**Leamer, 1983.**

1. *Explain the meaning of S and M in equation (4) at page 32.*

The M matrix represents the variance of the bias terms due to omitted variables, measurement error, or simultaneity. It captures the degree to which the model is misspecified. The S matrix represents the sample covariance matrix of the OLS estimates.

*2) When is M zero?*

The M matrix is zero only for the perfectly conducted experiment.

*3) What does Leamer say about the whimsical character of inference, how sensibly to base inferences on opinions when facts are unavailable?*

Leamer asserts that “objectivity” is a false idol. Researchers should embrace the whimsical character of inference by conducting sensitivity analyses to see how the results depend on their assumptions.

*4) What does Leamer mean by “The mapping is the message”?*

Researchers should limit themselves to reporting the mapping between assumptions and results, though they may still choose which assumptions are worth including in their report.

*5) What does Leamer mean when he says that experimental inference admits a conventional horizon in a critical dimension, namely the choice of explanatory variables?*

In a properly conducted experiment, it is reasonable to assume that the misspecification matrix M is approximately zero. Thus, one may include only outcome of interest and the sources of experimental variation.

*6) What is the purpose of Table 2?*

Table 2 suggests the degree of researcher freedom in specifying her model in an observational setting.

*7) What is the main takeaway of this article?*

Econometric evidence is presently unconvincing, so practitioners should perform sensitivity analyses. While they may sometimes rely on convention in determining the space of relevant assumptions, the work of testing that boundary is never complete.

**Angrist and Pischke, 2010.**

*1) What do A&P see as the main difference between their analysis and Leamer’s analysis?*

Angrist and Pischke emphasize finding a credible identification strategy instead of performing sensitivity analyses.

*2) Why do they want to discuss natural and quasi-experiments?*

The use of natural and quasi-experiments has grown dramatically and increased the credibility of empirical work.

*3) What does A&P mean when they say that the difference between a randomized trial and an observational study is one of degree?*

Under the right conditions, observational studies can approximate an RCT.

*4) How do they summarize “Lucas critique”?*

With forward-looking and optimizing agents, past policy changes are uninformative about the effect of new interventions.

*5) What is the point of Figure 1?*

US and Canadian homicide rates moved in parallel between 1950 and 2001 despite substantial policy variation. Thus, policy is unlikely to have substantial effects on homicide rates.

*6) A key passage is when A&P discuss the specific threats to validity that justify the use of randomized trials, or ivs, diff-in-diff, and regression discontinuity. Discuss.*

Convincing research designs, whether they employ an RCT, IV, diff-in-diff, or RD, must be defended with institutional details and empirical evidence. Such arguments often hinge on the possibility of omitted variable bias. Is the treatment variation as good as random? Can agents manipulate their treatment status? Is reverse causality possible? These technical questions often crowd out any discussion of the big picture. Even an airtight research design may obtain trivial results.

*7) When discussing research in industrial disorganization, A&P say that structural estimates should line up with those derived under weaker assumptions. Discuss.*

For A&P, one virtue of structural models is that they may clarify the economic mechanisms behind an outcome. However, illustrating the mechanism is only secondary to the primary goal of empirical work: making accurate predictions. There is a risk that an elaborate theoretical framework may obscure the assumptions that are necessary for identification of the key parameters.

*8) What do A&P say about the existence of a trade-off between internal vs external validity of empirical work? And what solution do they propose to increase the external validity of natural or quasi-experiments?*

A&P observe that all empirical evidence is specific to a particular context. Increasing the external validity of a result requires “accumulating empirical evidence” over a broad range of times and places.

*9) What is the main takeaway of this article?*

Econometrics became more credible by using quasi-experiments to estimate causal effects transparently. Macroeconomics and industrial organization should take note.

**Nevo and Whinston, 2010.**

*1) What do Nevo and Whinston mean by structural model?*

A model based in economic theory that describes how individual consumers and producers behave.

*2) What are two advantages of working with structural models, and how are they related to the trade-off between internal vs. external validity of empirical work?*

N&W emphasize two advantages of structural models: out-of-sample prediction and welfare analysis. Because reduced-form models typically estimate a linear relationship, they are likely to extrapolate poorly.

*3) What are the limitations that N&W ascribe to the analysis of Hastings? And how are they related to the trade-off between internal vs. external validity?*

The Hastings analysis is inherently limited because a “merger” is not a standard treatment. Industries can vary along many dimensions: which firms are merging? Are they vertically integrated? How large are the efficiencies? How many competitors are there? Is the industry in decline? The consensus among IO economists is that external validity is unlikely to hold for any analysis of a single merger.

*4) In what sense is extrapolation simpler using an economic model?*

Extrapolation from an economic model has several advantages: it does not require a natural experiment, it generates a welfare effect, and it can be performed under a variety of assumptions.

*5) What type of problems does the “optimal legal review” introduce in the analysis?*

The observed set of mergers suffers from a selection problem: all of them have been approved by the government. The average effect of a proposed merger and that of an approved merger are unlikely to be equal.

*6) What is the main takeaway of this article?*

The differences between the methods employed in industrial organization and those employed in labor are due to (1) culture and training, (2) the ability to extrapolate from a treatment of interest, and (3) data availability.

**Sims, 2010.**

1. *What does Sims say about the use by Donohue and Wolfers of over-indentifying restrictions and why are they useful?*

Donohue and Wolfers test over-identifying restrictions on models estimated with instrumental variables. They find that the results are sensitive to the set of instruments that are included.

1. *What are the ingredients of the economic model that Sims would like to see Donohue and Wolfers develop?*

Sims wants to capture all of the variation in homicides with a dynamic model that captures the interactions between policy, economic conditions, and crime.

1. *What concerns does Sims have about the work by Romer and Romer, and how does that concern affect other cases where government agencies take decisions? To what extent are those decisions exogenous to the model?*

The dates of contractionary monetary policy identified by Romer and Romer are not random: they can be predicted using past data. Thus, their estimates suffer from omitted variable bias. Did the policy change cause a particular effect, or was the policy change simply a response to some omitted variable that was ultimately responsible?

Because government agencies behave optimally, their actions are endogenous.

1. *How does Sims think that a school principal can extrapolate (again, external validation) the results from the study of Angrist and Lavy? Is the range of possible policy actions narrow or wide, and why?*

A school principal knows that students tend to learn better in smaller classes. However, she does not know about the trade-off between teacher quality and class size; the effect of a change in the distribution of class sizes, and whether the effect is the same for very large or very small class sizes. Thus, the range of policy actions is narrow.

1. *What is the main takeaway of this article?*

By estimating linear CEFs and using robust standard errors, economists obscure interesting variation and ignore the possibility of non-linear effects. These features are crucial for making policy recommendations.