

Third, while a similar analysis applies to the residuals in a linear regression, and while it is still true that standard formulas are likely to underestimate the standard errors, the formulas are not exactly the same, and I postpone discussion of the regression case until Section 2.1 below. This is a major issue that has been much neglected, not only in development economics, but in other applied fields using survey data.

1.1.9. General measurement issues

I do not believe that there is any reason to suppose that survey data are always and automatically of lower quality in LDCs, as if "backwardness" were a condition that applied equally to GDP and its measurement. While statistical services are sometimes poorly funded and staffed, especially in Africa, survey data are often relatively cheap to collect in poor countries, and responses are likely to be accurate where there is a high degree of literacy, and where the respondents have time to talk to the enumerators. There are also some very poor countries (such as India) where survey practice is (or at least was) second to none. Indeed, Indian statisticians have played a leading role in the development of sample surveys and of sampling techniques; the surveys of jute production in Bengal by the Indian Statistical Institute under the direction of Mahalanobis were among the first successful large-scale sample surveys [see Mahalanobis (1944, 1946)]. It is also true that respondents tend to be much more patient in LDCs, that they rarely refuse to participate in the survey, and that they will usually tolerate instruments that take several hours to administer. The differences in quality of survey data between poor and rich countries comes, not from survey administration, but from differences in the structure of income and employment. In particular, difficulties in estimating income arise, not because of respondent unwillingness or because of fear that enumerators will pass information to the fiscal authorities, but because a large fraction of poor people in LDCs are self-employed, mostly in agriculture. Self-employment incomes are notoriously difficult to estimate in developed economies, and if income estimates in general are of lower quality in LDCs, it is because self-employment income is a larger fraction of the total.

The problems are easily seen. Self-employed traders or farmers typically have no need of any concept that corresponds to economists' definitions of income. Direct questions about income or profitability cannot therefore generate useful answers, especially for individuals whose personal and business transactions are not clearly separated. Instead, it is necessary for surveys to ask detailed questions about business or agricultural operations, about sales and purchases, about quantities and prices, about taxes and transfers, about multiple business activities, about transactions in kind, and about assets. **From** this detailed information, an income measure has to be built up by imposing an