Empirical, methodological and theoretical issues in multilevel modelling in demography

Local context and marital fertility in Burkina Faso

A multilevel longitudinal analysis

Contact address

Bruno SCHOUMAKER
Institute of Demography, Université catholique de Louvain
1/17 Place Montesquieu - B-1348 Louvain-la-Neuve – BELGIUM
Phone : ++ 32 10 474136 --- Fax : ++ 32 10 472952
Email : schoumaker@demo.ucl.ac.be
Introduction and objectives

The influence of community-level factors on fertility has received increased attention over the last twenty years. Since Freedman’s WFS occasional paper discussing the effects of community-level variables on fertility (Freedman 1974), various authors have insisted on the role of the local context in explaining fertility behaviour (Bilsborrow and Guilkey 1987; Casterline 1985; Kravdal 2002; McNicoll 1988; Potter 1983; Schoumaker 1999; Smith 1989). Many explanatory approaches in fertility research (institutional, diffusion, micro-economic…) indeed point to potential influences of community-level characteristics on individual fertility behaviour (Schoumaker 2001a).

The development of multilevel models since the early 1980s, but especially from the mid 1990s, has somewhat stimulated research in that field. The number of articles on fertility determinants including community-level variables has substantially increased over the last twenty years (Schoumaker 2001a), and the statistical models used in multilevel fertility research have greatly improved, notably with the development of multilevel structural analyses (DeGraff, Bilsborrow and Guilkey 1997) and multilevel longitudinal analyses (Angeles, Guilkey and Mroz 1998). Yet, some of the “problems” in the multilevel literature on fertility identified a few years ago by DeGraff et al. (1997) are still valid.

One of the problems identified by these authors refers to the incorporation of only family planning services and the exclusion of other potential community influences on fertility. Indeed, even though other community-level variables have also been incorporated in multilevel analyses of fertility in recent years (population density, see Filmer and Pritchett 1996; educational level, see Kravdal 2002; child mortality, see Rosero-Bixby 1998), the effect of family planning services on fertility and contraceptive use is the most frequently studied community effect (for a review, Schoumaker 2001a). One obvious reason for this is the lack of rich community-level data sets, since most multilevel studies use the DHS service availability module as the source of community data.

Another issue that was mentioned by DeGraff et al. (1997) refers to the appropriate time frame for the dependent and explanatory variables. More specifically, most community-level variables used in multilevel studies of fertility refer to the time of the survey, while fertility behaviour (the dependant variable) often refer to past conditions, either because of the nature of the dependant variable (cumulative fertility) or because of the use of retrospective birth histories in event-history models. The major reason for this discrepancy is the lack of retrospective data on communities that can be merged with individual birth histories (Axinn,
Barber and Ghimire 1997). While other problems affect multilevel research of fertility, these two problems can thus be traced to the lack of rich and retrospective community-level data.

In this paper, we use a unique multilevel longitudinal data set to test new hypotheses on the effects of community-level variables on marital fertility in Burkina Faso. The two surveys used in this study contain detailed retrospective information on both women of reproductive age and on the communities in which they lived at every point in time. In this research, we also propose a new method for multilevel longitudinal analyses of birth histories. Our statistical model, a multilevel longitudinal adaptation of the Rodriguez-Cleland model of marital fertility (Rodriguez and Cleland 1988), allows us to test the effects of fixed and time-varying variables on fertility behaviour (Schoumaker 2001b).

In short, the originality of this paper lies in (1) the uniqueness of the multilevel longitudinal data and (2) the methodology proposed for multilevel analyses of birth histories.

The data

The data used in this study come from two surveys. The individual life history data come from a nationally-representative retrospective survey entitled “Migration Dynamics, Urban Integration and Environment Survey of Burkina Faso”. The survey was conducted in 2000 by the University of Ouagadougou, the University of Montreal and the CERPOD (Poirier et al. 2001). The female sample comprises 4 568 women aged 15-64 at the time of the survey. The questionnaire covered several topics such as migration, employment, marital and birth histories. The dependant variable of this study comes from the birth history; explanatory variables will include time-constant variables as well as time-varying variables from the employment and migration histories.

The community-level data come from a retrospective community survey conducted of 600 villages and small towns in 2002 (Schoumaker, Dabire and Gnoumou-Thiombiano 2004). The survey was designed to be linked with the individual retrospective survey described above. It comprises all the villages in which people lived at the time of the survey and a large sample of villages in which they lived in the past. This allows us to link individual life histories with community-level retrospective data, even for people who have migrated. Each village for which at least three spells of residence were reported in the migration histories was covered by the survey (600 villages of a total of 1800 villages in which at least one spell of residence was reported). Retrospective data was collected from groups of community informants,
consisting of village chiefs, administrative representatives and other knowledgeable informants. It covered a broad range of topics, including land availability, availability of schools and health centres, transportation, agriculture, and employment opportunities. Efforts were made to obtain retrospective information since 1960 on most village characteristics.

**Method**

The statistical model used in this study is a multilevel longitudinal adaptation of the Rodriguez-Cleland model of marital fertility. The Rodriguez-Cleland model was first used in a comparative study on the effect of instruction on fertility in developing countries (Cleland and Rodriguez 1988; Rodriguez and Cleland 1988). The starting point of the model is that in the absence of deliberate fertility control, fertility rates follow a natural fertility schedule, i.e. fertility rates at each age are the product of age-specific natural fertility rates and a constant. According to the Rodriguez-Cleland model, when there is some degree of fertility control, observed fertility rates depart from natural fertility rates as a function of duration of marriage in the following way:

\[
f(a,d) = \theta . n(a) . \exp(\beta . d)
\]

where \(f(a,d)\) is the observed marital fertility rate at age \(a\) and marriage duration \(d\), \(n(a)\) is the natural fertility rate at age \(a\), \(\theta\) is a parameter representing the level of natural fertility, and \(\beta\) is a parameter measuring the extent to which marital fertility departs from natural fertility as a function of duration of first marriage (Rodriguez and Cleland 1988).

Rodriguez and Cleland (1988) have shown that the parameters of this model can be estimated using a Poisson regression in which the exposure and the natural fertility rates are absorbed in the offset. They assume that the number of births (at age \(a\) and marriage duration \(d\)) over a given time period follows a Poisson distribution with mean \((\mu_i)\) equal to the product of exposure time \(t_i(a,d)\) and of a theoretical marital fertility rate \(f_i(a,d)\).

\[
\mu_i = t_i(a,d).f_i(a,d)
\]

Replacing the theoretical fertility rates by the model of equation 1, equation 2 becomes:

\[
\mu_i = t_i(a,d).\theta . n(a_i) . \exp(\beta . d_i)
\]

Taking the logarithm on both sides of equation 3, we obtain the following model:
\[
\ln(\mu_i) = \ln[t_i(a,d)] + \ln[n(a_i)] + \alpha + \beta d_i
\]  
Eq. 4

Where \( \alpha \) is equal to the natural logarithm of 0. The right-hand side of the equation contains an offset, equal to the sum of the logarithm of exposure time and of the logarithm of the natural fertility rates. The rest of the equation is a simple linear function of marriage duration.

Using this model, marital fertility rates can then be summarised with only two parameters. The first (\( \alpha \)) is a parameter representing the level of natural fertility and is interpreted as a spacing parameter. The second (\( \beta \)) measures the extent to which marital fertility rates depart from natural fertility rates and is a control parameter. This model is thus very parsimonious and its parameters have a simple demographic interpretation.

Rodriguez et Cleland (1988) have shown that the fit of the model was excellent for most of the 38 countries included in their study. The great interest of their model, however, is that it can be used to evaluate the effect of explanatory variables on marital fertility (Cleland and Rodriguez 1988). Explanatory variables (e.g. variable \( x \)) can have direct effects (influence on the degree of spacing) or interact with duration of marriage (influence on the degree of control):

\[
\ln(\mu_i) = \ln[t_i(a,d)] + \ln[n(a_i)] + \alpha + \delta_1 x_i + \beta d_i + \delta_2 x_i d_i
\]  
Eq. 5

In this example (Eq. 5), \( \delta_1 \) measures the effect of variable \( x \) on child spacing, and \( \delta_2 \) measures its effect on the degree of fertility control.

As we have shown before (Schoumaker 2001b), the Rodriguez-Cleland model can easily be adapted into a multilevel model of marital fertility. For example, for the empty model, one simply needs to add two random terms at the community-level in equation 4. This allows the spacing (\( \alpha \)) and limiting (\( \beta \)) parameters to vary between communities. The random terms are assumed to follow a bivariate normal distribution:\(^1\)

\[
\begin{pmatrix}
    u_{0i} \\
    u_{1i}
\end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{pmatrix}\right)
\]

\[
\ln(\mu_{ij}) = Offset + \alpha + u_{0j} + (\beta + u_{1j}) d_{ij}
\]  
Eq. 6

\(^1\) Index \( i \) represents individuals, index \( j \) represents communities, and index \( t \) represents time periods.
where $\sigma_{u0}^2$ is the variance of the spacing parameter, $\sigma_{u1}^2$ is the variance of the limiting parameter, and $\sigma_{u01}$ is the covariance between these two parameters.

An interesting application of this multilevel adaptation of the Rodriguez-Cleland model is to allow to estimate the heterogeneity of the spacing and limiting parameters among communities. Both individual and community-level explanatory variables can also be included in the multilevel Rodriguez-Cleland model, and their effects on the heterogeneity of the degree of spacing and of fertility control can thus be estimated. Another originality of our adaptation of the Rodriguez-Cleland model is to allow the inclusion of time-varying explanatory variables. This is possible through episode-splitting, an approach similar to the one used in estimating piecewise constant rate models (Schoumaker 2004).

In this study, the parameters of the multilevel Poisson regression models are estimated with MIXPREG, a software developed by Hedeker and Gibbons for two-level Poisson regression models (Hedeker 1999).

**Country setting, hypotheses and expected results**

Burkina Faso is one of the poorest countries in the world: it ranked 159th of 162 countries in the UNDP’s human development index, and its gross domestic product (GDP) per capita was approximately $230 US at the end of the 1990s. According to the 1996 census, its population was around 10.3 million inhabitants and was still growing at an average rate of around 2.5 %. Burkina Faso is also one of the least urbanized countries in the world, with slightly more than 20 % of its population living in urban areas (Beauchemin, Beauchemin and Le Jeune 2002). Although fertility has remained fairly stable over the last thirty years at the national level, urban areas have experienced a significant fertility decline since the 1980s. According to the latest DHS (1998/98), total fertility was still at 6.8 children per woman in the country as a whole, but was around 4.1 children in urban areas.

In the first part, the heterogeneity of the spacing and control parameters will be estimated with the multilevel adaptation of the Rodriguez-Cleland model. We expect that, as a result of the onset of fertility transition in Burkina Faso (especially in more developed villages and small towns), the degree of fertility control will vary across communities. We also expect that the degree of spacing (natural fertility) will significantly vary across villages, for both biological
and volitional reasons (varying levels of sterility, variety of ethnic groups and postpartum practices,…).

Both individual and contextual factors will be taken into account to explain fertility differentials. In addition to classical individual-level determinants (education, level of living, ethnic group, activity), we will test the effects of a range of community-level variables on fertility. More specifically, we will look at the effects of the availability of health and family planning centres in the community, the availability of schools in the village, the existence of employment opportunities in and around the community, the presence of an all-season road and the degree of agricultural development of the community. We expect that the degree of fertility control will be higher in more developed communities, where the demand for children should be lower (children’s work less needed, diffusion of values through more frequent contacts,…) and the ability to control should be higher due to better health and family planning services.

References


