Georgii Zherebilov
Universitat Pompeu Fabra, Department of Economics and Business
C/ Ramon Trias Fargas 25–27, Barcelona, Spain
+34 744 649 299 | georgii.zherebilov@upf.edu

I am a 6th-year PhD student in Economics at Pompeu Fabra University, Spain. I completed two years of PhD coursework at the Barcelona School of Economics and UPF, and I am currently working on my dissertation under the supervision of Ruben Enikolopov and Maria Petrova. I expect to be on the job market in fall 2028. My research interests lie in Applied Microeconomics and Political Economy, with a focus on digital platforms and political communication. Across my dissertation projects, I study how social media algorithms, automated accounts, and AI tools shape information environments, user behavior, and the diffusion of political content. My work combines large-scale online data, causal inference, and modern computational tools, with the goal of generating policy-relevant insights on digital communication and the use of AI to fight misinformation.

In my job market paper, I study whether AI can efficiently combat misinformation on social media. I use X Community Notes data to show that AI notes are accepted by the community 17% more often than human notes on the same post. I show that AI notes receive more favorable votes from raters who usually disagree with similar notes. The note rating algorithm values such votes the most, driving up the acceptance of AI-written notes.

In a companion online survey run on Prolific, I show participants potentially misleading X posts and randomly show them AI-written or human-written fact-checking notes. To separate author perceptions from note quality, I also randomize whether respondents see who wrote the notes. Respondents rate AI notes as more informative and trustworthy and are more willing to share these notes with other participants. Disclosing authorship makes AI notes seem less biased, but Republican participants view them less favorably. The project provides new evidence on how AI can be used to correct polarizing misinformation online, and on whether people trust the information and judgment of AI in a highly polarized environment.

In a second project, joint with Alexey Makarin, Andrey Simonov, Rafael Jiménez-Durán, and Sergey Sanovich, we study how coordinated pro-government bot activity shapes political communication on social media. We partner with Botnadzor, a Russian grassroots organization that develops algorithms to detect bot accounts, and use a unique dataset containing comments from over 100,000 identified bots across thousands of VK groups. Using variation from bot working hours, we show that a higher share of pro-government bots in the early comments significantly increases the number of subsequent pro-government comments by regular users under the same post. To identify the effects of bot exposure on political attitudes and online behavior, we conduct a randomized experiment in which participants install a browser extension that highlights bot comments and tracks browsing behavior across the Internet. This project provides some of the first causal evidence on how pro-government bots influence political beliefs, online engagement, and user behavior. It also evaluates the effectiveness of tools designed to identify and counter coordinated manipulation on social media, offering new insights into both the impact of political bots and strategies to combat online misinformation. A companion paper studies the rapid adoption of AI by political bots: using Pangram, an AI-detection tool, we classify bot accounts that extensively use AI. We compare them with human-operated bots in terms of communication strategies, content production, and patterns of activity. We also exploit disruptions at AI service providers as an instrument for AI bot activity to estimate its causal effects on user engagement and sentiment on social media.

Finally, in a joint project with Juan S. Morales and Annalí Casanueva Artís, we exploit the transition to the National Incident-Based Reporting System (NIBRS), which led to widespread disruptions in crime reporting to the FBI, as a quasi-exogenous shock to the availability of reliable crime statistics. Using annotated crime-related tweets by members of Congress, we investigate whether reduced access to verifiable information changes how politicians make factual claims and frame narratives about crime.