
Subject: Which recode file to choose IR-KR-BR
Posted by [Alanood](#) on Mon, 27 Feb 2023 16:53:51 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi

I repst my question here in " Dataset use in Stata" b/c my question is related to dataset and stata , I hope i will get answered.

I am a PhD student in economics, I am writing my first paper in the impact of women employment on child nutrition - for Egypt - 2008 and 2014

My outcome variable is child health indicators (HAZ - WHZ - WAZ), I produced binary variables: (stunted wasted overweight underweight)
The main independent variable is women employment , yes/no
Also I have controled for other variables such as child characteristics , mother characteristics, Household characteristics.

In the begging I started using the KR file since the child is my outcome variable.I cleaned the data and started analyzing it, however, when I tried to replicate the tables and analysis for one paper which is published in Q1 journal.

The authors used Egypt 2014 DHS dataset.

The result for the replicating frequency tables were not the same, so I tried to use the IR file. and then I replicated these tables successfully.

So to confirm, The frequency tables for maternal employment comes from IR file.

For the OLS, I couldn't replicate their work yet but i got significant coefficient as them.

```
*****stata comand*****  
gen strata=v023  
gen psu=v021  
gen sampwt=v005/1000000  
svyset psu, strata(strata) weight(sampwt) vce(linearized)  
gen id = sgov  
  
*****OLS*****  
svy: reg stunted maternalemployed ib1.gender bage100 bagesq100 i.twin i.birthsize mage100  
magesq100 mothertotedu100 mothertotedusq100 i.marital nofchildern i.wealth nofadult i.id
```

They stated " We include a governorate dummy to control for regional differences and governorate fixed effect" so I include the i.id (the governate fixed effect)

```
*****Result*****
```

They have 12,502 observation , R2= 0.0760 , mother employed coefficient : 0.0316 (10% significant)

for my analysis i got

In the IR -	Number of obs = 9,809	R-squared = 0.0825	, mother employed
coefficient :	..0336397 (5% significant)		
In the BR -	Number of obs = 13,415	R-squared = 0.0751	, mother employed
coefficient :	.0248322 (10% significant)		
In the KR -	Number of obs = 11,774	R-squared = 0.0751	, mother employed
coefficient :	-.1192583 (1% significant)		

I used the same variables as they used and i did the weighting, strata, cluster , and fixed effect for the governate.

as they stated in their paper

" all the regression analyses have corrected for the survey design, i.e. the sampling weight, the cluster, and the strata were all taken into account."

My questions here are:

1- - what do you think which file should i go with ? As you can see the coefficient of the Maternal employment in KR file is negative where it's positive in the IR file. the relationship between the dependent variable and independent variable will be different according to the which recode file i am using.

2- is it possible to do the Description table from one file (IR) and the analysis from another file (KR)
or the paper should use only one kind of recode file. (IR BR KR)

I appreciate your feedback and assistance.

Thank you