The unifying theme of my research is *heterogeneity*, which is common in different types of data and has nontrivial impacts on estimation, statistical inference and modeling.

My projects start with concrete examples or evidence of heterogeneity. I then tractably incorporate the relevant mechanism into a model that imposes interpretable assumptions and can be disciplined by empirical evidence. My work involves an extensive use of optimization algorithms, so I am interested in writing code that exploit features of the objective functions for performance.

The rest of this research statement summarizes my job market paper and other papers that explore the implications of heterogeneity. Ongoing and planned projects are highlighted.

Bias Correction for Panel Models with Fixed Effects

Adding individual fixed effects is a popular way to control for time–invariant unobserved heterogeneity in panel data, but fixed effect estimators are biased in nonlinear and dynamic linear models due to the incidental parameter problem. In my job market paper **Indirect Inference for Nonlinear Panel Models with Fixed Effects**, I propose a new debiased estimator for a class of nonlinear models such as binary responses models.

The method is so called *indirect fixed effects estimator*. It combines data simulations and parameter calibration by equating fixed effect estimators obtained from observed and simulated data. I provide theoretical properties of this method and illustrate with an application to female labor force participation. I plan to build on this project and explore dynamic discrete choice models with fixed effects.

Nonstationary Variables: in many applications, the explanatory variables of interest such as age and policy treatments are nonstationary. In the *ongoing* project **Crossover Jackknife Bias Correction for Nonlinear Panel** (with Victor Chernozhukov, Iván Fernández-Val, Hiroyuki Kasahara and Paul Schrimpf), we propose a new jackknife bias correction method to handle this type of nonstationarity that arises in settings such as difference–in–differences designs.

Dynamic Linear Panel: The Arellano–Bond estimator is subject to bias due to many instruments when the panel is long. In **Mastering Panel Metrics: Causal Impact of Democracy on Growth** (with Victor Chernozhukov and Iván Fernández-Val, *AEA P&P*, 2019), we show that sample splitting alleviates the bias problem.

Robust Inference in Incomplete Models

Strong assumptions complete economic models and predict unique outcomes, but they mask heterogeneity in data. In **Robust Tests of Model Incompleteness in the Presence of Nuisance Parameters** (with Hiroaki Kaido), we focus on a particular class of discrete choice models in

which the model is complete under the null hypothesis and incomplete under the alternative, and develop a testing procedure for key parameters of interest despite being agnostic about the selection mechanism. We propose a hybrid procedure that treats our test statistic as a specification test, but a formal analysis is deferred to future work.

Productivity Dispersion and Innovation Heterogeneity

Whether the U.S. industry–level productivity dispersion is caused by resource misallocation or firms' rational behaviors is important for policymakers. In **R&D Heterogeneity and Countercyclical Productivity Dispersion** (with Yang Ming), we provide a new theory featuring firms' heterogeneous R&D responses to industry–level *negative profit shocks* (NPS).

Using public data, we construct an index of the shocks and find that both productivity and R&D intensity dispersions enlarge immediately after a shock occurs and gradually dissipate. We build a duopolistic technology–ladder model with heterogeneous R&D costs and NPS. A quantitative model predicts that the low–cost firm responds to a shock by increasing R&D effort while the high–cost firm does the opposite. We provide empirical support for this mechanism and plan to explore (1) the source of these shocks and (2) implications for entry–exit dynamics.

Econometrics Software

In nonlinear models the partial effects vary with control variables, but usually only average partial effects are reported. Chernozhukov et al. (2018) proposed the sorted effect methods to estimate heterogeneous partial effects and quantify the uncertainty of the estimation. In **SortedEffects: Sorted Causal Effects in R** (with Victor Chernozhukov, Iván Fernández-Val and Ye Luo, *The R Journal*, 2020), we introduce the R package *SortedEffects* that implements the methods therein. I plan to develop packages for all of my current and future methodology papers.