Assignment 9

Fulin Guo

1. This is my referee report for the paper “The Impact of Machine Learning on Economics” (Athey, 2018):

   The author adequately defines the question. As the title of the paper “The Impact of Machine Learning on Economics” (Athey, 2018, p1) indicates, the research question of the paper is how machine learning contributes to economic research. Since some impacts have already happened (Athey, 2018), the research question can be more clearly summarized as how machine learning has influenced economic research and how it will have other impacts on economics in the future. The author defines the question in the title, and she also says in the first paragraph that “the impact of ML on economics is already well underway” (Athey, 2018, p1), so it is clear that the author’s objective is to summarize and analyze the influences that have happened as well as predict other impacts in the future. We could see that the whole paper focuses on this objective. Theoretically, we could regard this research question just like other research questions that have the type “How A influences B?”, or “What is the impact of A on B?” (like how education level influences people’s salary?). Given the characteristics or features of economic research and machine learning, we could regard that “the impact of machine learning on economics” (Athey, 2018, p1) is a deterministic law that needs to be discovered by researchers(like given the characteristics of education system and job market, how education level influences people’s salary is a certain law (the influences are random, but the law could be viewed as certain) that needs to be found and concluded using econometrical models and data by researchers). Therefore, the objective of the author of this paper is to find the law of the “impacts of machine learning on economics” (Athey, 2018, p1).

   The author answered the research question adequately and compelling. In the second part of paper, the author defines machine learning and showed how machine learning can be utilized in economics easily (like the estimation of the consumer demand model (Athey, 2018)). Then, she summarizes how researchers have utilized machine learning in prediction problems (Athey, 2018) in policy and how economists can utilize this feature. Therefore, one influence of ML on economics is that the feature of the strong prediction power of machine learning can make economists do predictive work more efficiently. Then, the author summarizes new issues that have been put forward due to the contribution of ML on policy problems, like fairness, stability (Athey, 2018) That is, one impact of ML on economics is that it will result in more research in these new debates. Then, the author focuses on the connection between machine learning and causal inference, which is a topic that economists are interested in. Causal inference is different from prediction. In prediction tasks, the only indicator of testifying the validity of a model is whether this model generates satisfied prediction

outcomes, while causal inference focuses on whether one variable has partial effect on another variable. Sometimes the econometrical models for causal inference can have very low goodness of fit, but as long as the coefficients of partial effect are unbiased, the model is valid. Given the big difference between prediction and causal inference, we need method to make machine learning be more suitable for causal inference. Therefore, one influence of ML on economic research is the combination of ML and causal inference. The author summarizes what research has been done in this field and what issues regarding causal inference has been solved. In conclusion, the author answers the question in the following aspects: (1) How the supervised learning and unsupervised learning (the two classification of machine learning) can benefit economics (2) How the strongest power of machine learning (the prediction ability) can contribute to economic research (3) What are some issues or debates that would arise due to these impacts (4) the key problem between machine learning and economics (the difference between casual inference and prediction).

Although I think we can regard the research question of the paper as a normal question with the type "what is the impact of A on B?", the method used in this research is not very similar to a typical type of research that we usually see. The usual types of method in a research paper are theoretical deduction, and/or data analysis (empirical data or experiment data), like we can construct mathematical models to describe how price influences consumers' behavior and then utilize real data to testify the models. However, in this paper, constructing mathematical models or doing data analysis are not possible for this research question, but the author still needs to combine theory and practice (just like combining mathematical models and data analysis in the consumer behavior research). By theory I mean that predicting "the impact of machine learning on economics" (Athey, 2018, p1) without data (without reading related papers), given that I am familiar with the feature of economics and machine learning. By practice I mean we need to testify that some of our prediction is correct by reading related papers (The influences that have happened). In this paper, the author combines theory and practice efficiently. First, she defines what machine learning is and analyzes the characteristics of both ML and economics (like ML can be classified as supervised learning and unsupervised learning, economists focus on causal inference, etc. (Athey, 2018)). The theory of how ML can contribute to economics should possibly come from the characteristics of ML and economics research, like unsupervised learning could be used to the problem of consumer demand (Athey, 2018) because unsupervised learning could find many variables that are used in this economic analysis. Besides the deduction of the research question, the author also reads a lot of papers that related to machine learning or economics and analyzes how the researchers utilized ML on economic research or how they conducted methodology research that can influence the efficiency of using machine learning in economics. For example, in the third part, the author describes how ML has been utilized in policy problems. In conclusion, the author appropriately and sufficiently combines the theory (the definition and/or features of machine learning

and economics) and practice (other researchers' papers related to ML and economics).

For the research question itself, there are not many papers discussing how machine learning has influenced economics or how it will influence economics in the future, so the author might not be able to refer other papers that have the similar research question. However, this paper does need to refer a lot of papers. The papers the author cites could be classified as the following types: (1) the papers defining machine learning or about some common applications of machine learning (2) the papers of economic research using tradition method (i.e. econometrical models) (3) the papers about the methodology of traditional econometrical models, including some key problems that are usually discussed, like the identification problem, the prove of assumption behind a model, etc. (4) the papers that utilize machine learning to analyze economic problem (5) the papers about methodology that aims to improve the efficiency of combining machine learning with economic research. Based on my analysis in the last paragraph, the (1), (2), and (3) are papers that can tell us theoretically how machine learning contributes to economics. From (1), we could know what features of machine learning and economics would lead to the contribution. From (2), we could know some key disadvantages or problems in traditional methods, which lets us be aware the demand for utilizing ML in economic research. From (3), we could be aware some key issues in empirical research in economics, so we would know what issues we need to pay much attention to when combining machine learning and economics. Then, the (4) and (5) are practice, (4) is the applications that have happened. Since the application cannot be directly used in some aspects (like causal inference), there must be some papers related to the improvement of the efficiency of these application. Therefore, the author does very well in citation. She sufficiently considers many aspects of the research question and does well in citation in each aspect. In the last paragraph of my referee report, I will describe one extension for this research, so the author could cite papers in that aspect.

In the introduction part, the author writes: "A first theme is that ML does not add much to questions about identification, which concern when the object of interest" (Athey, 2018, p1). I think "much to" is a typo; changing the "much to" to "too much" would be better.

One extension of the paper is that the author could more specifically describe how machine learning has influenced and will influence the paradigm of economic research. Although the author said in the paper that ML causes more data-driven paradigm, it is valuable to discuss it in more detail. In the past, economic research is highly theory-driven. That is, researchers construct theories first, and in the second step, they use data to test the theory. This is a very popular way for a long time. For example, in the famous neoclassical economic models, many theories came from very complex mathematical deduction based on simple assumptions about humans. Many times, these assumptions might not be realistic, which might decrease the validity and

efficiency of those models. Now, we are in the era of big data, and there are efficient tools (like machine learning) for us to realize the data-driven paradigm. I think the author could analyze in the paper how machine learning impacts the paradigm in economic research. In particular, (1) How machine learning changes the popularity of mathematical deduction in economic research? (2) Is the neoclassical economic models (and other famous models based on mathematical deduction and assumptions of humans) still important? (3) How machine learning can testify the assumptions in neoclassical economic models? (4) How machine learning increases the efficiency of economic research?

Reference:

Athey, Susan, "The Impact of Machine Learning on Economics," in Joshua Gans Ajay K. Agrawal and Avi Goldfarb, eds., The Economics of Artificial Intelligence: An Agenda, National Bureau of Economic Research https://www.nber.org/chapters/c14009.pdf 2018, forthcoming.