Alexander Goryunov

Data science, data analysis, econometrics, causal inference

Mobile: +972-53-571-0490
Email: alex.goryunov@gmail.com
Repo: https://github.com/agoryuno
LinkedIn: Alexander Goryunov
Blog: https://agoryuno.github.io

Education

2002 - PhD in International Economics from Economic research institute
1999 - degree in Finance from Khabarovsk State Technical University

Employment history

2019-2022: Federal comptroller's office of Russia, Russia - *chief inspector / senior data scientist*

- Came in as part of a small team of analysts and researchers charged with introducing data driven and modern causal inference statistical and ML methods into government auditing.
- Moved the team over to standard software development practices of using version control, testing, and writing modular code.
- Designed and implemented a new micromodel of household incomes, which is being used to assess the impact of social transfers on the poverty level and income distribution.
- Designed and oversaw the implementation of a system giving the agency access to Finance Ministry's data on federal budget planning and operational execution, along with a system of expert rules used for auditing of federal programs and expenses.
- My modelling work allowed the implementation in March 2022 of a new federal subsidy for poor families with children aged 7-18 years old .
- Introduced modern causal inference methods and their applications to policy evaluation to over 200 agency staff through a series of lectures and webinars.
- Co-authored agency-level guides on data transparency and reliability, and the use of quantitative methods for policy evaluation.

2012-2019: Economic research institute, Russia - *deputy director for science*

- Co-edited the report on a multidisciplinary project of foresight prediction of industrial development of Eastern Russian regions.
- Designed and read a course on modern data analysis methods with applications in R for the non-techie research staff.
- Secured three long-term research grants for a total of 1.5 mln USD.
- Published over 30 scientific articles and monographs in the areas of regional development and international economics.

2006-2009: Economic research institute, Russia - *head of economic theory department*

- Increased the department's grant financing threefold: from 100,000 USD annually to 300,000 USD.
- Estimated the demand for tourism services and total required investment in the tourism sector of the economy of the city of Vladivostok. The estimate was a major part of a research report which formed the basis of a federal program for the preparation for the APEC Summit of 2012.

2002-2006: Economic research institute, Russia - *researcher*

- Estimated regional economic effects of Russia's WTO accession. The estimates were used by several regional governments as basis for their adjustment claims on the federal government in preparation for the accession in 2008.
- (continued)

Tech stack

Analysis

- Git
- Cython
- Numpy
- Pandas
- Scikit Learn
- PyTorch
- Statsmodel
- PyMC3
- R

Engineering

- SQLAlchemy
- SQL
- Apache Beam
- Google Cloud Platform

Visualization

- Matplotlib
- Seaborn

Web interfaces

- Starlette
- Flask
- React.js

Alexander Goryunov

Data science, data analysis, econometrics, causal inference

Mobile: +972-53-571-0490
Email: alex.goryunov@gmail.com
Repo: https://github.com/agoryuno
LinkedIn: Alexander Goryunov
Blog: https://agoryuno.github.io

Employment history (continued)

1999-2002: Economic research institute, Russia - *post-graduate student*

- Defended a PhD thesis on the estimation of welfare effects of China's WTO accession

Recent projects

A database of real estate prices in major Russian cities from one of the two main open sources of classifieds. XHR calls to the website's internal API via Selenium used to scrape data, cloud-hosted Postgres with SQLAlchemy for data modeling used to store data. A simple binary tree-based method used to partition API queries used to improve reliability of the scraping service: https://github.com/agoryuno/binrange

Estimation of quality of high schools in 20 Russian regions based on the students' academic outcomes. Data was received from a Russian official agency and covered several years of anonymized standardized exam results for every student in the regional sample. A mixed effects model was chosen to extract the school's effect on individual outcomes with controls for age, gender and a proxy for extra-curricular activities of students. A quarter of an average student's result was attributable to their school with extreme regional inequality in access to education. The estimation method was proposed to the government agency as an alternative way of analyzing access to education.

A microsimulation model for long term projections of the Russian poverty level. The model was built in Python, based on data from the national household incomes survey. The model includes modules responsible for different scenarios of demographic change, and income and labour participation effects of the ongoing pension reform. It allows for scenario-based estimation of social transfers' income and poverty reduction effects.

Estimation of direct fiscal cost of additional subsidies for poor families with children. In the absence of data on family incomes, as opposed to household incomes, as well as a discrepancy in the definitions of monetary poverty used by the statistics and social support agencies, meant that neither total number of families with children, nor the number of poor families with children were directly known from official statistics. This gap in data was bridged by an MCMC estimation (PyMC3) utilizing partial information on existing family subsidies from the responsible agency. The total expenditure calculation was supplemented with sensitivity analysis for the possible margin of error, as well as estimates of the labor substitution effect of the new subsidy. The results were accepted by the government and the new subsidy was instituted about 6 months later.

Finding irregularities in the voting patterns of a suspected rigged gubernatorial election in a Russian region in 2018. Data scraping of official websites of polling stations for voting results and station coordinates (mechanize + BeautifulSoup). PCA, spatial autoregression, and visual analysis in R showed extreme spikes of votes for the winning candidate in two large cities, which were also tightly spatially correlated within the cities, suggesting that "networks" of administrators on the ground were involved in the doctoring of results as opposed to centralized vote counting fraud.