RESEARCH STATEMENT

Boyuan Zhang

I am interested in a variety of topics at the intersection of econometrics and machine learning, such as grouped heterogeneity in panel data models and density forecast combinations. I also work on climate forecasting and applied macroeconomics using econometric models.

In my job market paper "*Incorporating Prior Knowledge of Latent Group Structure in Panel Data Models*," I develop a constrained Bayesian grouped estimator that exploits researchers' prior beliefs on groups in a form of pairwise constraints, which represent the preference of assigning two units to the same group or distinct groups. I propose a prior to incorporate the pairwise constraints with varying degrees of confidence. The whole framework is built on the nonparametric Bayesian method, which implicitly assigns distribution over the group partitions, and so the posterior analysis takes the uncertainty of the latent group structure into account. Furthermore, this framework allows for specifying grouped structures in both slope coefficients and error variances of the cross-sectional units. Monte Carlo experiments reveals that adding prior knowledge yields more accurate estimates and scores predictive gains over unconstrained grouped estimators and estimators that omit the group structure by assuming homogeneity or full heterogeneity. I apply my method to two empirical applications. An application to forecasting inflation of the U.S. CPI sub-indices demonstrates that the suggested predictor yields more accurate density predictions. The better forecasting performance is mostly attributable to three key characteristics: the nonparametric Bayesian prior, prior belief on group structure, and grouped cross-sectional heteroskedasticity. In a second application, I revisit the relationship between a country's income and its democratic transition. The proposed framework recovers a group structure with a moderate number of groups. Each group has a clear and distinct path to democracy. In addition, I identify heterogeneous income effects on democracy and, contrary to the initial findings, show that a positive income effect persists in some groups of countries, though quantitatively small.

In "*On the Aggregation of Probability Assessments: Regularized Mixtures of Predictive Densities for Eurozone Inflation and Real Interest Rates*" (joint with Francis X. Diebold and Minchul Shin), we extend the idea of regularized forecast combination to the density forecast case and propose methods for constructing regularized mixtures of density forecasts. We explore a variety of objectives and regularization penalties, and we use them in a substantive exploration of Eurozone inflation and real interest rate density forecasts. All individual inflation forecasters (even the ex post best forecaster) are outperformed by our regularized mixtures. The average number of forecasters selected after regularization

is always small, regardless of the regularization method. From the Great Recession onwards, the optimal regularization tends to move density forecasts' probability mass from the centers to the tails, correcting for overconfidence.

In "*Optimal Combination of Arctic Sea Ice Extent Measures: A Dynamic Factor Modeling Approach*" (joint with Francis X. Diebold, Maxmilian Göbel, Philippe Goulet Coulombe, and Glenn Rudebusch), we propose and estimate a dynamic factor model that combines four of measures of Arctic sea ice extent in an optimal way and accounts for their differing volatility and cross-correlations. We then use the Kalman smoother to extract an optimal combined measure of Arctic sea ice extent. It turns out that almost all weight is put on the NSIDC Sea Ice Index, confirming and enhancing confidence in the Sea Ice Index and the NASA Team algorithm on which it is based.

In "*When Will Arctic Sea Ice Disappear? Projections of Area, Extent, Thickness, and Volume*" (joint with Francis X. Diebold, Maxmilian Göbel, Philippe Goulet Coulombe, and Glenn Rudebusch), we provide point, interval, and density forecasts for four measures of Arctic sea ice: area, extent, thickness, and volume. Importantly, we enforce the joint constraint that these measures must simultaneously arrive at an ice-free Arctic. We apply this constrained joint forecast procedure to models relating sea ice to cumulative carbon dioxide emissions and models relating sea ice directly to time. The resulting "carbon-trend" and "time-trend" projections are mutually consistent and predict an effectively ice-free summer Arctic Ocean by the mid-2030s with an 80% probability. Moreover, the carbon-trend projections show that global adoption of a lower emissions path would likely delay the arrival of a seasonally ice-free Arctic by only a few years.

In "*Measuring International Uncertainty: the Case of Korea*" (joint with Minchul Shin, Molin Zhong, and Dong Jin Lee), we leverage a data rich environment to construct and study a measure of macroeconomic uncertainty for the Korean economy. We provide several stylized facts about uncertainty in Korea from 1991M10-2016M5. We compare and contrast this measure of uncertainty with two other popular uncertainty proxies, financial and policy uncertainty proxies, as well as the U.S. measure constructed by Jurado et al. (2015). We find that neither financial nor policy uncertainty proxies capture economy-wide uncertainty. Unlike our measure or financial uncertainty, policy uncertainty does not have much effect on real variables in Korea.