



# An Overview of Current Results and New Methods for Estimating Heterogeneous Program Impacts

Till von Wachter

University of California Los Angeles

# Lessons Learned from Analysis of Impact Heterogeneity in SSA's Demonstrations

1. **Clarify the purposes** of analysis of impact heterogeneity & briefly review results from non-experimental literature
2. **Take stock of findings** of heterogeneity in estimated program impacts for eight demonstrations
3. **Review recent approaches** to estimate and use heterogeneity in treatment impacts

# Demonstrations Covered

## 1. SSDI or Concurrent Beneficiaries

- a) Benefit Offset National Demonstration (BOND)
- b) Mental Health Treatment Study (MHTS)
- c) Project NetWork, Accelerated Benefits (AB), Benefit Offset Pilot Demonstration (BOPD)
- d) Promoting Opportunity Demonstration (POD)

## 2. SSI Beneficiaries

- a) Structured Training and Employment Transitional Services (STETS)
- b) Transitional Employment Training Demonstration (TEDT, U.S. Department of Labor)

➔ **Chosen since they had at least some impact estimates by subgroups**

# Use-Case for Estimates of Impact Heterogeneity

## Program Improvements

- Diagnostics: For which groups could program be improved?
- Targeting: Should program be targeted to certain groups?

## Program Evaluation

- Give certain populations larger weights in cost/benefit analysis

## Analysis of Mechanisms

- Understand mechanisms through which a program operates

# Non-Experimental Findings on Impact Heterogeneity

## Econometric Estimates of Employment Potential

- Employment potential of individuals currently receiving SSDI and SSI is low.
- Some differences in employment potential:
  - *More employment potential: Younger, suffering from less severe impairments, on program for fewer years (e.g., Hemmeter & Bailey 2016)*

## Characteristics of Program Participants

- SSDI & SSI programs serve broad populations: range of potential barriers & opportunities to engage in work. Some changes in populations over time.

# Takeaways from Demonstrations' Subgroup Impacts

1. Subgroup effects reflect main effects
2. In other cases, null subgroup impacts despite main effect
3. In some cases, null main effect masks subgroup impacts
4. **Some key recurring subgroups:**
  - Age (BOND, MHTS, TETD, or TTW)
  - Impairment (BOND, MHTS, Project NetWork, TETD, STETS)
5. **Analysis of single subgroups suggests:**
  - Race may need oversampling (STETS, TETD)
  - Availability of other transfers payments or Medicaid (STETS, BOPD)

# Some Practical Implications From Comparing Heterogeneous Impact Estimates of Demonstrations

- 1. Worth considering a common definition of subgroups**
  - *Consider core set of subgroups: gender, race, age, education, etc.*
- 2. Worth considering a common set of program outcomes**
  - *SSA-administrative data base outcomes available to all demonstrations*
- 3. Worth considering a standard for reporting statistical results**
  - Use an interacted model to estimate subgroup impacts
  - Test for subgroup impact differentials (e.g., men vs. women)
  - Test for differences between main and subgroup impacts

# Ex-Ante Approaches: Design Demonstration to Maximize Chances to Detect Impact Heterogeneity

## 1) Stratified randomization & possible oversampling

- *E.g., by predicted employment potential or predicted program take up*

➔ **Target high-impact groups & oversample low-impact groups**

## 2) Cross-classified randomization with additional treatment

- *E.g., manipulate the likelihood of program take up*

➔ **Induce both those with higher & lower valuations to enter the program**



# Ex-Post Approaches: Use Statistical Algorithms to Search for Relevant Subgroups

Algorithms chose subgroups with largest impact heterogeneity:

1. Semi-parametric approaches: extensions of standard approach but consider large range of subgroups
2. Non-parametric approaches: *directly search data for groups for which differences in treatment effects is largest*

**Obtain Conditional Average Treatment Effect (CATE):**

- Analyze distribution of CATE
- Reassess program impact with alternative targeting strategies

# Proposed Blueprint for New Demonstration

1. Decide on Stratified or Cross-Classified Design
  - *E.g., stratify by estimates of employment potential*
2. Implement Algorithms to estimate impact differentials (i.e., CATEs)
3. Analyze distribution of impact differentials (diagnostics)
4. Compare results to traditional approach
5. Use impact differentials to assess effects of targeting strategies
6. Reverse engineer targeting-tool for real-life application

# Practical & Ethical Considerations

## 1. Program Diagnostics

- Do differential treatment effects arise because program is less effective, or for other reasons (e.g., barriers to access)?
- Additional qualitative analysis likely to be important

## 2. Targeting

- For practical implementation, CATEs may be too complex.
  - *This can be addressed depending on the data & program*
- Sampling error may have to be taken into account
- Risk of perpetuating biases in program take up or effect

# Conclusion

## 1. Evaluation results indicate:

- that impacts vary
- some substantive lessons
- some methodological lessons

## 2. Design and implement next demonstrations accordingly

## 3. Explore new statistical approaches to analyze impact heterogeneity

# Heterogeneity: Subgroup Findings

Discussant: Howard Goldman, University of Maryland



# Heterogeneity: Subgroup Findings

Discussant: Nick Hart, The Data Foundation



# Lessons Learned from SSA Demonstrations: A State of the Science Meeting



## Panel C: Question & Answer

# Lessons Learned from SSA Demonstrations: A State of the Science Meeting



We are on a break.  
Content will resume shortly.