
Subject: how to test parameter stability in stata and pool DHS DATA

Posted by [Robert](#) on Sat, 25 Apr 2015 09:18:52 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi every one. In the process of estimating the effect of fertility on maternal and child welfare in Tanzania using DHS for 2010, 2004 and 1996, i have been told to test for the parameter stability using chow test in stata. I have also been told to pool the DHS data set for Tanzania for the years 2010, 2004 and 1996. Could any one have an idea?

Robert

Subject: Re: how to test parameter stability in stata and pool DHS DATA

Posted by [user-rhs](#) on Wed, 29 Apr 2015 01:20:36 GMT

[View Forum Message](#) <> [Reply to Message](#)

Robert, what software are you using? If you pool the DHS datasets, you need to first rescale the weights (see discussion here: http://userforum.dhsprogram.com/index.php?t=msg&th=982&am;goto=2032&S=ee7daaf64f6e1a50d26d8f9bc002826a#msg_2032) and be sure to create a survey-year-specific cluster variable, because the same cluster numbers may have been used from year to year, but these clusters may not correspond to the same geographic boundaries between survey years.

The Chow test is done to test structural changes over time. Basically, what you do is interact your time dummies with the predictor(s) of interest and test the joint significance of those interaction terms. The Chow test is actually just a regular F-test.

From Wooldridge's Introductory Econometrics: A Modern Approach (2003) (link here:

<http://www.amazon.com/Introductory-Econometrics-Modern-Approach-Economics/dp/1111531048>), emphasis added:

Quote:

For many time periods and explanatory variables, constructing a full set of interactions can be tedious. First, estimate the restricted model by doing a pooled regression allowing for different time intercepts; this gives SSR_r . Then, run a regression for each of the, say, T time periods and obtain the sum of squared residuals for each time period. The unrestricted sum of squared residuals is obtained as $SSR_{ur} = SSR_1 + SSR_2 + \dots + SSR_T$. If there are k explanatory variables (not including the intercept or the time dummies) with T time periods, then we are testing $(T - 1)k$ restrictions, and there are $T + Tk$ parameters estimated in the unrestricted model. So, if $n = n_1 + n_2 + \dots + n_T$ is the total number of observations, then the df of the F test are $(T - 1)k$ and $n - T - Tk$. We compute the F statistic as usual: $[(SSR_r - SSR_{ur})/SSR_{ur}][(n - T - Tk)/(T - 1)k]$.

If you are using Stata, you can accomplish this by using the `-test-` command post-estimation (just as you would a "regular" F-test in Stata).

Reference

Wooldridge, J. 2003. Introductory Econometrics: A Modern Approach. 2nd ed. New York: Thomson Learning.

hth,
rhs
