

# HYOSEOK (STEVE) KIM

hk2562@columbia.edu ◇ 347-978-6904 ◇ Homepage

## EDUCATION

---

**Columbia University, Graduate School of Arts and Sciences**

Ph.D., Economics (Military Leave of Absence: 01/2021 - 12/2022)

New York, NY

09/2017 - 2025 (Expected)

**Columbia University, Graduate School of Arts and Sciences**

M.A., Economics

New York, NY

09/2016 - 08/2017

**Seoul National University, College of Social Sciences**

M.A., Economics

Seoul, South Korea

09/2014 - 08/2016

**Columbia University, Columbia College**

B.A., Chemistry, Economics-Mathematics

New York, NY

09/2009 - 05/2013

## FIELDS OF SPECIALIZATION

---

Primary Fields: Econometrics, Macroeconomics

Secondary Field: Asset pricing (Coursework: Asset Pricing 1, Asset Pricing 2, Empirical Asset Pricing)

## REFERENCES

---

Serena Ng

Edwin W. Rickert Professor

Department of Economics

Columbia University

serena.ng@columbia.edu

Jushan Bai

Professor

Department of Economics

Columbia University

jb3064@columbia.edu

Martin Uribe

Robert A. Mundell Professor

Department of Economics

Columbia University

mu2166@columbia.edu

## JOB MARKET PAPER

---

**Title: Macroeconomic forecasting with outlier-robust VAR**

**Abstract:** The vector autoregressive (VAR) model is one of the key models for macroeconomic forecasting and structural analysis. However, the emergence of COVID-19 has rendered the estimation and inference of VARs in the post-pandemic period challenging and has highlighted the need for outlier-robust models. The outlier-robust VAR model assumes that the reduced-form error is a sum of the regularized error and the outlier. Then, the PLS estimator of the VAR coefficients is obtained by imposing an  $L_1$  penalty on the outlier. It is demonstrated that the robust estimator of the VAR coefficients is equivalent to the multivariate Huber's M-estimator of the VAR coefficients. Subsequently, the asymptotic normality of the robust estimator is established on the foundation of the asymptotic normality of the multivariate M-estimators. To illustrate the methodology, we consider a five-variable VAR of macroeconomic variables, using monthly data from March 1959 to June 2024 sourced from the FRED-MD database. The results demonstrate that the out-of-sample forecasts for the post-Covid period from the robust VAR are generally more accurate than those from the standard VAR and the outlier-augmented stochastic-volatility Bayesian VAR proposed by Carriero, Clark, Marcellino, and Mertens (2022).

### **Title: Outlier-robust low-frequency trend in time-series**

**Abstract:** A significant number of interesting questions in macroeconomics are concerned with the analysis of trends or long-term variations in macroeconomic variables. Traditional methods for identifying trends in time series include the Beveridge-Nelson decomposition (Beveridge and Nelson, 1981), the HP filter (Hodrick and Prescott, 1997), and the band-pass filter (Baxter and King, 1999). A recent contribution to this literature is the concept of the low-frequency trend, as proposed by Muller and Watson (2008). This is defined as the fitted value of a linear regression of a time series on a small number of low-frequency cosine functions. Although the ordinary least squares (OLS) estimator is a common method for constructing the low-frequency trend, the recent pandemic has highlighted the potential benefits of an alternative estimator that is more robust to extreme observations. An outlier-robust low-frequency trend in a time series is obtained via penalized least squares (PLS) estimation of the regression coefficients, which imposes an  $L_1$  penalty on the error term of the regression. It is shown that the robust estimator is identical to Huber's M-estimator of the regression coefficients. Sufficient conditions for the asymptotic normality of the robust estimator are proposed based on the work of Wu (2007) on the M-estimator of linear regression with dependent errors. For empirical application, the levels and growth rates of U.S. real GDP per capita and U.S. real consumption per capita are decomposed into outlier-robust low-frequency trends, cycles, and outliers. Furthermore, identified outliers are discussed in the context of major historical events such as the Great Depression, World Wars, and pandemics.

### **Title: Bayesian Forecasting with High-Dimensional Macroeconomic Data**

**Abstract:** This study investigates the forecasting performance of several Bayesian models in the context of high-dimensional macroeconomic data. Two linear models, the Bayesian Lasso model and the spike-and-slab model, and one tree-based model, the Bayesian Additive Regression Trees (BAT), are considered. The Bayesian Lasso imposes a shrinkage prior, resulting in a dense representation, while the spike-and-slab model imposes a mixture prior that explicitly admits a sparse representation. Using 120 monthly macroeconomic and financial variables from the FRED-MD database, these models are evaluated based on their ability to forecast the one-month-ahead growth rate of U.S. industrial production. The result shows that all models tend to select variables commonly included in traditional vector autoregressive (VAR) models. Among these models, the BART delivers the best in-sample fit, due to its flexible non-parametric structure. However, in out-of-sample forecasting, linear models, especially the spike-and-slab model, outperform the BART. Furthermore, assuming t-distributed errors improves forecast accuracy, highlighting the importance of accounting for fat tails in macroeconomic data. These findings offer empirical guidance for selecting forecasting tools in the presence of model uncertainty and non-Gaussian disturbances.

### **Title: Estimation of mixed-frequency factor model via principal component analysis and maximum likelihood**

**Abstract:** This study presents a procedure for the maximum likelihood (ML) estimation of mixed-frequency factor models and conducts a simulation exercise that compares the ML estimators and the principal component analysis (PCA) estimators. The simulation exercise considers a variety of mixed-frequency models that allow for the following: (i) serially correlated common and group-specific factors, (ii) cross-sectionally correlated group-specific factors, (iii) cross-sectionally correlated idiosyncratic errors, and (iv) heteroskedastic idiosyncratic errors. The performance of the two estimators is compared by computing canonical correlations between the true factors and their estimates, as well as between the true factor loadings and their estimates. The simulation exercise demonstrates that the ML estimators for factors are more efficient than the PCA estimators for factors in terms of average canonical correlations. Similarly, the ML estimators for factor loadings are more efficient than the PCA estimators in the majority of cases. As an empirical application, we apply both methods to a panel dataset that consists of the quarterly growth rates of U.S. industrial production (IP) sectors and the annual growth rates of non-IP sectors.

## RESEARCH AND WORK EXPERIENCE

---

### **Amazon**

New York, NY

Economist Intern

June 2025 - August 2025

- Worked as an economist intern on the Amazon SCOT-AIM (Supply Chain Optimization Technologies – Automated Inventory Management) team.
- Developed an unbalanced Bayesian Dynamic Factor Model to identify latent drivers of Amazon's growth from 25+ key business metrics spanning topline performance and supply chain operations, producing both unconditional and conditional forecasts to support policy evaluation.

**Columbia University, Department of Economics**

New York, NY

Research Assistant for Prof. Jushan Bai

2018 - 2019

- Conducted Monte Carlo simulations for mixed-frequency factor models.

Research Assistant for Prof. Stepanie Schmitt-Grohe and Prof. Martin Uribe

June 2019 - August 2019

- Assisted in conducting simulations and preparing manuscripts.

**Seoul National University, Department of Economics**

Seoul, South Korea

Research Assistant for Prof. Myung Hwan Seo

2015 - 2016

- Conducted simulations for Lasso regression and bootstrapped confidence intervals.
- Performed structural break tests on GDP growth data across multiple countries.

Research Assistant for Prof. Simon Lee

2014-2016

- Conducted research on the partial identification of changes in the social preferences of North Korean refugees.

**Columbia University, Department of Economics**

New York, NY

Research Assistant for Prof. Macartan Humphreys and Prof. Guy Grossman

2012 - 2013

- Assisted in monitoring communication between Members of Congress and voters in Uganda.
- Analyzed petition records to evaluate whether legislators fulfilled their campaign commitments.

**TEACHING EXPERIENCE**

---

**Columbia University, Department of Economics**

New York, NY

Teaching assistant

2018 -

- Money and Banking (undergraduate), Spring 2025
- Introduction to Econometrics (undergraduate), Fall 2023
- Intermediate Macroeconomics (undergraduate), Spring 2023
- Corporate Finance (undergraduate), Fall 2020
- International Macroeconomics (undergraduate), Spring 2020
- Game Theory (undergraduate), Fall 2019
- Financial Economics (undergraduate), Fall 2018 and Spring 2019

**Seoul National University, Department of Economics**

Seoul, South Korea

Teaching assistant

2014 - 2015

- Advanced Econometrics (graduate), Fall 2015
- Introduction to Modern Economy (undergraduate), Fall 2014

**HONORS, SCHOLARSHIPS, AND FELLOWSHIPS**

---

- Dissertation Fellowship, Columbia University 2024
- Program For Economic Research (PER) Summer Research Fellowship, Columbia University 2020, 2023
- Woytinski Fellowship, Columbia University 2020  
Awarded to one student in the Economics department interested in the field of Statistics
- Dean's Fellowship, Columbia University 2017 - 2018
- Brain Korea 21 Plus Scholarship (Full scholarship), Seoul National University 2014 - 2016
- Undergraduate Scholarship (Full scholarship), Kwanjeong Educational Foundation 2009 - 2013

**MISC.**

---

- Programming skills: MATLAB, R, Stata, SQL, and Python
- Nationality: South Korea