

Michael Kevane
Dept. of Economics
Econ 135
Spring 2005

Data analysis assignment 2: Due Monday April 25

data is available on website, filename timealloc.xls

The data contains information on the allocation of time for women. There are 105 women in the sample, each has an ID number. Each line represents a record for the woman in one of four time periods (mkpeak, wkpeak, wkoff, mkoff, defined below and represented by dummy variables). So there are 420 records. The time allocation data are expressed as the percent of working time, during a particular time period (say off-peak agricultural time and week days), that is spent on that activity.

ID ID of household
husb percent time spent on husband's field during time period
incom percent time spent on own income activities during time period
home percent time spent at home cooking and leisure during time period
mkpeak time period of peak agricultural activities, market day
wkpeak time period of peak agricultural activities, week day
wkoff time period of off-peak agricultural activities, week day
mkoff time period of off-peak agricultural activities, market day (omitted category)
oxen number of oxen
eth ethnicity (5=bwa, 1=mossi, 2,3,4 = djula)
bd year of birth of woman
birthgrp birth decade group of woman
sin is woman single? 0-1
mar is woman married? 0-1
wid is woman widow? 0-1
yarohan is woman yarohan? 0-1
wives# number of co-wives
ord what order is this woman among co-wives
husbabs is husband absent?
husbbd year of birth of husband

Run three interesting crosstabs using the data above, as a preliminary attempt to explain the pattern of time allocation. Don't run a cross-tab with a group that has more than 3 categories. If a variable has more than three categories, regroup the categories by creating a new variable with fewer categories. Explain what you have done in your comments.

Run two regressions explaining the pattern of time allocation, using the explanatory variables, and creating interaction terms (e.g. bwa*oxen, oxen*oxen). The dependent variable is the percent time in one activity (either home, income or husb). The independent variables should include mkpeak, wkpeak and wkoff. The independent variables should not include the other time allocation categories (i.e. in explaining percent time on husband's field, percent time at home is not an explanatory variable. By definition the three categories have to add up to one- so more time at home will be negatively correlated with time on husband's field).

Interpret your results in short typed comments after every table and regression.