
Subject: How to arrive at final wealth index score using PCA
Posted by [annabellebronson](#) on Mon, 19 Feb 2024 10:10:10 GMT
[View Forum Message](#) <> [Reply to Message](#)

smash karts

I need to arrive at a final wealth index score for my analysis using PCA in stata. I am working with the Demographic and Health Survey dataset which provides only an asset-based wealth index for measuring socio-economic welfare. I am unsure about the variable normalisation process to arrive at the final wealth index score for each household. I have followed the variable preparation process using the Filmer and Pritchett /Demographic and Health Survey and coded each variable as a dummy. I have always assumed that if I predict the scores using the command `predict score1` for example and apply the survey weights to generate break the scores up into quintiles, I arrive at quintiles representing different categories of wealth from the poorest to wealthiest. I, however, have been unable to duplicate the survey provided wealth index and quintile distribution using this approach. They provide the code in SPSS, and I am working in stata. I need to be sure what I am doing is right and not necessarily arrive at the very same index.

I did some more reading used the following

document;<http://www.psi.org/equity-wealth-quintileguide> from the Population Service International (PSI). They show how to construct a wealth index based on a survey. In the document which is a guide they use the PCA command and then generate a factor score using the `predict` command, it appears they do not use this as the final wealth index score. They, however, go on to standardise each variable using the mean and standard deviation. They then proceed to create a final wealth index score for each household by multiplying the standardised variable by the factor scores obtained from the PCA and summing these up for each household. I guess my questions are;

1. Is predicting the relevant factor score using the `predict` command in stata alone adequate to arrive at a wealth Index score based on which to rank households.
2. Is ranking the households using quintiles and survey weight command sufficient to reflect the distribution of wealth based on this index.
3. is the process of generating the index using the standardised variable as described above same or similar to the stata `predict` command, ie. another way to obtain the score?

In the interest of brevity my commands are as follows;

```
pca n1, n2, n3 ... ni [aweight = weight], means
predict wealhscore1 wealhscore2
sort wealhscore1
xtile quintile= wealhscore1 [pweight=weight], nq(5)
```

Some clarity as soon as possible will be greatly appreciated. regards,