Dialogue on Causality

Between James Heckman and Ian Shrier
(Prepared for *Observational Studies*)

James J. Heckman

Econ 312, Spring 2022

THE UNIVERSITY OF
CHICAGO

General Perspective

THE UNIVERSITY OF
CHICAGO

Interviewer: What is your general approach to science and causality?

- The current literature on causality is filled with monologues of various participants hawking their wares, ignoring what others have to say.

- I have read the literature in statistics and computer science closely.

- It has had a big influence on many fields, including on my own, and often to their detriment.

THE UNIVERSITY OF
CHICAGO

- As an example, discussions about "preferred estimators" show its malign influence.
- Like much of the current "causal inference" literature, the emphasis on "preferred estimators" conflates estimation methods with conceptual definitions of causality.

THE UNIVERSITY OF
CHICAGO

- I am an economist trained (at the college level) in physics and mathematics.

- I switched to economics out of interest in social and economic issues and in the belief that economics can be a science and can contribute to understanding and resolving important policy debates.

- The specificity of the policy problems addressed and the generality of the analytical framework used to address them attracted me to the field.

THE UNIVERSITY OF
CHICAGO

- I seek non-tautological, rigorously justified models derived from theory and verified on rigorously justified data.

- Measurement and theory together are the hallmarks of good science.

- No serious scientist pretends to "let the data speak for themselves," nor would they impose models on data that wouldn't support them.

THE UNIVERSITY OF
CHICAGO

- I have long practiced abductive inference, building and testing models using all sources of data (quantitative and qualitative), which may include censuses, cross-section surveys, experimental data, observer reports, newspaper accounts, interviews, ethnographic studies, and controlled and natural experiments.

- No particular source of data is uniquely privileged, although more objective, carefully collected and documented data are preferred if they are available.

- The goal is always to obtain evidence as free of personal belief as possible.

THE UNIVERSITY OF
CHICAGO

- A rigorous scientific approach to any investigation requires admitting failure when it happens, revising models in the light of failures when they occur, extracting new implications of old models, and testing the revised models so derived on fresh samples if possible.

- The key to learning from data is being honest, determining and admitting multiple interpretations of the evidence if they appear.

- One should constantly play devil's advocate and challenge one's own work.

- The search for, care for, and consideration of alternative explanations is the key to rigorous empirical work.

THE UNIVERSITY OF
CHICAGO

- A central feature of hard science and rigorous economics is that it addresses well-posed problems and presents qualified answers.

- It is not about seeking the estimand of a "preferred" estimator.

- It is about having well-posed scientific questions and separating anecdote and bias from hard evidence.

- It is a public activity that invites scrutiny and challenge and not about any particular statistical procedure.

THE UNIVERSITY OF
CHICAGO

On Causality

Interviewer: What do you consider the core issues in causality?

- Before one can meaningfully talk about causality one must define counterfactuals, which lie at its core.

- Counterfactuals that describe possible outcomes under different conditions are an expression of human imagination and creativity.

- Different groups judge the quality of the counterfactuals created by different standards.

- Counterfactuals are *imagined* outcomes: *mental* constructs defined according to some set of rules, which may be implicit.

THE UNIVERSITY OF
CHICAGO

- There are no hard and fast rules for generating counterfactuals.
- Possibilities are only limited by the imaginations of their creators.
- Counterfactuals are products of the mind whose plausibility depends on how well they respect the rules invoked by their users.
- Scientific communities emerge to establish rules for creating counterfactuals from rigorous theory and to verify their construction.

THE UNIVERSITY OF
CHICAGO

- Science doesn't stop with possibilities.

- It is about the verification *in objective data* of counterfactual predictions.

- Verification is the process of checking the ingredients, both theory and evidence, and holding them up for public scrutiny, including replicability.

- Public scrutiny is not a popularity contest, although it is often interpreted this way by networks of like-minded individuals who ratify each other and ignore outsiders.

THE UNIVERSITY OF
CHICAGO

- While there may be many possible worlds, their plausibility depends on whether they rely on credible ingredients. Scientific research demands careful measurement and rigorous testing.

- It also asks that the counterfactuals generated are grounded in scientific principles previously established in the body of previous research.

- Science is a cumulative process that seeks consilience across studies.

THE UNIVERSITY OF
CHICAGO

- The acts of conceiving counterfactuals and their relationships—thought experiments—and the acts of estimating and testing the validity of these imagined relationships are fundamentally distinct.

- However, they are often confused, especially in fields without guiding abstract principles and cumulative knowledge.

- Purely statistical approaches often ignore the crucial point that science is an iterative process.

- Scientists build models, test them, adapt them to fresh data and/or on new data, and/or examine further implications of proposed new models to see if they hold up.

- It is by iteration and public argument that scholars learn from data and build models to explain them.

- This feature is often absent from many applications of "causal inference" in statistics.

THE UNIVERSITY OF CHICAGO

- There is no "correct" way to generate counterfactuals unless a set of generator rules is postulated.

- If counterfactuals claimed to be generated by such rules are in fact inconsistent with them, they are "incorrect" in terms of the announced rules.

- But there are no absolutes in this business—just agreement with prior knowledge up to the point of discovery. Arguments that counterfactuals are "nearest possible worlds" flounder on the lack of any metric (or topology) for "closeness" or enumeration of the set of possible worlds.

THE UNIVERSITY OF
CHICAGO

- When two (or more) counterfactuals are compared, a causal effect is obtained.
- Some counterfactuals may be rooted in fact, others in fancy.
- Thus, if we compare US history if the South had won the Civil War to US history if it had lost (as actually happened), we define a "causal parameter" (really a causal scenario) holding all else the same.
- Of course, the two counterfactuals differ greatly in terms of their anchor in fact.
- The first opens up many possibilities, none of which necessarily occurred, unless the two histories are identical.
- The second is a topic studied by historians.
- Our beliefs about the quality of causal effects are grounded in the quality of the counterfactuals underlying them.

THE UNIVERSITY OF
CHICAGO

Econometric Framework

- Economic theory is grounded in abstract principles.

- Economic models are models of possible outcomes.

- They are thought experiments.

THE UNIVERSITY OF
CHICAGO

- Consider a simple linear model:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + U. \tag{1}$$

- (X_1, X_2, U) are not necessarily random variables or generated by stochastic processes.

- These can also be constants when used in functional relationships; e.g., letting $(Y, X_1, X_2, U) = (y, x_1, x_2, u)$, we can write (1) as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u. \tag{2}$$

THE UNIVERSITY OF
CHICAGO

- Equation (1) is an example of a function $g : (X_1, X_2, U) \to Y$.
- Equation (2) is an example of g evaluated at the specified values.
- Economists study how hypothetical variation in the arguments (X_1, X_2, U) affect Y when the other arguments are fixed (held constant).
- This is a general functional relationship.
- One thought experiment imagines an increase in x_1, holding x_2, u at fixed values.
- Of course in the real world, one may not be able to randomly vary X_1 without also varying X_2 and U.
- That is an entirely different matter.

THE UNIVERSITY OF
CHICAGO

- In Equation (1), β_1 is the effect of a unit increase in X_1, holding X_2 and U fixed.

- Each Y obtained by varying X_1 holding ($X_2 = x_2, U = u$) is a possible ("potential") outcome over the support of (1).

- Such relationships are at the core of economic theory and have been for over 130 years, and have been at the core of mathematics proper for centuries.

- Marshall (1890) referred to β_1 as the *ceteris paribus* effect of X_1 on Y.

- *Ceteris paribus* means "everything else the same."

- Precisely the same types of qualifications are given when examining the pull of gravity on a feather.

THE UNIVERSITY OF
CHICAGO

- Thought experiments often illuminate real world policy debates.

- Thus, if Y is the consumption of cigarettes and X_1 is the price of cigarettes, the *ceteris paribus* variation in X_1 is informative about the possible effect of a tax (which increases the price) on smoking.

- Hard science and rigorous economics are grounded in thought experiments like these.

- A variety of potential outcomes can be obtained by varying $X_1, X_2,$ and U in different ways.

THE UNIVERSITY OF
CHICAGO

- I have deliberately used a linear model as my example of an abstract theoretical model.

- I distinguish it from a familiar linear regression model, as is usually taught in statistics.

- That model starts with a collection of random variables (Y, X_1, X_2, U).

- Up to now, I have not had to specify whether Y, X_1, X_2, U are proper random variables or not.

THE UNIVERSITY OF
CHICAGO

- Under normality of (X_1, X_2, U) (alternatively of (Y, X_1, X_2)) and taking conditional expectations of (1),

$$E(Y|X_1 = x_1, X_2 = x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + E(U|X = x_1, X_2 = x_2).$$
(3)

- If $E(U|X = x_1, X = x_2) = 0$, one obtains

$$E(Y|X_1 = x_1, X_2 = x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$
(4)

THE UNIVERSITY OF CHICAGO

- Evaluating Equation (1) at the values $X_1 = x_1, X_2 = x_2, U = u$ gives Equation (2), which is very similar to but conceptually distinct from Equation (4).

- Setting U in (2) at the expected value of $U(u = 0)$, one obtains

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \tag{5}$$

which is exactly Equation (4).

THE UNIVERSITY OF
CHICAGO

- The fact that the right-hand sides of Equations (4) and (5) are mathematically identical obscures a crucial difference: Equations (1), (2), and (5) are the result of *thought experiments*.

- The analyst *conceptually sets* the values of X_1, X_2 and U.

- No data are in sight, although the thought experiment may have been motivated by empirical studies.

- No properties of random variables need to be specified.

THE UNIVERSITY OF
CHICAGO

- In contrast, Equations (3) and (4) are generated by another type of thought experiment: by taking conditional expectations used to describe data.

- They are ingredients frequently used in statistical methods.

- Equation (5) is a causal relationship that determines the value of outcome y.

- The values of x_1, x_2, u are *hypothetically* assigned to (X_1, X_2, U).

- Equation (4) is defined by a statistical operation that might be used to estimate parameters from observed data over the support of the data.

- Equation (5) is defined without invoking any statistics.

THE UNIVERSITY OF
CHICAGO

- Many statisticians work in fields in which there is no formal methodology for describing "setting" (or fixing) inputs like (x_1, x_2, u).
- The Kolmogorov (1956) axioms do not need to be invoked in defining (5), although the laws describing mathematical functions are required.
- Abstract models like (1), (2), and (5) are outside of the Kolmogorov framework.
- They are nonstochastic.

- This observation helps explain the rise of interest in Pearl's (2009) do-calculus in some quarters of statistics and social science.

- His "do operator" works by setting inputs in a formal structure like (1) and creating special rules outside of formal statistics to interpret operations like "fixing" or "setting."

- However, his calculus is not needed to accomplish this task.

THE UNIVERSITY OF
CHICAGO

- Elsewhere, Heckman and Pinto (2015) formally extend conventional probability theory to include "fixing" or "setting" in the Kolmogorov system by introducing a new class of random variables.

- They develop an intuitive framework that enables analysts to express causal operations using only standard probability and statistical theory without the elaborate non-statistical rules used in the do-calculus.

THE UNIVERSITY OF
CHICAGO

- Representations like Equation (1) arise from thought experiments.

- Building on our cigarette example, raising a price (X_1) likely reduces cigarette demand.

- Holding everything else constant, the law of demand suggests that $\beta_1 < 0$, i.e., as $X_1 \uparrow$, $Y \downarrow$ holding X_2, U fixed.

- One can *set* $X_1 = x_1$ and vary it to create different counterfactual worlds.

THE UNIVERSITY OF
CHICAGO

- It may occur that in actual data, Equation (4) coincides with, or is very "close" to, Equation (5), but that is a separate issue.

- In the cigarette example, as in many other studies of demand, the correspondence between theory and evidence is quite close.

- In the language of econometrics, this would correspond to Equation (5) (or (1)) being identified.

- The basic distinction between a thought experiment generating different values of y by varying x_1, keeping x_2 (and u) constant, and conditional expectations ((3) and (4)) is at the heart of what economists bring to the causality table.

- This point was first formalized by Haavelmo (1943) and more recently advanced in Heckman and Pinto (2015).

THE UNIVERSITY OF
CHICAGO

- The mathematics is trivial.
- The conceptual distinction between (1) (or (5)) and (4) is not trivial and is the source of enormous confusion in statistics and in quarters of economics that follow and implement statistical frameworks that ignore thought experiments.

THE UNIVERSITY OF
CHICAGO

- Causal model (1) is defined independently of any estimator.
- In Equation (1), β_1 may or may not be an estimand of a regression of Y on X_1 and X_2.
- If the estimand from Equation (4) based on observed data coincides with Equation (5), we are in company with Wigner (1960) in admiring the unreasonable effectiveness of theory (1) in predicting reality.
- But theory and inference are logically different.
- See Table 1.

THE UNIVERSITY OF
CHICAGO

Table 1: Two Distinct Tasks that Arise in the Analysis of Causal Models

Task	Description	Requirements
1	Defining the set of hypotheticals or counterfactuals	A well-specified theory
2	Identifying and estimating causal parameters from data	Mathematical analysis of point or set identification in identification in hypothetical samples along with estimation and testing theory

THE UNIVERSITY OF
CHICAGO

- Appreciating the distinction between functions (like (1) or (5)) and estimands (like (3) and (4)) is key to understanding the contribution of economics to the study of causality.
- A function g (like (1) or (5)) maps $x \to y$.
- Formally,

$$g : x \to y.$$

- Functions are *defined* to be stable maps between X and Y over their entire support, regardless of the variation of its inputs.
- No special language or additional conditions such as those in "SUTVA" are required.

THE UNIVERSITY OF
CHICAGO

- Moreover, systems of interdependent (non-recursive) equations are readily formulated. Mathematics and economic theory are replete with systems of equations:

$$F : Z \to W, \qquad (6)$$

where Z is a collection of variables, W is a set of outcomes of any dimension, and F is a vector of functions that may include g from our previous example.

- State space equations are widely used in engineering, economics, and chemistry to list only a few applications.

- Simultaneity and interactions are readily characterized by such systems of functions.

- Many of the approaches current in "causal" analyses in statistics ignore the benefits of abstract theoretical models and thought experiments.

- They limit the range of causal questions that can be investigated by their practitioners.

THE UNIVERSITY OF
CHICAGO

- This simple example readily generalizes to considering a definition of causality.
- Abstract theoretical equations need not be linear.
- Normality is not an essential feature of them.
- Equation (1) can be replaced by a system of simultaneous equations such as (6).
- "SUTVA" is irrelevant.
- Causal effects are by *definition* manipulations of input variables (by a thought experiment) that generates hypothetical outputs.
- They are defined independent of any estimation technique or availability of data.
- Simultaneous causality is readily defined and poses no necessary paradox or contradiction to logic.

THE UNIVERSITY OF
CHICAGO

- Human knowledge is produced by constructing counterfactuals and theories.

- Blind empiricism unguided by theoretical frameworks for interpreting facts leads nowhere.

- Many statisticians are uncomfortable with counterfactuals.

- Their discomfort arises in part from the need to specify abstract models to interpret and identify counterfactuals.

- Many statisticians are not trained in science or social science and adopt as their credo that they "should stick to the facts."

THE UNIVERSITY OF
CHICAGO

- Economists since the time of Haavelmo (1943, 1944) have recognized the need for precise models to construct counterfactuals and to answer causal questions and more general policy evaluation questions, including making out-of-sample forecasts.

- The econometric framework is explicit about how counterfactuals are generated and how interventions are conducted (i.e., the rules of assigning "treatment").

- The sources of unobservables, in both treatment assignment equations and outcome equations, and the relationship between the unobservables and observables are studied.

THE UNIVERSITY OF
CHICAGO

- Rather than leaving the rules governing selection of treatment implicit, the econometric approach explicitly models the relationship between the unobservables in outcome equations and the choice of outcome equations to identify causal models from data and to clarify the nature of identifying assumptions.

- The theory of structural modeling in econometrics is based on these principles.

- Modeling choice also enables analysts to distinguish objective counterfactuals (e.g., did the drug work?) and subjective counterfactuals (e.g., what is the pain and suffering experienced by users of the drug?).

- Answers to both questions are valuable in evaluating the impact of any treatment.

THE UNIVERSITY OF
CHICAGO

- Ambiguity in model specification implies ambiguity in the definition of counterfactuals and hence in the notion of causality.

- The more complete the model of counterfactuals, the more precise the definition of causality.

- The ambiguity and controversy surrounding discussions of causal models are consequences of analysts wanting something for nothing: a definition of causality without a clearly articulated model of the phenomenon being described (i.e., a model of counterfactuals).

- They want to describe a phenomenon as being modeled "causally" without producing a clear hypothetical model of how the phenomena being described are generated or what mechanisms select the counterfactuals that are observed in hypothetical or real samples.

THE UNIVERSITY OF
CHICAGO

- In the words of Holland (1986), they want to model the "effects of causes" without modeling the causes of effects.

- Science is all about constructing models of the causes of effects.

- Such models are essential in analyzing policy problems, as in our cigarette example.

- Economic (and scientific) problems dictate the choice of an abstract model and the definitions of causal parameters—not some estimand from one or another procedure.

- In summary, causality is a property of a model of hypotheticals.

- A fully articulated model of the phenomena being studied precisely defines hypothetical or counterfactual states.

- A definition of causality drops out of a fully articulated model as an automatic by-product.

- A model is a set of possible counterfactual worlds constructed under some rules.

- The rules may be the laws of physics, the consequences of utility maximization, or the rules governing social interactions, to take only three of many possible examples.

- A model is in the mind.

- As a consequence, causality is in the mind.

THE UNIVERSITY OF
CHICAGO

Policy Evaluation

- In empirical studies of causality, there are some standard problems:

  1. selection bias; and
  2. that for any causal question, at most one of many counterfactuals is known empirically (i.e., what is actually observed).

- These problems have been formalized at least since the time of Cox (1958), if not before.

- There are three different policy evaluation problems that are fruitfully distinguished but often conflated:

    P1 *Evaluating the causal impacts of actual interventions, including their impact in terms of welfare.*

- This is the problem of identifying a given treatment effect or a set of treatment effects in a given environment.

- This is the policy question usually addressed in the epidemiological and statistical literatures on causality.

- A drug trial for a particular patient population is the prototypical problem in that literature where investigators typically focus on objective outcomes, e.g., the effect of a drug treatment on health rather than the subjective wellbeing of the patient.

THE UNIVERSITY OF
CHICAGO

- However interesting that may be, most policy evaluation is designed with an eye toward the future and toward decisions about new policies and application of old policies to new environments.

- I distinguish a second task of policy analysis:

  P2 *Forecasting the impacts (constructing counterfactual states) of interventions implemented in one environment in other environments, including their impacts in terms of welfare (subjective wellbeing).*

- Included in these interventions are policies described by generic characteristics (e.g., tax or benefit rates, etc.) that are applied to different groups of people or in different time periods from those studied in previous implementations of policies.

- This is the problem of *external validity*: transporting a structural parameter or a set of parameters estimated in one environment to another environment.

- The "environment" includes the characteristics of individuals and their social and economic setting.

- This is the forecasting problem long studied in economics.

THE UNIVERSITY OF
CHICAGO

- Finally, the most ambitious problem is forecasting the effect of a new policy, never previously experienced:

    P3 *Forecasting the impacts of interventions (constructing counterfactual states associated with interventions) never historically experienced to current or different environments, including their impacts in terms of welfare.*

THE UNIVERSITY OF
CHICAGO

- This problem requires that one uses past experience to forecast the consequences of new policies.
- It is a fundamental problem in knowledge.
- It requires that one use abstract models to connect the future to ingredients from the past.
- This is the focus of structural estimation.
- P3 has been consistently ignored by statisticians because they cannot or will not deal with abstract models of counterfactuals like Equation (1).

THE UNIVERSITY OF
CHICAGO

- Holland (1986) boasts that statistics deals with the "effects of causes"—in the terminology of this paper, outcomes of comparisons of counterfactuals—rather than the causes of the effects (models like (1) that determine the counterfactuals studied).

- A prime example of the limits to this approach is the claim that one can never learn the causal effect of race on outcomes because race cannot be randomly assigned.

- An estimand (the outcome of an RCT) is used to define causal parameters.

THE UNIVERSITY OF
CHICAGO

- In contrast, the abstract theory-based approach investigates which factors (X in Equation (1)) are important for producing outcomes.

- Having determined this, one can conceptually equalize the factors across persons of different race groups.

- That may not be an easy empirical task, but there is a large scholarly literature examining the factors leading to various outcomes (e.g., productivity in a job).

- Whatever the quality of the empirical work, the thought experiment remains valid even if a reliable estimate of the causal effect is difficult to obtain.

THE UNIVERSITY OF
CHICAGO

On Dominance of Specific Models

Interviewer: What are the advantages and disadvantages of the recent interest in randomized trials within economics?

- Experiments are useful for certain problems but are far from being a panacea and are frequently corrupted in practice, producing misleading inference.

- A good example is the analysis of a recent Head Start experiment (Kline and Walters, 2016).

THE UNIVERSITY OF
CHICAGO

- Families randomized out of the program at one location went to other Head Start locations or even better programs.
- The inconclusive treatment effect reported from the simple mean difference RCT "causal effect" was a result of uncontrolled "contamination bias."
- The "control group" participated in programs as good or better than the program being evaluated, so the "gold standard" estimated treatment effects were negative, when in fact estimates that correct for contamination bias show strong positive effects.
- Often, clean RCTs do not exist, and we always have to judge the causal claims of each study on a case by case basis.

THE UNIVERSITY OF
CHICAGO

Interviewer: What are your thoughts on the current dominance of propensity scores, diffs-in-diffs, regression discontinuity, and synthetic controls?

- There are many estimation methods out there, each justified by different assumptions.

- There is no universal best estimator.

- The conditional nature of (causal) knowledge alarms some analysts who seek absolute knowledge of "causal effects."

- One needs to understand the science behind a phenomena before one can select appropriate estimators.

THE UNIVERSITY OF
CHICAGO

- However, this is not often done in many statistical studies.
- An "effect" is estimated without a clear interpretation of what precisely it captures and how it is generated.
- Heckman and Robb (1985, 1986) compare the identifying assumptions underlying a large array of cross-section, repeated cross-section, and panel data estimators, including "diffs-in-diffs."

THE UNIVERSITY OF
CHICAGO

- Many users of the methods you describe do not have a clear problem in mind and instead chant "ATE, TOT," or whatever sells that day.

- It's like a fashion show in Milan with different inferential techniques coming down the runway at different times.

- Most users have no clear question in hand, just an "estimand" that will get them published.

- The estimands are well defined; the problems they address often much less so.

- Whether the estimands are relevant for solving the stated problems is an entirely different issue that is rarely addressed.

THE UNIVERSITY OF
CHICAGO

- Many users do not have a question in hand.
- They just want an estimate of a not so clearly defined something.
- A good example of this phenomenon is the literature in economics on estimating the "return" to schooling.
- It focuses on estimating the coefficient of schooling in an equation relating log earnings (y) to factors determining earnings.
- A close reading of the literature reveals its focus on one element of β in Equation (4) free of any bias arising from correlations between U and the associated X.

THE UNIVERSITY OF
CHICAGO

- The recent "credibility" revolution in economics focuses on "credible" estimators for β.

- It emphasizes easily computed, easily replicated statistical methods.

- These are surely desirable features of any inferential procedure.

- Missing, however, is any clear statement of why βs so obtained answer relevant economic questions.

- Heckman et al. (2006) demonstrate that the "return" featured in the "credibility revolution" answers interesting economic questions only under very special circumstances and that a deeper analysis is required to estimate the economically interesting return to schooling.

- The credibility revolutionaries create clean estimators for generally uninterpretable estimands.

- They conflate tasks 1 and 2 of Table 1.

THE UNIVERSITY OF CHICAGO

Interviewer: How do you see the trade-offs between more parametric approaches (SEM) versus less parametric approaches such as DAGs?

- I don't agree with the premise of the question.
- Structural models (SEM) estimating abstract theoretical models can be and often are nonparametric or at least semiparametric.

- Heckman and Pinto (2015) and Pinto and Heckman (2022) show that Pearl's DAGs and the rules of do-calculus do not accommodate instrumental variable methods or selection bias models.

- They account for simultaneity only by "shutting down equations" without regard for the properties of systems so generated, while structural econometric models readily accommodate simultaneity, as Matzkin establishes in her many papers.

On Treatment Effects

Interviewer: How would you describe the differences between average causal effects and your "marginal treatment effect" (Heckman and Vytlacil, 1999)?

- The choice of a parameter should be dictated by the problem one seeks to address, not by some arbitrary convention.

- The marginal treatment effect (MTE) is useful for determining benefits or costs to people at margins of choice.

- MTE is a building block from which all other conventional treatment effects can be constructed under appropriate support conditions.

- It does not replace treatment effects—it is a device for unifying them.

- It also links choice equations to outcome equations.

- Heckman and Vytlacil (2005) show that a variety of policy relevant causal parameters outside the basic toolkit can be formed from marginal treatment effects.

THE UNIVERSITY OF
CHICAGO

- This approach has its origins in the insights of calculus: the relationship between derivatives and integrals.

- The marginal treatment effect can be interpreted in some settings as the marginal willingness to pay for a good for a subset of people who are indifferent between buying it or not.

- The common treatment effects aggregate over all people who buy the good with their different intensity of preferences.

THE UNIVERSITY OF
CHICAGO

Interviewer: What are the relative advantages and disadvantages of local treatment effects?

- Local treatment effects are the building blocks from which all treatment effects can be constructed.

- They can be used to establish the relationships among the various treatment effects using a common conceptual tool.

- It unifies knowledge.

- Local treatment effects characterize choices at margins which are central to economic analysis.

THE UNIVERSITY OF
CHICAGO

On Mediation

Interviewer: Do you find mediation/path-specific inference a useful or promising area of causal research and application?

- Yes, it is.
- But note that "mediation" analysis has long been conducted in econometrics and is very useful for many questions.
- Adelman and Adelman (1959) simulated the dynamic multiperiod Klein-Goldberger (1955) model of the US economy using what is now called "mediation."
- They based their analysis on "dynamic impact multipliers," which chart the impact of policy changes in one period on the output, consumption, and investment of future periods.

THE UNIVERSITY OF
CHICAGO

- Long ago, Sewall Wright (1934) pioneered the mediation approach, calling it path analysis, and there is a long tradition of using it in social science.

- Mediation analysis enables analysts to understand causes of effects rather than stopping at just reporting effects.

- It is essential for policy analysis that aims to improve outcomes.

- It tells analysts which levers to pull to make effective policies, and it is valuable for the intellectually curious who seek to know why treatment outcomes occur.

THE UNIVERSITY OF
CHICAGO

On Bounds and Sensitivity Analyses

THE UNIVERSITY OF
CHICAGO

Interviewer: For informing policy decisions, do you find bounding or sensitivity analyses useful?

- Yes. But there is usually much more information available than the numerical information on the supports of random variables that is featured in recent work on bounds.

- This additional information is useful in doing sensitivity analysis (e.g., newspaper accounts, common sense, etc.), but it is rarely used.

- It often requires too much subject matter knowledge for most "causal analysts."

THE UNIVERSITY OF
CHICAGO

On Limiting Assumptions

Interviewer: How do you approach the common assumptions in causal inference that appear limiting given real data?

- Many standard methods for analyzing social interactions and general equilibrium effects, simultaneity, and feedback violate the "causal framework" protocols of statistics but answer interesting policy questions outside that straightjacket.

- The SUTVA assumption is a good example.

THE UNIVERSITY OF
CHICAGO

- One aspect of "SUTVA" is "structural invariance" or "autonomy" developed by Frisch (1938) and Hurwicz (1962).

- It characterizes functional relationships like Equations (1) or (5).

- It is a core idea applied more than eighty years ago in econometrics.

- The other aspect of "SUTVA" is the assumption of no interaction among agents.

- That rules out a lot of the study of interesting social phenomena.

THE UNIVERSITY OF
CHICAGO

- Simultaneity and interaction have been studied for decades.

- For instance, much research in economics focuses on general equilibrium environments, and a sizeable literature in economics is devoted to the study of peer-effects.

- The goal of this literature is precisely the evaluation of the "interference" (really interaction) among randomization units, which is ruled out by SUTVA.

- Interactions are treated as a problem in statistics but are, in fact, a source of information in economics and social science more broadly.

- SUTVA is an artifact of the obsession of many statisticians and their followers on single-agent RCTs.

THE UNIVERSITY OF
CHICAGO

- I believe statisticians' ignorance of the econometrics literature with respect to simultaneity and social interactions has harmed the advance of knowledge.

- Information about the mechanisms producing "interference" is useful for analyzing the propagation of disease, economic shocks, network externalities, etc.

THE UNIVERSITY OF
CHICAGO

On Machine Learning

Interviewer: Do you think machine learning will be useful in causality?

- Machine learning is a useful tool but it's only a computational device for prediction, estimation, and establishing empirical relationships.

- It offers no insight about causality per se, other than what is learned from careful descriptions of phenomena.

- It is useful to have good descriptions of phenomena as material on which to build interpretative causal models.

THE UNIVERSITY OF
CHICAGO