Załącznik 3
do Wniosku o przeprowadzenie postępowania habilitacyjnego

# Autoreferat przedstawiający opis dorobku i osiągnięć naukowych

dr Fryderyk Falniowski

21 marca 2025

# Spis treści

# 1 Informacje podstawowe

**Imię i nazwisko:** Fryderyk Falniowski
**ORCiD:** https://orcid.org/0000-0002-1994-9019

## Posiadane dyplomy, stopnie naukowe

**doktor nauk matematycznych, specjalizacja: układy dynamiczne**

Podmiot wydający: Rada Wydziału Matematyki i Informatyki Uniwersytetu Jagiellońskiego

Data wydania: 24 kwietnia 2014 r.

Tytuł rozprawy: *Uogólnione entropie dynamiczne i oparte na nich niezmienniki*

Promotor: prof. dr hab. Wojciech Słomczyński[1]

Recenzenci: prof. dr hab. Tomasz Szarek, prof. dr hab. Jacek Tabor

**Praca została wyróżniona przez Radę Wydziału Matematyki i Informatyki UJ**

**magister matematyki** (w ramach Indywidualnych Studiów Matematyczno-Przyrodniczych)

Podmiot: Wydział Matematyki i Informatyki Uniwersytetu Jagiellońskiego

Tytuł pracy: *Zbieżność entropii*

Promotor: prof. dr hab. Wojciech Słomczyński

Rok: 2005

## Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych

W październiku 2005 roku rozpocząłem pracę w Katedrze Matematyki Akademii Ekonomicznej w Krakowie (aktualnie Uniwersytet Ekonomiczny w Krakowie) na stanowisku asystenta. Od 2014 roku pracuję w tej jednostce na stanowisku adiunkta.

[1] Tytuły naukowe podaję zgodnie ze stanem na rok składania wniosku.

# 2 Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.)

## 2.1 Główne osiągnięcie naukowe — cykl powiązanych tematycznie artykułów naukowych

### 2.1.1 Tytuł osiągnięcia

W pracy naukowej po uzyskaniu stopnia doktora koncentrowałem się na zagadnieniach związanych z teorią gier i optymalnością funkcji dobrobytu społecznego. Rezultatem moich badań jest cykl artykułów pod zbiorczym tytułem

**„Złożona dynamika zachowań agentów w sytuacjach decyzyjnych z bodźcami zewnętrznymi a optymalizacja dobrobytu społecznego — analiza przy użyciu narzędzi algorytmicznej i ewolucyjnej teorii gier"**

Prace te wskazuję jako osiągnięcia naukowe, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.).

### 2.1.2 Lista publikacji składających się na osiągnięcie naukowe

Na cykl składa się dziewięć przedstawionych poniżej publikacji naukowych. Dla każdej z nich podałem 5-letni współczynnik Impact Factor z 2023 roku (IF23, Clarivate 2024) oraz punktację z wykazu czasopism naukowych oraz recenzowanych materiałów z konferencji międzynarodowych opublikowaną przez Ministerstwo Nauki i Szkolnictwa Wyższego (MNiSW) dnia 5 stycznia 2024 r. Dodatkowo uwzględniłem najnowsze dane dotyczące cytowań (bez autocytowań) z baz Scopus oraz Web of Science (WoS) na dzień 20 marca 2025 r. (na podstawie raportu cytowalności sporządzonego przez Bibliotekę Uniwersytetu Ekonomicznego w Krakowie). Dla każdej publikacji wskazałem również mój procentowy udział w jej tworzeniu. Szczegółowy opis mojego wkładu w przygotowanie poszczególnych artykułów znajduje się poniżej listy publikacji, a także w Załączniku 4. Stosowne oświadczenia współautorów dotyczące mojego udziału w pracach zamieściłem w Załączniku 5.

**Część 1: Złożona dynamika zachowań agentów w grach**

**A1.** T. Chotibut, F. Falniowski, M. Misiurewicz and G. Piliouras. Family of chaotic maps from game theory. *Dynamical Systems*, 36(1), 48-63, 2021.
MNiSW: 40p., IF23: 0.5, liczba cytowań: Scopus: 16, WoS: 8

Mój wkład w przygotowanie publikacji oceniam na 30%.

**A2.** T. Chotibut, F. Falniowski, M. Misiurewicz and G. Piliouras. The route to chaos in routing games: When is price of anarchy too optimistic? *Advances in Neural Information Processing Systems*, **33**, 766-777, 2020.

MNiSW: 200p., IF23: materiały konferencyjne[2], liczba cytowań: Scopus: 17, WoS: 7

Mój wkład w przygotowanie publikacji oceniam na 25%.

**A3.** J. Bielawski, T. Chotibut, F. Falniowski, G. Kosiorowski, M. Misiurewicz and G. Piliouras. Follow-the-regularized-leader routes to chaos for online learning in congestion games. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *PLMR*, pages 925–935, 2021.

MNiSW: 200p., IF23: materiały konferencyjne[3], liczba cytowań: Scopus: 16, WoS: 9

Mój wkład w przygotowanie publikacji oceniam na 20%.

**A4.** J. Bielawski, T. Chotibut, F. Falniowski, M. Misiurewicz and G. Piliouras. Memory loss can prevent chaos in games dynamics. *Chaos* 34(1): 013146, 2024.

MNiSW: 140p., IF23: 2.7, liczba cytowań: Scopus: -, WoS: 1

Mój wkład w przygotowanie publikacji oceniam na 25%.

**A5.** F. Falniowski and P. Mertikopoulos. On the discrete-time origins of the replicator dynamics: from convergence to instability and chaos. *International Journal of Game Theory*, 54(1), 7, 2025.

MNiSW: 70p., IF23: 0.6, liczba cytowań: Scopus: -, WoS: -

[2]Ze względu na specyfikę *Advances in Neural Information Processing Systems* nie posiadają IF. Jednak, zgodnie np. z Google Scholar, jego indeks h5 (czasopism najczęściej cytowanych w ostatnich 5 latach) plasuje go w TOP10 czasopism (niezależnie od dyscypliny). Według *2022 CS Conference Impact Factors*, dostępnym na stronie health-nlp.com, impact factor tej konferencji jest szacowany na 23.27, zaś Impact Score z 2023r., wg research.com, to 40.6.

[3]Według wyliczeń Google Scholar indeks h5 plasuje je w TOP20 czasopism (niezależnie od dyscypliny). CS Conference Impact Factors wskazuje impact factor tej konferencji na 26.66, Impact Score z 2023r. to 30.1.

Mój wkład w przygotowanie publikacji oceniam na 55%.

**A6.** F. Falniowski. Entropy-based measure of statistical complexity of a game strategy. *Entropy*, 22(4), 470, 2020.

MNiSW: 100p., IF23: 2.2, liczba cytowań: Scopus: 2, WoS: 2

**Część 2: Optymalizacja dobrobytu społecznego a mikrobodźce graczy**

**B1.** O. Stark, M. Jakubek and F. Falniowski. Reconciling the Rawlsian and the utilitarian approaches to the maximization of social welfare. *Economics Letters* 122(3), 439-444, 2014.

MNiSW: 100p., IF23: 2.1, liczba cytowań: Scopus: 3, WoS: 2

Mój wkład w przygotowanie publikacji oceniam na 33,3%.

**B2.** O. Stark, F. Falniowski and M. Jakubek. Consensus income distribution. *Review of Income and Wealth* 63(4), 899-911, 2017.

MNiSW: 100p., IF23: 1.9, liczba cytowań: Scopus: 3, WoS: 3

Mój wkład w przygotowanie publikacji oceniam na 33,3%.

**B3.** O. Stark, J. Bielawski and F. Falniowski. A class of proximity-sensitive measures of relative deprivation. *Economics Letters* 160, 105-110, 2017.

MNiSW: 100p., IF23: 2.1, liczba cytowań: Scopus: 16, WoS: 16

Mój wkład w przygotowanie publikacji oceniam na 33,3%.

W ramach każdej z prac wieloautorskich razem ze współautorami: zaproponowałem hipotezy badawcze, stworzyłem modele matematyczne, udowodniłem twierdzenia, opracowałem wnioski, dokonałem przeglądu literatury. W każdej z prac, z wyłączeniem pracy **B1**, napisałem pierwszą wersję artykułu oraz uczestniczyłem w jego edycji. W artykułach **A3, A4, A5** byłem inicjatorem badań i autorem pierwszych rezultatów. W badaniach, których efektem była praca **B2** zauważyłem, że analogiczne wyniki jak w pracy **B1** powinny być prawdziwe dla pochodzących od Atkinsona [7] użyteczności izoelastycznych. Główne pomysły pracy **B3** były konsekwencją dyskusji z dr. Bielawskim. Byłem autorem korespondencyjnym artykułów **A4** i **A5**.

Prace w ramach omawianego cyklu mają charakter badań podstawowych i dotyczą interakcji pomiędzy wyborami graczy optymalizujących zachowanie w warunkach ograniczonej racjonalności a optimum dobrobytu społecznego.[4]

[4]Wyniki uzyskane w pierwszej części cyklu mają bezpośrednie zastosowanie w problemach związanych z obciążeniem/przeciążeniem rynków oligopolistycznych [2], modelowania migracji

W pracach **A1-A6**, w których wraz ze współautorami skupiam się na ewolucji zachowań uczących się graczy i ewolucji całego systemu, wykorzystuję narzędzia teorii gier (w szczególności algorytmicznej i ewolucyjnej teorii gier), dyskretnych układów dynamicznych i teorii ergodycznej. Narzędzia te pozwalają na precyzyjną analizę długoterminowej ewolucji zachowań graczy i ewolucji całej populacji. Co więcej, ze względu na interdyscyplinarną specyfikę problemu, przydatne okazują się również narzędzia wywodzące się z teorii informacji (praca **A6**).

W pracach **B1-B3**, w których głównym celem była analiza ewolucji optimum przy zmianie intensywności oddziałujących bodźców, używam klasycznych narzędzi optymalizacyjnych.

Mój wkład w dyscyplinę ekonomii i finansów opisuję w części 2.2 i podsumowuję w części 2.3.

## 2.2 Szczegółowy opis osiągnięć

Po uzyskaniu stopnia doktora w pracy naukowej koncentrowałem się na (teoretycznej) analizie dynamiki zachowań graczy (agentów, podmiotów gospodarczych) o ograniczonej racjonalności oraz związku pomiędzy optymalnymi wyborami graczy i maksymalizacją dobrobytu społecznego.

Przed przejściem do szczegółowego opisu osiągnięć zwrócę uwagę na kluczowe rezultaty przeze mnie uzyskane oraz ich kontekst.

1. Przeprowadziłem szczegółową analizę ewolucji populacji graczy w systemie ekonomicznym, gdzie dynamika gry zadawana jest poprzez proces uczenia się uczestników gry. W pracach **A1-A5** skupiłem się na grach zatorowych (ang. *congestion games*). Wyniki te znajdują zastosowanie w problemach, gdzie uczestnicy gry (uczestnicy rynku, podmioty gospodarcze) dysponując niepełną informacją (lub będąc ograniczenie racjonalnymi) konkurują z innymi podmiotami o zasoby. W szczególności można je wykorzystać w tak zróżnicowanej problematyce ekonomicznej jak: modele oligopoli [1, 2, 16, 32, 35, 78], rynki kryptowalut [74], problem zmowy cenowej algorytmów [19, 23, 71], czy gospodarka miejska (*urban economics*) [3, 91, 110, 117]

2. Pokazałem, że wzrost tempa uczenia się graczy (bądź zwiększenie wagi przykładanej przez nich do wyników z przeszłości) w grach zatorowych może do-

[15, 107], rynkach kryptowalut [74] czy problemach zatorowych w ruchu ulicznym [32]. Zastosowanie tych wyników nie ogranicza się do ekonomii znajdując swoje zastosowanie w innych naukach społecznych czy przyrodniczych — w obszarach, gdzie modeluje się ewolucję populacji agentów — podmiotów gospodarczych, osób, użytkowników internetu czy genów. Warto zauważyć, że prace **A2**, **A3** zostały wskazane w technicznym raporcie dotyczącym zagrożeń dla wieloagentowych modeli zaawansowanej AI [54].

prowadzić do destabilizacji systemu i chaotycznych zachowań zarówno jednostek, jak i całej populacji graczy. Może to skutkować wysokimi kosztami dla uczestników gry i nieefektywnością systemu.

3. Wskazałem problem nieprzewidywalności dynamik zadawanych przez wieloagentowe modele zaawansowanej sztucznej inteligencji (patrz [54], str. 30-31).

4. Uzyskałem wyniki, które rzuciły nowe światło na zależność pomiędzy optymalnymi wyborami egoistycznych graczy a zapotrzebowaniem systemu w sytuacji gdy gracze są w ograniczonym stopniu racjonalni. Potwierdza to potrzebę precyzyjnej analizy dynamiki gry wykraczającej poza standardową analizę równowag, nawet gdy równowagi te są Pareto optymalnymi rozwiązaniami gry.

5. Pokazałem, że w niektórych problemach teorii gier, np. opisujących tzw. przeciążony oligopol [2], modele konkurencji na sieciach [32], modele ruchu drogowego czy aukcji online [18, 29], możemy obserwować średnie (po czasie) zachowanie systemu (populacji) sugerujące jego optymalność (zbieżność do równowagi Nasha) przy jednoczesnych wysokich kosztach wynikających z ewolucji systemu *dzień po dniu*.

6. Pokazałem, że zapominanie (dyskontowanie) przeszłości może mieć pozytywne własności predykcyjne dla zachowań graczy. Wynik ten sugeruje, że ujednolicenie funkcji kosztów (wypłat) ze strategii, np. poprzez transfery pieniężne, może zwiększać dobre własności predykcyjne systemu.

7. Zaproponowałem i przeanalizowałem model prostej gry populacyjnej z homogenicznymi graczami (agentami) uczącymi się przez wzmocnienie. Model ten w naturalny sposób rozszerza się na przypadek agentów o różnorodnych przekonaniach i różnym warunkowaniu swoich zachowań na podstawie historycznych danych [13]. Pozwala on wyjaśnić tzw. *El Farol bar problem* [6] i ma bezpośredni związek z tzw. *complexity economics* [9, 25, 27, 37, 45].

8. Zaproponowałem teorioinformacyjną miarę złożoności strategii wykorzystywanych przez graczy.

9. Wykazałem, że optymalne wybory planistów izoelastycznych (czyli np. utylitarystycznych [55] czy rawlsowskich [94, 95]) mogą być zgodne jeśli agenci uwzględniają w swoich funkcjach użyteczności bodźce zewnętrzne wynikające z nierówności dochodowych. Presja nierówności dochodowych (i awersja jednostek do nierówności) pozwala zunifikować wybory planistów. Stanowi to wyjaśnienie (w duchu Harsanyi'ego [57]) zgodności wyborów planistów konkuren-

cyjne do pochodzącego od Arrowa wytłumaczenia bazującego na awersji do ryzyka [5].

10. Na podstawie naturalnych aksjomatów, wyprowadziłem klasę uogólnionych miar względnego niedostatku w systemie ekonomicznym, wrażliwych na zmiany dochodowe wśród bogatszych. Są one z powodzeniem wykorzystywane do pomiaru nierówności dochodowych [40, 41, 43, 50, 68, 108, 109].

W badaniach, których wynikiem były prace **A1-A6** skoncentrowałem się na analizie długookresowej dynamiki zachowań graczy (agentów) w grach rozgrywanych wielokrotnie. W takich grach agenci mogą w trakcie gry uczyć się jej warunków, wypłat czy zachowań innych uczestników gry. Dopuszczenie w modelu możliwości uczenia się pozwala osłabić restrykcyjne założenia klasycznej teorii gier takie jak racjonalność graczy. W serii artykułów napisanych ze współpracownikami z wiodących ośrodków zagranicznych analizowałem i szczegółowo opisałem, w jaki sposób uczenie się graczy, a co za tym idzie ewolucja reguł gry i ograniczeń z nich wynikających, może wpłynąć na złożone zachowania graczy i ewolucję systemu (lub stan populacji). W pracach **A1-A4** pokazałem ze współautorami, że w grach z potencjałem [82], które uznawane są za gry o prostej dynamice, zachowania graczy mogą być bardzo skomplikowane. W szczególności badałem konsekwencje wzrostu tempa uczenia się, zapotrzebowania systemu czy wielkości populacji na dynamikę gry i wybory graczy. Zrobiłem to dla modeli (i ich modyfikacji) pochodzących z uczenia maszynowego, gier behawioralnych i ewolucyjnej teorii gier. Ponadto zbadałem implikacje zapominania (dyskontowania) przez graczy przeszłości (algorytmy sEWA [49, 66]) i konsekwencje przejawiania różnej awersji do ryzyka (algorytmy FTRL [77]). Pokazałem, że nawet gdy analiza średniego zachowania populacji sugeruje zbieżność do równowagi Nasha, zachowania graczy, jak i ewolucja systemu (czy całej populacji graczy) *dzień po dniu* mogą być dowolnie skomplikowane i nieprzewidywalne.[5] Uzupełniając to podejście, w pracy **A5** napisanej z Panayotisem Mertikopoulosem (CNRS, Francja) w kontekście gier populacyjnych porównałem dynamikę wprowadzaną w prostej grze antykoordynacyjnej przez uczących się graczy z dynamiką wprowadzoną przez imitacyjny protokół rewizji [63, 106]. Wykazałem, że ten model wykazuje chaotyczne zachowania. Pokazałem również, że zastosowania modeli ewolucyjnej teorii gier w ekonomii fundamentalnie różnią się od jej biologicznego pierwowzoru. Wreszcie, w pracy **A6**, zaproponowałem nową miarę złożoności strategii, którą można wykorzystać do badania złożoności strategii

[5] Uzyskane wyniki można przenieść na grę z heterogenicznymi agentami, gdzie ich heterogeniczność jest konsekwencją zarówno ich różnorodnych przekonań, jak i różnej szybkości uczenia się (intensywności wyboru) [13].

używanych przez graczy w grach powtarzanych.[6]

Graczem analizowanym w pierwszej części cyklu może być dowolny podmiot: człowiek, firma, podmiot gospodarczy. Teoria gier pozwala badać interakcje pomiędzy graczami. Ta interakcja może przeciwdziałać otrzymaniu (i utrzymaniu) optimum społecznego (mierzonego np. poprzez utylitarystyczną funkcję dobrobytu społecznego). W pracach składających się na pierwszą część przedstawianego cyklu pokazałem, że nawet gdy optimum społeczne jest tożsame z jedyną równowagą Nasha populacja może być daleko od tej równowagi, generując znaczne koszty społeczne i podważając wnioski płynące z analizy systemu w równowadze. W drugim podcyklu, na które składają się prace **B1-B3**, skupiłem się na wpływie bodźców na graczy. W szczególności ze współautorami analizowałem jak uwzględnienie w wypłatach jednostek porównań dochodowych w populacji, może wpłynąć na wybory planistów społecznych. Pokazałem, że jeśli członkowie populacji są odpowiednio czuli na nierówności dochodowe, to planista będzie dążyć do wyrównania ich dochodów niezależnie od własnego poziomu awersji do nierówności (prace **B1**, **B2**). W ostaniej pracy cyklu (praca **B3**) zaproponowaliśmy uogólnioną miarę względnego niedostatku czułą na różnice dochodowe wewnątrz populacji.

### 2.2.1 Złożona dynamika zachowań agentów w grach

Przechodzę teraz do szczegółowego opisu moich osiągnięć.

**Postawienie problemu i jego kontekst.** Teorię gier tradycyjnie rozwijano jako teorię strategicznych interakcji między agentami, którzy są racjonalni. Odejście od paradygmatu racjonalności podmiotów gospodarczych było możliwe m.in. dzięki ewolucyjnej teorii gier, która — motywowana koncepcjami zaczerpniętymi z ewolucji biologicznej — poprzez analizę zachowań graczy (kolejnych ich pokoleń), którzy nie muszą być racjonalnymi ani nawet świadomymi decydentami, stara się wyjaśnić, w jaki sposób równowaga może powstać i być dobrym odzwierciedleniem rzeczywistych zachowań graczy w dłuższej perspektywie. Gdzieś pośrodku między klasyczną a ewolucyjną teorią gier znajdują się modele algorytmicznej teorii gier. Bazują one zazwyczaj na koncepcji uczenia się uczestników gry, które uwzględniają zachowania adaptacyjne zorientowanych na cel (niekoniecznie wysoce racjonalnych) graczy Różnorodność opisywanych w ten sposób podmiotów obejmuje graczy od konsumentów do decydentów na Wall Steet, a modele mogą służyć do analizy rozmaitych struktur od banków po rząd. Modele o których mowa nie opierają się na założeniu, że gospodarka zbiega do z góry określonego stanu równowagi. Za-

[6]Prace **A1-A3** oraz **A6** były rezultatem grantu *Miary statystycznej złożoności i nieprzewidywalności strategii oparte na pojęciu entropii w grach ekonomicznych*, którego byłem kierownikiem. Był on realizowany w ramach konkursu Sonata 11 w Narodowym Centrum Nauki.

miast tego w dowolnym momencie każdy podmiot działa zgodnie ze swoją obecną sytuacją, stanem otaczającego go świata i regułami determinującymi jego zachowania [46]. W każdym stanie układu agenci wybierają swoje działania według jakiegoś algorytmu i aktualizują swoje strategie, korzystając z ustalonej reguły (uczenia się, reagowania na otrzymywane bodźce). Dynamiczna perspektywa pozwala zatem lepiej uzasadnić podstawy teorii gier, koncepcje równowagi, oraz modelować empirycznie zaobserwowane zachowania [97]. Kluczowym pytaniem z perspektywy ekonomii jest to, jakie działania podejmują gracze pod wpływem danych bodźców. Podejście polegające na uwzględnieniu w modelu uczenia się graczy zapewnia ścisły kontekst i opis takich pytań, poprzez skoncentrowanie uwagi na wpływie uczenia i adaptacji na ewolucję ich zachowań.

Podstawowym rozwiązaniem w teorii gier niekooperacyjnych jest równowaga Nasha, spełniająca wymóg, by każdy agent wybrał strategię, która jest optymalna, biorąc pod uwagę wybory pozostałych. Pomimo centralnej roli tej koncepcji w zastosowaniach teorii gier w ekonomii, tradycyjne, racjonalistyczne uzasadnienie jej zastosowania nie jest szczególnie przekonujące. Opiera się na nierealistycznych założeniach dotyczących uczestników gry. W przypadku dużych populacji powtarzanie gry może jednak prowadzić do dynamiki gry zbieżnej do równowagi, nawet jeśli informacje i możliwości uczestników gry są dość ograniczone [17, 70, 84, 104]. Jednakże począwszy od klasycznych wyników Harta i MacCollela [59–61] wiadomo, że nie istnieją procedury uczenia zadające dynamiki zbieżne w każdej grze. Zatem nie można liczyć na ogólne twierdzenia dotyczące zbieżności [10, 59, 61, 62].

W badaniach, których wynikiem są prace **A1-A4** koncentrowałem się na analizie długoterminowego zachowania populacji graczy, którzy dokonują wyboru strategii zgodnie z algorytmami uczenia poprzez wzmacnianie (ang. *reinforcement learning*), powszechnie stosowanymi do modelowania ludzkich zachowań w kontekstach ekonomicznych.[7] Najnowsze rezultaty ( [13,28,79,90,105,113], prace **A1-A4**) wskazują, że złożone, charakteryzujące się brakiem stabilizacji, zachowanie graczy stosujących uczenie przez wzmacnianie są raczej regułą niż wyjątkiem. Wówczas wybory graczy mogą być nieoptymalne i skutkować dużymi kosztami dla agentów i społeczeństwa. Eksperymenty opisywane w pracach przynależnych do nurtu teorii gier behawioralnych [65–67] potwierdzają wyniki prac **A1-A4**.

[7] Wybór uczenia poprzez wzmacnianie jest naturalny, a model ugruntowany i potwierdzony badaniami. Erev i Roth [42, 97] wykazali, że algorytmy uczenia się przez wzmacnianie mają lepszą moc predykcyjną niż standardowa analiza równowagi i dokładnie opisują zachowanie podmiotów gospodarczych. Niedawno Mäs i Nax [75] wykazali, że modele uczenia się przez wzmacnianie często odpowiadają zachowaniom osób w grach eksperymentalnych, ponieważ osoby z większym prawdopodobieństwem wybiorą określone działania po otrzymaniu dużej korzyści z ich wykonania. Coraz więcej badań wskazuje, że w rzeczywistych scenariuszach uczenie się przez wzmacnianie przewyższa klasyczną dynamikę adaptacyjną [23, 81].

Aby zrozumieć różnorodność możliwych zachowań agentów, warto analizować możliwie proste gry. Prace **A1-A5** skupiają się na grach, w których gracze (populacja graczy) mają do wyboru dwie (czyste) strategie (którymi są zasoby, ścieżki itp.). Koszt realizacji każdej strategii rośnie wraz z liczbą (odsetkiem) uczestników gry używających danej strategii. Grami, które modelują takie sytuacje są gry zatorowe (ang. *congestion games*).[8] Gry zatorowe to klasa gier wprowadzona przez amerykańskiego ekonomistę Rosenthala [96], gdzie koszt strategii zależy od całkowitej liczby (odsetka) wybierających ją agentów. Gry te znajdują zastosowanie w różnych kontekstach aplikacyjnych, jak choćby w analizie problemu oligopolu [2], modelowania sieci komunikacyjnych, zatorów komunikacyjnych czy procesów ekologicznych w siedliskach przyrodniczych (np. problem wspólnego pastwiska). Obok szerokiego wachlarza zastosowań, wybór gier zatorowych jako przedmiotu analizy wynika z faktu, że każda taka gra posiada swoją funkcję potencjału (jest więc tzw. grą z potencjałem, [82]), której minimami są równowagi Nasha. Takie gry są bardzo dobrze zbadane i uznawane zwykle za klasę gier o przewidywalnej dynamice. Ze względu na możliwość dokładnej analizy i klarowność wniosków rozpatrywane w pracach gry posiadają dokładnie jedną równowagę Nasha. Dzięki takiemu wyborowi przedmiotu badań mogliśmy skupić się na bezpośrednich implikacjach dynamiki.

**Przykład poglądowy.** Zanim formalnie omówię otrzymane rezultaty, opiszę je obrazowo za pomocą przykładu. Pozwoli on lepiej zrozumieć ekonomiczne konsekwencje uzyskanych wyników. Rozpatrzmy oligopol, w którym gracze dysponują ograniczoną informacją. Firmy maksymalizują swoje wypłaty (lub minimalizują koszty) używając pewnego rodzaju algorytmów uczenia się.[9] Firmy mają dwie dostępne strategie (czyste) — mogą wybierać jeden z dwóch zasobów, z których wypłata (koszt) zależy od tego jak wiele firm z niego korzysta. Wtedy kontekst (i rezultaty) prac **A1-A5** można rozumieć jako:

A1. duopol, w którym obie firmy mogą wybierać strategię mieszaną. Analizujemy konsekwencje wzrostu tempa uczenia zgodnie z algorytmem uczenia ze wzmocnieniem (wykorzystywanym szeroko w ekonomii [23, 36, 42, 56, 98]) na zachowania firm i ewolucję systemu.

A2. oligopol z kontinuum firm (populacją graczy), w którym firmy używają jedynie strategii czystych, a odsetek firm używających danej strategii zadaje

[8] Mimo olbrzymiej literatury w języku angielskim znacząca część pojęć wykorzystywanych w przedkładanym opisie nie ma ugruntowanego w środowisku nazewnictwa w języku polskim. Dlatego wprowadzając pojęcie po raz pierwszy podaję tłumaczenie własne wraz z angielskim oryginałem. W przypadku algorytmów podaję również akronim, który następnie stosuję w opisach.

[9] Takie podejście jest szczególnie istotne w gałęziach gospodarki, w których ceny aktualizowane są często, np. w sposób ciągły, a ludzi zastępują wyspecjalizowane algorytmy [22, 73].

stan populacji. Analizujemy zachowania systemu w reakcji na wzrost zapotrzebowania systemu (np. wzrost wielkości popytu).

A3. analizujemy czy wyniki z **A2** można przenieść na przypadek gdy firmy mają inną (niż w pracy **A2**) awersję do ryzyka (używają innych regularyzatorów niż entropia Shannona).

A4. badamy jaki wpływ na wyniki pracy **A2** miałoby zapominanie (dyskontowanie) przeszłości — gdy najistotniejsze dla graczy są najnowsze dane.

A5. oligopol z pracy **A2**, w którym firmy nie używają z góry zadanego algorytmu. Zamiast tego aktualizują swoje wybory na podstawie porównywania uzyskiwanych wypłat z wypłatami innych członków populacji.

## Złożona dynamika zachowań agentów o ograniczonej racjonalności — uczenie ze wzmocnieniem

**Dynamika gry zatorowej z dwoma agentami wykorzystującymi algorytm multiplikowanych wag.** W pracy **A1** analizowana jest dynamika gry zatorowej, w której dwóch graczy ma dwie dostępne strategie czyste i obaj stosują algorytm aktualizowanych multiplikatywnych wag (ang. *Multiplicative Weights Update (MWU)*)[10] wybierając przy jego pomocy strategie mieszane w kolejnej grze.

Niech $x$ będzie prawdopodobieństwem, że pierwszy gracz wybierze pierwszą strategię (drugą strategię wybierze więc z prawdopodobieństwem $1-x$) i niech $y$ będzie prawdopodobieństwem, że drugi gracz wybierze pierwszą strategię (zatem $(x, 1-x)$ jest strategią mieszaną używaną przez pierwszego gracza, a $(y, 1-y)$ przez drugiego). Zakładamy, że koszt strategii jest proporcjonalny do jej obciążenia. Niech $c(i,j)$ będzie (oczekiwanym) kosztem gracza $i$ wykorzystującego strategię o numerze $j$. Aktualizacja (mieszanych) strategii graczy w chwili $n+1$ w wyniku stosowania MWU z tempem uczenia się $\varepsilon \in (0,1)$ jest zadana przez równania

$$x_{n+1} = \frac{x_n(1-\varepsilon)^{c(1,1)}}{x_n(1-\varepsilon)^{c(1,1)} + (1-x_n)(1-\varepsilon)^{c(1,2)}},$$

$$y_{n+1} = \frac{y_n(1-\varepsilon)^{c(2,1)}}{y_n(1-\varepsilon)^{c(2,1)} + (1-y_n)(1-\varepsilon)^{c(2,2)}}.$$

Jaka jest dynamika zachowań graczy w tej grze? Pokazujemy, że zależy ona od tego, czy zaczynają od tej samej strategii, czy nie. W artykule wykazano, że jeśli obaj gracze używają tej samej strategii mieszanej (ale nie czystej), wówczas dynamika

[10] MWU to algorytm zaproponowany jeszcze przez Hannana [56] o szerokich zastosowaniach w uczeniu maszynowym, ekonomii czy biologii [4, 26, 38].

gry zależy od tempa uczenia się graczy. Tak więc, dopóki jest ono wystarczająco małe, widzimy zbieżność do (jedynej) równowagi Nasha. Następnie, gdy tempo uczenia przekroczy pewien próg, stabilność równowagi zostanie utracona. W przypadku symetrycznym gdy funkcje kosztu obu strategii są takie same, obserwujemy zbieżność, z wyjątkiem co najwyżej przeliczalnej liczby warunków początkowych, dla których iteracje trafiają w odpychającą równowagę Nasha, trajektorii układu (czyli wybieranych przez graczy strategii mieszanych) do orbity okresowej o okresie 2. W przeciwnym razie, gdy koszty są różne, istnieje progowa wartość tempa uczenia się, po przekroczeniu której obserwować będziemy zachowania chaotyczne[11], zachowania okresowe o dowolnym okresie, skomplikowane zachowania trajektorii w układzie dynamicznym i brak zbieżności do równowagi Nasha. Wynik uzyskany w pracy **A1** jest jednym z pierwszych wykazujących występowanie chaosu w grach zatorowych. Co więcej, chaotyczne zachowania graczy mają miejsce, gdy *średnio* gracze są blisko równowagi Nasha, tzn. średnie po czasie zachowanie graczy zbiega do równowagi Nasha (niezależnie od ich tempa uczenia się). Z kolei gdy agenci zaczną od różnych strategii początkowych, dynamika jest zbieżna do jednej z czystych równowag, czyli $(1, 0)$ lub $(0, 1)$. Jeżeli pierwszy gracz częściej będzie korzystał z pierwszej strategii, niż gracz drugi, wówczas dynamika gry zbiega do profilu strategii, w którym pierwszy gracz zawsze będzie korzystał z pierwszej a drugi gracz z drugiej strategii. Jeśli drugi gracz częściej stosuje pierwszą strategię niż gracz pierwszy, to (w granicy) zamienią się rolami.

Wyniki pracy **A1** implikują, że gdy gracze są homogeniczni i wybierają początkowo tą samą strategię mieszaną, odpowiednio szybka nauka prowadzi do złożonych i nieprzewidywalnych zachowań uczestników gry. Praca ta dotyczyła gry z dwoma uczestnikami. Pytanie co się wydarzy gdy w grze bierze udział więcej graczy? Co gdy gracze dokonują różnych wyborów strategii początkowych w takiej sytuacji? Odpowiedzi na te pytania można znaleźć w kolejnej pracy (praca **A2**).

**Model z kontinuum graczy i problem utraty efektywności.** Fundamentalnym problemem teorii gier, który wywarł ogromny wpływ na nauki społeczne, jest zależność pomiędzy egoistycznymi wyborami jednostek (odzwierciedlanymi poprzez równowagi Nasha) a rozwiązaniami społecznie optymalnymi (gdzie maksymalizowana jest utylitarystyczna funkcja dobrobytu społecznego [57,58]). Jak zmierzyć cenę jaką

[11] W literaturze znaleźć można wiele definicji chaosu. W przedstawianym osiągnięciu pisząc o chaosie będę miał na myśli **chaos w sensie Li-Yorke'a**, który gwarantuje istnienie nieprzeliczalnego zbioru punktów, które tworzą tzw. pary splątane — każde dwie trajektorie punktów z tego zbioru zbliżają i oddalają się od siebie nieskończenie wiele razy. W kontekście omawianym w pracach implikuje on inne chaotyczne cechy układu, takie jak orbity okresowe dowolnych okresów czy dodatnią entropię topologiczną. Lokalnie można też zaobserwować czułą zależność dynamiki od warunków początkowych.

ponosi populacja w wyniku samolubnych decyzji poszczególnych graczy? Odpowiedź na to pytanie można uzyskać porównując najgorszy stacjonarny scenariusz długoterminowy (koszt w najgorszej, najbardziej kosztownej, równowadze Nasha) z kosztem w optimum społecznym systemu. Służy do tego miara zwana ceną anarchii (ang. *Price of Anarchy*) [72], zdefiniowana jako:

$$\mathrm{PoA} = \frac{\sup_{x \in \mathrm{NE}} SC(x)}{\min SC(x)},$$

gdzie koszt społeczny profilu $SC(x)$ jest sumą kosztów graczy (zawsze dodatnią), a NE zbiorem równowag Nasha tej gry, lub w kategoriach funkcji dobrobytu społecznego

$$\mathrm{PoA} = \frac{\inf_{x \in \mathrm{NE}} SW(x)}{\max SW(x)},$$

gdzie $SW$ jest utylitarystyczną funkcją dobrobytu społecznego (czyli sumą wypłat graczy, zawsze dodatnią).

Niska cena anarchii (bliska 1) oznacza, że wszystkie równowagi Nasha są bliskie optimum społecznego, a zatem jakakolwiek zbieżna dynamika uczenia się wystarczy, aby (asymptotycznie) działanie (a co za tym idzie efektywność) systemu było bliskie optimum. Oznacza to, że niska cena anarchii jest bardzo korzystna w problemach ekonomicznych z dużą liczbą podmiotów gospodarczych, gdzie ze względu na złożoność problemu (i ograniczony wpływ na długofalowe zachowania graczy) możemy mieć jedynie pewne gwarancje zbieżności dynamiki do równowagi, ale bez informacji do której. Jednym z osiągnięć badań nad ceną anarchii było opracowanie ścisłych (górnych) ograniczeń na PoA (przynajmniej w przypadku gier zatorowych), które są niezależne od topologii sieci lub ilości graczy [99, 100]. Jednocześnie uzyskane w ostatnich latach wyniki [30,31,47] sugerują, że wzrost wielkości populacji (wzrost zapotrzebowania) może zmniejszyć cenę anarchii. Pokazano, że jeśli zapotrzebowanie systemu rośnie, to najgorszy stabilny scenariusz (równowaga o największym koszcie/najniższej wypłacie) nie jest odległy od optimum. Stąd analiza równowag sugeruje, że efektywność systemu jest bliska optymalnej dla dużych populacji. Jednak dopiero praca **A2** była pierwszą, w której przeprowadzono systematyczne badanie wpływu zwiększenia całkowitego zapotrzebowania na zachowania uczących się graczy. W konsekwencji wskazany został niepożądany efekt wzrostu wielkości zapotrzebowania systemu: nawet w przypadku najprostszych nietrywialnych gier — gier zatorowych z dwiema dostępnymi strategiami czystymi, gdy całkowite zapotrzebowanie systemu rośnie, dynamika uczenia nieuchronnie staje się niestabilna, zazwyczaj chaotyczna.

Wynik przez nas uzyskany został opublikowany w materiałach jednej z najważniejszych konferencji w dziedzinie uczenia maszynowego (patrz np. conference-rankings)

— *Conference on Neural Information Processing Systems.* W pracy analizuję ze współautorami zachowanie graczy używających algorytmu MWU w grach zatorowych z nieprzeliczalną populacją. Dla dowolnej bezatomowej[12] gry zatorowej i dowolnego tempa uczenia się, pokazujemy, że istnieje progowa pojemność systemu (wielkość populacji) taka, że gdy całkowite zapotrzebowanie (popyt) ją przekroczy, system staje się niestabilny, a wszelkie wnioski bazujące na analizie stanu równowagi nie są aplikowalne do tej sytuacji. Udowodnione zostało istnienie zarówno zachowań okresowych, jak i chaotycznych, i podano formalne gwarancje dotyczące warunków, w jakich one powstają. Pomimo tej nieprzewidywalności, średnie przepływy (między strategiami czystymi) w populacji wykazują regularność zbiegając do równowagi niezależnie od zapotrzebowania systemu. Jednocześnie system jest nieefektywny, a średni koszt w czasie może być dowolnie wysoki. Dzieje się to w sytuacji gdy cena anarchii jest optymalna, czyli PoA = 1.

**Model z pracy A2.** Rozważamy bezatomową grę zatorową o dwóch strategiach z kontinuum graczy (agentów), gdzie każdy z nich stosuje MWU do zaktualizowania swoich strategii. Wielkość całkowitego przepływu (zapotrzebowania) wszystkich graczy wynosi $N$, a każdy gracz kontroluje jego infinitezymalną część. Oznaczmy odsetek graczy przyjmujących pierwszą strategię w chwili $n$ jako $x_n$. Przyjmuje się, że koszt każdej strategii (zasobu, połączenia, ścieżki) jest proporcjonalny do obciążenia tej strategii. Oznaczając przez $c(j)$ koszt wyboru strategii $j$ (kiedy frakcja $x$ graczy wybiera pierwszą strategię), przy współczynnikach proporcjonalności $\alpha, \beta > 0$, otrzymujemy

$$c(1) = \alpha N x, \qquad c(2) = \beta N(1-x). \tag{1}$$

W momencie $n+1$ zakładamy, że gracze znają koszt strategii z chwili $n$, czyli znają odsetek graczy używających każdej ze strategii (odpowiednio $x_n$ i $1-x_n$). Mając tę wiedzę agenci aktualizują swoje wybory. Ponieważ jest ich kontinuum, zrealizowany przepływ (podział) jest dokładnie opisany przez rozkład $(x_n, 1-x_n)$. Aktualizacji prawdopodobieństw dokonujemy za pomocą MWU, który zadaje odsetek graczy używających w chwili $n+1$ strategii pierwszej jako

$$x_{n+1} = \frac{x_n(1-\varepsilon)^{c(1)}}{x_n(1-\varepsilon)^{c(1)} + (1-x_n)(1-\varepsilon)^{c(2)}} = \frac{x_n}{x_n + (1-x_n)(1-\varepsilon)^{c(2)-c(1)}} \tag{2}$$

dla ustalonego tempa uczenia się $\varepsilon \in (0,1)$.

W pracy pokazałem, że dla małych wartości zapotrzebowania systemu nadal mamy zbieżność do przyciągającej równowagi Nasha.[13] Niemniej jednak w przy-

[12] Czyli takiej, w której zmiana zachowania któregokolwiek gracza nie może wpłynąć na zmianę wypłat innych graczy.

[13] Jedynym istotnym parametrem jest odsetek graczy stosujących pierwszą strategię w stanie

padku dużego zapotrzebowania zachowanie systemu jest znacznie bardziej skomplikowane. Rozważmy najpierw przypadek, gdy koszty obu strategii opisuje ta sama funkcja kosztu ($\alpha = \beta$), a jedyną równowagą Nasha jest $(\frac{1}{2}, \frac{1}{2})$. Wtedy, wraz ze wzrostem zapotrzebowania systemu równowaga Nasha, tożsama z optimum społecznym (PoA = 1), traci stabilność, co podważa analizy dotyczące optymalnej długoterminowej efektywności systemu. Po utracie stabilności układ pozostaje jednak w pewnym stopniu przewidywalny — wszystkie trajektorie (z wyjątkiem przeliczalnie wielu wpadających w odpychającą równowagę Nasha) są przyciągane przez orbitę okresową o okresie 2 symetryczną względem $\frac{1}{2}$. Jednocześnie średni koszt społeczny jest bliski najgorszego przypadku dla gry jednoetapowej. Sytuacja może być więc tak zła (tak kosztowna), jak to tylko możliwe.[14] A co gdy funkcje kosztu się różnią? Wtedy wzrost zapotrzebowania systemu doprowadzi do okresowych zachowań o dowolnych okresach, chaotycznych zachowań graczy (chaos w sensie Li-Yorke'a), a także (lokalnie) czułej zależności od warunków początkowych. Pociąga to za sobą konsekwencje w grach bezatomowych. Gdy gra jest asymetryczna (koszty są asymetryczne), czyli gdy równowaga wewnętrzna nie jest podziałem 50%-50%, zwiększenie całkowitego zapotrzebowania systemu nieuchronnie doprowadzi do chaotycznego zachowania, niezależnie od postaci funkcji kosztu. Zatem przewidzenie zachowania populacji graczy staje się prawie niemożliwe. Dodatkowo, choć przy dużym obciążeniu systemu (dużej populacji) można zaobserwować pewne zachowania okresowe, to nawet niewielkie zmiany wartości parametrów mogą prowadzić do zupełnie innego zachowania.[15] Co zaskakujące, wszystko to dzieje się, gdy zachowanie średnie w czasie wydaje się optymalne — średnia po czasie zbiega do (jedynej) równowagi Nasha niezależnie od szybkości uczenia się stosowanej przez graczy, funkcji kosztów, na które patrzą, lub stanu początkowego systemu. Za-

równowagi. Pierwszą zaletą tego sformułowania jest to, że odsetek graczy stosujących każdą ze strategii w stanie równowagi jest niezależna od poziomu $N$. Drugą zaletą jest fakt, że cena anarchii tych gier wynosi dokładnie 1, niezależnie od $\alpha, \beta$, i $N$. Dlatego **nasz model stanowi naturalny punkt odniesienia** dla porównania analizy w stanie równowagi, która sugeruje optymalny koszt społeczny, ze średnim (po czasie) kosztem społecznym wynikającym z braku zbieżności dynamiki uczenia się, który, jak pokazujemy, może być tak duży jak to tylko możliwe.

[14] Upodmiotowiając uzyskany wynik powróćmy do przykładu z oligopolem. W tej sytuacji kontinuum podmiotów gospodarczych (firm) tworzy populację graczy. Omawiany wynik oznacza, że jeśli zapotrzebowanie systemu jest odpowiednio duże, to wybory populacji będą dalekie od optimum (czyli podział $50\% - 50\%$). Dla ustalenia uwagi jeśli przyciągająca orbita to $\{(\frac{1}{3}, \frac{2}{3}), (\frac{2}{3}, \frac{1}{3})\}$, w oligopolu populacja będzie na zmianę realizować stan bliski pierwszemu bądź drugiemu z nich. Prowadzić to będzie do suboptymalnych wypłat dla każdej z firm.

[15] W kontekście oligopolu z kontinuum graczy oznaczać to będzie, że wzrost zapotrzebowania (popytu) może zdestabilizować zachowanie populacji firm, powodując złożone zachowania graczy, chaotyczną ewolucję populacji firm. Spowoduje to znaczące straty. Jednocześnie analiza średniego zachowania sugerować będzie dobre zachowanie dynamiki.

tem patrząc z poziomu makro, na średnie zachowanie populacji, można sądzić, że zachowanie systemu jest bliskie optymalnemu. Co więcej, średnie zachowanie populacji jest podobne do średniego zachowania w grach o sumie zero, gier uznawanych za skrajnie odmienne od gier zatorowych. Niemniej jednak przyjrzenie się bliżej (rzeczywistej) dynamice daje zupełnie inne wnioski. Wymienione rezultaty można rozszerzyć na przypadek wielu strategii (zasobów, ścieżek) jak również na nieliniowe koszty (patrz Appendix pracy **A2**). Praca ta dała podstawy do dalszej analizy. W szczególności odpowiedzi wymagały pytania czy na chaotyczność systemu ma wpływ wybór algorytmu postępowania bądź założenie homogeniczności populacji? W kolejnych pracach badaliśmy jaki wpływ na dynamikę zachowań graczy (i populacji) ma odejście od tych założeń. W dalszej części opisu omówię konsekwencje wyboru innych algorytmów (praca **A3**) czy zapominania przeszłości (praca **A4**). Zanim to zrobię warto jednak nadmienić, że wyniki tu przedstawione można uogólnić na przypadek heterogeniczny. W artykule [13], przyjętym do publikacji w *Proceedings of National Academy of Sciences*, formalnie (analitycznie) udowodniliśmy zjawisko pojawienia się „inteligencji zbiorowej" będącej konsekwencją dynamiki uczenia się w dużej populacji graczy w grze inspirowanej problemem *El Farol Bar* [6].[16] W naszym modelu każdy gracz uczy się i dostosowuje swoje strategie w dyskretnym czasie, zgodnie ze standardową klasą MWU. Model dostarcza dwa źródła heterogeniczności w populacji graczy. Po pierwsze, każdy gracz może zacząć od różnych przekonań co do tego, jakie działanie powinien wybrać. Aby uwzględnić tę zmienność każdemu graczowi przypisaliśmy własny typ. Po drugie, każdy typ gracza ma swoją własną intensywność adaptacji, czyli szybkość uczenia się; niektórzy gracze mogą szybko dostosować się do sygnałów dotyczących kosztów (szybko uczący się), podczas gdy inni mogą wykazywać większą cierpliwość w aktualizowaniu swoich przekonań (wolno uczący się gracze). Stan populacji opisuje rozkład prawdopodobieństwa przekonań wśród wszystkich typów graczy, których może być skończenie lub nieskończenie wiele, co odpowiada heterogenicznym zachowaniom związanym z uczeniem się. Wedle naszej wiedzy to pierwszy taki model z różnymi źródłami heterogeniczności graczy. Jednocześnie analitycznie pokazaliśmy w nim, że skomplikowana, nieprzewidywalna dynamika uczenia się w wysoce heterogenicznej

[16] Jest to tematyka o fundamentalnym znaczeniu dla ekonomii, biologii matematycznej czy sztucznej inteligencji. Model wieloagentowego uczenia się ze wzmocnieniem (MARL) jest kluczowy dla właściwej analizy zachowań graczy w skomplikowanych środowiskach. Niestety mnogość wyników dotyczących MARL nie poprawia znacząco zrozumienia prawdopodobnie najbardziej znanego przykładu wieloagentowej algorytmicznej adaptacji w grach zatorowych, problemu El Farol Bar [6]. Problem ten, sformułowany przez amerykańskiego ekonomistę Williama Arthura (będący pierwotnie częścią dyskusji z Kennethem Arrowem i Paulem Krugmanem na temat podstaw teorii ogólnej równowagi) ma u swoich podstaw pomóc analizować w kontekście nauk społecznych takie koncepcje jak rozumowanie indukcyjne czy ograniczona racjonalność.

populacji może wygenerować porządek makroskopowy (z perspektywy średniego zachowania), który jest zgodny z klasycznymi koncepcjami równowagi w teorii gier (choć postać równowagi zależy m.in. od tempa uczenia się graczy). Jednak każdy uczestnik gry ponosi wysokie koszty. Wynik ten pomaga w zrozumieniu dynamiki zachowań graczy w problemie *El Farol Bar*.

### Złożona dynamika zachowań a awersja do ryzyka

Kolejne rozważane przeze mnie zagadnienie dotyczyło złożonego zachowania graczy o zróżnicowanej awersji do ryzyka. Choć algorytm MWU zadaje dynamikę gry będącą dyskretnym odpowiednikiem dynamiki replikacyjnej jego waga (i uniwersalność zastosowań) wynika z jego optymalnych własności minimalizujących żal [56, 101]. Stąd w sytuacjach gdy dostęp do informacji jest ograniczony gracze mają motywację do jego używania. Problemem wymagającym analizy było czy destabilizacja systemu obserwowana dla MWU jest konsekwencją użytego sposobu aktualizowania zachowań graczy. W tym celu w pracy **A3** analizowałem ze współautorami szeroką klasę algorytmów podążania za uregularyzowanym liderem (ang. *Follow-the-regularized-leader (FTRL)*). Pokazujemy w niej, że wyniki uzyskane dla MWU możemy rozszerzyć na sytuację gdy gracze inaczej warunkują swoje decyzje (np. poprzez inne spojrzenia na ryzyko). Była to kolejna praca, która znalazła miejsce na czołowej konferencji z zakresu uczenia maszynowego, tym razem *International Conference on Machine Learning*.

FTRL to klasa algorytmów uczenia online, wykorzystywanych np. w aukcjach, analizie ryzyka kredytowego czy modelowaniu genetycznym. Dynamika FTRL obejmuje w szczególnych przypadkach takie algorytmy jak MWU czy algorytm *gradient descent* [102]. W ramach FTRL strategia w każdej iteracji jest wybierana poprzez minimalizację ważonej (według szybkości uczenia się) sumy całkowitego kosztu wszystkich działań wybranych przez graczy i składnika regularyzującego. Załóżmy, że gracze w chwili $n+1$ znają koszt strategii w chwili $n$ na bazie $(x_n, 1-x_n)$ i aktualizują swoje wybory zgodnie z algorytmem FTRL. Mianowicie, w kontekście opisanym w omówieniu pracy **A2**, w momencie $n+1$ gracze wybierają pierwszą strategię z prawdopodobieństwem $x_{n+1}$ takim, że:

$$x_{n+1} = \arg\min_{x\in(0,1)} \left(\varepsilon \sum_{j\leq n} \left[c_1(x_j)\cdot x + c_2(1-x_j)\cdot(1-x)\right] + R(x, 1-x)\right), \qquad (3)$$

gdzie $c_1(x_j)\cdot x + c_2(1-x_j)\cdot(1-x)$, to całkowity koszt ponoszony przez populację graczy grających przeciwko mieszance $(x, 1-x)$ w momencie $j$, podczas gdy $R\colon (0,1)^2 \mapsto \mathbb{R}$ jest stromą w zerze ($r'(0) = -\infty$, gdzie $r(x) = R(x, 1-x)$) symetryczną funkcją

wypukłą, zwaną regularyzatorem, która reprezentuje „karę za ryzyko”: ten składnik zadaje karę za nagłe zmiany strategii na podstawie niewielkiego przyrostu informacji z poprzedniej iteracji gry.

Dynamika generowana przez algorytmy FTRL zależy od cech regularyzatora [80]. Strome regularyzatory są kluczowymi składowymi algorytmu gwarantującymi, że proces nie zbiegnie do jednej z czystych strategii. Reprezentują one albo naturalną niechęć gracza przed rezygnacją z pierwotnie wybranej strategii, albo celową ochronę przed wybraniem strategii zbyt ryzykownych. Zauważmy, że gdyby w problemie (3) nie było regularyzatora, populacja z prawdopodobieństwem 1 wybierałaby tę samą strategię czystą, co skutkowałoby wysokimi kosztami. W algorytmach klasy FTRL regularyzator jest powiązany z awersją graczy do ryzyka, tzn. im bardziej stromy (czyli $r$ rośnie szybciej w otoczeniu zera) jest regularyzator, tym bardziej gracz jest zniechęcany do podejmowania ryzykownych zachowań. Wiadomo, że dynamika FTRL z perspektywy czasu jest konkurencyjna względem najlepszej stałej akcji (jest tzw. algorytmem typu *no regret* [56, 101]), pod warunkiem, że jest on wykonywany z wysoce zoptymalizowaną szybkością uczenia się. W szczególnym przypadku gdy wszyscy agenci używają (ujemnej) entropii Shannona jako regularyzatora, otrzymujemy algorytm MWU. Czy możemy rozwinąć nasze zrozumienie dynamiki gry i zachowań z MWU na bardziej ogólną dynamikę FTRL? Czy wyniki są jakościowo podobne?

Odpowiedzią na te pytania zająłem się w pracy **A3**, w której pokazałem ze współautorami, że nawet w prostych bezatomowych grach zatorowych z liniowymi funkcjami kosztu, z dwoma zasobami i dowolną stałą szybkością uczenia się (z wyjątkiem sytuacji gdy gra jest w pełni symetryczna) zwiększenie wielkości populacji lub skali kosztów powoduje, że dynamika uczenia się staje się niestabilna i ostatecznie chaotyczna w sensie Li-Yorke'a. Ponadto, udowodniliśmy istnienie nowych, niestandardowych zjawisk, takich jak współistnienie stabilnych równowag Nasha i chaosu w tej samej grze. Dynamika zachowań w tej grze zależy więc od stanu początkowego populacji — zależnie od niego obserwować możemy zbieżność do równowagi bądź zachowania chaotyczne. Zatem (długoterminowy) koszt społeczny zależy w sposób krytyczny od warunków początkowych. Na koniec, chociaż dynamika FTRL może być skomplikowana, udowodniliśmy, że podobnie jak dla MWU, średnia po czasie wciąż zbiega do równowagi Nasha dla dowolnego wyboru szybkości uczenia się, dowolnych liniowych funkcji kosztów i dowolnego stanu początkowego. Zatem obserwujemy skrajnie skomplikowane zachowania graczy przy *porządnym* zachowaniu średnim.

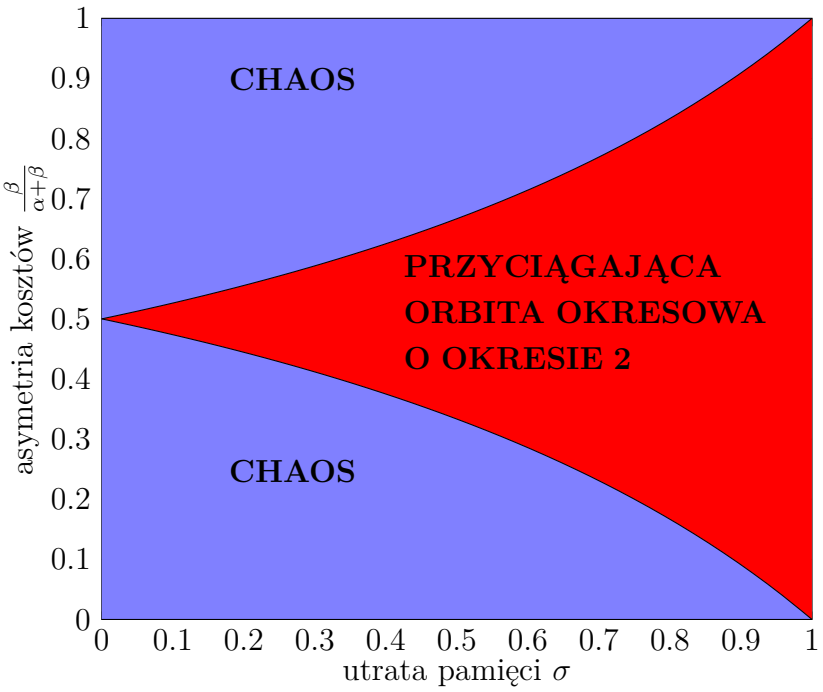
## Złożona dynamika zachowań a zapominanie, model z behawioralnej teorii gier

Kolejnym zagadnieniem, które podjąłem było zbadanie dynamiki populacyjnej gry zatorowej z agentami używającymi algorytmów sEWA (ang. *simplified Experience Weighted Attraction*, sEWA), ważnego narzędzia ekonomii eksperymentalnej [24, 49, 66, 67, 75]. Wyniki z wcześniej omówionych artykułów opisują dynamiki gry, w której uczący się gracze kształtują swoje zachowanie jednakowo oceniając koszty ze wszystkich poprzednich okresów. Mniej jasne jest to, jak dobrze to robią i czy zadane ograniczenia zgadzają się z naszym najbardziej naturalnym rozumieniem tego, jak ludzie faktycznie zachowują się, gdy stają przed strategicznymi decyzjami w praktyce. W szczególności czy podejmując decyzję korzystają z pełnej historii, czy może najważniejsze są ostatnie doświadczenia.

Pytanie, w jaki sposób ludzie uczą się modyfikować swoje strategie w grach w świecie rzeczywistym, jest przedmiotem badań teorii gier behawioralnych [24, 65–67]. *Przyciąganie ważone doświadczeniem* (ang. *Experience Weighted Attraction*) jest kanonicznym modelem uczenia się w teorii gier behawioralnych. Pierwotnym celem wprowadzenia EWA było ujednolicenie modeli opartych na uczeniu się przez wzmacnianie i modeli opartych na przekonaniach (*belief-based models*). Jednakże, model ten zawiera zbyt wiele wolnych parametrów [24, 103], dlatego ostatnie prace — [49, 90], praca **A4** — skupiły się na jej okrojonej wersji którą nazywamy uproszczonym EWA (sEWA). Model ten dopuszcza tylko dwa parametry: intensywność wyboru i parametr utraty pamięci, który odgrywa rolę podobną do roli stopy wykładniczego dyskontowania przeszłych wypłat.[17] W artykule **A4** rozważyłem ze współautorami ten model, w którym gracze używają logitowej najlepszej odpowiedzi na oszacowanie historycznego wyniku każdego działania. Wpływ przeszłych wypłat maleje tu wykładniczo. Co ważne, prace eksperymentalne w tej dziedzinie sugerują, że powszechna jest duża intensywność wyboru [24]. Ta informacja zestawiona z wcześniej uzyskanymi przeze mnie i moich współautorów wynikami wskazywała na możliwość wystąpienia konfliktu pomiędzy standardowymi założeniami stabilności modelu a przetestowanymi eksperymentalnie prawidłowościami behawioralnymi. W jaki sposób wzajemne oddziaływanie intensywności wyboru i utraty pamięci wpływa na wyniki zbieżności w prostych grach zatorowych?

Analizę interakcji pomiędzy utratą pamięci a asymetrią kosztów przeprowadziliśmy w dobrze znanym z poprzednich prac kontekście bezatomowych gier zatorowych z dwiema dostępnymi strategiami czystymi. Głównym wynikiem jest następujące twierdzenie (patrz Rysunek 1):

[17] Współczynnik utraty pamięci można postrzegać również jako kolejną składową ograniczonej racjonalności.




Rysunek 1: Zachowanie układu dla dużych wartości intensywności wyboru. Tak długo jak $\frac{\beta}{\alpha+\beta} \in \left(\frac{1-\sigma}{2-\sigma}, \frac{1}{2-\sigma}\right)$, prawie wszystkie trajektorie są przyciągane przez orbitę okresową okresu 2. Poza tym przedziałem obserwujemy zachowanie chaotyczne.

**Twierdzenie.** *Dla ustalonego współczynnika utraty pamięci $\sigma \in [0,1]$ (gdzie $\sigma = 0$ oznacza brak utraty pamięci, a przy $\sigma = 1$ agenci używają po prostu najlepszej odpowiedzi na ostatnią akcję) dla dużych wartości intensywności wyboru mamy dwa odmienne zachowania zależne od asymetrii stosunku kosztów $\frac{\beta}{\alpha+\beta}$:*

1. *Jeśli funkcje kosztów obu strategii nie różnią się znacząco, tzn. $\frac{\beta}{\alpha+\beta} \in I_\sigma = (\frac{1-\sigma}{2-\sigma}, \frac{1}{2-\sigma})$, to po utracie stabilności system pozostaje przewidywalny zbiegając zawsze do orbity okresowej o okresie 2.*

2. *Jeśli $\frac{\beta}{\alpha+\beta}$ znajduje się poza $I_\sigma$ (czyli różnica w funkcjach kosztu jest znacząca), to dla wystarczająco dużej intensywności wyboru system staje się nieprzewidywalny i chaotyczny (w sensie Li-Yorke'a).*

*Co więcej, wraz ze wzrostem utraty pamięci (wartości parametru $\sigma$) system staje się przewidywalny dla szerszego zakresu funkcji kosztu.*

Powyższe twierdzenie daje zasadnicze rozróżnienie długoterminowego zachowania systemu: w pierwszym przypadku chociaż dla dużych wartości intensywności wyboru system nie stabilizuje się, pozostaje w miarę przewidywalny — nie ma więc

znaczenia początkowy stan systemu (populacji), bo finalnie układ zbiega do tego samego dwuokresowego zachowania. Gdy funkcje kosztu są wystarczająco podobne, zapominanie (dyskontowanie) przeszłości powoduje, że koszty są nie do odróżnienia z punktu widzenia gracza. W takim wypadku, kiedy intensywność wyboru jest duża, populacja graczy (podmiotów gospodarczych) zmiarza do przyciągającej okresowej orbity o okresie 2. Po pewnym czasie każda parzysta (nieparzysta) iteracja umieści ją blisko swojej poprzedniej pozycji. Z kolei gdy funkcje kosztu istotnie się różnią, system dla odpowiednio dużej intensywności wyboru jest chaotyczny. System jest w nieprzewidywalnym reżimie, z okresowymi orbitami o różnych okresach i skomplikowaną, potencjalnie czułą na zaburzenia dynamiką — populacja ewoluuje w sposób nieprzewidywalny.[18]

Komentarza wymaga kwestia równowagi w tej dynamice. Razem ze współautorami pokazałem, że ze względu na zaburzenie systemu, równowagi Nasha nie mogą być stałymi punktami dynamiki. Jedyny punkt stały, który może być przyciągający jest tzw. równowagą logitową.[19] Równowaga ta leży między równym podziałem a równowagą Nasha, do których zbliża się dla dwóch skrajnych wartości intensywności wyboru. Można to uzasadnić następująco: kiedy intensywność wyboru dąży do zera, gracz jest obojętny na swoją wypłatę, a co za tym idzie, który zasób wybrać. Ponieważ oba wybory są równie prawdopodobne, wybiera podział $(\frac{1}{2}, \frac{1}{2})$. Z drugiej strony, gdy intensywność wyboru zmierza do nieskończoności, mała przewaga „historyczna" danego wyboru sprawia, że wybór ten staje się bardziej prawdopodobny. Stąd równowaga logitowa zbliża się wtedy do równowagi Nasha. Wzrost intensywności wyboru przybliża nas więc do równowagi Nasha. Dzieje się to jednak kosztem utraty stabilności. Razem ze współautorami pokazałem, że tak długo, jak stały punkt przyciąga lokalnie, przyciąga również globalnie. Jednakże istnieje dolne ograniczenie na intensywność uczenia, powyżej którego równowaga staje się odpychająca. Od tego momentu nie możemy liczyć na zbieżność dynamiki, a analiza stanu równowagi staje się bezprzedmiotowa. Ponadto wskazaliśmy, że ta wartość progowa maleje wraz ze wzrostem współczynnika dyskontującego $\sigma$.

W pracy **A4** wskazałem ze współautorami jeszcze jedno ważne rozróżnienie pomiędzy przypadkiem pełnej pamięci a sytuacją gdy następuje jej utrata. W przypadku pełnej pamięci, przy dużej intensywności wyboru można tak zmienić warunki gry (zróżnicować koszty zasobów/dróg), aby wymusić zbieżność do równowagi. Jednakże, gdy zapominanie przeszłości wpływa na wybór graczy ($\sigma > 0$), wystarczająco duża intensywność wyboru nieuchronnie zdestabilizuje system i żadna zmiana warunków gry go nie ustabilizuje. Przy utracie pamięci istnieje wartość progowa in-

[18]Ten wynik może sugerować, że okresowe zachowania, które widzimy w praktycznych problemach mogą być efektem ograniczonej pamięci, a zapominanie ma własności stabilizujące system.

[19]Równowaga logitowa jest ważnym narzędziem ekonomii eksperymentalnej [76].

tensywności wyboru, po której przekroczeniu, gra staje się niestabilna bez względu na koszty. Natomiast dla gier z pełną pamięcią ($\sigma = 0$), dla dowolnej intensywności można znaleźć gry zatorowe, gdzie dynamika będzie zbieżna.

Finalnie zauważmy, że w przypadku $\sigma > 0$ planista społeczny mający wpływ na modyfikację wypłat (np. poprzez opodatkowanie czy transfery pieniężne) może dążyć do zbliżenia wypłat w celu większej przewidywalności systemu (trafieniu w zakres parametrów dla których mamy w najgorszym razie zbieżność do orbity okresowej o okresie 2, patrz Rysunek 1). Zatem choć dyskontowanie przeszłości wyklucza możliwość modyfikacji wypłat w taki sposób, aby wymusić zbieżność, daje ono alternatywny sposób na zagwarantowanie przewidywalnej dynamiki: zbliżenie kosztów zasobów może wyprowadzić system z chaosu.

### Mikroekonomiczne a biologiczne podstawy ewolucyjnej teorii gier

W ewolucyjnej teorii gier w zależności od kontekstu, dynamikę gry wyprowadza się zwykle w jeden z następujących sposobów: (i) z biologicznego modelu ewolucji populacji, zwykle formułowanego w kategoriach zdolności reprodukcyjnej danego gatunku; (ii) na podstawie zestawu mikropodstaw ekonomicznych, które wyrażają tempo wzrostu typu (lub strategii) w populacji za pomocą protokołu rewizji (model ekonomiczny określający skłonność gracza do przejścia na strategię o wyższych wypłatach); lub (iii) z algorytmu uczenia się zaprojektowanego w celu optymalizacji krótkowzrocznego kryterium wydajności (takiego jak minimalizacja żalu gracza) w skądinąd agnostycznym otoczeniu, w którym gracze nie znają gry, w którą grają.

Jedną z najczęściej badanych dynamik gry jest dynamika replikatorów Taylora i Jonkera [111], prawdopodobnie *spiritus movens* ewolucyjnej teorii gier. Najpierw wyprowadzona jako model ewolucji populacji biologicznych pod presją selekcyjną w duchu Morana [83], dynamika replikatorów została następnie przeniesiona do teorii ekonomii poprzez mechanizm znany jako imitacja proporcjonalna parami (ang. *pairwise proportional imitation*, PPI), pierwotnie za sprawą Helbinga [63].[20] Mniej więcej w tym samym czasie pokazano, że MWU w przypadku ciągłym również prowadzi do dynamiki replikatorów.

W pracy **A5** razem z Panayotisem Mertikopoulosem postawiłem następujące, naturalne pytanie: *Czy koncepcyjnie różne modele leżące u podstaw dynamiki replikatora prowadzą do jakościowo różnych wyników w czasie dyskretnym? A jeśli tak, to w jakim stopniu?* Szerzej: *Czy ekonomiczne i biologiczne podstawy ewolucyjnej*

[20] Model ten ma swoje korzenie w teorii rewizji protokołów (ang. *revision protocols*). Każdy uczestnik gry czasami otrzymuje możliwość zmiany działań i w takich chwilach rozważa zmianę strategii, porównując swoją wypłatę z wypłatą losowo wybranej osoby w populacji i zmieniając ją z prawdopodobieństwem proporcjonalnym do korzyści wynikających ze zmiany [106].

*teorii gier są zgodne?*

Można z dużym prawdopodobieństwem oczekiwać, że odpowiedzi na te pytania będą pozytywne w przypadku dużych (ze względu na liczbę graczy czy liczbę dostępnych strategii) gier, w których na dłuższą metę pojawia się wiele zróżnicowanych zachowań; z drugiej strony, w prostych grach zakres pojawiających się zachowań byłby prawdopodobnie podobny jakościowo i różniłby się jedynie na poziomie ilościowym (takim jak tempo zbieżności do równowagi itp.). Pokazaliśmy, że to założenie jest zbyt optymistyczne, nawet w klasie gier z potencjałem (które posiadają najsilniejsze gwarancje zbieżności w ramach dynamiki replikatora), a nawet w przypadkach, gdy gra jest symetryczna, a gracze mają do dyspozycji jedynie dwie strategie (to najmniejsza sensowna gra). W szczególności rozważaliśmy przypadek symetrycznego losowego dopasowywania w grze antykoordynacyjnej $2 \times 2$ i pokazaliśmy, że różne źródła dynamiki replikatora w czasie dyskretnym wykazują jakościowo różne zachowania:

1. W biologicznym modelu konkurencji wewnątrzgatunkowej dynamika zbiega do równowagi Nasha dla dowolnej wartości kroku czasowego $\delta > 0$.

2. W ekonomicznym modelu bazującym na protokole postępowania PPI istnieją pewne konfiguracje równowagi, które globalnie przyciągają dla dowolnej wartości $\delta$, inne, dla których równowaga gry jest odpychająca dla pewnego zakresu wartości $\delta$, i jeszcze inne, które prowadzą, poprzez utratę stabilności równowagi i podwojenie okresu, aż do pojawienia się chaosu Li-Yorke'a (dla innego zakresu wartości $\delta$).

3. Wreszcie, w przypadku algorytmu MWU, wszystkie konfiguracje równowagi stają się niestabilne dla odpowiednio dużego kroku czasowego $\delta$ i jeśli korzyści z porzucenia najbardziej obciążonego wyboru nie są równe, chaos Li-Yorke'a pojawia się za każdym razem, gdy krok czasowy przekracza pewien próg zależny od dokładnej pozycji równowagi gry.

Pokazałem więc, że wyniki ekonomicznego modelowania charakteryzują się w tej sytuacji nieprzewidywalnością, co stanowi wyraźny kontrast z uniwersalnie zbieżnym krajobrazem dla omawianych gier, który pojawia się w ciągłym przypadku (i który podziela jedynie model biologiczny). Ten wynik jest szczególnie intrygujący, ponieważ dostarcza konkretnej, ilościowej przestrogi co do zakresu, w jakim przy pomocy dynamiki replikatorów można prognozować długoterminowe zachowania jego dyskretnych odpowiedników.

Co więcej, wynik ten wskazuje na odrębność mikropodstaw zastosowań i interpretacji ewolucyjnej teorii gier w różnych kontekstach aplikacyjnych. W modelu biologicznym widzimy zbieżność do równowagi Nasha niezależnie od długości cyklu

ewolucyjnego. W modelach ekonomicznych zbieżność do równowagi Nasha zachodzi dla takiego samego zakresu szybkości rewizji/szybkości uczenia się. Jednakże ich zachowania po utracie stabilności mogą być diametralnie różne. Na koniec warto nadmienić, że wynik z pracy **A5** można uogólnić np. pokazując, że chaotyczne zachowania mogą występować w grach jeśli tylko gracze wykorzystują protokoły imitacyjne (których reprezentantem jest PPI) [12].

### Miara złożoności strategii w grach iterowanych

Optymalizacja kryterium wydajności, jak minimalizacja żalu czy rewizja własnych wyborów poprzez porównanie swoich wypłat z innymi dostępnymi w populacji (protokoły rewizji), dają możliwość (dynamicznej) analizy ewolucji populacji, której dokonałem we wcześniej omówionych pracach. Jednakże, pojedynczy gracze zamiast używać algorytmu mogą starać się używać najlepszej odpowiedzi na (potencjalne) wybory innych uczestników gry ucząc się efektywnie ich zachowań. Gracze w grach iterowanych mogą używać skomplikowanych reguł [8, 69]. W jaki sposób gracz powinien reagować na ich zachowania? *W jaki sposób odróżnić zachowania (calkowicie) losowe od tych ustrukturyzowanych wynikających z jakiejś reguły?* Można założyć, że gracze preferują wykorzystywanie *prostych* strategii. *Ale co takie sformułowanie oznacza?* W pracy **A6** zaproponowałem narzędzie pozwalające analizować takie problemy — miarę (statystycznej) złożoności strategii.

W grach powtarzanych strategia to zbiór planów działania zależnych od historii. W literaturze przyjmuje się zazwyczaj, że gracze mogą realizować dowolną strategię z określonego zestawu strategii. Choć założenie to może wydawać się niewiążące w modelu, w którym każdy gracz ma dostęp do niewielu strategii, jest ono nierealistyczne w przypadku bardziej złożonych modeli, w których zestaw strategii zawiera dużą liczbę możliwości, z których wiele jest na tyle skomplikowanych, że trudno je wykorzystać. Dzieje się tak zwykle w przypadku gier powtarzanych – na każdym etapie gracze grają w grę (nie)zależną od historii i dlatego zbiór wszystkich strategii, a także równowag Nasha, może być ogromny [48]. Wielość równowag skłoniła badaczy do skupienia uwagi na wybranych. Takie ograniczenia są raczej normą niż wyjątkiem w modelach dynamicznych w ekonomii. Przeciwnik (konkurent) może stosować złożone strategie, które są trudne do przewidzenia. Gracz może zatem stanąć przed problemem wyboru najlepszej strategii wobec strategii przeciwnika, której uczy się w trakcie powtarzanej rozgrywki. Im bardziej ustrukturyzowana (złożona) jest strategia, tym trudniej jest przewidzieć, co zrobi przeciwnik i tym trudniej znaleźć najlepszą odpowiedź. , Gracz potrzebuje więc sposobu na wykrycie złożonych strategii.

Aby zapewnić wgląd w strukturę i złożoność strategii, w artykule **A6** przyjrza-

łem się tempu zbieżności entropii w grze powtarzanej. Powszechnie uważa się, że powolne tempo zbieżności entropii jest oznaką złożonej struktury procesu [34]. Oczywiście bardziej złożona strategia jest trudniejsza do przewidzenia. Dlatego strategie o wolnym tempie zbieżności są uznawane za mniej przewidywalne. Stąd ograniczenie zestawu możliwych strategii do zestawu strategii o ograniczonej nieprzewidywalności może dać podstawy do dokładniejszego wglądu w teorię konkurencji długoterminowej.

W pracy **A6** zaproponowałem wprowadzenie nowej miary statystycznej złożoności strategii opartej na teorioinformacyjnej koncepcji nadwyżki entropii [34, 52, 53] i tempa zbieżności [33]. Patrząc na tempo zbieżności entropii strategicznej (ang. *strategic entropy rate*), pojęcia wprowadzonego przez Neymana i Okadę [85–88], mierzy ona strukturę, regularność i przewidywalność rozgrywanych strategii. Wielkość tą nazwałem nadwyżką strategicznej entropii (nadwyżką s-entropii, ang. *excess s-entropy*). Wprowadzona przeze mnie miara służy do pomiaru złożoności statystycznej i nieprzewidywalności strategii. Naturalne interpretacje tego miernika narzucają jego zastosowanie np. w przypadku modeli ograniczonej pamięci (ang. *bounded recall*) i reputacji.[21]

Ideą stojącą za, wprowadzoną przeze mnie, nadwyżką s-entropii jest obserwacja, że gdy zmiany wyborów są częste, powinniśmy spodziewać się nieskończonej nadwyżki s-entropii, natomiast jeśli zmiany są rzadkie, nadwyżka s-entropii może być skończona. To spostrzeżenie pozwala doprecyzować pojęcie prostej strategii — to ta, której nadwyżka s-entropii jest skończona.[22] Do analizy nadwyżki s-entropii możemy wykorzystać metody pochodzące z teorii informacji i teorii układów dynamicznych. Patrząc przez pryzmat teorii informacji, nadwyżka s-entropii jest właściwym narzędziem do wykrywania i analizowania wzorców wytwarzanych w procesie. Potrafi rozróżnić wzorce, które mają zbliżone cechy strukturalne. Nadwyżkę s-entropii można wykorzystać do wykrywania i ilościowego określenia wzorców wytwarzanych podczas gry. Z kolei z perspektywy układów dynamicznych możemy badać tempo zbieżności entropii np. przy użyciu narzędzi dynamiki symbolicznej.

[21] Podjęte badania i uzyskane wyniki wpisują się w rosnący obszar zastosowań koncepcji entropijnych w ekonomii, [11, 20, 21, 64].

[22] Oczywiście jeśli gracze wyjściowo znajdują się w równowadze Nasha, nadwyżka s-entropii (i złożoność strategii) będzie równa zero, ponieważ gracz nie będzie chętny do zmiany swojej strategii. Jeśli jednak gracz nie znajduje się w równowadze Nasha od początku, może wybrać jedną ze strategii mieszanych, która nie jest zbyt skomplikowana — na przykład taką, która nie zmusza go zbytnio do zmiany wyboru w danym etapie gry. Zatem ograniczenie zbioru strategii do tych z ograniczoną nadwyżką s-entropii wydaje się naturalnym wyborem. Nawet jeśli gracz chce „zbliżyć się do równowagi Nasha", może preferować te równowagi Nasha, do których zbieżne są jego sekwencje strategii w grach jednoetapowych, stosując strategie z ograniczoną nadwyżką s-entropii.

Ponieważ gracze mogą preferować mało skomplikowane strategie (jako prostsze w implementacji) w przyszłych badaniach nasuwa się potrzeba skoncentrowania uwagi na strategiach z małą nadwyżką s-entropii, co narzuci dużą przewidywalność strategii mieszanej i wpłynie na możliwość wyboru najlepszej odpowiedzi przez gracza z nieograniczonym zbiorem strategii.

Obszarem badań, gdzie nadwyżka s-entropii może znaleźć zastosowanie są też modele reputacji. Pogląd, że zaangażowanie jest wartościowe (tzw. *commitment strategies*), wywarł głęboki wpływ na teorię gier niekooperacyjnych i wiele dziedzin nauk społecznych. Naturalne wydaje się oczekiwanie, że agenci grający w długoterminowej rywalizacji będą opierać swoje decyzje na reputacji przeciwnika. Co więcej, jeśli przeciwnik jest bardziej przewidywalny (rozkład prawdopodobieństwa, którego używa, jest przewidywalny), wówczas gracz może znaleźć lepszą odpowiedź na przewidywaną strategię lub może preferować grę z bardziej przewidywalnym graczem. Warto dodać, że miary entropijne znajdują coraz szersze zastosowania w modelach reputacyjnych [39, 44, 51].

### 2.2.2 Optymalizacja dobrobytu społecznego a mikrobodźce graczy

W pracach omówionych w poprzedniej części analizowałem zachowania graczy, populacji (i ewolucję tych zachowań) w obliczu bodźców pochodzących z przeszłych wypłat oraz tego jak bardzo istotne są one dla agenta. Analizy dokonałem w sytuacji gdy stabilne rozwiązania gry były tożsame z optimum społecznym mierzonym utylitarystyczną funkcją dobrobytu społecznego (PoA = 1). W teorii ekonomii nadal żywa jest dyskusja dotycząca wyboru miary dobrobytu społecznego. Na przeciwległych krańcach spektrum znajdują się utylitarystyczna funkcja dobrobytu społecznego (zadana jako suma użyteczności/wypłat [57] ) oraz rawlsowska funkcja dobrobytu społecznego (gdzie dobrobyt mierzony jest zamożnością najbiedniejszego [94]). *Czy da się pogodzić te skrajności? Czy cele planistów społecznych, którzy mogą dokonywać transferów w populacji (poprzez np. opodatkowanie najbogatszych) będą zgodne?* W pracach **B1** i **B2** dałem ze współautorami odpowiedź twierdzącą pod warunkiem, że jednostki (członkowie populacji) uwzględniają w swych wypłatach nierówności dochodowe.

Kluczowym punktem odniesienia dla tej części przedstawianego cyklu publikacji jest pojęcie względnego niedostatku. Zostało ono wprowadzone przez Yitzhakiego [116] i bierze pod uwagę nierówności dochodowe (lub w bogactwie) w populacji. Rozpatrzmy skończoną populację agentów $\{1, \ldots, n\}$. Niech $\mathbf{x} = (x_1, \ldots, x_n)$ będzie wektorem dochodów (bogactwa) populacji agentów, gdzie $x_i \geq 0$ dla każdego $i \in \{1, \ldots, n\}$. Wtedy względny niedostatek doświadczany przez jednostkę $i$ w populacji

o rozkładzie dochodów $\mathbf{x}$ zadany jest jako

$$RD_i(\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} \max\{x_j - x_i, 0\}. \tag{4}$$

Oznacza to, że agent $i$ doświadcza nierówności dochodowych poprzez porównywanie się jedynie z bogatszymi od niego.

W pracach **B1** i **B2** badamy w jaki sposób zmienia się optimum różnych funkcji dobrobytu społecznego w zależności od intensywności oddziaływania bodźców zewnętrznych, a dokładniej porównań dochodowych w populacji. Planista społeczny, który dysponuje możliwością transferów wewnątrz populacji, optymalizuje wybory zależnie od różnych funkcji dobrobytu społecznego: utylitarystycznej [57], rawlsowskiej (gdy optymalizujemy dochód najbiedniejszego [94, 95]) czy Bernoullego-Nasha [7]. Korzystając z narzędzia jakim jest miara względnego niedostatku Yitzhakiego, dokonujemy analizy, która wskazuje, że intensywność porównań dochodowych między jednostkami może istotnie wpłynąć na rozwiązania dokonywane przez planistów.

*Czy da się pogodzić podejście rawlsowskie z utylitarnym?* W artykule **B1** przedstawiamy rozwiązanie tego dysonansu poprzez uwzględnienie w funkcji użyteczności agentów niechęci do niskich dochodów względnych. Pokazujemy, że gdy waga niechęci jednostki do niskich dochodów względnych (mierzonych względnym niedostatkiem) jest większa od wartości krytycznej, co zależy od kształtu funkcji użyteczności jednostek i początkowego rozkładu dochodów, rozwiązanie problemu maksymalizacji utylitarystycznej funkcji dobrobytu społecznego pokrywa się z maksymalizacją użyteczności jednostki znajdującej się w najgorszej sytuacji. W artykule przedstawiliśmy zatem wyjaśnienie w duchu Harsanyi'ego[23] i Rawlsa, godzące rawlsowskie i utylitarne kryteria maksymalizacji dobrobytu społecznego. Nasz artykuł idzie o krok dalej, pokazując nie tylko, że redystrybucja wzrasta gdy uwzględni się niechęć jednostek do niskich dochodów względnych, ale także że możliwe jest, że w takiej sytuacji cele utylitarnego planisty społecznego i planisty rawlsowskiego są zgodne.

Tym razem zawężamy się do sytuacji gdy agenci są jednostkami przynależącymi do skończonej populacji. Na ich wypłaty składają się użyteczność wynikająca z otrzymywanego dochodu oraz składnik, względny niedostatek, wynikający z porównań dochodowych z jednostkami bogatszymi. Niech $f\colon [0,\infty) \mapsto \mathbb{R}$ będzie dwukrotnie rózniczkowalna, ściśle rosnąca i ściśle wklęsła. Wypłatę osoby $i$ zadajemy jako

$$u_i(\mathbf{x}) = (1-\beta)f(x_i) - \beta RD_i(\mathbf{x}), \tag{5}$$

[23] W ramach badań nad dobrobytem społecznym, Harsanyi [57] przypisuje znaczącą rolę porównaniom interpersonalnym w funkcji dobrobytu społecznego.

gdzie $x_i \geq 0$ jest dochodem agenta $i$. Upodobanie jednostki do dochodu bezwzględnego jest ważone przez $1-\beta$ , $\beta \in [0,1)$, a niechęć jednostki do niskich dochodów względnych przez $\beta$. Oczywiście jeżeli danej osobie nie przeszkadza niski względny dochód, to $\beta = 0$. Odpowiada to przypadkowi standardowego planisty utylitarystycznego.

Niech wektor początkowych dochodów $n$ osób będzie zadany przez $\mathbf{e} = (e_1, \ldots, e_n)$ gdzie $0 < e_1 \leq \ldots \leq e_n$. Planista społeczny może transferować dochody by uzyskać to, co według niego stanowi optymalny rozkład dochodów populacji. Niech $x_i$ oznacza możliwy dochód jednostki $i$ po transferze (lub opodatkowaniu) oraz niech $t = \sum \max\{e_i - x_i, 0\}$ oznaczają całkowity dochód, jaki przyjmuje planista socjalny od osób fizycznych (zwany dalej „podatkiem"). Ze względu na koszty transferu przekazywana jest tylko część podatku. Ułamek ten oznaczamy przez $\lambda \in (0,1]$. Zatem zbiorem, na którym szukamy rozwiązania problemu planisty społecznego, jest zbiór

$$\Omega(\mathbf{e}, \lambda) = \{\mathbf{x} \colon x_i \geq 0 \text{ dla dowolnych } i, \text{ oraz } \lambda \sum \max\{e_i - x_i, 0\} = \sum \max\{x_i - e_i, 0\}\}.$$

Przeszukując zbiór dochodów, które można osiągnąć w ramach początkowej alokacji $\mathbf{e}$ poprzez opodatkowanie niektórych osób, otrzymujemy podatek w wysokości $t$ i rozdzielamy go pomiędzy pozostałe osoby w taki sposób, że transfer wynosi $\lambda t$. Planista rawlsowski, mający na celu maksymalizację na zbiorze $\Omega(\mathbf{e}, \lambda)$ wielkości $\min_{i \in \{1,\ldots,n\}} u_i(\mathbf{x})$, zdecyduje się opodatkować i przetrasferować podatek tak, aby wyrównać wszystkie dochody. Z kolei utylitarysta nie biorący pod uwagę nierówności dochodowych dokona trasferu jedynie gdy $\frac{f'(e_n)}{f'(e_1)} < \lambda$ i będzie opodatkowywał najbogatszych oraz transferował dochód do nabiedniejszych tak długo aż $x_1^{U*} = \ldots x_i^{U*} = \underline{e}$, $x_{i+1}^{U*} = e_{i+1}, \ldots, x_j^{U*} = e_j$ i $x_{j+1}^{U*} = \ldots = x_n^{U*} = \overline{e}$, gdzie $\lambda = \frac{f'(\overline{e})}{f'(\underline{e})}$. Co za tym idzie wybierze plan zgodny z wyborem planisty rawlsowskiego tylko gdy transfer jest bezstratny ($\lambda = 1$). Co zmienia uwzględnienie w wypłacie (równanie (5)) względnego niedostatku? W pracy **B1** pokazaliśmy, że dla dowolnej funkcji $f$ i progu $\lambda$ istnieje $\beta(\mathbf{e}, \lambda) < 1$ taki, że utylitarystyczny planista czuły na nierówności dochodowe wybierze transfer prowadzący do wyboru rawlsowskiego jeśli $\beta > \beta(\mathbf{e}, \lambda)$.

Wynika stąd, że optymalna polityka podatkowa utylitarysty może pokrywać się z optymalną polityką podatkową planisty patrzącego jedynie na dobrobyt najbiedniejszego w populacji, gdy użyteczność zależy nie tylko od własnych dochodów jednostki, ale także od dochodów innych osób. Innymi słowy, gdy utylitarny planista społeczny uwzględni niechęć jednostek do niskich względnych dochodów, może w rezultacie dokonać wyborów zgodnych z planistą społecznym Rawlsa w zakresie optymalnej polityki podatkowej i transferowej. Ta zgodność oferuje konkurencyjną do tradycyjnej drogę pogodzenia dwóch interpretacji polityki fiskalnej, które wypływałyby z obiektywnej perspektywy jednostki znajdującej się za *rawlsowską zasłoną*

*ignorancji.* Arrow [5] sugerował, że oczekiwana użyteczność mierzona (utylitarną) funkcją dobrobytu społecznego Harsanyi'ego zbiega do rawlsowskiej funkcji dobrobytu społecznego, jeśli wklęsłość użyteczności jednostki z dochodu jest wystarczająco duża. Praca **B1** wskazuje jednak, że tę zgodność można przypisać zmianie stanowiska utylitarnego planisty społecznego i nie musi być ona uzależniona od szczególnej niechęci jednostek do ryzyka. Wynik ten pozwala sądzić, że włączenie „użyteczności porównawczej" do optymalnych modeli podatkowych może urealnić wyniki. Choć coraz więcej ekonomistów uznaje, że takie porównania stanowią istotny aspekt rzeczywistości, jak dotąd narzędzia porównawcze miały jedynie niewielki wpływ na to, jak myślimy o porównywaniu optymalnych polityk podatkowych. Nasz artykuł wskazuje, że zaniedbanie porównań może znacząco wpłynąć na błędy benchmarków.

W pracy **B2** rozwijam pomysły z pracy **B1**. Skupiając pod jednym izoelastycznym „dachem" wszystkich kluczowych planistów społecznych, określam wraz ze współautorami warunki gwarantujące, że planiści wybiorą równą dystrybucję dochodów. Klasa izoelastycznych funkcji dobrobytu społecznego [7] pozwala nam przedstawić różny stopień niechęci planistów społecznych do nierówności w podziale dochodów ludności jako przypadki szczególne. Ze względu na atrakcyjne podstawy aksjomatyczne i elastyczność w przyjmowaniu podstawowych kryteriów równości [7] funkcja ta stała się popularną miarą dobrobytu społecznego w różnych dziedzinach, począwszy od optymalnego opodatkowania po ekonomię zdrowia oraz ekonomię środowiska. Naszym celem było odkrycie warunku, w którym wszyscy kluczowi „izoelastyczni planiści społeczni" – utylitarysta, rawlsowianin, Bernoulli-Nash czy jakikolwiek planista „pośredni" – wybiorą ten sam rozkład dochodów. Otrzymaliśmy wynik silnej zgodności: gdy funkcje użyteczności jednostki odzwierciedlają wystarczająco duże obawy związane z niskim dochodem względnym, optymalna polityka podatkowa wszystkich planistów społecznych jest zgodna. Jednomyślność ta obowiązuje dla całej klasy izoelastycznych funkcji dobrobytu społecznego z parametrem niechęci do nierówności $\alpha \in [0, \infty)$, (określonym poniżej we wzorze (6)) Charakteryzujemy konsensus optymalnej dystrybucji dochodów – czyli rozkładu równych dochodów – i stwierdzamy, że intensywność obaw jednostek związanych z niskim dochodem względnym wypiera preferencje dotyczące dystrybucji dochodów wyznawane przez poszczególnych planistów społecznych. Co więcej identyfikujemy krytyczną intensywność obaw jednostek związanych z niskim dochodem względnym, poniżej którego każdy izoelastyczny planista społeczny inny niż rawlsowski wybierze inną, nierówną dystrybucję. Formułujemy warunek konieczny i wystarczający pogodzenia wszystkich izoelastycznych planistów społecznych. Co więcej, chociaż im wyższe $\alpha$, tym bardziej izoelastyczna funkcja dobrobytu społecznego przechyla się na korzyść wyrównywania dochodów, stratny transfer zakłóca tę „skłonność" do dowolnego $\alpha$.

Włączenie jednostek obawiających się posiadania względnego niskiego dochodu przywraca „władzę” (wzmacnia mandat) planisty społecznego w zakresie wyrównywania dochodów.

Dokładniej, niech będzie dana populacja $n \geq 2$ osobników (gdzie n jest liczbą naturalną). Udowodniliśmy następujące twierdzenie.

**Twierdzenie.** *Niech izoelastyczna funkcja dobrobytu społecznego będzie zdefiniowana jako*

$$SWF_\alpha(\mathbf{x}) = \begin{cases} \left[\frac{1}{n}\sum\limits_{i=1}^{n} u_i^{1-\alpha}(\mathbf{x})\right]^{\frac{1}{1-\alpha}}, & gdy\ \alpha \geq 0, \alpha \neq 1, \\ \sqrt[n]{\prod\limits_{i=1}^{n} u_i(\mathbf{x})}, & gdy\ \alpha = 1, \end{cases} \tag{6}$$

*gdzie* $\mathbf{x} = (x_1, \ldots, x_n) \in \Omega(\mathbf{e}, \lambda)$, $x_i \geq 0$ *jest dochodem agenta i, użyteczności* $u_i(\mathbf{x}) > 0$ *dla dowolnego* $\mathbf{x} \in \Omega(\mathbf{e}, \lambda)$, *a* $\alpha \in [0, \infty)$ *jest parametrem niechęci do nierówności stosowanym przez planistę społecznego.*[24] *Wtedy w optimum społecznym planista dokona równego podziału wtedy i tylko wtedy, gdy jeden z poniższych warunków jest spełniony:*

1. *podział początkowy jest równy, tzn.* $e_1 = \ldots = e_n = x^*$
2. *planista jest planistą rawlsowskim, tzn.* $SWF_\alpha(\mathbf{x}) = SWF_R(\mathbf{x})$, *gdzie* $SWF_R$ *jest rawlsowską funkcją dobrobytu społecznego*
3. $\beta \geq \beta^*(\mathbf{e}, \lambda)$, *gdzie* $\beta^*(\mathbf{e}, \lambda) < \frac{f'(x^*)(1-\lambda)}{f'(x^*)(1-\lambda)+\lambda+\frac{1-\lambda}{n}} < 1$ *i to ograniczenie nie zależy od wartości* $\alpha$.

Twierdzenie to wskazuje, że preferencje jednostek dominują nad gustem planisty społecznego. Dla całej klasy izoelastycznych funkcji dobrobytu społecznego istnieje jeden krytyczny poziom intensywności obaw jednostek związanych z niskim dochodem względnym, który prowadzi do równego rozkładu dochodów. Fakt, że krytyczny poziom intensywności obaw jednostek związanych z niskim względnym dochodem jest taki sam dla wszystkich poziomów izoelastycznego parametru awersji do nierówności, sugeruje, że stopień obaw jednostek związanych z niskim względnym dochodem odgrywa wyraźną i ważniejszą rolę w kształtowaniu optymalnej polityki redystrybucji niż intensywność niechęci planisty społecznego do nierówności.

Praca zamykająca omawiany cykl jest głosem w dyskusji nad miarami nierówności społecznych. W pracy **B3** wyprowadziłem (ze współautorami) nową klasę uogólnionych miar względnego niedostatku. Innowacyjną cechą tej klasy jest to, że w za-

[24] Funkcja dobrobytu społecznego zadana przez (6) w szczególnych przypadkach daje utylitarystyczną funkcję dobrobytu społecznego ($\alpha = 0$), funkcję dobrobytu społecznego Bernoullego-Nasha ($\alpha = 1$) i Rawlsa ($\alpha = \infty$).

leżności od wartości parametru bliskości $p$, klasa jest w stanie przyjąć zarówno malejącą wagę (przypadek $p > 1$), jak i rosnącą wagę (przypadek $p \in (0, 1)$) przypisywaną danym zmianom dochodów osób bogatszych od jednostki (wzorca), w zależności od ich bliskości w podziale dochodów z jednostką. Powszechnie przyjętym założeniem w pracach, które uwzględniały względną deprywację, jest to, że porównania z innymi osobami, które w podziale dochodów znajdują się na prawo (bogatszych) od jednostki, liczą się jednakowo niezależnie od rzeczywistych nierówności. Jednak badania [89, 93] wskazują, że ludzie przywiązują inną wagę do zmian w dochodach osób, które są dalej w podziale dochodów, niż do zmian w dochodach osób o zbliżonych dochodach. W artykule **B3** zakwestionowaliśmy konwencję równych wag. Zaproponowaliśmy ogólny sposób ważenia, oparty na założeniu, że agenci nie muszą przywiązywać tej samej wagi do zmian w dochodach osób znajdujących się w różnej (dochodowej) odległości od osoby, której mierzona jest względna deprywacja. Operacjonalizacja podejścia niedoboru dochodów za pomocą zestawu aksjomatów pozwaliła nam otrzymać klasę miar wrażliwych na ten czynnik.

Naszą miarę wprowadziliśmy poprzez zdefiniowanie relacji preferencji na zbiorze wszystkich możliwych grup referencyjnych. Proponując zestaw naturalnych aksjomatów, pokazaliśmy, że jedyną miarą je spełniającą jest miara zadana jako

$$RD_p(z, \mathbf{x}) = \left( \frac{1}{n} \sum_{i=1}^{n} (\max\{x_i - z, 0\})^p \right)^{\frac{1}{p}} \tag{7}$$

gdzie $z$ jest dochodem (lub bogactwem) jednostki $\omega$, a $\mathbf{x}$ oznacza wektor bogactwa wszystkich osób z jego grupy referencyjnej. Jednocześnie pokazaliśmy dwie cechy nowej klasy miar: wrażliwość na bliskie dochody i wrażliwość na transfery pośród bogatszych. Dla $p \in (0, 1)$ miara (7) spełnia własność rosnącej wrażliwości na bliskość, tzn. jednostka $\omega$ odczuje większą względną deprywację w wyniku wzrostu dochodów jednostki, która (umieszczona na prawo od $\omega$ w podziale dochodów) jest mu bliższa niż w wyniku takiego samego wzrostu dochodu osoby, która (umieszczona na prawo od $\omega$ w rozkładzie dochodów) jest dalej. Dla $p > 1$ mamy własność przeciwną — malejącej wrażliwości. Ponadto pokazaliśmy, że dla $p \in (0, 1)$ jednostka preferuje transfer wstępujący (pośród bogatszych od niego) — od biedniejszego do bogatszego. Z kolei dla $p > 1$ jednostka preferuje transfer zstępujący — od bogatszego do biedniejszego. Zaproponowana miara znalazła szerokie zastosowania [40, 41, 43, 50, 68, 108, 109].

## 2.3 Podsumowanie osiągnięć w dyscyplinie ekonomii i finansów

W tym podrozdziale podsumowuję swój wkład w rozwój dyscypliny ekonomia i finanse. Moje badania mają charakter badań podstawowych, które prowadzą do rozwinięcia ogólnej wiedzy w dyscyplinie ekonomii i finansów.

Pokazałem, że

1. złożone zachowania chaotyczne mogą pojawić się w najprostszych nietrywialnych grach, w których gracze nie dysponują pełną informacją. Ten wynik wskazuje na konieczność precyzyjnej analizy dynamiki gry wykraczającej poza analizę jej równowag i wpisuje się w problem wyboru rozwiązania gry.

2. efektywności systemu (mierzonej ceną anarchii) nie można bezstratnie poprawić poprzez zwiększenie wielkości populacji czy zapotrzebowania systemu. Zwiększenie zapotrzebowania może doprowadzić do destabilizacji systemu i nieprzewidywalnych zachowań graczy.

3. modele uczenia ze wzmocnieniem stosowane w ekonomii i finansach (np. MWU) jak również modele eksperymentalnej ekonomii (sEWA) mają istotne ograniczenia. W szczególności wymagają zazwyczaj dogłębnej analizy wykraczającej poza zbadanie cech stacjonarnych rozwiązań gry.

4. złożone zachowania obserwowane w mikroskali lub w zachowaniu *dzień po dniu*, może występować razem z pożądanym zachowaniem średnim również w grach z potencjałem. W problemach z jakimi można się spotkać np. w tzw. przeciążonych oligopolach [2], w modelowaniu ruchu drogowego czy aukcjach online [18, 29], możemy obserwować średnie w czasie zachowanie systemu sugerujące jego optymalność (zbieżność do równowagi Nasha) z jednoczesnymi wysokimi kosztami wynikającymi z zachowań systemu *dzień po dniu*.

5. zwiększona awersja graczy do ryzyka nie stabilizuje systemu (praca **A3**). Co więcej, możliwa jest koegzystencja stabilnych rozwiązań (przyciągającej równowagi Nasha) i chaosu — zachowanie populacji graczy zależy wtedy od stanu początkowego populacji (systemu).

6. algorytmy sEWA mogą prowadzić do chaotycznych zachowań populacji graczy. Są to pierwsze analityczne wyniki dla algorytmów sEWA, które potwierdzają eksperymentalne wyniki znane dla gier losowych [49, 90]. Ze względu na ścisły związek algorytmów sEWA ze stosowanymi szeroko metodami q-learning [115] stanowi to uzasadnienie dla pojawiania się cykli w problemach związanych np. z tzw. zmową algorytmów [19, 23, 71, 114].

7. zapominanie (dyskontowanie) przeszłości może mieć pozytywne własności predykcyjne dla systemu (praca **A4**); W szczególności, jeśli planista społeczny chce zwiększyć przewidywalność populacji agentów może starać się to uzyskać poprzez zbliżenie wypłat/kosztów zasobów (np. poprzez dodatkowy podatek związany z „lepszym” zasobem).

8. okresowe zachowania, które obserwujemy np. na rynkach finansowych mogą być również konsekwencją ograniczonej pamięci uczestników rynku.

9. mikroekonomiczne podstawy ewolucyjnej teorii gier są fundamentalnie różne od klasycznych pochodzących z nauk biologicznych i prowadzą do całkowicie odmiennych zachowań.

10. dysonans pomiędzy optymalnymi wyborami różnych planistów społecznych może zniknąć, gdy agenci są czuli na nierówności dochodowe (w bogactwie). Wykazałem, że optymalne wybory planistów izoelastycznych (czyli np. utylitarystycznych czy rawlsowskich) mogą być zgodne jeśli agenci uwzględniają w swoich funcjach użyteczności bodźce zewnętrzne wynikające z nierówności dochodowych. Pozwala to zunifikować wybory planistów i stanowi wyjaśnienie w duchu Harsanyi'ego konkurencyjne do pochodzącego od Arrowa wyjaśnienia z perspektywy awersji do ryzyka.

Zaproponowałem i przeanalizowałem

1. modele dyskretnej dynamiki opisujące zachowania uczących się graczy (populacji graczy) warunkujących swoje zachowania minimalizacją żalu (MWU, FTRL) lub porównaniami interpersonalnymi (PPI).

2. model dyskretnej dynamiki, w której gracze zapominają przeszłość (sEWA).

3. nową miarę statystycznej złożoności strategii, którą można stosować w modelach reputacyjnych.

4. aksjomatycznie wyprowadzoną miarę względnego niedostatku wrażliwą na odległość porównań i na transfery pieniężne wśród bogatszych.

Skonstruowane modele umożliwiają formułowanie zaleceń i rekomendacji np. w

1. analizie oligopoli (np. do *congested markets* [2] czy w oligopolach na sieciach [32]),

2. analizie tzw. zmowy algorytmów (ang. *algorithmic collusion* [19, 23, 71]),

3. problemie wyboru portfela online [112] (wyniki pracy **A3**),

4. optymalizacji wyborów w gospodarce miejskiej,

5. teoretycznych modelach systemów ekonomicznych,

6. wieloagentowych modelach sztucznej inteligencji [54].

Wyniki uzyskane dla gier zatorowych mają bezpośrednie zastosowanie w oligopolach, aukcjach online, rynkach kryptowalut, problemie algorytmicznej zmowy cenowej, czy szerzej w problemach gdzie algorytmy wykorzystywane są do wyboru strategii (drogi, zasobu, ceny). Ponieważ algorytmy wykorzystywane są coraz szerzej podczas dokonywania decyzji (i w procesach) na rynkach, dokładne ich zbadanie wydaje się być koniecznością.

# 3 Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej

## 3.1 Aktywność naukowa

Moja kariera naukowa związana była z dwiema uczelniami: studia magisterskie, a następnie doktoranckie odbyłem na Wydziale Matematyki i Informatyki Uniwersytetu Jagiellońskiego. Jednocześnie, po ukończeniu studiów magisterskich, podjąłem pracę w Akademii Ekonomicznej w Krakowie (aktualnie: Uniwersytet Ekonomiczny w Krakowie). Jeszcze w ramach studiów doktoranckich brałem udział w *Sieci układów dynamicznych i teorii ergodycznej* (lata 2007-2010), co skutkowało moim udziałem w szkołach w Polsce i za granicą (m.in. Holandia, Hiszpania). W 2014 roku, po nawiązaniu współpracy z prof. Odedem Starkiem (Uniwersytet Warszawski, Uniwersytety w Tuebingen, Bonn, Alpen-Adria University) brałem udział w *First Winter School on Migration* na Georgetown University, School of Foreign Service in Qatar w Doha.

Po uzyskaniu stopnia doktora w latach 2017-20 realizowałem, jako kierownik projektu, grant Narodowego Centrum Nauki uzyskany w ramach konkursu Sonata 11 pt. *Miary statystycznej złożoności i nieprzewidywalności oparte na pojęciu entropii w grach ekonomicznych*, nr 2016/21/D/HS4/01798. Jego rezultatem są publikacje **A1-A3** oraz **A6**. Aktualnie jestem kierownikiem grantu nr 2023/51/B/HS4/01343 pt. *Ewolucja zachowań podmiotów gospodarczych w warunkach ograniczonej racjonalności* (szerszy opis znajduje się w części 3.2).

Moja aktywność poza Uniwersytetem Ekonomicznym w Krakowie obejmuje następujące działania:

- prowadzenie badań w międzynarodowych zespołach badawczych,
- uczestnictwo w międzynarodowych i krajowych konferencjach naukowych,
- uczestnictwo w seminariach naukowych,
- recenzowanie artykułów naukowych w międzynarodowych i krajowych czasopismach.

## 3.2 Międzynarodowa współpraca badawcza

Od czerwca do września 2014 roku wizytowałem Uniwersytet Alpen-Adria w Klagenfurcie realizując projekt w którym analizowałem wpływ nierówności dochodowych na wybory agentów (konsekwencją tego stażu jest praca **B2**).

W latach 2017-22 byłem jednym z dwóch przedstawicieli Polski w Management Committee międzynarodowej grupy badawczej *European Network for Game Theory GAMENET* prowadzonej przez Maastricht University (Holandia), finansowanej przez COST Action. W ramach tej sieci brałem udział w pracach dwóch podgrup: *Learning in large-scale distributed networks* i *Stochastic methods in game theory*. W listopadzie 2021 roku wizytowałem CNRS Inria Grenoble realizując tam z prof. Panayotisem Mertikopoulosem projekt pt. *Game dynamics and chaos*. Jego wynikiem jest publikacja **A5**. Ponadto od 2017 roku współpracuję z Georgiosem Piliourasem (Singapore University of Technology and Design, Singapur; Google DeepMind, Wielka Brytania), Michałem Misiurewiczem (Indiana University, USA) oraz Thiparatem Chotibutem (Chulalongkorn University, Tajlandia). W ramach tego interdyscyplinarnego zespołu realizuję projekt *Non-equilbrium phenomena in game theory*. Współpraca ta zaowocowała pracami **A1-A4**, a także [13, 14]. Moim staraniem w Katedrze Matematyki UEK powstała i intensywnie rozwija się grupa badawcza skupiona na tematyce uczenia się w grach, oraz algorytmicznej i ewolucyjnej teorii gier.

Otrzymane wyniki pomogły uzyskać na lata 2024-27 grant NCN w ramach konkursu Opus 26 pt. *Ewolucja zachowań podmiotów gospodarczych w warunkach ograniczonej racjonalności* (grant nr 2023/51/B/HS4/01343, funkcja: kierownik 4-osobowego zespołu). Projekt ten będzie realizowany z wymienionym wyżej międzynarodowym zespołem poszerzonym o prof. Heinricha Naxa (UZH Zurich, Szwajcaria), Cesare Carissimo (ETH Zurich, Szwajcaria) oraz prof. Panayotisa Mertikopoulosa (CNRS, Francja). Pierwsze uzyskane w ramach tego grantu wyniki są już dostępne w formie preprintów [12, 14].

## 3.3 Uczestnictwo w konferencjach i seminariach naukowych

Prezentowałem wyniki swoich badań na ponad 30 konferencjach. Do najważniejszych należą: 18th International Symposium on Dynamic Games and Applications, Hurwicz Workshop on Mechanism Design Theory, 34th Conference on Neural Information Processing Systems, 38th International Conference on Machine Learning, XXX European Workshop on Economic Theory; Learning, Evolution and Games. Wygłosiłem też godzinny wykład na Games, Learning and Networks (Singapur, kwiecień 2023). Szczegółowa lista referatów znajduje się w Załączniku 4.

Swoje rezultaty prezentowałem na następujących seminariach:

- Chaos i informacja kwantowa, seminarium na Wydziale Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego, kilkanaście referatów w latach 2003-2020
- Matematyka Stosowana, seminarium na Wydziale Matematyki i Informatyki Uniwersytetu Jagiellońskiego, kilka referatów w latach 2014-2021
- Układy Dynamiczne, seminarium na Wydziale Matematyki i Informatyki Uniwersytetu Jagiellońskiego, kilka referatów w latach 2005-2021
- seminarium Zespołu Naukowego Modelowania Ekonometrycznego na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego, kilka referatów w latach 2020-24
- *Gry dynamiczne i informacja w grach* - seminarium robocze (grudzień 2017, UW; maj 2018, UŁ; wrzesień 2019, UEK)
- Seminarium CNRS Inria, Grenoble, listopad 2021
- posiedzenie naukowe Komisji Nauk Ekonomicznych Polskiej Akademii Umiejętności, 20 maja 2021r.
- zebrania Katedry Matematyki Uniwersytetu Ekonomicznego w Krakowie.

# 4 Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę

## 4.1 Osiągnięcia dydaktyczne

- Prowadzę, bądź prowadziłem, następujące kursy na Uniwersytecie Ekonomicznym w Krakowie:

  - Algebra i analiza matematyczna (kierunek: Informatyka stosowana, I stopień)
  - Algebra liniowa (kierunki: Informatyka i Ekonometria, Analityka gospodarcza, I stopień)
  - Algorytmiczna teoria gier (kierunek: Informatyka stosowana, I stopień)
  - Analiza matematyczna I, II (kierunki: Informatyka i Ekonometria, Analityka gospodarcza, I stopień)
  - Elementy teorii gier (kierunek: Analityka gospodarcza, I stopień)
  - Financial Mathematics (kierunek: Corporate Finance, I stopień)
  - Incentives in Computer Science (kierunek: Applied Informatics, I stopień)
  - Matematyka (kierunki: Ekonomia, Zarządzanie, I stopień)
  - Obliczenia symboliczne i numeryczne w Maple (kierunek: Informatyka i Ekonometria, I stopień)
  - Rynki na sieciach (kierunek: Informatyka stosowana, I stopień)
  - Teoria gier (kierunek: Analityka gospodarcza, I stopień)
  - Teoria gier w naukach społecznych (kierunek: Nauki społeczne stosowane, I stopień)
  - Wprowadzenie do matematyki (kierunek: Informatyka stosowana, I stopień)
  - Zaawansowane metody w analityce biznesowej (kierunek: Global Business Services, I stopień).

- współtworzyłem materiały pomocnicze do kursów *Zajęcia wyrównawcze — wprowadzenie do matematyki* oraz *Analiza matematyczna i algebra liniowa* na kierunku zamawianym *Informatyka stosowana.*

- w latach 2020-25 byłem opiekunem studentów II roku studiów licencjackich kierunku *Analityka gospodarcza.*

- byłem promotorem dwóch prac licencjackich (kierunek Analityka gospodarcza). Wyniki jednej z nich były prezentowane na konferencji studenckiej i wydane w materiałach pokonferencyjnych (E. Faracik *Algorytmy typu no-regret w oligopolach Cournot i Betrtranda* w [92]).

- w maju 2024 r. na zaproszenie prof. Dirka Hellbinga (ETH Zurich) i prof. Heinricha Naxa (UZH) miałem wykłady w ramach kursu *Controversies in Game Theory*, realizowanego wspólnie przez UZH i ETH.

## 4.2 Osiągnięcia organizacyjne

Czynnie uczestniczyłem też w pracach organizacyjnych środowiska naukowego. Byłem

- członkiem komitetu organizacyjnego *The XIXth European Workshop on General Equilibrium Theory* (Kraków, czerwiec 2010)
- przewodniczącym komitetu organizacyjnego *Andrzej Malawski Memorial Session* (październik 2017)
- przewodniczącym komitetu organizacyjnego konferencji *Gamenet Action Conference*, *Gamenet Conference Kraków 2018* (wrzesień 2018)
- przewodniczącym komitetu organizacyjnego szkoły dla doktorantów i młodych naukowców *Gamenet Training School on Recent Applications of Game Theory* (wrzesień 2018)
- członkiem komitetu organizacyjnego konferencji *Nauki społeczne - matematyczne czy matematyzowalne* pamięci prof. Andrzeja Malawskiego (wrzesień 2018)
- członkiem komitetu organizacyjnego *Conference on Dynamical Systems celebrating Michał Misiurewicz's 70th birthday* (czerwiec 2019)
- członkiem *Scientific Committee Gamenet Conference 2019* w Pradze (listopad 2019)

Biorę udział w Uczelnianym Zespole ds. Ewaluacji Jakości Działalności Naukowej.

# 5 Inne

## 5.1 Promotorstwo pomocnicze

Jestem promotorem pomocniczym w przewodzie doktorskim mgr Elżbiety Pliś pt. *Różnorodność a ewolucja systemu ekonomicznego* napisanej pod kierunkiem dr hab. Agnieszki Lipiety, prof. UEK. Praca uzyskała pozytywne recenzje, a obrona planowana jest na 28 marca 2025 r.

## 5.2 Recenzje

Recenzowałem prace naukowe w czasopismach takich jak

1. Central European Journal of Economic Modelling and Econometrics, MNiSW: 70p., IF23: 0.2;

2. Applied Mathematics and Computation, MNiSW: 100p., IF23: 3.5;
3. Journal of Difference Equations and Applications, MNiSW: 70p. , IF23: 1.1;
4. Reports on Mathematical Physics, MNiSW: 70p., IF23: 1;
5. Entropy, MNiSW: 100p., IF23: 2.4;
6. Graphs and Combinatorics, MNiSW: 70p., IF23: 0.6;
7. Journal of Classification, MNiSW: 140p., IF23: 1.8;
8. Chaos, MNiSW: 140p., IF23: 2.7;
9. Discrete and Continuous Dynamical Systems - series B, MNiSW: 100p., IF23: 1.3;
10. Electronic Commerce Research and Apllications, MNiSW: 100p., IF23: 5.9.
11. Annales Universitatis Paedagogicae Cracoviensis

Systematycznie zamieszczałem też recenzje w *Mathematical Reviews* prowadzonym przez American Mathematical Society.

Signed by /
Podpisano przez:

Fryderyk
Falniowski

Date / Data:
2025-04-14 08:52

# Referencje


[1] D. Acemoglu, K. Bimpikis, and A. Ozdaglar. Price and capacity competition. *Games and Economic Behavior*, 66(1):1–26, 2009.

[2] D. Acemoglu and A. Ozdaglar. Competition and efficiency in congested markets. *Mathematics of Operations Research*, 32(1):1–31, 2007.

[3] R. Arnott and K. Small. The economics of traffic congestion. *American Scientist*, 82(5):446–455, 1994.

[4] S. Arora, E. Hazan, and S. Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[5] K. J. Arrow. Rawls's principle of just saving. *The Swedish Journal of Economics*, 75(4):323–335, 1973.

[6] W. B. Arthur. Inductive reasoning and bounded rationality. *American Economic Review*, 84(2):406–411, 1994.

[7] A. B. Atkinson. On the measurement of inequality. *Journal of Economic Theory*, 2(3):244–263, 1970.

[8] R. Axelrod. Effective choice in the prisoner's dilemma. *Journal of conflict resolution*, 24(1):3–25, 1980.

[9] R. L. Axtell and J. D. Farmer. Agent-based modeling in economics and finance: past, present, and future. *Journal of Economic Literature*, pages 1–101, 2022.

[10] Y. Babichenko. Completely uncoupled dynamics and Nash equilibria. *Games and Economic Behavior*, 76(1):1–14, 2012.

[11] G. Bavly and R. Peretz. Limits of correlation in repeated games with bounded memory. *Games and Economic Behavior*, 115:131–145, 2019.

[12] J. Bielawski, Ł. Cholewa, and F. Falniowski. The emergence of chaos in population game dynamics induced by comparisons. *arXiv preprint arXiv:2412.06037*, 2024.

[13] J. Bielawski, T. Chotibut, F. Falniowski, M. Misiurewicz, and G. Piliouras. Heterogeneity, reinforcement learning and chaos in population games. to appear in Proceedings of National Academy of Sciences.

[14] J. Bielawski, T. Chotibut, F. Falniowski, M. Misiurewicz, and G. Piliouras. Interval maps mimicking circle rotations. *arXiv preprint arXiv:2411.01495*, 2024.

[15] J. Bielawski and M. Jakubek. The interplay between migrants and natives as a determinant of migrants' assimilation: A coevolutionary approach. *Central European Journal of Economic Modelling and Econometrics*, 13:213–251, 2021.

[16] K. Bimpikis, S. Ehsani, and R. Ilkılıç. Cournot competition in networked markets. *Management Science*, 65(6):2467–2481, 2019.

[17] K. Binmore and L. Samuelson. An economist's perspective on the evolution of norms. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, pages 45–63, 1994.

[18] A. Blum, V. Kumar, A. Rudra, and F. Wu. Online learning in online auctions. *Theoretical Computer Science*, 324(2-3):137–146, 2004.

[19] Z. Y. Brown and A. MacKay. Competition in pricing algorithms. *American Economic Journal: Microeconomics*, 15(2):109–156, 2023.

[20] A. Cabrales, O. Gossner, and R. Serrano. Entropy and the value of information for investors. *American Economic Review*, 103(1):360–77, 2013.

[21] A. Cabrales, O. Gossner, and R. Serrano. A normalized value for information purchases. *Journal of Economic Theory*, 170:266–288, 2017.

[22] H. Cai, K. Ren, W. Zhang, K. Malialis, J. Wang, Y. Yu, and D. Guo. Real-time bidding by reinforcement learning in display advertising. In *Proceedings of the tenth ACM international conference on web search and data mining*, pages 661–670, 2017.

[23] E. Calvano, G. Calzolari, V. Denicolo, and S. Pastorello. Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110(10):3267–3297, 2020.

[24] C. F. Camerer. *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press, 2011.

[25] A. Charpentier, R. Elie, and C. Remlinger. Reinforcement learning in economics and finance. *Computational Economics*, pages 1–38, 2021.

[26] E. Chastain, A. Livnat, C. Papadimitriou, and U. Vazirani. Algorithms, games, and evolution. *Proceedings of the National Academy of Sciences*, 111(29):10620–10623, 2014.

[27] S.-H. Chen and U. Gostoli. Coordination in the el farol bar problem: The role of social preferences and social networks. *Journal of Economic Interaction and Coordination*, 12:59–93, 2017.

[28] Y. K. Cheung and G. Piliouras. Vortices instead of equilibria in minmax optimization: Chaos and butterfly effects of online learning in zero-sum games. In *Conference on Learning Theory*, PMLR, pages 807–834, 2019.

[29] D. Coey, B. J. Larsen, K. Sweeney, and C. Waisman. Scalable optimal online auctions. *Marketing Science*, 40(4):593–618, 2021.

[30] R. Colini-Baldeschi, R. Cominetti, P. Mertikopoulos, and M. Scarsini. The asymptotic behavior of the price of anarchy. In *International Conference on Web and Internet Economics*, pages 133–145. Springer, 2017.

[31] R. Colini-Baldeschi, R. Cominetti, and M. Scarsini. Price of anarchy for highly congested routing games in parallel networks. *Theory of Computing Systems*, 63(1):90–113, 2019.

[32] R. Cominetti, J. R. Correa, and N. E. Stier-Moses. The impact of oligopolistic competition in networks. *Operations Research*, 57(6):1421–1437, 2009.

[33] T. M. Cover. *Elements of information theory.* John Wiley & Sons, 1999.

[34] J. P. Crutchfield and D. P. Feldman. Regularities unseen, randomness observed: Levels of entropy convergence. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 13(1):25–54, 2003.

[35] B. De Borger and K. Van Dender. Prices, capacities and service levels in a congestible Bertrand duopoly. *Journal of Urban Economics*, 60(2):264–283, 2006.

[36] A. Dolgopolov. Reinforcement learning in a prisoner's dilemma. *Games and Economic Behavior*, 144:84–103, 2024.

[37] D. Easley, J. Kleinberg, et al. *Networks, crowds, and markets: Reasoning about a highly connected world*, volume 1. Cambridge University Press, 2010.

[38] O. Edhan, Z. Hellman, and D. Sherill-Rofe. Sex with no regrets: How sexual reproduction uses a no regret learning algorithm for evolutionary advantage. *Journal of theoretical biology*, 426:67–81, 2017.

[39] M. Ekmekci, O. Gossner, and A. Wilson. Impermanent types and permanent reputations. *Journal of Economic Theory*, 147(1):162–178, 2012.

[40] A. A. El Anshasy, M. Shamsuddin, and M.-S. Katsaiti. Proximity-sensitive relative deprivation and international migration intentions. *Migration Letters*, 19(6):833–841, 2022.

[41] A. A. El Anshasy, M. Shamsuddin, and M.-S. Katsaiti. Financial wellbeing and international migration intentions: Evidence from global surveys. *Journal of Happiness Studies*, 24(7):2261–2289, 2023.

[42] I. Erev and A. E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, pages 848–881, 1998.

[43] L. Esposito, S. M. Kumar, A. Villaseñor, and R. F. Souza. Relative deprivation and Hey and Lambert's motivation: Mixed methods evidence from Rio de Janeiro. *Journal of Income Distribution*, 2023.

[44] E. Faingold. Reputation and the flow of information in repeated games. *Econometrica*, 88(4):1697–1723, 2020.

[45] J. D. Farmer. *Making Sense of Chaos: A Better Economics for a Better World*. Yale University Press, 2024.

[46] J. D. Farmer and D. Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, 2009.

[47] M. Feldman, N. Immorlica, B. Lucier, T. Roughgarden, and V. Syrgkanis. The price of anarchy in large games. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 963–976, New York, NY, USA, 2016. ACM.

[48] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. In *A long-run collaboration on long-run games*, pages 209–230. World Scientific, 2009.

[49] T. Galla and J. D. Farmer. Complex dynamics in learning complicated games. *Proceedings of the National Academy of Sciences*, 110(4):1232–1236, 2013.

[50] K. Gero, A. Yazawa, N. Kondo, M. Hanazato, K. Kondo, and I. Kawachi. Comparison of three indices of relative income deprivation in predicting health status. *Social Science & Medicine*, 294:114722, 2022.

[51] O. Gossner. Simple bounds on the value of a reputation. *Econometrica*, 79(5):1627–1641, 2011.

[52] P. Grassberger. Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics*, 25:907–938, 1986.

[53] P. Grassberger. Information and complexity measures in dynamical systems. In *Information dynamics*, pages 15–33. Springer, 1991.

[54] L. Hammond, A. Chan, J. Clifton, J. Hoelscher-Obermaier, A. Khan, E. McLean, C. Smith, W. Barfuss, J. Foerster, T. Gavenčiak, et al. Multi-agent risks from advanced AI. *arXiv preprint arXiv:2502.14143*, 2025.

[55] P. J. Hammond. Dual interpersonal comparisons of utility and the welfare economics of income distribution. *Journal of Public Economics*, 7(1):51–71, 1977.

[56] J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.

[57] J. C. Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321, 1955.

[58] J. C. Harsanyi. Rule utilitarianism and decision theory. *Erkenntnis*, 11(1):25–53, 1977.

[59] S. Hart. Commentary: Nash equilibrium and dynamics. *Games and Economic Behavior*, 71(1):6–8, 2011.

[60] S. Hart and A. Mas-Colell. Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review*, 93(5):1830–1836, 2003.

[61] S. Hart and A. Mas-Colell. Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economic Behavior*, 57(2):286–303, 2006.

[62] S. Hart and A. Mas-Colell. *Simple adaptive strategies: from regret-matching to uncoupled dynamics*, volume 4. World Scientific, 2013.

[63] D. Helbing. A mathematical model for behavioral changes by pair interactions. *Economic evolution and demographic change: formal models in social sciences*, pages 330–348, 1992.

[64] Z. Hellman and R. Peretz. A survey on entropy and economic behaviour. *Entropy*, 22(2):157, 2020.

[65] T.-H. Ho and C. Camerer. Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity, and time-variation. *Journal of Mathematical Psychology*, 42:305–326, 1998.

[66] T.-H. Ho and C. Camerer. Experience-weighted attraction learning in normal form games. *Econometrica*, 67:827–874, 1999.

[67] T.-H. Ho, C. F. Camerer, and J.-K. Chong. Self-tuning experience weighted attraction learning in games. *Journal of Economic Theory*, 133:177–198, 2007.

[68] H. O. Hounkpatin, A. M. Wood, and G. D. Brown. Comparing indices of relative deprivation using behavioural evidence. *Social Science & Medicine*, 259:112914, 2020.

[69] J. H. Kagel and A. E. Roth. *The handbook of experimental economics, volume 2*. Princeton university press, 2016.

[70] E. Kalai and E. Lehrer. Rational learning leads to Nash equilibrium. *Econometrica*, 61(5):1019–1045, 1993.

[71] T. Klein. Autonomous algorithmic collusion: Q-learning under sequential pricing. *The RAND Journal of Economics*, 52(3):538–558, 2021.

[72] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *(STACS)*, pages 404–413. Springer-Verlag, 1999.

[73] R. Kumar, J. Schneider, and B. Sivan. Strategically-robust learning algorithms for bidding in first-price auctions. *arXiv preprint arXiv:2402.07363*, 2024.

[74] S. Leonardos, B. Monnot, D. Reijsbergen, E. Skoulakis, and G. Piliouras. Dynamical analysis of the eip-1559 ethereum fee market. In *Proceedings of the 3rd ACM Conference on Advances in Financial Technologies*, pages 114–126, 2021.

[75] M. Mäs and H. H. Nax. A behavioral study of "noise" in coordination games. *Journal of Economic Theory*, 162:195–208, 2016.

[76] R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995.

[77] B. McMahan. Follow-the-regularized-leader and mirror descent: Equivalence theorems and l1 regularization. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 525–533, 2011.

[78] E. Melo. Price competition, free entry, and welfare in congested markets. *Games and Economic Behavior*, 83:53–72, 2014.

[79] P. Mertikopoulos, C. Papadimitriou, and G. Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2703–2717. SIAM, 2018.

[80] P. Mertikopoulos and W. H. Sandholm. Riemannian game dynamics. *Journal of Economic Theory*, 177:315–364, 2018.

[81] P. Milgrom and J. Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6):1255–1277, 1990.

[82] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, pages 124–143, 1996.

[83] P. A. P. Moran. The statistical processes of evolutionary theory. 1962.

[84] J. Newton. Conventions under heterogeneous behavioural rules. *The Review of Economic Studies*, 88(4):2094–2118, 2020.

[85] A. Neyman and D. Okada. Strategic entropy and complexity in repeated games. *Games and Economic Behavior*, 29(1-2):191–223, 1999.

[86] A. Neyman and D. Okada. Repeated games with bounded entropy. *Games and Economic Behavior*, 30(2):228–247, 2000.

[87] A. Neyman and D. Okada. Growth of strategy sets, entropy, and nonstationary bounded recall. *Games and Economic Behavior*, 66(1):404–425, 2009.

[88] A. Neyman and J. Spencer. Complexity and effective prediction. *Games and Economic Behavior*, 69(1):165–168, 2010.

[89] T. Obloj and T. Zenger. Organization design, proximity, and productivity responses to upward social comparison. *Organization Science*, 28(1):1–18, 2017.

[90] M. Pangallo, T. Heinrich, and J. D. Farmer. Best reply structure and equilibrium convergence in generic games. *Science Advances*, 5(2):eaat1328, 2019.

[91] E. Pels and E. T. Verhoef. The economics of airport congestion pricing. *Journal of Urban Economics*, 55(2):257–277, 2004.

[92] A. Prędki. *Narzędzia analityczne – teoria i zastosowania w naukach ekonomicznych*. Mfiles.pl, 2023.

[93] C. Quintana-Domeque and J. Wohlfart. "Relative concerns for consumption at the top": An intertemporal analysis for the UK. *Journal of Economic Behavior & Organization*, 129:172–194, 2016.

[94] J. Rawls. Two concepts of rules. *The Philosophical Review*, 64(1):3–32, 1955.

[95] J. Rawls. Some reasons for the maximin criterion. *American Economic Review*, 64(2):141–146, 1974.

[96] R. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[97] A. E. Roth and I. Erev. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8(1):164–212, 1995.

[98] A. E. Roth and I. Erev. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1):164–212, 1995.

[99] T. Roughgarden. The price of anarchy is independent of the network topology. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 428–437, 2002.

[100] T. Roughgarden. *Selfish routing and the price of anarchy*. MIT press, 2005.

[101] T. Roughgarden. *Twenty lectures on algorithmic game theory*. Cambridge University Press, 2016.

[102] S. Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[103] T. C. Salmon. An evaluation of econometric models of adaptive learning. *Econometrica*, 69(6):1597–1628, 2001.

[104] L. Samuelson and J. Zhang. Evolutionary stability in asymmetric games. *Journal of Economic Theory*, 57(2):363–391, 1992.

[105] J. B. T. Sanders, J. D. Farmer, and T. Galla. The prevalence of chaotic dynamics in games with many players. *Scientific Reports*, 8(1):1–13, 2018.

[106] W. H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2010.

[107] O. Stark. Stressful integration. *European Economic Review*, 63:1–9, 2013.

[108] O. Stark. On a tendency in health economics to dwell on income inequality and underestimate social stress. *Economics & Human Biology*, 49:101232, 2023.

[109] O. Stark, J. Bielawski, and F. Falniowski. Measuring income inequality in social networks. *Journal of Economic Inequality*, 22(2):333–356, 2024.

[110] M. Sweet. Does traffic congestion slow the economy? *Journal of Planning Literature*, 26(4):391–404, 2011.

[111] P. D. Taylor and L. B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1–2):145 – 156, 1978.

[112] T. Van Erven, D. Van der Hoeven, W. Kotłowski, and W. M. Koolen. Open problem: Fast and optimal online portfolio selection. In *Conference on learning theory*, pages 3864–3869. PMLR, 2020.

[113] E.-V. Vlatakis-Gkaragkounis, L. Flokas, T. Lianeas, P. Mertikopoulos, and G. Piliouras. No-regret learning and mixed Nash equilibria: They do not mix. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 1380–1391. Curran Associates, Inc., 2020.

[114] L. Waltman and U. Kaymak. Q-learning agents in a Cournot oligopoly model. *Journal of Economic Dynamics and Control*, 32(10):3275–3293, 2008.

[115] C. J. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8:279–292, 1992.

[116] S. Yitzhaki. Relative deprivation and the Gini coefficient. *The Quarterly Journal of Economics*, 93(2):321–324, 1979.

[117] C. S. Yoo. Network neutrality and the economics of congestion. *Geo. LJ*, 94:1847, 2005.