścia z najwyższej do najniższej klasy produktywności i odwrotnie, zachodzą na marginalnym poziomie. W naszej próbie 2326 profesorów tytularnych w ciągu ostatnich czterech dekad znalazło się tylko 35 naukowców, którzy radykalnie zmienili swoją klasę produktywności w dół, i 65, którzy przesunęli się radykalnie w górę (czyli w sumie grupa ta objęła tylko 4,3% obecnych profesorów tytularnych). Ponadprzeciętną mobilność zaobserwowano w dyscyplinach BIO, MATH i PHYS, natomiast najmniejszą w PHARM.

Być może najciekawsze jest pytanie, dlaczego wzorzec „raz wysoce produktywni, zawsze wysoce produktywni” jest tak wszechobecny w polskim szkolnictwie wyższym. Spośród kilku możliwych wyjaśnień jedno jest zgodne z dwiema tradycyjnymi teoriami produktywności: teorią iskry bożej i teorią kumulacji przewag. Pierwsza zakłada, że istnieje niewielka grupa naukowców, którzy zawsze będą osiągać ponadprzeciętne wyniki, ponieważ mają ową iskrę, której brakuje innym, są z natury wysoko zmotywowani, dobrze zorganizowani, kreatywni i uzdolnieni. Druga teoria wskazuje na grupę naukowców, którzy – posiadając ową iskrę lub jej nie posiadając – gromadzą przewagi od samego początku swojej kariery.

Ich przewagi wynikają z socjalizacji do umiędzynarodowionego środowiska pracy naukowej, specyficznej kultury pracy i nawyków pracy dostępnych głównie w elitarnych instytucjach lub na elitarnych wydziałach; wynikają z obecności promotorów, którzy stanowią dla nich wzorce do naśladowania, oraz z zasobów dostępnych dzięki finansowaniu badań naukowych, w tym długoterminowych stypendiów międzynarodowych. Teoria kumulacji przewag tłumaczy wysoką produktywność zestawem czynników wzmacniających, które – połączone – nieustannie napędzają kariery akademickie (z coraz lepszym dostępem do wszelkiego rodzaju zasobów: czasu na badania, infrastruktury, finansowania, sieci międzynarodowych, publikacji w prestiżowych czasopismach, zewnętrznie finansowanych doktorantów i postdoków itp.)

Inną użyteczną linią teoretyczną wyjaśniającą odkryte prawidłowości jest cykl wiarygodności w karierze akademickiej (Latour i Woolgar 1986: 200–208), w którym prestiżowe prace przekładają się na uznanie, co prowadzi do skutecznych wniosków o granty, które z kolei przekładają się na nowy sprzęt, dane, oprogramowanie, argumenty naukowe i kolejne prestiżowe publikacje. Być może cykl wiarygodności jest szybszy w przypadku naukowców podlegających temu mechanizmowi na wczesnym etapie kariery: po uzyskaniu finansowania, z doskonałymi publikacjami, mają oni większe szanse na ponowne finansowanie i szybszy awans na wyższe stanowiska, co odzwierciedla kon-

cepcję, że każdy element cyklu wiarygodności w karierze akademickiej „jest tylko częścią niekończącego się cyklu inwestycji i konwersji” (Latour i Woolgar 1986: 200). Przewagi już zdobyte szybciej prowadzą do przyszłych przewag, jak w każdej konkurencji pozycyjnej mającej charakter gry o sumie zerowej, ponieważ „to, co wygrywają zwycięzcy, tracą przegrani” (Hirsch 1976: 52). Powyższe mechanizmy teoretyczne mają silniejszy wpływ na systemy o ograniczonych zasobach, takie jak polski, w którym historycznie rzecz biorąc, ze względu na permanentne niedofinansowanie nauki można było otrzymać finansowanie dzięki niewielkiej różnicy w stosunku do konkurentów (Kwiek 2015a; Kwiek 2020).

6. Wykorzystanie Big Data do analiz kadry akademickiej

Z szerzej perspektywy wyniki naszych badań wskazują na możliwości, jakie dają ustrukturyzowane Big Data (w tym przypadku zbiór surowych danych pochodzących z bazy Scopus, w zasadzie jak dotąd niemal całkowicie niedostępnych). Przeanalizowaliśmy wszystkich obecnych polskich profesorów tytularnych pracujących w obszarze STEM i posiadających przynajmniej jedną publikację w bazie Scopus, ale dane, które wykorzystaliśmy, zostały zebrane z dwóch dużych baz danych. Nasze Obserwatorium Polskiej Nauki zawiera pełne dane biograficzne i administracyjne prawie 100 000 polskich naukowców oraz pełne metadane ich 380 000 publikacji zawartych w bazie Scopus z lat 2009–2018. Drugi zbiór danych zawiera metadane pochodzące z bazy Scopus dotyczące prawie miliona ($N = 935\,167$) polskich publikacji z ostatnich 50 lat.

Co zatem w naszym przypadku nie byłoby możliwe do przeprowadzenia bez użycia surowych metadanych z bazy Scopus (lub WoS)? (1) Zdefiniowanie dyscypliny: przeanalizowaliśmy wszystkie publikacje pochodzące z całego życia każdego naukowca, aby określić dyscyplinę modalną (dominującą) każdego profesora. (2) Pomiar produktywności znormalizowanej do prestiżu czasopisma: wszystkie publikacje w historii publikacyjnej wszystkich profesorów zostały powiązane z prestiżem czasopism wyrażonym w rangach percentylowych czasopism z bazy Scopus, po czym obliczono odpowiednio czteroletnią produktywność. (3) Powiązanie każdego artykułu z trzema etapami kariery akademickiej wszystkich profesorów: tylko baza Scopus (lub WoS) zawiera wszystkie artykuły wszystkich profesorów z całego okresu ich życia. (4) Ustalenie wieku akademickiego (doświadczenia akademickiego) wszystkich profesorów: data pierwszej publikacji była niezbędna w modelach regresji.

Połączenie kilku baz danych umożliwiło nam stworzenie nie tylko aktualnych klas produktywności, do których przypisano wszystkich profesorów, ale także stworzenie klas produktywności retrospektywnie. Co najważniejsze, każdy profesor został porównany pod względem produktywności badawczej jako doktor i jako doktor habilitowany ze swoimi odpowiednikami (obecnymi profesorami tytularnymi), kiedy znajdowali się

na tych samych, wcześniejszych etapach kariery akademickiej w tej samej dyscyplinie. Retrospektywnie przeanalizowaliśmy ich kariery akademickie na wszystkich trzech etapach.

Wzorce mobilności między klasami produktywności w trakcie całej kariery akademickiej mają daleko idące implikacje dla polityki naukowej, zwłaszcza w zakresie zatrudniania i awansowania. Zatrudnianie i przedłużanie okresów zatrudnienia zarówno naukowców o niskiej, jak i wysokiej produktywności może nieść ze sobą długofalowe konsekwencje dla poszczególnych instytucji i ich wydziałów oraz dla systemu krajowego w zakresie średniego oczekiwanego poziomu produktywności. Kariera naukowa jest bowiem w Polsce zazwyczaj długa. Po wejściu do systemu i osiągnięciu stabilności zatrudnienia naukowcy w Polsce (gdzie wskaźnik odchodzenia z zawodu jest niezwykle niski), jak i w innych krajach, zwykle pozostają w systemie przez lata, jeśli nie dekady (zob. Abramo et al. 2017 omawiający gwiazdy nauki i naukowców nieproduktywnych we Włoszech). Naukowcy uwzględnieni w naszym badaniu, z których wszyscy są obecnie profesorami tytularnymi reprezentującymi 14 dyscyplin STEMM i są obecni w bibliometrycznej bazie Scopus, pozostają w systemie od 20–40 lat. Indywidualne decyzje o zatrudnieniu i awansie podejmowane na poziomie wydziałów czy uczelni mają zatem długofalowy wpływ na produktywność na zagregowanym poziomie krajowym, obejmującym kilkadziesiąt lat.

I wreszcie, bardziej ogólnie, ustrukturyzowane duże dane (Big Data) oferują zupełnie nowe możliwości badania profesji akademickiej zarówno w skali krajowej, porównawczej międzynarodowej, jak i globalnej. Big Data, gromadzone i przechowywane przez różne podmioty (np. przez rządy i korporacje, jak w naszym przypadku) w celach innych niż akademickie, mogą być analizowane przez badaczy profesji akademickiej jako nowe, uzupełniające źródła danych, które dopełniają źródła tradycyjne, takie jak badania ankietowe czy wywiady pogłębione. Ankiety i wywiady stosowaliśmy od dekady – Big Data stosujemy od niedawna. Dzięki temu można uzyskać równowagę między badaniami na małą skalę i na dużą skalę (z małymi i dużymi N), co z kolei może mieć pozytywny wpływ na całą dziedzinę badań nauki i badań polityki naukowej. Kluczowym słowem jest dla nas komplementarność: nowe źródła danych uzupełniają, a nie zastępują tradycyjne źródła.

Nowe dane muszą być jednak ponownie przetworzone (Salganik 2018) i mają swoje własne ograniczenia. Jednak ilość dostępnych danych i ich wzdłużny charakter (umożliwiający analizę zmian w karierze akademickiej na przestrzeni czasu) otwierają nowe horyzonty, w tym między innymi – umożliwiają ujęcia globalne, którymi się od roku zajmujemy (zob. Kwiek i Szymula 2023). Ze zbiorów danych, które są olbrzymie i złożone, możemy wydobywać tylko przydatne informacje dotyczące naukowców i ich dorobku zarówno w przeszłości, jak i obecnie. Możemy badać potężne ilości danych, aby odkryć

wzorce, które w innym przypadku pozostawałyby niezauważone, analizować wartości skrajne, odchylenia i szczególne przypadki oraz przeprowadzać analizy oparte na bezprecedensowej liczbie obserwacji. Podczas gdy Big Data w znacznym stopniu pogłębiają nasz wgląd w społeczeństwo w ogóle (Selwyn 2019), konkretne części ustrukturyzowanych, archiwizowanych i wiarygodnych Big Data (takich jak komercyjne zbiory danych bibliometrycznych) mogą radykalnie poprawić nasz wgląd w profesję akademicką, umożliwiając jej badanie za pomocą nowych analiz czasowych, tematycznych, geograficznych i sieciowych (zob. Börner 2010: 62-63). Można bowiem badać różne wymiary pracy akademickiej z coraz większą precyzją i na niezwykłym poziomie szczegółowości.

Wykorzystanie dobrze przygotowanych, rozbudowanych źródeł danych pozwala badać profesję akademicką na przestrzeni lat, w różnych krajach (instytucjach, miastach), w różnych dyscyplinach akademickich, na różnych poziomach granulacji oraz w odniesieniu do zespołów badawczych i poszczególnych naukowców, ich wieku, płci i dyscypliny. Niewielka liczba obserwacji uzyskiwana w tradycyjnych badaniach profesji akademickiej ogranicza możliwości analityczne i osłabia zdolność do wyciągania z badań implikacji dla polityki publicznej. Badania na małą skalę są przydatne i teoretycznie inspirujące, ale w globalizującym się świecie, coraz szerzej opartym na danych, mogą nie być przekonujące dla wspólnoty akademickiej, decydentów i agencji grantowych.

Kilka czynników zwiększa presję na badanie profesji akademickiej przy użyciu Big Data: po pierwsze, coraz większa dostępność danych cyfrowych dotyczących nakładów i efektów pracy naukowej na poziomie indywidualnym (finansowanie, publikacje, współpraca badawcza, mobilność); po drugie, coraz większa dostępność mocy obliczeniowych umożliwiających analizę ogromnych zestawów danych w chmurze; i po trzecie, nacisk na zapewnienie zarówno społeczeństwu, jak i społeczności naukowej bardziej skwantyfikowanego, opartego na danych, solidnego i przekonującego obrazu zmian zachodzących w szkolnictwie wyższym, nauce i w profesji akademickiej. Upraszczając: jeszcze niedawno napisanie prezentowanego tu studium polskich profesorów tytularnych i wyciągnięcie wniosków na temat mobilności (lub jej braku) w ramach klas produktywności na podstawie dorobku publikacyjnego ich całego życia zawodowego – byłoby nie tylko niewykonalne, ale i trudne do wyobrażenia.

Podziękowania

Jesteśmy niezwykle wdzięczni gospodarzom i słuchaczom zaproszonych seminariów prof. Marka Kwieka zorganizowanych na Uniwersytecie Oksfordzkim (CGHE, Center for Global Higher Education) i na Uniwersytecie Stanforda (METRICS, Meta-Research Innovation Center at Stanford) w czerwcu 2022 roku. Jesteśmy również wdzięczni Panu

mgr. Łukaszowi Szymuli z CSPP UAM za poprawienie wizualizacji naszych wyników. Wyrażamy wdzięczność za pomoc laboratorium Elseviera International Center for the Studies of Research (ICSR Lab) w ramach wieloletniej umowy o współpracy, a zwłaszcza Kristy James, Senior Data Scientist. Ponadto dziękujemy za wsparcie udzielone w ramach grantu MEIN NDS nr NdS/529032/2021/2021.

Bibliografia

- Abramo G., D'Angelo C.A., Caprasecca A. (2009). The contribution of star scientists to overall sex differences in research productivity. *Scientometrics*, 81(1), 137–156.
- Abramo G., D'Angelo C.A., & Murgia G. (2016). The combined effects of age and seniority on research performance of full professors. *Science and Public Policy*, 43(3), 301–319.
- 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, 867–877.
- 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.
- Allison P.D., Stewart J.A. (1974). Productivity differences among scientists: Evidence for accumulative advantage. *American Sociological Review*, 39(4), 596–606.
- Boekhout, H., van der Weijden, I., Waltman, L. (2021). Gender differences in scientific careers: A large-scale bibliometric analysis. <https://arxiv.org/abs/2106.12624>
- Börner K. (2010). *Atlas of Science. Visualizing What We Know*. The MIT Press.
- Cole J.R., Cole S. (1973). *Social stratification in science*. 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., van Leeuwen T.N., Bordons M. (2010). Self-citations at the meso and individual levels: Effects of different calculation methods. *Scientometrics*, 82, 517–537.
- 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.
- Fox M.F. (1983). Publication productivity among scientists: A critical review. *Social Studies of Science*, 13 (2), 285–305.
- Fox M.F., Mohapatra, S. (2007). Social-organizational characteristics of work and publication productivity among academic scientists in doctoral-granting departments. *Journal of Higher Education*, 78(5), 542–571.
- Fox M.F., Nikivincze I. (2021). Being highly prolific in academic science: Characteristics of individuals and their departments. *Higher Education*, 81, 1237–1255.

- Hermanowicz J. (2012). The sociology of academic careers: Problems and prospects. [W:] J.C. Smart, M.B. Paulsen (red.), *Higher education: Handbook of theory and research* 27 (s. 207–248).
- Hirsch F. (1976). *Social limits to growth*. Harvard University Press.
- Katz D.A. (1973). Faculty salaries, promotions, and productivity at a large university. *American Economic Review*, 63(3), 469–477.
- Kwiek M. (2015a). *Uniwersytet w dobie przemian. Instytucje i kadra akademicka w warunkach rosnącej konkurencji*. Warszawa: PWN.
- Kwiek M. (2015b). Academic generations and academic work: Patterns of attitudes, behaviors and research productivity of Polish academics after 1989. *Studies in Higher Education*, 40(8), 1354–1376.
- Kwiek M. (2015c). The unfading power of collegiality? University governance in Poland in a European comparative and quantitative perspective. *International Journal of Educational Development* 43, 77–89.
- Kwiek M. (2016). The European research elite: A cross-national study of highly productive academics across 11 European systems. *Higher Education*, 71(3), 379–397.
- Kwiek M. (2018). High research productivity in vertically undifferentiated higher education systems: Who are the top performers? *Scientometrics*, 115(1), 415–462.
- Kwiek M. (2019). *Changing European academics. A comparative study of social stratification, work patterns and research productivity*. London and New York: Routledge.
- Kwiek M. (2020). Internationalists and locals: International research collaboration in a resource-poor system. *Scientometrics*, 124, 57–105.
- Kwiek M. (2022). *Globalna nauka, globalni naukowcy*. Warszawa: PWN.
- Kwiek M., Szymula Ł. (2023). Young Male and Female Scientists: A Quantitative Exploratory Study of the Changing Demographics of the Global Scientific Workforce. Preprint: <https://arxiv.org/abs/2301.06196>
- Kwiek M., Roszka, W. (2021a). Gender disparities in international research collaboration: A large-scale bibliometric study of 25,000 university professors. *Journal of Economic Surveys*, 35(5), 1344–1388.
- Kwiek M., Roszka, W. (2021b). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, 15 (3), article 101171. 1–38.
- Kwiek M., Roszka, W. (2022a). Are female scientists less inclined to publish alone? The gender solo research gap. *Scientometrics*, 127, 1697–1735.
- Kwiek M., Roszka, W. (2022b). Academic vs. biological age in research on academic careers: a large-scale study with implications for scientifically developing systems. *Scientometrics*, 127, 3543–3575.
- Kyvik S. (1990). Age and scientific productivity: Differences between fields of learning. *Higher Education*, 19, 37–55.
- Latour B., Woolgar S. (1986) *Laboratory Life. The Construction of Scientific Facts*. Princeton University Press.
- Merton R.K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.

- Mihaljević-Brandt H., Santamaría L., Tullney M. (2016). The effect of gender in the publication patterns in mathematics. *PLOS ONE*, 11 (10), e0165367.
- Piro F.N., Rørstad K., Aksnes D.W. (2016). How do prolific professors influence the citation impact of their university departments? *Scientometrics*, 107(3), 941–961.
- Ruiz-Castillo J., Costas R. (2014). The skewness of scientific productivity. *Journal of Informetrics*, 8(4), 917–934.
- Salganik M.J. (2018). *Bit by bit. Social research in a digital age*. Princeton University Press.
- Selwyn N. (2019). *What is digital sociology?* Polity Press.
- Stephan P. (2012). *How economics shapes science*. Harvard University Press.
- Stephan P.E., Levin S.G. (1992). *Striking the mother lode in science: The importance of age, place, and time*. Oxford University Press.
- Sugimoto C., Larivière V. (2018). *Measuring research: What everyone needs to know*. Oxford University Press.
- 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. (1988). The sociology of science. [W:] N.J. Smelser (red.), *Handbook of sociology* (s. 511–574). Sage.

Zawsze wysoce produktywni? Zawsze nisko produktywni? **Dynamiczne ujęcie dorobku naukowego profesorów tytularnych** **wykorzystujące klasy produktywności**

W pracy poddano analizie trwałość przynależności do klas produktywności badawczej na poziomie indywidualnym. Do analiz wykorzystano dane pochodzące z prowadzonego przez nas Obserwatorium Polskiej Nauki (100 000 naukowców, 380 000 publikacji z ostatniej dekady) oraz metadane z bazy Scopus dotyczące niemal miliona ($N=935\,167$) polskich artykułów z ostatnich 50 lat. Przeanalizowano przebieg kariery naukowej 2326 polskich profesorów tytularnych. Zbadano daty kolejnych awansów naukowych i liczbę publikacji ($N=79\,027$ artykułów) między awansami w ciągu 20–40 lat pracy naukowej w 14 dyscyplinach (nauki ścisłe, techniczne, inżynieryjne, matematyczne i medyczne, czyli w obszarze STEMM). Interesowało nas przemieszczanie się między trzema klasami produktywności – najwyższą (górną 20%), przeciętną (60%) i najniższą (dolną 20%). Połowa najbardziej produktywnych doktorów okazała się najbardziej produktywnymi doktorami habilitowanymi, a połowa najbardziej produktywnych doktorów habilitowanych – najbardziej produktywnymi profesorami tytularnymi (52,6% i 50,8%). Przechodzenie od najwyższej do najniższej i od najniższej do najwyższej klasy produktywności dotyczyło tylko 100 (4,3%) naukowców. W modelach regresji logistycznej dwoma silnymi predyktorami przynależności do najwyższej klasy produktywności wśród profesorów tytularnych okazały się wysoka produktywność w okresie bycia doktorem i w okresie bycia doktorem habilitowanym (zwiększając szanse średnio

niemal dwukrotnie i czterokrotnie, o 179% i 361%). Podsumowanie obejmuje wykorzystanie Big Data do badania karier akademickich.

Słowa kluczowe: Big Data w analizie karier akademickich, polscy naukowcy, produktywność badawcza, stratyfikacja społeczna w nauce, podejście longitudinalne

Always Highly Productive, Always Low Productive? A Longitudinal and Class-Based Approach to Research Output of Full Professors

Lifetime biographical and publication histories of 2,326 full professors were examined. A combination of administrative, biographical, and bibliometric data was used. Retrospectively constructed productivity, promotion age and speed classes were examined. About 50% of current top productive professors have been top productive throughout their academic careers, over 30–40 years. Top-to-bottom and bottom-to-top transitions in productivity classes over academic careers are very rare. We used prestige-normalized productivity in which more weight is given to articles in high-impact than in low-impact journals, recognizing the highly stratified nature of academic science. The combination of biographical and demographic data with raw Scopus publication data from the past 50 years ($N=935,167$ articles) made it possible to assign all full professors retrospectively to different productivity, promotion age, and promotion speed classes. In logistic regression models, there were two powerful predictors of belonging to the Top productivity class for full professors: being highly productive as associate professor and as assistant professor (increasing the odds by 180% and 360%). Neither gender nor age (biological or academic) emerged as statistically significant. Our findings have important implications for hiring policies as scientists stay in Polish academia usually for several decades.

Key words: Big Data in academic careers, Polish scientists, research productivity, social stratification in science, longitudinal approach

