VALENTINA M. SEMENOVA

89 Kingston Road • Oxford, OX2 6RJ • 07937969214 • semenova@maths.ox.ac.uk

EDUCATION

OXFORD UNIVERSITY, PHD CANDIDATE, Mathematics (expected) Jun 2023

- Student of Professor J. Doyne Farmer in the Complexity Group focused on better understanding the dynamics of complex systems
- Student of Professor Xiaowen Dong in the Oxford Man Institute developing graph signal processing techniques, including Graph Neural Networks

COLUMBIA UNIVERSITY, MASTER OF SCIENCE, Operations Research Dec 2018

- Student of Professor Yuri Faenza researching discrete optimization
- Winner of *Robert F. Gartland Fellowship* for academic excellence and professional promise

DARTMOUTH COLLEGE, BACHELOR OF ARTS, Majors: Mathematics, Economics Jun 2013

SELECTED RESEARCH

REDDIT'S SELF-ORGANIZED BULL RUNS *(with Julian Winkler)*

- This paper develops an empirical and theoretical case for how 'hype' among retail investors can drive large asset price fluctuations. We use text data from discussions on WallStreetBets (WSB), an online investor forum with over eleven million followers as of February 2022, as a case study to demonstrate how retail investors influence each other, and how social behaviours impact financial markets. We document that WSB users adopt price predictions about assets (bullish or bearish) in part due to the sentiments expressed by their peers. Discussions about stocks are also self-perpetuating: narratives about specific assets spread at an increasing rate before peaking, and eventually diminishing in importance -- a pattern reminiscent of an epidemiological setting. To consolidate these findings, we develop a model for the impact of social dynamics among retail investors on asset prices. We find that the interplay between 'trend following' and 'consensus formation' determines the stability of price returns, with socially-driven investing potentially causing oscillations and cycles. Our framework helps identify components of asset demand stemming from social dynamics, which we predict using WSB data. Our predictions explain significant variation in stock market activity. These findings emphasise the role that social dynamics play in financial markets, amplified by online social media.
- ***Coverage: Financial Times, LSE Business Review***
- ***Winner of second place prize in Rebuilding Macroeconomics: Complexity in Macroeconomics competition, 2021***
- ***Link: https://arxiv.org/abs/2104.01847***

DEBAGREEMENT *(with John Pougue-Biyong)*

- In this paper, we introduce DEBAGREEMENT, a dataset of 42,894 comment-reply pairs from the popular discussion website Reddit, annotated with agree, neutral or disagree labels. We collect data from five forums on Reddit: BlackLivesMatter, Brexit, climate, democrats, Republican. For each forum, we select comment pairs such that they form altogether a user interaction graph. DEBAGREEMENT presents a challenge for Natural Language Processing (NLP) systems, as it contains slang, sarcasm and topic-specific jokes, often present in online exchanges. We evaluate the performance of state-of-the-art language models on a (dis)agreement detection task, and investigate the use of contextual information available (graph, authorship, and temporal information). Since recent research has shown that context, such as social context or knowledge graph information, enables language models to better perform on downstream NLP tasks, DEBAGREEMENT provides novel opportunities for combining graph-based and text-based machine learning techniques to detect (dis)agreements online.
- ***Published in NeurIPS 2021***
- ***Link: https://openreview.net/forum?id=udVUN__gFO***

FROM MICRO TO MACRO: UNDERSTANDING THE SOCIAL DYNAMICS BEHIND POLITICAL AND ECONOMIC CHANGE *(with John Pougue-Biyong)*

- This paper outlines and tests a novel signed, temporal clustering algorithm on static and temporal, synthetically-generated data. The algorithmic performance justifies its use on noisy social media data to better understand the dynamics of the Brexit discussion. We apply the algorithm on a hand-annotated dataset of agreements and disagreements on the *r/Brexit* discussion forum on Reddit. The algorithm highlights periods of social turmoil (July-September 2019) and of relative stability (May-August 2020). Our proposed metric for community overlap over time correlates with unemployment and the GBP/USD exchange rate, linking the data on social discussions to the macroeconomy. We use NLP techniques to offer qualitative insights into the discussion.
- ***Winner of first place prize in Rebuilding Macroeconomics: Complexity in Social Macroeconomics competition, 2022***

WORK EXPERIENCE

PALANTIR TECHNOLOGIES New York, NY
Deployment Strategist 2015 – 2017

- *Drove Palantir's expansion in the Asia Pacific region*
  - Part of a 9-person team executing "proof of concept" project with our first commercial client in the Australia
  - Led the development of a software product targeted at better storage and analysis of customer data
  - Implemented several complex calculations which improved the accuracy of loan payments data
- *Developed Palantir's joint venture with Credit Suisse targeted at mitigating financial operational risk*
  - Identified traders who pose high "operational risk" through numerically encoding behavior and outlier detection modeling
  - Developed software to capture user feedback for machine learning improvements to outlier detection model

GOLDMAN SACHS New York, NY
Algorithmic Volatility Trader 2013 – 2015

- *Proposed and executed successful trading strategies for a portfolio of options on over 1200 stocks*
  - Trading strategies targeted preemptively identifying toxic flow, systematically preventing unfavorable trading and modeling market trends in order to enhance future returns
  - Studied and implemented advanced statistical tools to improve our group's modelling accuracy

SELECTED CONFERENCES AND PRESENTATIONS

- NeurIPS 2021, Royal Economic Society (organizer of special session), Networks 2021, Alan Turing Institute – Economic Data Science Interest Group, Santa Fe Institute – Annual Risk Meeting, NetSci 2022, Institute of Mathematics and its Applications – the Mathematical Challenges of Big Data 2022

SKILLS, ACTIVITIES AND INTERESTS

ENTREPRENEURSHIP: Co-founder of a biotech startup targeted at increasing human lifespan (2017)
LANGUAGES: fluent in Russian; conversational in French, Spanish
OTHER ACTIVITIES AND INTERESTS: Alumna of the Dartmouth Varsity Sailing team, painting, yoga (certified yoga instructor)