

Estimation of income groups in the LITS2

Users of the 2010 Life in Transition Survey (LITS2) will find in the dataset released to the public a variable that partitions the population in each country into three income groups. This variable was constructed based on the raw data of the LITS2 and the purpose of this note is to explain how it was estimated.

Assessments on living standards of the population routinely use the consumption aggregate of the household as an economic measure of welfare. Although consumption does not cover all aspects of the welfare of a person, it does capture a central component of any evaluation of living standards and arguably it is a preferred welfare measure than income.

The LITS2 was not designed to collect accurate information about the consumption aggregate, but it included a few questions that allow the estimation of a crude measure of consumption. The fact that the analysis does not use that variable per se but a partition of the population into three equal groups according to the consumption distribution should lessen concerns about the precision of this variable. These three groups based on a simple measure of consumption will be used as a proxy for three income groups: poor, middle and rich. The implicit assumption is that the partition of the population in each country into three equal groups based on this simple nominal consumption aggregate correlates well with a similar (unobserved) partition of the population based on their true welfare levels.

Consumption aggregates are typically calculated based on extensive household surveys that devote entire modules to capture its different components. The LITS2 has different objectives than surveys whose main goal is to measure living standards of the population but includes a few questions to collect information about consumption. In particular, data on eight different components were gathered: food, beverages and tobacco; utilities (electricity, water, gas, heating, fixed line phone); transportation (public transportation, fuel for car); rent (actual or imputed); education (including tuition, books, kindergarten expenses); health (including medicines and health insurance); clothing and footwear; and durable goods (e.g., furniture, household appliances, car, etc.).

Expenses on rent were excluded from the total consumption mainly because of the high rate of non-response and the corresponding difficulty in imputing rents for those with missing information. The LITS2 asks about actual rents paid to households that are renting their dwelling and asks about imputed rents to households that are homeowners. Imputed rents are not always reliable because these are hypothetical amounts that homeowners would pay if they were to rent a dwelling like theirs. The LITS2 sample is split between 13% of renters and 87% of homeowners, with large variations across countries. However the rate of non-response is rather large: around 10% of renters and 40% of homeowners do not report rent, which means that more than one third of the sample has missing information on the value of housing. Imputing rents for those with missing data would have probably resulted in an extremely imprecise estimation, thus it was decided to exclude this component from the consumption aggregate. If imputed rents cannot be estimated, for consistency purposes actual rents must be excluded from the consumption aggregate.

Different reference periods are employed to capture the seven broad consumption categories that will be included in the consumption aggregate, a feature that is related with the frequency of consumption and purchases: the last month for food, beverages and tobacco; utilities; and transportation, while the last 12 months for education, health, clothing and footwear, and durable goods. Total consumption was constructed per month, thus all annual figures were converted into monthly expenses in order to have a common reference period.

Nominal consumption needs to be adjusted for temporal and spatial price differences to get a measure of real consumption. Temporal price differences are associated with the length of the fieldwork, while spatial differences are associated with geographical price differences, say, urban and rural prices. The measure of consumption based on the LITS2 does not make any price adjustment, which means that the proxy for welfare will be nominal consumption rather than real consumption. A temporal adjustment was not considered necessary because the fieldwork of the survey was fairly short. On the other hand, the spatial adjustment would have been more relevant but it was difficult to implement, thus the assumption is that geographical price differences within countries are not significant enough to alter the ranking of the population in terms of consumption.

The consumption aggregate of the household was divided by household size to obtain a measure of individual rather than household welfare. Last, the consumption aggregate per person is used to rank in ascending order the population in each country covered in the LITS2 and three equal groups are constructed: the lower 33%, the middle 33% and the upper 33%. Notice that the LITS2 gathers consumption data in the local currency of each country, thus consumption levels across countries are not directly comparable. Yet this does not affect the estimation of the income groups because they are calculated per country.