

2020

## 36<sup>th</sup> IARIW General Conference

Paper Prepared for the 36<sup>th</sup> IARIW General Conference, Oslo, Norway, August 24-28, 2020

The Corrected Gini Coefficient: Combining Administrative and Survey Data

Minki Hong

The Gini coefficient is calculated from the sample household survey. Sample household surveys generally have limitations that do not capture high income, so inequality indicators calculated from household surveys may not reflect real inequality well. The measure of inequality is heavily influenced by high income (Cowell and Flachaire, 2007), which has greatly affected the overall income inequality trend since the 1980s, with the share of the top income increasing around the U.S. and U.S. regions (Burkhauser, et. al, 2012). When the top income share increases, it is more likely that the income inequality indicators measured from the survey data will differ from the actual inequality.

This study aims to correct the estimates of the Gini coefficient by combining household survey data with national tax statistics that capture high income. There are two challenges to be solved in the study. First, there should be a reference distribution to compare how much of the high income the sample survey is underestimate. This study combines national tax statistics with household surveys to form the distribution of personal income. Second, although the national tax statistics show personal income, it is necessary to know the number of households and household income in order to calculate the household income Gini coefficient. Although the personal income is observed, household and other household sources income is unobserved, can be regarded as missing data. This study adopts the Bayesian data augmentation method to make complete data by imputation of missing data, and to estimate the Gini coefficient repeatedly.

Compared with the Gini coefficient from survey data, the corrected Gini coefficient estimates was 10~18% higher. Since sample surveys are underestimating high income, the effects of the correction have been great as the top income shares have increased recently. The Gini coefficient from survey data has been declining since 2010, but the corrected Gini coefficient is rather increasing. Although income inequality seems to decrease after 2010, adjusted Gini coefficient shows that income inequality is increasing in recent years.

Individual income inequality is very high in Korea compared to OECD countries, but household income inequality is low. The main reason is that the income correlation between household members of Korea is lower than that of OECD countries.

