

# TAPS 2002-2006 PANEL DATA DICTIONARY

## THE PURPOSE OF THIS WORD DOCUMENT

This document describes and explains the Tsimane' Amazonian Panel Study (TAPS) Panel Data Dictionary Only variables in all years. Background information on TAPS -- its purpose, sample, scope, publications, etc. -- can be found elsewhere (<http://people.brandeis.edu/~rgodoy/>). The TAPS working paper # 37 (<http://people.brandeis.edu/~rgodoy/working%20papers/TAPS-WP-37.pdf>) and the introductory web pages of TAPS give an overview of the project.

## WHAT IS THE TAPS PANEL DATA DICTIONARY?

The TAPS Panel Data Dictionary is an Excel file that names and describes all the variables used in TAPS in all years from 2002 to 2006. The Panel Data Dictionary also shows differences between the definitions of variables (if any) across years.

## HOW IS THE FILE ORGANIZED?

The Excel file contains two worksheets:

1. All Years Combined
2. Suffixes Used

## WORKSHEET 1: ALL YEARS COMBINED

This worksheet list all the variables used from 2002 until 2006 (inclusive). For each variable, this worksheet gives the following information in three broad categories:

1. Information on the Variables (*columns A – C*)
  - a. The Final Variable Name. (*column A*)
  - b. The suffix used (*column B*)
  - c. The definition of the variable (*column C*)
2. Definition differences (*columns F – G*)

### 1a. INFORMATION ON THE VARIABLES - Final Variable

#### *Column A*

Column A gives a list of all the variables used in TAPS from 2002 to 2006.

### 1b. INFORMATION ON THE VARIABLES – Suffixes Used

#### *Column B*

Each variable has a suffix added to it that give two pieces of information:

- I. What years the variable was collected.
- II. Differences in the variable definition between years.
- III. Unusual values

- I. What year the variable was used is given by the suffix Y followed by a number for the year: 2=2002, 3=2003, 4=2004, etc. Each variable in this Excel sheet was used in all years. Thus, they all have the suffix Y23456.
- II. Differences in the definition of variables between years are given the suffix D followed by a number for the year: 2=2002, 3=2003, 4=2004, etc. Thus, if 2004 has a different variable definition than all other years, we would use the suffix D4. If the variable has identical definitions in all the

years it was used, there is no D suffix added. The D suffix always appears after Y23456.

- III. Unusual Values** Some variables have unusual values. We have flagged these variables with a suffix *\_f*.

In the STATA file, the Final Variable Names are combined with the suffixes. Thus, the variable *hagriceseed* (row 11) has the suffix Y23456D23 (cell B11). The STATA file will call this variable *hagriceseed Y23456D23*. For more information, the worksheet Suffixes Used defines all of the suffixes used.

### **1c. INFORMATION ON THE VARIABLES – Definition of the Variable**

#### ***Column C***

This worksheet column defines each of the variables. If the variable has different definitions in different years, this column will give the main (2005) definition (to see any alternate definitions, see columns F – G).

### **2. DEFENITION DISCREPENCIES**

#### ***Columns F – G***

If the variables have different definitions in different years (as defined with the suffix D, in column B) the alternate definitions appear in columns F – G. Column F gives the 2004 definition ONLY if it is different than the main definition. Column G gives the 2002 – 2003 definition ONLY if it is different than the main (2005) definition.

For example, refer to the variable *hagriceseed* (cell A11). The suffix used is Y23456D23 (cell B11). The “D23” in the suffix alerts us that the 2004 definition is different from the main (2005) definition. Because the 2004 definition is different than the 2005 definition, we will find the 2002-2003 definition in column F (cell G11).

### **3. WARNING ON THE USE OF THE VARIABLE AGE**

The variable for reported age (*idageY23456*, *idageaproxY23456*, *idagefracyearY23456*) contains errors because Tsimane’ do not have birth certificates and guess when reporting age. During most annual survey years we asked people afresh for their age and noted their responses, so there is much inconsistency in the age variable from year to year; some people age by about one year from one survey to the next, but others show a decrease in age. We have retained the original reported answers to be faithful to what people said, but users should be sensitive that the variable contains considerable measurement error. We are currently trying to provide improved estimates of age, and these should be available to the public as soon as we have the estimates.

## **WORKSHEET 2: SUFFIXES USED**

This worksheet gives a reference to all the suffixes used. Suffixes can be found in two places:

1. Worksheet 1: ALL YEARS COMBINED (column B)
2. The final STATA file.

For more information on how the suffixes work, refer to “1b. INFORMATION ON THE VARIABLES – Suffixes Used” above.