Proposed Research Medha Bulumulla

Poverty and malnutrition remain a critical issue in the Global South. According to the World Bank Group for the first time in 20 years, poverty is expected to increase due to the COVID-19 pandemic (World Bank, 2020). As a result, the ability to measure poverty and malnutrition and predict changes is even more important. Poverty and malnutrition rate estimation often requires household surveys of their respective locations. The most common surveys available for the global south are Living Standard and Measurement Surveys (LSMS) and Demographic and Health Surveys (DHS). These surveys are often expensive and time-consuming. They require survey managers, and administrators on-site. Right now, in-person surveys are not even possible due to the travel restrictions and social distancing guidelines in place. Given potential delays in the roll-out of COVID-19 vaccines due to cost and accessibility as well as countries struggling to control the virus, it will be a long time before in-person surveys can resume. So, alternative mechanisms for estimates and projections of poverty are more critical than ever.

The project, "Innovations in Feed the Future Monitoring and Evaluation: Harnessing Big Data and Machine Learning to Feed the Future", has developed a model to identify markers for poverty and malnutrition rates (Barrett et al, 2021). It uses DHS Malnutrition and Asset Poverty Data, applies statistical techniques like the Random Forest Model, and has identified markers to predict future malnutrition and poverty. These markers are available via open source as well as materials that can be accessed timely or in real-time to minimize costs. The model is used based on 11 United States Aid for Development (USAID) Feed the Future (FTF) priority countries (Bangladesh, Ethiopia, Ghana, Guatemala, Honduras, Kenya, Mali, Nepal, Nigeria, Senegal, and Uganda). While the focus is on high-need countries, it can be applied to many countries in the Global South to predict poverty and malnutrition.

My contribution to this project would include several tasks. I would document poverty and malnutrition estimation models (originally created using R and/or python software) to make them available for public use. The documentation is geared towards novice users so they can modify and replicate it. Then, I will post the documentation on a public site. Specifically, I will work with International Food Policy Research Institute (IFPRI), the primary partner for the project, to post the documentation and other data products on the IFPRI website. Finally, I hope to work with IFPRI to develop training modules to explain the data, the estimation routine, and how analysts at government, multilateral agencies, and Non-Governmental Organizations (NGOs) can replicate the model. Specifically, it will explain how they can apply to data available to them and potential applications.

Making this documentation available and accessible will be beneficial to countries with high poverty and malnutrition rates, especially FTF countries, in several ways. The ability to predict before a crisis strikes can help prevent or minimize the impact. Relief programs can be targeted and prevent those in need from falling into poverty traps. Given this model would allow regional(as opposed to national) poverty and malnutrition rate estimates and projections in the Global South, programs can be better targeted to those in need. Then the impact of the relief programs can be measured. Long term it may contribute to reducing poverty and malnutrition rates in the global south.

There are several areas where I hope to expand my skills of analysis, investigative study, problem-solving, and data management through this project. I hope to get a deep understanding of the statistical models used, specifically the Multivariate Random Forest method and the ability to replicate them. In doing so, I hope to expand my R programming knowledge. This will enable me to get an

understanding of the literature and analysis of poverty and malnutrition. I will have skills in documenting data sets. Additionally, I will be able to expand my webmaster experiences in posting data sets and posting documentation. Finally, given the interdisciplinary nature of this project that cuts across economics, statistics, international development, meteorology, health, nutrition, geography, statistics, and agriculture, this will allow me to broaden my knowledge base in these areas as well as how they intersect.

References:

Poverty Overview. (n.d.). Retrieved January 17, 2021, from https://www.worldbank.org/en/topic/poverty/overview#1

Barrett, C., Browne, C., Matteson, D., McBride, L., Hu, L., Liu, Y., . . . Wen, J. (2021, January 3). *Multivariate Random Forest Prediction of Poverty and Malnutrition Indicators*∗.