Using Administrative Data to Impute Income Non-Response in Household Surveys

**Author:**V. Kerry Smith, Michael P. Welsh, Richard Carson, Stanley Presser

Income is simultaneously one of the most important variables used by economists and the variable most likely to be missing due to item non-response. While observations that are missing income responses are often dropped from analyses, such treatment is usually inappropriate. More appropriate solutions rely on imputation based on either covariates (e.g., age and education) measured in the survey or on spatial estimates (most often for zip codes) from the American Community Survey. We describe a new spatially-based alternative using publicly available Internal Revenue Service tax data that allows estimates of zip code’s income distribution.

**Url:**<https://www.nber.org/papers/w30420>