Statistical Modeling Discussion

Andreea L. Erciulescu, Westat

Federal Committee on Statistical Methodology (FCSM)
Research and Policy Conference 2022
October 26, 2022

Session Papers

- John Angle: Applying a Model of Income Distribution, Adopted as Econophysics, to Estimating Non-Compliance with the Individual Federal Income Tax
  - **mathematical model**
- Noé Nava: A Bayesian Model of the Probabilistic Role of Weather Variations on Crop Yield Potential
  - **statistical model; Bayes inference**
- Daniel Lin: Methods and Assumptions of the CPS ASEC Tax Model and Imputing 2020 Stimulus Payments
  - **black-box model**
- Anil Rupasingha: Place-Based Tax Incentives and Minority Employment: Evidence from New Market Tax Credit (NMTC) Program
  - **statistical model; matching**
- Cesar Montalvo: Skill Profiles and Portability of Credentials for the Technical Workforce
  - **statistical model; frequentist inference**

Modeling Income Distribution; Angle, J.

Contributions

- Mathematical model known as 'inequality process' (IP)
- Estimation of 'tax gap' (i.e., noncompliance with the federal income tax)

Discussion points

- The model is presented as 'simple' and accounting for 'a wide scope of patterns in statistics of income and wealth' - fifteen categories of empirical income and wealth phenomena
  - Needs: explicit model specification, along with assumptions; details on estimation approaches
- The model is validated using data from the Current Population Survey from multiple years - distribution of income conditional on education, for age group 25+ year old
  - Needs: other variables to condition on; finer age groups; other socio-demographic groups

Modeling Weather Variations on Crop Yield; Nava, N.

Contributions

- Bayesian formulation of the Ricardian approach that integrates agronomic evidence on the yield-weather interaction
- Crop yield projections from Midwestern and Eastern U.S. counties for 2022, 2027, and 2032

Discussion points

- The model is presented as hierarchical Bayes model
  - Needs: definition of the Ψ function in the data level; definition of the Ψ function in the link level; prior distribution for σ in the link level; hyper-priors for M and Σ in the prior for β; level for Ψ
- The model is applied to crop yield, using three weather variables
  - Needs: selection of weather variables; details on model fit and validation; assumptions on weather variables 5 and 10 years into the future

Modeling Tax and Imputing Stimulus Payments; Lin, D.

Contributions

- Tax model
- Complement to the Current Population Survey Annual Social and Economic Supplement (CPS ASEC), including the first and second Economic Impact Payments for the 2021 CPS ASEC

Discussion points

- The model assumptions, steps, inputs, and outputs are described in some detail
  - Needs: inclusion of error from statistical matching step; description of methods used
- The model is validated by comparing aggregated model estimates against IRS benchmarks
  - Needs: alignment of underlying population and reference time in the two sources

Modeling Tax Incentives and Minority Employment; Rupasingha, A.

Contributions

- Matching estimation approach
- Effectiveness of the New Market Tax Credit (NMTC) program on minority and rural populations

Discussion points

- The methods described are supported by the literature; three matching methods are compared
  - Needs: assumptions; inclusion of error from statistical matching
- The variables used for matching are enumerated, along with their reference time
  - Needs: possible confounding variables

Modeling Technical Workforce; Montalvo, C.

Contributions

- Regression - weighted least squares estimation, and network analysis
- Describe how technical skills and experience shape the salary returns for skill-intensive occupations and identify nondegree credentials that allow skilled technical workers to redefine career pathways in manufacturing

Discussion points

- The regression equation for salary return is defined at the individual level and includes three covariates
  - Needs: data sources; reference time
- The results are presented by major occupation group, and other groups
  - Needs: different model specification for each analysis; multicolinearity tests for models with multiple covariates

Thank you!

AndreeaErciulescu@westat.com