Structural vs. Reduced Form Approaches to Empirical Research

Introduction

Internal Validity

The degree to which the research design yields estimates that reflect the true casual effect for the population, time, and/or location under investigation.

 Threats

- Functional form (linearity, interactions, etc.)
- Omitted variable bias, endogeneity, selection bias
- Measurement error in regressors

External Validity

The degree to which the research design yields estimates that reflect the true casual effect for other population, times, and/or locations of interest that lie outside of the population, times, and/or locations that were studied.

Consider the model for individual $i$ from group $p$, location $l$, and time $t$:

$$ y_{iplt} = \alpha + (\beta_0 + \beta_1 x_{iplt}) T_{iplt} + \omega_{iplt} \theta + \epsilon_{iplt} $$

where $x_{iplt}$ are characteristics that relate to the treatment effect (interactive confounders), and $\omega_{iplt}$ are characteristics that affect the outcome directly (additive confounders).
Threats

- Lack of internal validity
- Interaction of setting and treatment

Example. Consider an experiment in which $T_{iplt}$ is randomly assigned (RA), then our estimate will be internally valid. The estimate will reflect the average treatment effect for the distribution of $x$ that is involved in the experiment. Whether the estimates are externally valid depends on:

- the extent to which the population of interest is reflected in the study.
- the degree to which treatment effects are heterogeneous across $p, l, and t$.
- internal validity.

Structural Form vs Reduced Form

The battleground:

- The credibility of theoretical vs statistical assumptions.
  - Structural econometricians tend to make more theoretical assumptions, while reduced from econometricians tend to make more statistical assumptions.
- Achieving external validity.
- The types of questions that can and should be asked.

Structural Approach

- It was developed to bridge the gap between theory and empirics.
- Theory is often used as a starting point, which allows for interdependencies among outcomes.
- Specifies cause and relationships
  - Some variables are endogenous, which are determined within the model, while others are exogenous
- Equations are generated that correspond to individual choices and/or aggregates.
- The data is used to estimate parameters.
- Maximum likelihood estimation of non-linear simultaneous equations is quite common.
• The parameters present in the model are “fundamental” in that they are arguably not context dependent (e.g. risk aversion, preferences, technology).

• These primitives allow for generalizability, i.e., out of sample prediction.
  – We can simulate behavior responses to policies that haven’t happened.
  – We can conduct welfare calculations

Weaknesses

• Identification: there is usually weak support for the exogeneity assumptions and/or instruments.

• The theoretical assumptions may be overly restrictive, but this is hard to assess.

• Due to convergence at local extrema, estimates can be highly sensitive to starting values (Knittel & Meta, RESTAT).

Reduced Form Approach

• This method is also known as the “design-based” or the “quasi-experimental” approach.

• It doesn’t try to uncover the “true model” that is generating the data.

• Its primary focus is on research design.
  – The main objective is to obtain credible causal estimates of the parameters of interest. Note that these are not necessarily going to be “fundamental” parameters.

• Reduced form econometricians see actual experiments as the gold standard since they are intuitive and straight forward.

• The conceptual framework highlights the source of variation.
  – If experiments are not possible, then we must search for circumstances where it is reasonable to assume that unobservable variables are unrelated to the variable of interest. These circumstances may be driven by chance, forces of nature, or institutions.
  – Since the variation is clearly articulated, then we know what threats need to be examined, i.e., we can make a data driven case that we have exogeneity.
Weaknesses, Criticisms, Responses

- External validity: if there is substantial heterogeneity, then this is a very serious issue.

- The reduced form approach is divorced from theory.
  - It can’t speak for mechanisms, which are often examined empirically for plausibility.
  - Economic theory guides interpretation of the evidence and the questions that are asked.
  - Studies often try to test competing theoretical models which offer different predictions.
  - “In a design based framework, economic theory helps us understand the picture that emerges from a constellation of empirical findings, but does not help us paint the picture.”

- Major data requirements: if the data does not exist, then we cannot proceed.

- Difficult to say anything conclusive about welfare, efficiency, or behavioral responses to policies that haven’t occurred.
  - It does, however, tell us which policy levers we can pull to get the types of results that we want to see.
  - We still have to worry about unintended consequences with respect to outcomes that haven’t been thoroughly investigated.
  - Similar issues are arise for variables that are “outside the model” of structural work.

Clarification

Structural approaches involve modeling the system in which a collection of endogenous variables are simultaneously determined by one another, by exogenous variables, and of course, by agents maximizing utility and/or profits. The researcher makes several choices here. What to specify as endogenously determined vs exogenously determined? What functional forms to use? The internal validity of the estimates depend on these theoretical assumptions.

Reduced form approaches don’t involve such choices. They "bypass estimation of the underlying economic model" and instead focus on the "reduced form" effects of variables of interest that are arguably exogenous in the circumstances under consideration. The internal validity of the estimates depends on whether a variable of interest is truly uncorrelated with other variables that affect outcomes in this set of circumstances.

Example. Suppose we wish to examine the effect of class size on test scores. The two approaches provide different outlooks on how to proceed.

- **Reduced form:** With this approach, we must document the effect of class size reduction. The reduced form method starts with the parameter of interest and usually employs a linear function.

\[ y = \alpha + \beta \text{ class size} + \varepsilon. \]
● **Structural:** With this approach, we create a model with several equations and variables, and we specify the interactions between them. In this example, there is a relationship between class size and the following variables: teacher quality, teacher effort, parental effort, and student effort. The structural approach models all of these interactions, and the listed variables all affect test scores.