Technical Note No. 6

Note on Some Derived Variables

NOTE ON SOME DERIVED VARIABLES

Background:

In this note, we explain how three variables that are in the Yemen dataset were constructed. These variables refer to the topographic information about the location of households, the poverty category of a household, and the Social Welfare Fund beneficiary status of households.

Constructed variables:

(i) Topographic information: `topography`.
   In our dataset, the topography of Yemen is assumed to have four categories: Mountainous, Plateau & Desert, Costal Area – Arabian Sea, and Costal Area – Red Sea. A dataset containing information on these topographic characteristics of locations of Yemen, provided by CSO/MOPIC was merged with our dataset. We had topographic information for each combination of governorate, sector number, section number, and area number.

(ii) Poverty category: `poor`.
   The variable `poor` is constructed based on the proxy-means test that was developed by the SWF with technical assistance of the World Bank to target the beneficiaries of the SWF and using data from each round of the NSPMS. The PMT was also calculated using household level information of each round. The PMT weights used to assess the poverty status of the families were based on a multivariate linear regression of the logarithm of the per-capita expenditure on a series of household-level variables, head of household personal characteristics and area of residence and governorate. Data used in the regression were taken from the 2005-2006 Household Budget Survey (HBS) and the same variables used in the regression to get the weights of the PMT were also collected by the Comprehensive Social Survey (CSS). Thus, in order to predict the value of the household per-capita expenditure using the CSS database of actual or potential SWF beneficiaries, these weights were multiplied by the same HBS variables used in the regression analysis. Based on the predicted per-capita expenditure, households were classified into groups A,B,C and D (poor) and E and F (non-poor), according to cut-off points that vary by both governorate and area of residence (rural and urban), in order to take differences in the cost of living into account.

(iii) Beneficiary status: `swf`, `bitreat`, and `benefcase`.
   a. The construction of the variable `swf` is based on the round information provided by variables `p6b_q_15` and `p6b_q_16`. Variable `p6b_q_15` allows us to know whether the household receives income from the Social Welfare Fund: if it equals 1 `swf` is set equal to 1, and if it equals 0 `swf` is set equal to 0. By using the information in the `p6b_q_16` variables, we define a third category for `swf`: it is set equal to 88 if someone not currently in the household receives income from the Social Welfare Fund.

b. The variable `bitreat` was constructed based on the information contained in the `swf` variable: if in any of the NSPMS rounds the `swf` variable is different from zero, `bitreat` equals 1; otherwise, `bitreat` equals 0.

c. Finally, in order to classify a household as an old or new beneficiary of the Social Welfare Fund, we use administrative information that was provided by the SWF administration. Through this procedure it was possible to classify 85 per cent of the SWF beneficiaries between old and new beneficiaries. The former corresponds to those beneficiaries who started receiving the SWF transfer before the CSS and the latter to those who became beneficiaries after the CSS. However, there were 734 SWF beneficiaries in the NSPMS sample that were not matched with the SWF administrative database due to insufficient information to perform the match (mostly lack of knowledge of the SWF card number or misreporting). These are the cases for which we have developed a methodology based on the total amount of SWF transfers received during round 1 (October-December 2012) and the self-reported year of accreditation into the program. Since around 50 per cent of the households in the NSPMS did not know when any of its beneficiaries started receiving the Social Welfare Fund benefit, we had to find another source of information to classify the remaining 15 per cent of beneficiaries into old and new categories. Old and new beneficiary households show a distinctive pattern in terms of the amount received in round 1: new beneficiaries received transfers above 30,000 Riyals due to the accumulated value of 5 payments in arrears whereas old beneficiaries received their normal payment.

In order to impute the type of beneficiary (old or new) to the households whose beneficiaries were not matched into the SWF administrative database (table SWF.4), the following procedure was adopted. Based on the self-reported definition of the SWF beneficiary in round 4 of the NSPMS, households where classified as new beneficiaries if their individual beneficiary had:

(1)   Received 30,000 Riyals or more in round 1;

(2)   Received less than 30,000 Riyals but became a SWF beneficiary after 2011 in round 4.

This imputation was restricted to households which stayed in the NSPMS for the four rounds (balanced sample), as we are interested in documenting the evolution of indicators for new and old beneficiaries separately. The variable `benefcase` is equal to 1 if there is any old beneficiary in the household; equal to 2 if all the beneficiaries in the household are new; and equal to 0 if no member is a beneficiary. Variables `oldtreat` and `newtreat` are directly derived from `benefcase` to indicate, respectively, old beneficiaries versus non-beneficiaries and new beneficiaries versus non-beneficiaries.