Subject: Re: Panel data set for cross country analysis
Posted by Reduced-For(u)m on Sat, 11 Jul 2015 00:32:46 GMT
View Forum Message <> Reply to Message

The unbalanced-ness (if that is a word) of the panel stems from, I think, you conceiving the panel as a survey-round specific panel. So the panel would be Country as the group variable and Survey Year as the time variable - an "observation" in the panel would be a single survey from a single country.

In that case, yes, you have a really unbalanced panel. You won't have very many observations per year (to use Year fixed effects or dummies) and you would only be able to use about 40 countries. But you could still look at changes over time within each country and aggregate that over all countries (a "within" estimator, or a "country fixed effect" estimator without time dummies). And it is not clear to me that the unbalanced panel problem is so big as to make any estimates worthless... you'd have to search through the literature on that, but it isn't always a huge problem.

Another option is to generate a country-by-cohort panel, so that for each country and each birth-year, you estimate the U-5 mortality rate for cohorts born in that year. You could do this using the birth history and generate a (mostly) balanced panel for about 2 decades from each country. The big issue here is that you have to use data from way back in birth histories, and measurement error about year of birth grows really fast as you ask mothers to look back over time.

An alternative to this would be to use U1 mortality. Then you could safely go back about 10 years or so from each survey date, having a reasonable U1 mortality rate estimate for each year.

One other problem: you want to look at SES determinants of child mortality. But things like, say, wealth index* or maternal education are measured in the survey year, and you only know U5 mortality rate for children born at least 5 years before the survey. So a household current SES measure is possibly a poor measure of the SES facing the household when that child died (or that caused that child to die). This problem has no obvious solution unless you can bring in local aggregate data from another source or use only time-invariant characteristics of countries (if any such exist).

*Also - wealth index is not comparable across countries or over time, so it would be a bad covariate choice in this case.