Literature Review

Yuwei Zheng, Weitong Yao

9th June

1 Introduction

To understand better how research methodology has evolved, a series of published papers are studied. This series paper can be seen as a debate between 5 groups of economists or econometrists; they are Angrist and Pischke, Keane, Sims, Nevo and Whinston, and Rust. Angrist and Pischke echoed Leamer who firstly questioned the credibility of traditional econometrics methods and argued that a good experimental research design is the best solution, however, lately, other scholars more defied than supported Angrist and Pischke's arguments. Controversies were focused on two questions: 1) What are the causes of the poor researches and 2) which improvement is better? This debate can be named as: are improvements of research credibility in Microeconomics, Macroeconomics, Industry Organization, and Labour fields credited to good research design or the use of structural model? This literature Review will illustrate scholars' view in each field with some examples they gave.

2 Microeconomics

Angrist and Pischke firstly analysed Leamer's critiques and provided a review of history on studying research credibility. They believed that research methods have been improved since the Leamer's critique and pushing for a better and more clearly designed research is the primary force for improving empirical works, which heavily influenced the microeconomics.

Two pieces of research were mentioned as a negative example. Can capital punishment deter crime? is a topic that US Supreme Court cared a lot and was primarily wrote by Ehrlich. Ehrlich sought to address problems of reverse causality and omitted variable bias by using IV in 2-stage OLS procedure, a quasi-experimental method, but he didn't explain why whose IV affected the endogenous variables of interest or why they constitute a "good experiment." This issue was recognised as "Lack of credible research design" by Angrist and Pischke. Sims in his paper also criticised Ehrlich's work in a way similar to what Angrist-Pischke mentioned, however he pointed out one more problem that Angrist-Pischke did not mention was over-identification problem that Ehrlich used too many IVs without clear commendations. Sims argued that the research can be improved if dynamic interactions among right-hand-side variables were considered into the model; the overly used IV method is partial because that they are easier comparing with modelling the dynamic interaction among variables that we don't even have a clue to start such reduced form estimation.

The second negative example is about school size. Coleman et al. (1966) studied the effect of school inputs, such as relations between class size and student achievement. However, their study failed to

separate interested variables and other confounding variations and suffers the problem of reverse causality and omitted variable bias. Looking into some good works, for example, works done by Card and Krueger (1992), Angrist and Pischke found that models with Random Trial, IV, DID, and Regression Discontinuity methods performed better. This fact supported Angrist-Pischke's argument that quasi-experiments are better than observational empirical works. However, again, Sims criticized that experimental researches are external invalid, looking at Tennessee's STAR experiment, because of the limits of modifiers.

3 Macroeconomics

When coming to macroeconomics, the random trail doesn't seem to be a good method. Lucas expressed that it is impossible and immoral to conduct a random trail study on macroeconomics, thus after that the contemporary studies have abandoned traditional empirical work entirely, economists turned to a new method called calibration which is achieved through computational experiment by changing model parameters within a well-established theoretical framework named as the dynamic stochastic general equilibrium framework.

Angrist and Pischke believed that the literature on empirical growth has long suffered from the lack of experimental design since Leamer's era, but in Romer's era, beginning from 2004, this situation has been improved in Macroeconomics field especially in studying fiscal and monetary policies. He highly affirmed the work conducted by Romer and Romer and Richardson et al. who conducted quasi-experimental researches. Angrist and Pischke also mentioned that the DSGE model structured by many theories have been countlessly studied, however, it produces no direct evidence on the magnitude or existence of causal effects. "An effort to put reasonable numbers on theoretical relations is harmless and may even be helpful. But it's still a theory".

Angrist and Pischke argued that the failure in Macroeconomic was due to the lack of experimental design which is nevertheless different from Sims' opinion that Romer and Romer's model identification was weak and mistaking: multivariate time series model including structural vector autoregression is the estimation method should be used after acknowledging complex interactions among variables when conducting monetary policy researches. Thus, Sims prefers DSGE models to experiments.

Rust also expressed his idea. He pointed out one challenge confronted by scholars is the exploration problem. When dealing with out-of-sample issues and conduct exploration forecasts, structural models are more reliable. Because "A structure is defined as something which remains fixed when we undertake a policy change, and the structure is identified if we can estimate it from the given data". When dealing with some policies that haven't been applied before, a good experimental design seems rootless, but researches started with economic theories help to calculate utility and welfare which portray the human nature and help to conduct counterfactual studies. Thus, Sims and Rust both favour structural models more than experiments in Macroeconomics field.

4 Industrial Organisation

Angrist and Pischke, Keane and Nevo have mentioned methodology improvements in industrial organisation and specifically focused on merger and price issue. Angrist and Pischke believe that studies done by Ashenfelter and Nevo didn't fully explain the causal effect. He applied findings of Ashenfelter in 2009 that only 20 studies evaluated the price effects of consummated mergers directly. Causal studies have been largely negligent in merger analysis. He then criticised the structural models Nevo used, which includes identifying the demand equation. Angrist and Pischke claimed that the simulation step typically focuses on a single channel by which mergers affect prices through the reduction in the number of competitors and in this framework, it's hard to see precisely which features of the data drive the ultimate results. However, the DID method Hasting used to estimate gasoline market was affirmed. Angrist and Pischke claim that a good research design that demonstrates economic mechanism, as well as causal effects, is more compelling than the conclusion derived from a theoretical approach. However, he explicitly said at the end of the topic that more model comparisons have to be done before drawing a clear conclusion, indicating that he is not very confident about his position.

Keane disagreed with Angrist's idea that progress in IO has been hindered by reliance on the structural model. He argued that the problem rooted in the IO field is not using the structural model but using static models which exaggerated price elasticity very much. The bad static models lead to the misspecification in the demand side is a consensus; therefore, a structural model is innocent and shouldn't be blamed, and we need to focus on improving the quality of data because for example, data set has exogenous variation helped to find long-run evidence in the television market.
Nevo demonstrated different opinions and questioned the external validity of Hasting's DID methods. He believes that DID methods was good but cannot be used for estimating other types of mergers, especially mergers trigger by endogenous factors. However, even structural models are heavy assumed, "elaborate", and "non-transparent" they are better in terms of forecasting extrapolation such as novel policies.

5 Labour

Nevo, to a large extent, agreed with Keane's statement that good data can improve empirical work and claimed that Structural models appear more in IO but less in labour because of the data availability. "Industrial organization economists seem far more concerned than labour economists that environmental changes are heterogeneous so that useful estimates of average treatment effects in similar situations are not likely to be available." Furthermore, data available to labor economists may be more likely than that in industrial organization to contain many examples of similar changes, as well as a richer set of directly observable controls, for example, a policymaker approaching a labor question probably has no more information than does an outside researcher." For those reasons, it is acceptable that experiments are more practical than structural models. Nevo's second half of the statement kindly supported Angrist and Pischke's argument, but Keane strongly disagreed it. He acutely criticised Angrist's claim that labour economics has developed wide consensus as estimations converge to the same results by rebelling every single example Angrist used to support his argument. Keane then stated that marketing is a field where broad consensus has been reached and external

validation has been accredited due largely to the availability of consumer panel data, however structural models are required for estimating utility functions of customers. Here Keane explained why structural models are necessary for welfare estimation, which is important in Labour economy.

6 Conclusion

Five groups of scholars had expressed their positions clear. Angrist and Pischke favour reduced form estimations and experiments based empirical works because of transparency and clear indicated casual inference. They focused on demonstrating the advantages of using Random trial and Quasi-experiments such as DID, IV, and regression discontinuity. The rest scholars somehow disagreed with Angrist and Pischke. Keane, Sims, Nevo and Whinston, and Rust believe that structural models are innocent and not the cause of poor estimations. They defied Angrist and Pischke from Macroeconomics, Industrial organisation, and labour economics fields by elucidating that structural models are good in leaning extrapolations, counterfactual problems, and dynamic decision makings; moreover, high-quality data improves casual inference when using structural models.
I am convinced by Angrist and Pischke that good design can improve empirical work especially in microeconomics field but also convinced by Keane, Sims, Nevo and Whinston, and Rust that theories, the prior knowledge, are important for any questions we study. Lastly, I would like to cite Rust's argument to conclude:

"The main advantage of structural models— that they can be used to simulate counterfactuals rapidly and cheaply—complements the main weakness of the experimental approach, which is that, whether done in the field or in the laboratory, experiments are much more time and resource intensive. At the same time, the main advantage of experiments is that they come much closer to predicting what the true impact of a policy intervention really is, and this complements the main weakness of structural models, which is that they can be wrong and produce incorrect policy forecasts. But by working together, experiments can help structural econometricians develop better models and discard inappropriate assumptions, while at the same time, the structural models can help experimentalists design more well-focused and productive experiments."

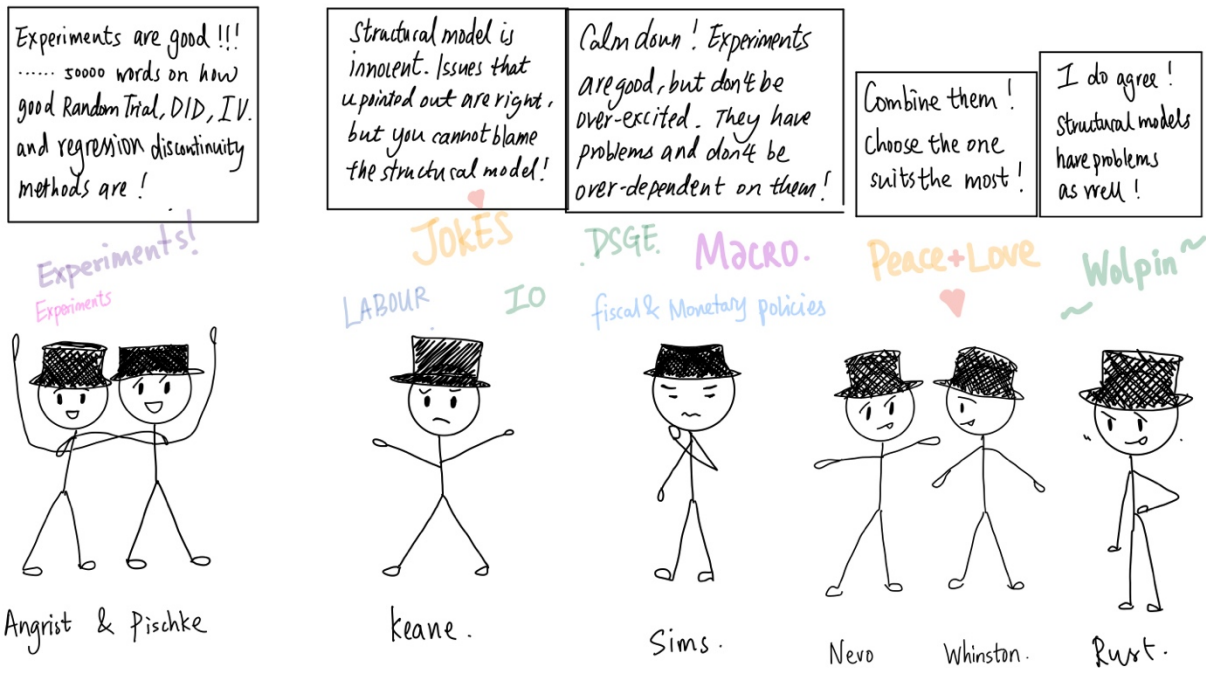

Figure 1: by yuwei