THE CDS/TA MAPFILE

This Excel file has one row per CDS child, beginning with the traditional PSID 'primary key' ID for individuals, their 1968 ID and Person number as a pair. As noted in Tutorial #7, the use of the FIMS system is essential for obtaining the full range of mappings between one group of individuals in the PSID with those who are in a relation of interest – such as children to grandparents, or to their cousins, or to their siblings, or to other relatives. The purpose of the mapfile or of FIMS is to identify a subset of observations of interest based on relationships to others in the PSID file and in the case of TA to identify whether the individual still is 'at home' (dependent on the parents) or on their own – a head or wife permanently living away from the parents.

For the TA and CDS data, there are several relatives of obvious interest in that they have characteristics which, by their relation to a given TA/CDS child, are likely to be individuals for whom the researcher will want variables describing them as they relate to the CDS/TA child – for example, the education of the child's Primary Caregiver gives rise to a interest in knowing the 68 ID and person number of that person and similarly for the OCG (Other Care Giver).

The TA/CDS data are a cohort based sub-panel in their own right, so we often want to know if the original CDS child from 1997 was successfully followed in 2002/03 and beyond into CDS III (2007/08), or into TA05 and beyond, depending on the year in which the child 'ages out' of CDS (at age 18 they are out of the CDS waves and become interviewed in the TA phase) and whether the child was successfully followed in TA. Another perspective on the mapfile is to note that the CDS created a new type of relationship in the PSID. There have always been categories of relationship to head – wife, son, daughter, and many others. The PCG is a relationship to child and is defined by the person so identified at each CDS wave, with priorities as noted in the CDS documentation http://psidonline.isr.umich.edu/CDS/cdsi_userGD.pdf#page=17.

An important element in the mapfile is the record of both successful interviewing and the different age based eligibility for TA. Consider the 14th Excel or (13th data row). There was a successful interview with the child in both 1997 and 2002. For 2007 the codes for the CDS 07 are blank. That information by itself could be considered an indication that a 2009 CDS interview was a refusal. On the other hand, as of CDSI, the age range was 0-12, so by 2007 the age range was approximately 12-24, so many of those in CDSII could be eligible for TA07 (and some also for TA05, as for data row 12, when the ages would have been approximately 8-20). For row 13 we can see that there is a value of '1' in the column for TA07KID and TA09KID. This indicates that there was a successful administration of the TA module for this individual child in both 2007 and 2009.

An important distinction for any TA observation is whether, in that year, the TA person was a member of their parental family (referred to as an Other Family Unit Member or 'OFUM' in PSID terminology, or had become 'split off', forming their own individual family in the larger PSID sample. This family status is indicated in a respondent status variable, TA_RESP. Was the information collected from the TA as a HEAD, a WIFE ("WIFE") or as a dependent child (OFUM)? The codes are HEAD = 1, WIFE = 2 and OFUM = 3. Was the TA living in rented or owner-occupied housing? That information would be in the family record for the given PSID wave. Here, it is in the cell for TA07_PSID. Specifically it is 2007 family '1154'. So the 2007 family with that ID for the main survey would have the information on whether the family owned or rented, for example. It could be that of the parents for OFUM's or for the TA as the main survey participant – if they were on their own as head or wife.

Across the CDS waves, there is always a PCG if the interview was competed, but who is the PCG can and often does change and, in addition, the OCG (other caregiver) can and does change. Further, it was often not possible to identify an OCG in any wave of the CDS for a given child. If so the data entry will be blank. From the perspective of the child, this changing identity of the PCG and OCG can be seen from examining a few rows of the CDS/TA mapfile. Take the very first row of the file. The 68 ID/PN pair is 4

and 37. The child has a parent in the Parent ID file (most do), and if in the wave (all are in CDSI) the child has a PCG with a 68ID and PN (all do). For 2002/03 the child and PCG/OCG were not successfully re-interviewed. As a result the 68 ID and PN of the 2002/03 PCG and OCG are blank, and similarly for the next row. The third row is for a CDS child where a successful re-interview occurred in 2002/03. There we can see that the relation of the PCG to the child was that of the maternal grandmother (MM with a numeric code value of 3 and note there are separate codes for each of four possible grandparents). There are cases where the PCG is the father (F with a numeric code value of 2). Suppose one wanted to know what factors lead to a PCG who is MM. Here there are some cases – and maybe not enough to support a study of when is MM the PCG. But suppose a project was studying the education of the PCG. Then it would be the mother's education in most cases, but in others it would be the education of the maternal grandmother (MM) and in still other cases it may be the paternal grandfather (FF). Depending on the research question, even though the number of cases is small, one could link to the education of these different individuals to create a variable 'education of the PCG' for all CDS kids as of 1997. Of course here there are judgments, too. The PCG changes in many cases, so is it the education of the PCG as of 1997 or 2002/03 or some average if the PCG changes?

Another value of the mapfile is to underscore the variability on care arrangements over the early childhood. Relatively few rows have an unchanging PCG and OCG across even the first two CDS waves. This can give rise to a research agenda on the role of family stability in the early childhood. Note that as additional TA waves are posted in the Data Center we will be extending the rows to accommodate the added information on whether the individual continued as a respondent (TA11RESP, TA13RESP,…). Finally, we can see that many of the rows are blank for the TA05KID column. This is because either the child was not successfully followed into TA05 or the child had not yet migrated into the TA age range. There is an annual flow of CDS 'alumni' into the TA ranks and the TA sample thus builds up over time in this fashion.