CUHK SOCIOLOGY 中大社會系

NEWSLETTER

Three Things I Learned While Doing Fieldwork For My Dissertation

YONAHA Yvan Ysmael Tobias

Image from https://blog.prototypr.io

Doing any kind of research is hard work. Whether quantitative or qualitative, each kind presents unique challenges for the researcher to overcome. Below, I share some of the things I learned while doing qualitative fieldwork for my dissertation.

1. *Plan for delays in your research timeline*

Social scientists often enter the field with good intentions and well-made plans for implementing their research designs. Driven by the desire to advance knowledge of a particular field or advocate for a specific sector, we might think getting respondents for our research is fairly quick and easy. But this will not always be the case. My experience coordinating with my chosen case resulted in failure (read: catastrophe) because they immediately rejected my request for assistance in finding interview respondents. Luckily, they suggested another group that was willing and made the process easier for me. But it still caused some delays before I could start doing fieldwork. Plan these incidents into your timeline by scheduling some wiggle room in your timeline.

2. *Be emphatic but also cultivate emotional resilience*

Doing qualitative work entails relating deeply with people as they give you an account of their experience, thoughts, and feelings on your topic. This could be emotionally taxing, especially for those dealing with sensitive topics. My research involves conversing with nurses who cared for the sick during the COVID-19 pandemic and delivery riders who braved epidemiological risks to support their families. Sometimes, I hear stories that are heavy on the heart because of the oppressive practices rained upon these essential workers. As researchers, it is thus imperative to cultivate mental resilience not only to remain analytically sharp but also so we do not get bogged down by negative emotions when the going gets tough.

3. *Seek out surprises*

Timmermans and Tavory (2022) argued for the pivotal role of surprises in qualitative research. While they made good methodological points about how surprises can push theoretical boundaries forward, one thing that stayed with me is their existential reason for seeking it. They said:

> *"...life is finite, research takes a lot of time. Are you going to spend the good years of your life saying something most people already know, or are you aiming for things that inspire others – not to mention yourself – to look at the world differently? (Timmermans and Tavory 2022:15)".*

While doing my fieldwork now, I am constantly looking for new things to reframe or revisit how we look at nurses and delivery workers. While this work is difficult because of the number of new materials produced every year, I think it is worthwhile.

Work Cited: Timmermans, Stefan and Iddo Tavory. 2022. *Data Analysis in Qualitative Research: Theorizing with Abductive Analysis*. Chicago, Illinois, USA: University of Chicago Press.

Computational Social Science and Causal Inference

CAO Ji

Image from https://doi.org/10.1002/widm.1449

The efforts that scholars dedicated to causal inference can be divided into two levels: causal identification and effect estimation. It goes without saying that the identification strategy is the harder part of the job. The computational techniques that are inspired by machine learning or deep learning, no matter how promising they are, can only help with the effect estimation, since causal identification (design) is at another level. Recognizing this limitation could manage our expectations and use the relevant techniques more effectively.

One important application is the estimation of the conditional average treatment effect (CATE), which is the average treatment effect for certain subgroups divided by certain characteristics. The thing that CATE varies across the value of covariates rather than always equal to ATE, is the technical definition of treatment effect heterogeneity. It has been of substantial interest for a long time. The conventional econometric approach to identify CATE is through the estimation of interaction terms in the regression model, or through the regression decomposition. The choice of the interaction term is theory-guided. But in some cases, scholars are motivated to seek assistance from machine learning.

First, sometimes scholars are highly suspicious about the functional form of interaction terms--the concern of model misspecification. In this case, scholars adopt nonparametric estimation based on machine learning methods, in order to find a more precise CATE. Therefore, even though existing theories informed appropriate moderators (e.g., family socioeconomic status for the return to higher education), we have to address these problems with new techniques. Second, a different motivation comes from intersectionality. That is to say, sometimes the treatment effect heterogeneity is not unidimensional, thus, one simple interaction term might not help to catch the major CATE. However, adding a large number of complex interaction terms to the regression model causes cumbersome interpretation, and lowers the precision of estimation. The last motivation is that sometimes scholars don't really know what moderates the treatment effect.

Several recent machine-learning based methods have been developed to address all three problems, and there are some applications since 2017. For example, Liu (2021) applied two kinds of tree-based methods (Athey and Imbens 2016; Green and Kern 2012) and the ensemble method (Kunzel et al. 2019) to estimate the heterogeneous effect of father absence on children's educational outcomes, across a variety of family characteristics. Davis and Heller (2020) applied the methods proposed by Athey and Imbens (2016), Wage and Athey (2018) to study the treatment effect of youth summer job employment on their criminal behaviors, conditioning on individual characteristics.

Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, *113*(27), 7353–7360.

Davis, Jonathan M. V., and Sara B. Heller. 2020. "Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs." *Review of Economics & Statistics* 102(4):664–77.

Green, D. P., and H. L. Kern. 2012. "Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees." *Public Opinion Quarterly* 76 (3):491–511. doi:10.1093/poq/nfs036.

Kunzel, Soren R., Jasjeet S. Sekhon, Peter J. Bickel, and Bin Yu. 2019. "Metalearners for Estimating Heterogeneous Treatment Effects Using Machine Learning." *Proceedings of the National Academy of Sciences of the United States of America* 116 (10):4156–4165. doi:10.1073/pnas. 1804597116.

Liu, R. (2020). Leveraging machine learning methods to estimate heterogeneous effects: father absence in China as an example. *Chinese Sociological Review*, 1-29.

ANNOUNCEMENTS

✓ **Job Ads**

- Assistant Professor, Department of Sociology and Social Policy, Lingnan University. **Deadline for Applications: March. 29, 2023** [See the ad]

- Research Assistant Professor, Populations and Wellbeing, Faculty od Social Sciences, Hong Kong Baptist University. **Deadline for Applications: March. 30, 2023** [See the ad]