Matching Points: Supplementing Instruments with Covariates in Triangular Models

Junlong Feng 1

1Hong Kong University of Science and Technology

Motivation & Objective

A Common Problem in Practice: In economic applications, it is common to have a discrete endogeneous variable (U) and an instrument (Z) taking on fewer values:

- Return to education: D: Multiple levels of education, Z: Whether lived near a college or not (binary).
- Program evaluation: D: Multiple training programs, Z: A lottery granting access to a certain program (binary).

No Sufficient Variation in IV: Identification is A Common Problem in Practice

- In economic applications, it is common to have a discrete endogenous variable (U) and an instrument (Z) taking on fewer values:
- Return to education: D: Multiple levels of education, Z: Whether lived near a college or not (binary).
- Program evaluation: D: Multiple training programs, Z: A lottery granting access to a certain program (binary).

Sufficient Variation in IV: No Identification

- Selection heterogeneity can be vector-valued.
- In economic applications, it is common to have a discrete endogeneous variable (U) and an instrument (Z) taking on fewer values:
- Return to education: D: Multiple levels of education, Z: Whether lived near a college or not (binary).
- Program evaluation: D: Multiple training programs, Z: A lottery granting access to a certain program (binary).

Main Idea

- The selection heterogeneity can be vector-valued.
- In economic applications, it is common to have a discrete endogeneous variable (U) and an instrument (Z) taking on fewer values:
- Return to education: D: Multiple levels of education, Z: Whether lived near a college or not (binary).
- Program evaluation: D: Multiple training programs, Z: A lottery granting access to a certain program (binary).

Model

A triangular model with S(D) = {1, 2, 3} and Z(D) = {0, 1}:

- Outcome eq: Y = \sum_{d \in D} \mathbb{1}(D = d) \frac{E(Y|Z = x)}{E(Z|Z = x)} + \epsilon

Selection eq: D = \mathbb{1}(h(y, x, Z, U) = 1)

- The outcome heterogeneity \epsilon is a scalar.
- The selection heterogeneity can be vector valued.
- The selection heterogeneity can be vector valued.
- The selection heterogeneity can be vector valued.

Main Idea

- The outcome heterogeneity \epsilon is a scalar.
- The selection heterogeneity can be vector valued.
- The selection heterogeneity can be vector valued.
- The selection heterogeneity can be vector valued.

Finding the Matching Point

Two different approaches depending on how much we know about Model-SL:

If the selection model is known/specified: Find the matching point by solving \( h(d, x, z, w) = h(d, x, z, w) \).

Example. Ordered choice with linear cutoffs.

- \( h(x, Z, V) = \mathbb{1}(V < x_k + Z^T \beta_k) \) and \( h(x, Z, V) = \mathbb{1}(V > x_k + Z^T \beta_k) \) for all \( k \). Then matching points can be obtained by solving \( a_u = (a, x_k) \). If the selection model is unknown/unsupervised: Obtain the matching points by matching the generlized propensity scores under the following assumptions:

Identification of Model-SP

- The term in red offsets the change caused by conditioning on a different value of the covariates.
- The term in red offsets the change caused by conditioning on a different value of the covariates.
- The term in red offsets the change caused by conditioning on a different value of the covariates.
- The term in red offsets the change caused by conditioning on a different value of the covariates.

Identification of Model-NSP

- The term in red offsets the change caused by conditioning on a different value of the covariates.
- The term in red offsets the change caused by conditioning on a different value of the covariates.
- The term in red offsets the change caused by conditioning on a different value of the covariates.
- The term in red offsets the change caused by conditioning on a different value of the covariates.

Case 1: The difference between two solution paths is continuous. Impossible because they must enter the local-neighborhood region by continuity.

Case 2: The difference is discontinuous. Then the alternative solution path must jump up for some \( k \). But by monotonicity and continuity of \( \Phi \), there must exist a \( k \) for which the alternative solution path jumps down to make the equation \( h(x, z, w) \) hold, violating monotonicity.