Forecasting Refugee Migration with High-Dimensional Covariate Space

Haodong Qi | Malmo University

Forecasting refugee migration is challenging, exacerbated by the high dimensional and dynamic nature of its drivers, such as climatic, economic, and political stressors. In this talk, I will introduce a novel forecasting framework which incorporates a time-varying regularization and a new model selection criterion: the Penalized Deviance Ratio (PDR). Unlike conventional metrics such as the Deviance Ratio (DR), which emphasize in-sample fit, PDR explicitly penalizes model complexity, enhancing generalization in high-dimensional covariate settings. We apply this framework to forecast asylum seeker rates (ASR) from Somalia to EU member states, leveraging a comprehensive set of district-level predictors. Extensive validation demonstrates that PDR-tuned models consistently outperform DR-based benchmarks in out-of-sample accuracy, reducing average point prediction errors by 40% and improving interval forecasts by 79%. Furthermore, we demonstrate how the framework supports explanatory insights at multiple levels -- origin district, destination, and temporal -- revealing both persistent and transient nature of migration drivers. The proposed methodology not only advances forecasting accuracy under high-dimensional covariate conditions but also enhances the interpretability of complex and evolving migration systems.

Haodong Qi is an associate professor of demography at Stockholm University and a project researcher at Malmo University, Sweden. He obtained his Ph.D. in Economic Demography from Lund University in 2016. His research lies at the intersection of economics, demography, and statistics, with a broader ambition to advance social science through cutting-edge computational methods while maintaining strong commitments to transparency, interpretability, and policy relevance. Haodong is currently the Principal Investigator of CLIMB - a Belmont Forum project investigating Climate-Induced Migration in Africa and Beyond using Big Data and Predictive Analytics.