Seminar Intro

Jheng Shao Yu(R06323010)

1   What is the Question?

Experiment often involves heterogeneous participants, and when such heterogeneity is patterned, researchers need to address types of participants in their project. The traditional way of measurement is via finite mixture models, but such models do not properly suit the setting of experiments: small samples and repeated observations(panel data). Furthermore, finite mixture model require research to presume the number types, which is not data-driven partitions. The author thus ask: is there any proper method to cope with patterned heterogeneity in experiment?

2   Why Should We Care

Experiment gets common in social science,especially economics. For example, the author also address many topics of law and economics through experimental methods. But when it comes to systematic heterogeneity, there still lacks a satisfactory solutions. The aurhor seeks to propose one.

3   What's the authors' answer?

The author proposes a two-step algorithm that is less data-hungry, but this new method can (1) find the number of types in data and (2)estimating results for panel data.

4   How did the author get there?

The key move of the author is the introduction of machine learning. For each participant, this method firstly estimate local regression on time-changing variables. Then a regression tree is used to train a classification algorithm. When types is assigned from data, then interact those types with treatment.

5   Real World Example

For example, if we want to experimentally measure whether "single-sex high school" has any effect on "being single(no relationship) in undergraduate program," because we conjecture that students in single-sex school tend to have less dating experience. But when we recruit participants, it is possible that we find there are more than one type of students(some may still have many dating experiences in high school), then we can apply this method to address how many types in our data.

6   Notations

1. y_i: observed choice

2. τ:type

3. $t \in 0, 1:$: treatment

4. $CRAT$: Classification and Regression Tree