Understanding the Geography of Women's Labor Market Outcomes

Joshua Merfeld
KDI School of Public Policy and Management, Republic of Korea & IZA
merfeld@kdis.ac.kr

David Newhouse
World Bank & IZA, Germany

Michael Weber
World Bank & IZA, Germany

Partha Lahiri
University of Maryland, United States

Better understanding the geography of women's labor market outcomes within countries is important to inform targeted efforts to increase women's economic empowerment. This paper investigates the extent to which combining simulated survey data with geospatial indicators from Google Earth Engine and OpenStreetMap improves state and municipal estimates of labor force participation (LFP) and unemployment rates in urban Mexico. Model-based estimates are generated separately for men and women, using a weighted nested-error linear model following an arcsin transformation, specified at the level of the Área Geoestadística Básica (AGEB). The model is estimated using a simulated simple random sample of individuals within sampled AGEBs, selected using proportional-to-size sampling. The resulting state and municipal estimates are compared against benchmarks from the full census. When generating state-level estimates, incorporating geospatial information substantially improves the accuracy of male and female LFP and unemployment rates, reducing mean absolute deviation by 50 to 62 percent for LFP and 25 to 52 percent for unemployment. Small area estimation also greatly improves municipal estimates of LFP, as mean absolute error falls by approximately one half and mean squared error falls by almost 75 percent when holding coverage rates constant. In contrast, the results for municipal unemployment rate estimates are not reliable, because unemployment rates are low and therefore poorly suited for linear models. The municipal results hold in repeated simulations of alternative samples. Models utilizing AGEB-level auxiliary information generate more accurate predictions than area-level models specified using the same auxiliary data. Overall, integrating survey data and publicly available geospatial indicators is feasible, and in this context greatly improves state-level estimates of male and female LFP and unemployment rates, as well as municipal estimates of male and female LFP.