

p13: General Comment on Addition of New Variables

Many of the 'proximate' contextual variables that are included in the models are what might be called second order proximate contextual variables. To me they all seem to be describing some opportunity structure (either women based or child based). I suppose in economic terms it might be expressed in terms of conditions which are faced by individuals or households which are trying to maximize their utility. Social contexts in which women's status is high indicate a context in which certain opportunities are available for women, which in turn shapes values related to other behavior, such as fertility (and these values affect all women in that context irrespective of their individual status). The Davis argument seems to me to be that as opportunities (for increased social status) develop there will be demographic responses which allow the opportunities to be accessed. The new variable 'agricultural pressure' is meant to measure the lack of opportunities within a particular setting, one problem identified in the text is that there are alternatives i.e migration which can be used to access other opportunities, or responses other than fertility that can be used to compensate for lower levels of opportunities (in the text you identify delayed marriage). However, what might also be important, are other opportunities within the community. In part this is reflected in some of the other contextual variables (such as the percent in non-agricultural occupation -- which while used here as an indicator of women's status probably also reflects to some extent the availability of non-agricultural occupations). In terms of the proposal these meanderings are probably not very helpful but if some measure of opportunities was available for a particular area it might provide a better indication of than agricultural pressure (for example in Central Java many kabupaten with high population densities also experience higher standards of living than kabupaten with lower densities precisely because their are other industries integrated into the agricultural household). I can't think of any better variable (although the in-migration rate in the five years prior to the census would be a good indirect measure of opportunities, assuming rationality of individuals, for both urban and rural areas), so I guess it might be just a matter of experimenting with different forms of agricultural density indicators.

p14: - Third Paragraph

One strategy is to develop latent variable models and estimate them with appropriate software such as LISREL or EQS (Bentler, 1985). In this framework, there is an unmeasured variables representing the theoretical concept which predicts the measured indicators. The model estimates the structural relationships between the unmeasured independent variables and observed measure of fertility, using all the information from the relationships of the measured variables. The advantage of this approach is that common covariance of the indicators is used to estimate the structural relationship, leaving out the unique effects that may be due to correlations with unknown variables. The result is a "purer" measure of the effect of the conceptual variable on fertility than can be obtained from a single indicator. The disadvantage is the loss of the potentially greater predictive power from the combination of separate indicators. In the preliminary analyses based on these models we will use confirmatory factor techniques to determine the applicability of adopting a latent variable approach. Our analysis will then compare latent variable models with models that rely only on observed variables. The only real change in the preceding paragraph is say that a latent variables is constructed in terms of its indicators (and other latent or observed variables in which there is a non-zero covariance). There are no covariances between dependent and independent variables (these are what are being explained see Bennet p 10-11). Also, as there is no structural path between indicators and a dependent variable not directly dependent on the indicators (ie. not the latent variable that produces the observed indicators) then there is no way that the values of the indicators vis-a-vis the dependent variable

can effect the latent variable. I suppose what I am saying, although not very well, is that the structural equation we are proposing is recursive and only the latent variable, in conjunction with shared correlation from other independent variables, can give rise to a particular shared covariance between the indicator and latent variable. If a model was non-recursive then I suppose that the shared predictive power of the indicators could be incorporated into the latent variable (although I can't see how this could be estimated in EQS). The advantage of the latent variable model is that it reduces the problems of measurement error. Any single indicator is an imperfect measure of the concept, by using several (the more the better) indicators and maximizing the common variance a better measure of the concept should be obtained.

p18: - Possible addition after third paragraph

Our cross-sectional analyses have suggested that at different stages of the fertility transition the influence of factors on fertility differ. With pooled data from the two time periods now available we are able to expand our analyses to formally test hypotheses relating to structural change (ie. whether there are significant differences in the **cross-sectional** determinants of fertility in two periods). This is equivalent to extending equation 3 above to include interactions between time and the structural variable(s) of interest. Where three or more time points are available we can test hypotheses about the effects on fertility of change in social contexts across time periods by including interactions between change variables and the year dummies. The construction of pooled microdata with data from several time periods will therefore allow us to test models of fertility determination which depend on changes at the societal (contextual) level as their central explanatory concept.

Charlie what I am saying above is a little convoluted but the essential line I am trying to get across is that we basically have two sets of hypotheses. The first focuses on differences in the year coefficient (this is your model 3) and the interpretations basically revolve around change in the coefficient for year after controlling for the levels of the independent variables. This is (for want of better term) a test of structural equivalency -- ie. any changes that occur in the coefficient for year are due to changes in the levels of the independent variables. Now it is quite possible to argue (as we did implicitly in the rural soc paper) that there could be no change in the levels of the independent variables yet a significant change in the year coefficient. This would be a structural change hypotheses, ie. the structural process operating at one period were different from those operating at another period. We make this interpretation in the rural soc paper by running the models separately and then examining the coefficients. A formal test of the hypotheses generated from that study would involve interactions with years (in a completely saturated model we would get exactly the same predicted values as we do for the separate equations (see Johnston, p207). The advantages of this framework is that it opens up a whole new area (as you point out in the last para of page 3 of the proposal).