Living standards and fertility in Indonesia: A Bayesian analysis

Alessandra Mattei
Fabrizia Mealli
Department of Statistics, University of Florence

Stephen Pudney
Institute for Social and Economic Research
University of Essex

January 2009

Abstract

We investigate the relationship between living standards and fertility, using a three-wave panel dataset from Indonesia to provide information on women's fertility histories and the levels of consumption expenditure in the households to which they belong. We adopt a Bayesian approach to estimation and exploit the dynamically recursive structure implied by gestation lags to identify causal effects of living standards on fertility and *vice versa*.

Keywords: Indonesia; fertility; living standards; Bayesian

JEL codes: D31, J13, P36

Contact: Steve Pudney, ISER, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK; tel. +44(0)1206-873789; email spudney@essex.ac.uk

This work was supported by the Economic and Social Research Council through the project "Poverty Dynamics and Fertility in Developing Countries" (award no. RES000230462) and MiSoC research centre (award no. RES518285001) and by the European Community - Research Infrastructure Action under the FP6 "Structuring the European Research Area" Programme.

1 Introduction

The strong empirical association between high population growth and low economic welfare remains a contentious research issue in economics and demography and has been the subject of a great deal of econometric research (Birdsall et al. 1988, Lanjouw and Ravallion 1995, McNicoll 1997, Livi-Bacci and de Santis 1998). There is a common presumption of a causal link between poverty and fertility based, at the macro level, on the assertion that a higher population growth rate depresses capital accumulation and wages. Poverty in turn is considered as a key factor in driving high fertility and population growth. Poverty is thus widely seen as a crucial element in delaying the demographic transition of countries from high- to low-fertility status. A similar argument applies at the micro-level: individual fertility behaviour adjusts to changes in perceived and actual costs and benefits of children, with economic forces, social organisations and cultural patterns influencing these incentives. However, these theoretical arguments are not simple: the structure of causality is not immediately obvious and not easily determined from available observational data.

Most existing studies of the relationship between fertility and living standards have relied on cross sectional survey data or aggregate time series. Instead, we revisit this long-standing issue by using longitudinal data. Cross sectional data, no matter what techniques are applied, are unlikely to provide robust causal inferences about the relationship between the occurrence of births and poverty and the role of other intermediate variables. Only longitudinal surveys can provide information on the timing and duration of poverty, implying that panels will provide much richer information on issues such as the permanent nature of the poverty, changes in poverty status of individuals over time and about the events related to entry into and escape from income poverty.

In this paper we study the dynamic relationships between poverty and fertility in Indonesia, a developing country with a large population which has experienced unprecedented economic growth and sharp fertility declines over recent decades. The analysis is based on data from the Indonesia Family Life Survey (IFLS), a longitudinal survey conducted by RAND in collaboration with UCLA and Lembaga Demografi, University of Indonesia.

Recent work on poverty dynamics and fertility in developing countries shows that longitudinal micro data used with appropriate econometric techniques (such as non-parametric matching, instrumental variable methods or simultaneous equation models with correlated random effects) can enhance our understanding of causal relationships between poverty and fertility, and possibly other related processes such as education, health, and employment. Results for Indonesia, also based on IFLS data, are provided by Kim et al. (2005), Kim and Prskawetz (2006), Aassve and Kim (2006), and Mattei (2007).

A drawback of these studies is that they focus on one-way direction of the causal relationship between poverty and fertility, rather than treating the two processes as dynamically inter-related. The statistical approaches used in these studies do not deal with the issue of time-varying covariates, nor do they exploit specific information on the timing of events. Instead, they assume only a single point of time at which a possible 'treatment' (e.g. birth of a new child, occurrence of idiosyncratic income shocks, etc.) or an outcome may occur. As Abbring and van den Berg (2003b) point out, treatment and outcome are characterized by their time of occurrence, and the timing of the treatment relative to the outcome can convey useful information on the treatment effect. Exploiting timing by modeling jointly dependent processes improves the identification position (Abbring and Van Den Berg, 2003a) and facilitates analysis of poverty persistence by different groups. By analysing dynamically the impact of fertility on poverty and of poverty on fertility, these events enter as endogenous time varying covariates. To allow for this, we need to estimate the equations describing living standards and demographic events as a simultaneous system.

Aassve et al. (2006) implement simultaneous and dynamic random effect models as a way of analysing causality issues related to poverty and fertility in Ethiopia. In this paper we follow a similar approach for Indonesia, but with some important differences. First, rather than treating poverty status as the primitive concept, with consequent discrete transitions between poor and non-poor status, we model consumption as a continuous variable and then draw conclusions about poverty dynamics from the estimated model. This has three major advantages. First, it gives much better robustness to measurement error in the consumption data since, under classical assumptions, error merely inflates the residual variance rather than generating classification error in the poverty variable. Second, it is more efficient statistically, since it avoids discarding the rich information contained in the continuous consumption variable in favour of a coarse binary poverty indicator. Third, it is considerably more convenient, since there is no need to re-estimate the statistical model when exploring the consequences of using different poverty lines or equivalence scales.

A second major difference from the conventional panel data approach of Aassve et al. (2006) is the way we deal with the timing issue. We develop a discrete-time consumption-fertility model with inter-related dynamic equations for living standard and demographic events and, to allow for the 9-month gestation period in a satisfactory way, we use a quarterly chronology. Thus any influence of family circumstances on fertility decisions must occur at least three quarters earlier than the consequent fertility event. In implementing the model, we face the problem of a high prevalence of missing data on consumption. The sequence of fertility outcomes is fully observed, but the consumption process is only observed at three points in time for each woman, so that the vast majority of potential consumption observations are 'missing' as a consequence of the survey design. We adopt a Bayesian approach to deal with missing consumption observations using Markov Chain Monte Carlo (MCMC) methods with data augmentation. This is a major methodological innovation in the literature.

The remainder of the paper is organized as follows. Section 2 describes the Indonesian survey data used for the analysis and the statistical model which forms the basis for estimation of the structure of the joint fertility-living standards process. Section 3 outlines the Bayesian approach to estimation, with technical details relegated to Appendix 1. Section 4 discusses the results and their implications for causal structure and fertility-consumption dynamics and section 5 concludes.

2 The observational process and statistical model

2.1 The Indonesian Family Life Survey (IFLS)

The IFLS comprises four waves, the first conducted in 1993/4, the second in 1997 and the fourth in 2000. The third wave, in 1998, which was specifically aimed at capturing the short-term effects of the financial crisis in Indonesia at the end of 1997, used a modified questionnaire and covered only 25% of the original sample. For this reason, we use only data from waves 1, 2 and 4. The survey was conducted by the RAND Corporation in collaboration with UCLA and Lembaga Demografi, University of Indonesia. The response rate for the IFLS is impressive: over 90% of the original 1993 sample remained in the survey at wave four. The survey contains a wealth of information collected at the individual and household levels, including consumption expenditure, which we use as the primary indicator of economic welfare. There is also detailed information on food prices, access to health and educational facilities, and the prices of services available at those facilities, allowing the construction of a high-quality real consumption variable. The questionnaire is sufficiently detailed to document relationships among co-residents and non-co-resident family members and inter-generational mobility. The sample is representative of about 83% of the Indonesian population and contains over 30,000 individuals living in 13 of the 27 provinces in the country.

The sample we use for estimation consists of all ever-married women for whom the following conditions are satisfied: (*i*) a complete fertility history is available (limited by the survey design to one woman per household, generally the spouse of the household head); (*ii*) aged 15-30 at the first wave of the IFLS; (*iii*) 'at risk' of fertility (i.e., has married or already conceived) by wave 1; and (*iv*) is recorded as marrying at or after age 14. This sample selection focuses attention on the core group of fertile women and excludes a small number of cases where there is grounds for concern about the accuracy of the recorded marital history.

Living standards are measured using the real monthly value of the total household consumption expenditure in constant 1993 Rupiah prices. Following the literature on developing countries, we argue that consumption is better suited than income as an indicator of living standard in Indonesia. Consumption is believed to vary more smoothly than income, both within any given year and across the life cycle, to be more readily observed, recalled and measured than income (at least in developing countries) and to suffer less from underreporting problems. In addition, life-cycle theories suggest that individuals will try to smooth their consumption across their low- and high-income years (in order to equalize their marginal utility of consumption across time), through appropriate borrowing and saving. This does not mean that consumption smoothing is perfect, in part due to imperfect access to commodity in credit markets and to difficulties in estimating precisely one's "permanent" or life-cycle income. Sometimes, consumption is also deemed to be a more direct indicator of achievements and fulfilment of basic needs. A caveat is, however, that consumption is indeed an outcome of individual free choice, an outcome which may differ across individuals of the same income and ability to consume.

The IFLS observational scheme yields essentially complete observation of the fertility process, which can be constructed from the composite fertility history reported by each sample woman. However, the family's consumption history can only be observed at three discrete points: the interview dates in the 1993/4, 1997 and 2000 waves of the IFLS panel. For different households, these three points span different segments of the family's history and, for any one household, the three observation points may occur in different seasons. Consequently, although we have identifying information on the whole fertility/living standards process and its seasonal structure, we are faced with a very large scale missing data problem.

2.2 The model

Our aim is to model the joint evolution of two processes: *fertility* is represented in a counting process $\{N_{it}\}$ recording the number of births to woman i up to time period t; and *living standards* are represented by a process $\{C_{it}\}$ defined as the log of real monthly total current consumption expenditure of the household to which woman i belongs. To allow for the 9-month gestation period in a satisfactory way, we use a quarterly chronology. Thus any influence of family circumstances on fertility decisions must occur at least three quarters earlier than the consequent fertility event.[1]

$$C_{it} = \rho C_{it-1} + \boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{h}_{it-1}\boldsymbol{\gamma} + u_i + \varepsilon_{it} \tag{1}$$

$$n^*_{it} = \lambda C_{it-3} + \boldsymbol{x}_{it}\boldsymbol{\delta} + \boldsymbol{f}_{it-1}\boldsymbol{\psi} + \eta u_i + \xi_{it} \tag{2}$$

$$\Delta N_{it} = \begin{cases} 0 & \text{if any birth in last 3 quarters or } n^*_{it} \leq 0 \\ \nu_{it} & \text{if no birth in last 3 quarters and } n^*_{it} > 0 \end{cases} \tag{3}$$

where: $\boldsymbol{x}_{it}$ is a vector of observed covariates; $\boldsymbol{h}_{it-1}$ and $\boldsymbol{f}_{it-1}$ are (vectors of) variables summarising the woman's fertility history up to time $t-1$; $\mathbb{1}(.)$ is the indicator function; u_i is an unobserved individual-specific random effect; ε_{it} and ξ_{it} are random innovations; and ν_{it}

[1] We neglect the possibility of pregnancy termination as an endogenous factor.

is the observed number of children born in any birth episode that occurs for woman i at time t. All other symbols represent parameters. We are particularly interested in the coefficient vector $\boldsymbol{\gamma}$, which represents the influence of past fertility on current economic welfare, and the parameter λ, which captures the anticipated negative influence of living standards on prospective fertility. Since there are theoretical reasons to expect random walk-like behaviour in consumption (Hall, 1978), we do not exclude the unit root case $\rho = 1$. Households are assumed to be sampled independently[2]. The number of birth outcomes at a birth event, ν_{it}, is treated as random, distributed independently of all other variables and is not modeled explicitly.

The woman's fertility process is assumed to begin at the date of her marriage or first conception, whichever is the earlier. Call this quarter 0. We specify the following model for the initial condition for consumption in period 0:

$$C_{i0} = \boldsymbol{w}_i \boldsymbol{\tau} + \alpha u_i + \zeta_i \tag{4}$$

where $\boldsymbol{w}_i$ is a vector of observed covariates, ζ_i is an unobserved variable and $\boldsymbol{\tau}$ and α are parameters.

We assume that $\{u_i, \varepsilon_{i1}, \ldots, \varepsilon_{iT_i}, \xi_{i1}, \ldots, \xi_{iT_i}, \zeta_i\}$ are a set of mutually independent variates and that:

$$\{u_i, \varepsilon_{i1}, \ldots, \varepsilon_{iT_i}, \xi_{i1}, \ldots, \xi_{iT_i}, \zeta_i\} \perp\!\!\!\perp \{\boldsymbol{x}_{i1}, \ldots, \boldsymbol{x}_{iT_i}, \boldsymbol{w}_i\}. \tag{5}$$

The structure (1)-(3) embodies causal links between the consumption and fertility processes in both directions, but has a dynamically recursive structure as a consequence of the gestation lag in the birth process.

[2]By design, the original IFLS survey over-sampled urban households and households in provinces other than Java, therefore sampling weights were estimated in order to make the sample of households or individuals representative of the 1993 population. However, since households and individuals are sampled only based on covariates, which are included in our models, design weights can be ignored when conducting inference (Rubin, 1976).

In order to make inferences about these causal links we introduce some additional assumptions on the distributions of the errors. We assume that for $i = 1, \ldots, N$, the unobserved variable ζ_i has a $N(0, \sigma_\zeta^2)$ distribution and ε_i has a multivariate $N(0_{T_i}, \sigma_\varepsilon^2 I_{T_i})$ distribution, where I_K is the $K \times K$ identity matrix. Finally, we assume that ξ_i has a standard logistic distribution.

3 The Bayesian estimation approach

3.1 The missing data problem

The main difficulty we face in implementing the model (1)-(3) empirically is the prevalence of missing data. Although we need to work in terms of a quarterly calendar to capture the effect of gestation lags, the consumption process is only observed at three points in time for each woman, so that the vast majority of consumption observations are 'missing'. However, they are missing as a consequence of the survey design and are therefore ignorably missing (Rubin, 1976 and 1987). In addition to this missing data problem we have an initial condition problem, since the start of the consumption process is endogenous and (for most cases) unobserved. There is no corresponding missing data / initial conditions difficulty for the fertility process, since we have a full fertility history for each woman in the sample, covering her whole life up to wave 3 of the IFLS.

We address complications due to missing consumption using a Bayesian approach. In recent years, the use of Bayesian statistics has conspicuously increased in econometrics and micro-econometrics (e.g., Athey and Imbens, 2004; Chamberlein and Hirano, 1999, Hirano, 2002; Koop, 2003; and Lancaster, 2004). One of the great advantages of a Bayesian approach to the analysis is the relative ease with which missing data can be allowed for, using Monte Carlo Markov Chain (MCMC) methods with the parameter and data augmentation method of Tanner and Wong (1987). Gilks, Richardson and Spiegelhater (1996) provide a review of these methods with references to the recent literature. Here we briefly discuss the general approach we take in this article. The Appendix contains more details on the specific implementation.

The data augmentation algorithm allows us to exploit the complete-data likelihood function, which depends on the observed data as well as the missing elements in the consumption process, $\{C_{it}\}_{t=0}^{T_i}$, $i = 1, \ldots, N$.

Let $\boldsymbol{W}_i$ denote the observed data for woman $i = 1, \ldots, N$:

$$\boldsymbol{W}_i = \{(N_{it}, \boldsymbol{x}_{it}, \boldsymbol{h}_{it-1}, \boldsymbol{f}_{it-1}, \boldsymbol{w}_{it}, C_{it_1}, C_{it_2}, C_{it_3}) \ , \quad t = 0, 1, \ldots, T_i, \}$$

where C_{it_1}, C_{it_2}, and C_{it_3} are the observed consumption level for woman i, and let $\boldsymbol{W}$ be the matrix stacking observations for all the N women. Moreover, let $\boldsymbol{\theta}$ be a vector including the common parameters: $\boldsymbol{\theta} = (\rho, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_\varepsilon, \lambda, \boldsymbol{\delta}, \boldsymbol{\psi}, \eta, \boldsymbol{\tau}, \alpha, \sigma_\zeta)$; and let $\boldsymbol{u}$ be the vector stacking the individual random-effects for all the N women: $\boldsymbol{u} = (u_1, \ldots, u_N)'$. Then, the complete-data likelihood function can be written as follows:

$$\begin{aligned} & \mathcal{L}\left(\boldsymbol{u}, \boldsymbol{\theta} | \boldsymbol{W}, \{C_{i0}\}_{i=1}^{N}, \left(\{C_{it}\}_{t=1}^{T_i}, i = 1, \ldots, N\right)\right) \\ = & \prod_{i=1}^{N} \frac{1}{\sigma_\zeta} \phi\left(\frac{C_{i0} - \mu_{i0}}{\sigma_\zeta}\right) \prod_{t=1}^{T_i} \frac{1}{\sigma_\varepsilon} \phi\left(\frac{C_{it} - \mu_{it}}{\sigma_\varepsilon}\right) \left[\pi_{it}^{\Delta N_{it}} \left(1 - \pi_{it}\right)^{1 - \Delta N_{it}}\right]^{\mathbb{1}(t \geq 3)} \end{aligned} \tag{6}$$

where: $\phi(\cdot)$ is the density function of the $N(0,1)$ distribution; $\mu_{i0} = \boldsymbol{w}_i \boldsymbol{\tau} + \alpha u_i$; $\mu_{it} = \rho C_{it-1} + \boldsymbol{x}_{it} \boldsymbol{\beta} + \boldsymbol{h}_{it-1} \boldsymbol{\gamma} + u_i$, $t = 1, \ldots, T_i$; and $\pi_{it} = [1 + \exp\{-\lambda C_{it-3} - \boldsymbol{x}_{it} \boldsymbol{\delta} - \boldsymbol{f}_{it-1} \boldsymbol{\psi} - \eta u_i\}]^{-1}$, $t = 3, \ldots, T_i$.

To implement the Bayesian approach, we need to specify prior distributions for the parameters. Treating the individual-random effects $\{u_i\}_{i=1}^{N}$ as parameters, the full parameter vector is $(\boldsymbol{u}, \boldsymbol{\theta}) = (\{u_i\}_{i=1}^{N}, \rho, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_\varepsilon, \lambda, \boldsymbol{\delta}, \boldsymbol{\psi}, \eta, \boldsymbol{\tau}, \alpha, \sigma_\zeta)$, which takes values in a high-dimensional space. This suggests using a hierarchical prior for the individual-random effects, such as $u_i | \mu_u, \sigma_u \sim i.i.d. N(\mu_u, \sigma_u^2)$ with a quite diffuse prior distribution for μ_u, σ_u. In our study we set $\mu_u = 0$ specifying a prior distribution only for the hyper-parameter σ_u^2 (see Gelman et al., 2004; Koop, 2003; and Lancaster, 2004). Specifically, the priors for σ_u^2, σ_ε^2, and σ_ζ^2 are Inverse-$\chi^2(1, 10)$. The prior for ρ is $N(1, 1/4)$, a choice which is mainly driven by the theoretical reasons in favour of random walk-like behaviour in consumption (Hall, 1978). The prior distributions on each element of α, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\tau}$ are $N(0, 20)$. Finally, the priors for λ, $\boldsymbol{\delta}$, $\boldsymbol{\psi}$, and η are (improper) uniform. All these parameters are assumed to be independent in the prior.

3.2 Model Checking

The posterior distributions were simulated using an MCMC algorithm (see Appendix 1), based on a single chain, which was run for 1800 iterations after a burn-in stage of 300 iterations[3]. To initialize the chain, any missing data on consumption was set equal to the previous observed value. Values of consumption for earlier quarters than the 1993 wave were initially set equal to 1993 consumption. Using the initialized consumption values, we chose initial values for the model parameters. Those for the two consumption equations were initialized to values of the linear regression model coefficients and the initial values for the fertility process were set to values drawn from a single run of the Sampling Importance Resampling (SIR) algorithm of Rubin (1988) (see Appendix 1).

We assessed the goodness of fit of the model by comparing the observed values of log consumptions with corresponding simulated values from the model. We simulated log consumption in two ways: by sampling from its conditional posterior distribution at the posterior mean of the parameters; and by computing the mean of the posterior predictive distribution of log consumption. We then selected the sub-sample of log consumption values corresponding to the observed data points. Means, standard deviations, and the four quintile points of

[3]Given the computational effort, these results are preliminary and will be updated once results from longer chains are available.

the distribution of the simulated and observed log consumption per head in each wave are reported in Table 1. The simulated data replicate the observed consumption data very well.

Table 1 Means, standard deviations and quintile points of the distribution of simulated and observed log consumption per head

	Log consumption draws from ...								
	Conditional posterior distribution at posterior mean of parameters			Posterior predictive distribution			Observed consumption distribution		
Wave	1993	1997	2000	1993	1997	2000	1993	1997	2000
Mean	5.861	6.082	6.203	5.860	6.086	6.244	5.916	6.105	6.204
Sd	0.663	0.652	0.671	0.795	0.743	0.814	0.778	0.643	0.656
Quintiles									
20%	5.348	5.576	5.721	5.230	5.503	5.624	5.280	5.599	5.690
40%	5.691	5.900	6.024	5.615	5.896	6.035	5.727	5.901	5.988
60%	5.954	6.206	6.348	5.983	6.239	6.432	6.091	6.239	6.311
80%	6.334	6.547	6.701	6.488	6.746	6.861	6.511	6.573	6.731

4 Empirical analysis

The model specification relates fertility and consumption-based living standards to the following set of factors, which are assumed to be strictly exogenous: area type, season, age of the woman and religion. The variables describing the past history of both processes are lagged values of real log consumption expenditure, household size and the number of children. The number of other household members, excluding children, is assumed strictly exogenous. The sample distribution of the covariates is summarised in Appendix Tables A1 and A2.

4.1 Estimates

The means, standard deviations and quantiles of the posterior distribution of the model parameters are given in Table 2.The two consumption equations are linear regression models and consequently the instantaneous marginal effects of the covariates and longer-term dynamic properties of living standards (for a given fertility realisation) can be read off directly from their coefficients.

Household size has a positive but very small impact on (unequivalised) consumption, whereas the effect of the number of children is large and negative. As a result, the comparative statics effect on equivalised consumption of an increased number of children is large and negative (corresponding to an impact elasticity of around -0.548 at the posterior mean and a long-run elasticity of approximately -0.572).[4] Thus, ignoring dynamic feedbacks between the consumption and fertility processes, there is evidence that fertility has a substantial negative effect on access to consumption resources of household members.

The dynamic structure of log consumption is far from the random walk suggested by some versions of the Hall (1978) consumption model. Despite use of a (mild) prior centered on unity, the posterior distribution for the autoregressive parameter ρ has a mean of only 0.329. If we accept the principle of Hall-type intertemporal consumption-smoothing, this suggests that Indonesian families face credit markets that are far from the perfect markets underlying the permanent income hypothesis. We do not find this a surprising result. Seasonality of consumption is estimated to be very important, with consumption being low in the first half of the year, especially so in quarter 1. Again, this suggests the existence of constraints on the ability of households to smooth consumption streams.

The family random effect plays a significant role and the estimates of σ_ε and σ_u imply an intra-class correlation of 0.466 at the posterior mean. The negative coefficient of u_i in the initial condition for consumption suggests that families with unobserved characteristics that predisposed them to poverty initially, nevertheless have tended to succeed in raising their living standards subsequently.

[4] These elasticities are calculated as $(a+b)K - K/N$ (impact) and $(a+b)K/(1-\rho) - K/N$ (long-run), where a and b are the coefficients of household size and $K = 2$ and $N = 4$ are the baseline levels for the numbers of children and the household size.

Table 2 The posterior distribution of model parameters: Individual Random effect and Birth Process

Variable	Mean	sd	2.5%	25%	50%	75%	97.5%
Random Effect							
Standard deviation (σ_u)	0.466	0.014	0.440	0.456	0.466	0.475	0.493
Birth Process							
Intercept	-2.597	0.362	-3.359	-2.848	-2.570	-2.338	-1.962
Calendar Time (Wave 2000)							
Before wave 1997	0.288	0.241	-0.180	0.133	0.279	0.440	0.777
Between wave 1997 and wave 2000	-0.085	0.251	-0.578	-0.253	-0.093	0.081	0.415
Province (Yogyakarta)							
North Sumatra	0.562	0.178	0.210	0.443	0.555	0.679	0.915
West Sumatra	0.859	0.205	0.477	0.716	0.853	0.993	1.281
South Sumatra	0.345	0.181	0.002	0.225	0.347	0.461	0.719
Lampung	0.916	0.265	0.440	0.732	0.891	1.086	1.487
Jakarta	0.093	0.211	-0.297	-0.054	0.083	0.223	0.554
West Java	0.084	0.158	-0.227	-0.027	0.086	0.192	0.395
Central Java	0.155	0.190	-0.204	0.021	0.144	0.283	0.534
East Java	-0.197	0.176	-0.541	-0.318	-0.208	-0.080	0.166
Bali	0.187	0.197	-0.202	0.055	0.188	0.315	0.575
West Nusa Tenggara	0.525	0.181	0.188	0.401	0.524	0.641	0.888
South Kalimantan	0.112	0.203	-0.287	-0.021	0.117	0.248	0.502
South Sulawesi	0.467	0.188	0.100	0.350	0.468	0.594	0.834
Area (Rural)							
Urban	0.205	0.126	-0.034	0.118	0.205	0.290	0.446
Season at quarter t (October – December)							
January – March	-0.053	0.105	-0.257	-0.124	-0.052	0.017	0.155
April – June	-0.168	0.106	-0.374	-0.237	-0.170	-0.098	0.042
July – September	-0.038	0.106	-0.241	-0.112	-0.037	0.032	0.169
Interaction between Area and Season							
Urban and January – March	0.151	0.145	-0.133	0.053	0.152	0.248	0.440
Urban and April – June	0.251	0.148	-0.040	0.152	0.254	0.347	0.530
Urban and July – September	0.015	0.147	-0.279	-0.082	0.011	0.114	0.305
Access to safe drinking water (Other)							
PAM/PUMP water	0.099	0.067	-0.031	0.051	0.100	0.145	0.231
In the village there are ...							
Systems of sewage channels/gutters	0.085	0.067	-0.042	0.041	0.086	0.132	0.211
Factories or cottage industries	-0.105	0.071	-0.242	-0.154	-0.107	-0.058	0.033
Medicinal posts or midwives	0.034	0.063	-0.088	-0.009	0.034	0.075	0.162
The farm business is owned by the Household	0.149	0.072	0.012	0.099	0.150	0.196	0.294
Household size at quarter $t-1$	0.045	0.018	0.009	0.034	0.045	0.057	0.077
Age at quarter t	-0.012	0.008	-0.029	-0.017	-0.012	-0.006	0.006
Muslim	-0.019	0.121	-0.254	-0.099	-0.017	0.058	0.222
No of children at quarter $t-1$	-0.231	0.034	-0.298	-0.252	-0.231	-0.211	-0.163
Log consumption at quarter $t-3$	0.158	0.024	0.115	0.141	0.156	0.173	0.208
Interaction between Log consumption $(t-3)$ and presence of children at quarter $t-1$	-0.257	0.019	-0.294	-0.270	-0.256	-0.242	-0.223
Individual Random effect	0.669	0.135	0.464	0.571	0.646	0.753	0.969

Table 2 (continued) The posterior distribution of model parameters: Consumption Process

Variable	Mean	sd	2.5%	25%	50%	75%	97.5%
Intercept	3.202	0.776	1.641	2.662	3.373	3.909	4.144
Calendar Time (Wave 2000)							
Before wave 1997	-0.097	0.066	-0.206	-0.153	-0.105	-0.036	0.012
Between wave 1997 and wave 2000	0.021	0.044	-0.066	-0.010	0.022	0.052	0.105
Province (Yogyakarta)							
North Sumatra	0.388	0.079	0.231	0.334	0.389	0.442	0.543
West Sumatra	0.691	0.093	0.527	0.622	0.688	0.753	0.875
South Sumatra	0.337	0.077	0.167	0.282	0.336	0.396	0.467
Lampung	0.905	0.101	0.772	0.832	0.881	0.954	1.156
Jakarta	0.710	0.090	0.519	0.665	0.709	0.758	0.885
West Java	0.338	0.074	0.177	0.294	0.343	0.395	0.455
Central Java	0.327	0.102	0.176	0.258	0.291	0.406	0.530
East Java	0.136	0.119	-0.048	0.035	0.131	0.208	0.386
Bali	0.048	0.095	-0.161	-0.007	0.043	0.111	0.229
West Nusa Tenggara	0.382	0.084	0.238	0.320	0.382	0.432	0.559
South Kalimantan	0.541	0.079	0.414	0.480	0.532	0.595	0.705
South Sulawesi	0.444	0.072	0.308	0.395	0.437	0.498	0.590
Area (Rural)							
Urban	0.369	0.043	0.280	0.339	0.370	0.399	0.448
Season at quarter t (October − December)							
January − March	-0.274	0.141	-0.577	-0.360	-0.264	-0.180	0.004
April − June	-0.067	0.098	-0.235	-0.142	-0.066	-0.001	0.124
July − September	0.106	0.101	-0.031	0.032	0.076	0.161	0.347
Interaction between Area and Season							
Urban and January − March	-0.057	0.162	-0.349	-0.213	-0.015	0.080	0.187
Urban and April − June	0.110	0.182	-0.128	-0.038	0.056	0.257	0.476
Urban and July − September	-0.117	0.088	-0.322	-0.163	-0.101	-0.056	0.014
Access to safe drinking water (Other)							
PAM/PUMP water	0.159	0.027	0.109	0.140	0.159	0.178	0.211
In the village there are ...							
Systems of sewage channels/gutters	0.043	0.021	0.001	0.029	0.044	0.058	0.082
Factories or cottage industries	0.163	0.029	0.104	0.143	0.164	0.186	0.215
Medicinal posts or midwives	0.040	0.021	-0.004	0.027	0.041	0.055	0.079
The farm business is owned by the Household	0.061	0.041	-0.015	0.035	0.052	0.086	0.158
Household size at quarter $t-1$	0.080	0.012	0.055	0.072	0.079	0.091	0.104
Age at quarter t	0.038	0.004	0.029	0.035	0.038	0.041	0.045
Muslim	-0.220	0.042	-0.301	-0.250	-0.219	-0.188	-0.145
No of children at quarter $t-1$	-0.104	0.009	-0.124	-0.110	-0.104	-0.098	-0.086
Log consumption at quarter $t-1$ (ρ)	0.329	0.086	0.212	0.253	0.318	0.393	0.501
Standard deviation (σ_ε)	0.443	0.009	0.430	0.436	0.442	0.448	0.462

Table 2 (continued) The posterior distribution of model parameters: Initial Consumption

Variable	Mean	sd	2.5%	25%	50%	75%	97.5%
Intercept	4.042	2.546	-1.931	2.359	4.169	5.976	8.346
Cohort (Before 1979)							
Between 1980 – 1984	3.724	1.828	-0.491	2.646	3.863	5.045	6.678
Between 1985 – 1989	5.003	1.761	0.682	4.093	5.024	6.177	8.243
Between 1990 – 1993	2.503	1.829	-1.496	1.442	2.587	3.727	5.601
Province (Yogyakarta)							
North Sumatra	-0.197	1.665	-2.917	-1.635	-0.262	1.108	3.061
West Sumatra	2.346	2.306	-2.832	0.845	2.847	3.918	6.255
South Sumatra	0.856	1.324	-2.000	0.047	1.027	1.726	3.432
Lampung	1.315	2.117	-2.007	-0.463	0.858	3.150	5.195
Jakarta	3.173	1.807	0.800	1.775	2.880	4.057	7.856
West Java	1.745	1.251	-0.736	0.868	1.841	2.643	4.067
Central Java	3.643	1.548	1.164	2.452	3.335	4.829	6.782
East Java	3.718	1.571	0.566	2.705	3.421	4.777	7.158
Bali	1.722	1.895	-1.716	0.490	1.387	2.627	6.040
West Nusa Tenggara	2.098	2.201	-3.108	1.274	2.557	3.571	5.337
South Kalimantan	1.909	1.449	-1.234	0.919	2.112	3.075	4.123
South Sulawesi	4.779	1.493	1.770	3.877	5.007	5.773	7.509
Area (Rural)							
Urban	1.091	1.193	-1.471	0.365	1.034	1.954	3.358
Season (October – December)							
January – March	2.649	1.012	1.064	1.897	2.482	3.276	4.977
April – June	0.443	1.278	-1.627	-0.552	0.367	1.190	3.650
July – September	-1.996	0.606	-3.169	-2.431	-2.017	-1.562	-0.859
Interaction between area and season							
Urban and January – March	1.335	2.771	-3.539	-0.962	1.811	3.679	5.405
Urban and April – June	-1.617	1.216	-4.227	-2.365	-1.691	-0.854	0.872
Urban and July – September	-0.889	2.108	-5.054	-2.551	-0.930	0.889	2.651
Access to safe drinking water (Other)							
PAM/PUMP water	-0.128	0.732	-1.296	-0.672	-0.147	0.245	1.395
In the village there are ...							
Systems of sewage channels/gutters	-0.527	1.103	-2.716	-1.110	-0.278	0.163	1.046
Factories or cottage industries	0.459	0.875	-1.776	0.102	0.614	1.142	1.555
Medicinal posts or midwives	0.843	0.954	-0.770	-0.003	0.784	1.675	2.431
The farm business is owned by the Household	-0.739	2.091	-4.305	-2.516	-0.648	1.174	2.418
Household size	0.075	0.159	-0.253	-0.058	0.091	0.209	0.319
Age	0.130	0.212	-0.138	-0.026	0.036	0.350	0.583
Muslim	-3.150	1.762	-5.827	-4.774	-3.447	-1.601	0.036
Individual random effect	-3.301	0.423	-4.070	-3.580	-3.317	-3.069	-2.291
Standard deviation (σ_η)	0.771	0.144	0.567	0.659	0.736	0.872	1.094

4.2 Dynamic effects

If we are interested in birth probabilities and the incidence of consumption poverty rather than the expected level of log consumption, it is not possible to give a meaningful interpretation of the parameters directly. Instead, we calculate summary measures of the probabilities of birth and poverty, using both short-run and long-run summaries.

4.2.1 Short-run effects

Define a short-run static summary measure $S(z^*, \boldsymbol{\theta}, u^*)$ where z^* is a set of reference values for the predetermined variables at some point in time and u^* is a fixed value for the individual effect u_i. If the posterior distribution of the model parameters is $P(\boldsymbol{\theta})$, then the posterior mean of the measure S is $\int S(z^*, \boldsymbol{\theta}, u^*) dP(\boldsymbol{\theta})$. This is estimated from the output of the MCMC algorithm as:

$$\hat{S}(z^*, \boldsymbol{\theta}, u^*) = R^{-1} \sum_{r=1}^{R} S(z^*, \boldsymbol{\theta}^{(r)}, u^*) \tag{7}$$

where $\boldsymbol{\theta}^{(r)}$ is the rth MCMC draw from $P(\boldsymbol{\theta})$.

For fertility, the probability of a birth during period t, with fixed levels of past consumption $C_{t-3} = C^*$, fertility history $\boldsymbol{f}_{t-1} = \boldsymbol{f}^*$, covariates $\boldsymbol{x}_t = \boldsymbol{x}^*$ and individual effect $u = u^*$ is given by the logistic function $\Lambda(\lambda C^* + \boldsymbol{x}^* \boldsymbol{\delta} + \boldsymbol{f}^* \boldsymbol{\psi} + \eta u^*)$. We choose the values $(C^*, \boldsymbol{f}^*, \boldsymbol{x}^*, u^*)$ to correspond to a reference family which is urban, Muslim, has lagged consumption at the sample mean level for 1997, comprises a 4-person household with 2 children and has the mean value 0 for the individual effect. The quarterly birth probability for this reference case is 3.9% and we calculate marginal influences by varying the family characteristics in the various ways set out in Table 3 below. For the poverty probability, we adopt an arbitrary poverty threshold of 250Bt per head. Given this poverty line, we construct the corresponding poverty threshold, Q, for unequivalised log consumption. Then the short-run poverty probability is $\Phi\left(\left[Q - \rho C^{*} - \boldsymbol{x}^{*}\boldsymbol{\beta} - \boldsymbol{h}^{*}\boldsymbol{\gamma} - u^{*}\right]/\sigma_{\varepsilon}\right)$. The poverty rate for the baseline family is 4.2% and the effects of variations in family characteristics are given in Table 3.

In terms of these short-run impact effects, the dominant influence on fertility is the existing stock of children. As can be expected, compared to a baseline 2-child family, childless families are predicted to have a much higher fertility rate: raised by almost 26 percentage points (p.p.). An additional two children yields a reduction of more than 1 p.p. Religious affiliation has a negligible impact on fertility rate and there is a small negative effect of the woman's age, with older women having a birth probability lower by almost 0.2 p.p. Consumption has a negative effect on the birth probability, but a huge consumption shock would be required to generate a fertility variation as large as those associated with individual heterogeneity or the stock of children.[5].

The unobserved individual effect, u, has an asymmetric impact with a fall of 1 standard deviation cutting the birth rate by 1 p.p. (at the posterior mean) and a 1 standard deviation rise increasing the birth rate by 1.4 p.p. The individual effect is also the primary influence on the family-specific poverty rate: a variation of $-\sigma_u$ in u_i raises the poverty rate for this representative individual by 16.3 p.p. and a rise of σ_u reduces the poverty risk by 3.8 p.p. These results suggest that persistent, unobservable factors like personal characteristics, local social norms, productivity, access to assets, etc. are responsible for much of the empirical association between poverty and fertility. However, fertility has a substantial influence on the poverty rate. If the stock of children is reduced from two to zero, the posterior mean poverty rate for the representative family falls by 4.1 p.p. and an additional two children raise it by 16.3 p.p.

[5] Note that consumption has a positive effect on the birth rate for a childless reference family comprising a 2-person household, which has a quarterly birth probability equal to 29.6% (see the second panel in Table 3).

The predicted effect of rural, rather than urban, residence is to raise the poverty rate by 6.6 p.p. This effect is slightly lower and equal to 4 p.p. whether the rural family owns the farm business. Age effects are substantial, with the poverty risk declining by 2.5 p.p. for an additional five years of age and rising by 4.8 p.p. for a correspondingly younger family. Muslim families are predicted to have a poverty rate higher by 2.5 p.p. than a similar non-Muslim family. There are again very large seasonal variations. The influence of lagged consumption on poverty is substantial, reflecting the consumption persistence in the dynamic regression (1).

4.2.2 Long-run effects

Our long-run dynamic measures are defined in general terms as the expectation of some functional of the sequence $(\mathbb{N}, \mathbb{C}) = (N_3, \ldots, N_T, C_0, \ldots, C_T)$, generated conditional on a fixed value u^*, initial environment $\boldsymbol{w}^*$, covariate sequence $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_T^*$ and parameters $\boldsymbol{\theta}$. Call this functional $\mathcal{F}(\mathbb{N}, \mathbb{C}; \boldsymbol{\theta})$ and then the posterior mean of the dynamic measure is $\int E\left(\mathcal{F}(\mathbb{N}, \mathbb{C}); \boldsymbol{\theta})\right) dP(\boldsymbol{\theta})$, which can again be estimated from the MCMC replications. This class of dynamic measures includes impulse-response functions, long-run steady-state equilibrium effects and dynamic multipliers. Because of the complex mixed discrete-continuous nature of the fertility-consumption process, the most convenient way of computing the expectation $E\left(\mathcal{F}(\mathbb{N}, \mathbb{C}; \boldsymbol{\theta})\right)$ is by dynamic stochastic simulation, so that two nested simulation loops are required.

We implement this approach for a simulation period of 10 years plus the initial quarter. This simulation is made for a hypothetical baseline family, formed in an urban area in 1993, as a Muslim married couple, with the wife aged 18. From simulations of the fertility-consumption process, we calculate five summary measures: the number of years spent in (absolute) poverty; the number of quarter-to-quarter transitions into or out of poverty; the

Table 3 The posterior distribution of marginal effects on short-run birth and poverty probabilities

	Birth probability		Poverty probability	
	Mean	Standard deviation	Mean	Standard deviation
Baseline probability	0.0394	0.0059	0.0420	0.0495
Calendar Time				
Between wave 1997 and wave 2000	-0.0119	0.0034	-0.0125	0.0104
Wave 2000	-0.0087	0.0073	-0.0065	0.0068
Province				
Yogyakarta	-0.0033	0.0077	0.3193	0.1821
North Sumatra	0.0220	0.0074	0.0944	0.0773
West Sumatra	0.0416	0.0101	0.0015	0.0103
South Sumatra	0.0106	0.0065	0.1190	0.0929
Lampung	0.0468	0.0154	-0.0269	0.0320
West Java	-0.0005	0.0058	0.1282	0.1122
Central Java	0.0024	0.0055	0.1082	0.0640
East Java	-0.0097	0.0051	0.2249	0.1165
Bali	0.0039	0.0081	0.2915	0.1665
West Nusa Tenggara	0.0199	0.0076	0.0905	0.0689
South Kalimantan	0.0009	0.0078	0.0582	0.0753
South Sulawesi	0.0168	0.0083	0.0914	0.1096
Rural	-0.0115	0.0043	0.0660	0.0287
Rural (The farm business is owned)	-0.0071	0.0044	0.0402	0.0302
Access to safe drinking water				
No PAM/PUMP water	-0.0036	0.0025	0.0359	0.0329
In the village there are not ...				
Systems of sewage channels/gutters	-0.0031	0.0025	0.0094	0.0116
Factories or cottage industries	0.0043	0.0030	0.0352	0.0277
Medicinal posts or midwives	-0.0014	0.0024	0.0089	0.0101
Season				
April - June	-0.0006	0.0038	-0.0382	0.0518
July - September	-0.0043	0.0038	-0.0385	0.0502
October - December	-0.0035	0.0039	-0.0389	0.0497
The farm business is owned	0.0062	0.0035	-0.0140	0.0196
2 fewer Children	0.2568	0.0506	-0.0415	0.0487
1 less Child	0.0077	0.0021	-0.0332	0.0373
1 extra Child	-0.0065	0.0016	0.0677	0.0599
2 extra Children	-0.0119	0.0029	0.1633	0.1226
Age − 5	0.0022	0.0016	0.0476	0.0430
Age + 5	-0.0021	0.0015	-0.0250	0.0270
No Muslim	0.0010	0.0047	-0.0247	0.0253
Consumption − 20%	0.0009	0.0003	0.0183	0.0206
Consumption − 10%	0.0004	0.0002	0.0080	0.0092
Consumption + 10%	-0.0004	0.0001	-0.0064	0.0074
Consumption + 20%	-0.0007	0.0003	-0.0116	0.0135
$u_i - \sigma_u$	-0.0103	0.0029	0.1629	0.1075
$u_i + \sigma_u$	0.0139	0.0048	-0.0379	0.0433

Table 3 (continued) The posterior distribution of marginal effects on short-run birth and poverty probabilities for a childless family

	Birth probability		Poverty probability	
	Mean	Standard deviation	Mean	Standard deviation
Baseline probability	0.2962	0.0555	0.0005	0.0009
Consumption − 20%	-0.0072	0.0013	0.0005	0.0009
Consumption − 10%	-0.0034	0.0006	0.0002	0.0004
Consumption + 10%	0.0031	0.0006	-0.0001	0.0002
Consumption + 20%	0.0060	0.0011	-0.0002	0.0004

number of children born; the probability of four or more children being born; and the number of child-years of poverty. The posterior means of these long-run outcomes are calculated for the baseline family and marginal effects are calculated by varying the baseline characteristics and repeating the simulation. The results will be available at the conference.

5 Concluding remarks

In this paper, we have demonstrated the feasibility of a new approach to modelling the inter-relationship between fertility and poverty. This approach is based on a simultaneous dynamic model of births and a log-linear model of consumption equation. The implications of the model for poverty dynamics can then be deduced from the parameters of the model after estimation. The model is specified in discrete time, using a quarterly chronology so that the human 9-month gestation period can be accommodated. Although the fertility process is fully observed, consumption can only be observed at interview dates, so that most potential consumption observations are missing. A Bayesian approach is used to deal with this large-scale missing data problem. We also examine the short-run and longer-run dynamic impacts of family circumstances on fertility and poverty probabilities.

Our findings are striking. High fertility is strongly related to poverty but there is little evidence of any negative feedback of improved living standards on fertility. Thus our results do not support the idea that increasing economic welfare is, in itself, enough to generate a demographic transition towards a low-fertility outcome. However, there is evidence of a large persistent unobservable effect specific to each individual family which jointly promotes high fertility and low living standards. Such unobservables are generally taken to represent factors such as cultural norms, beliefs, skills and access to various assets and credit markets. Together, these two findings suggest that movement out of the high-fertility, high-poverty 'trap' requires more fundamental changes than would be brought by simple income growth.

References

[1] Abbring, J. H. and G. J. van den Berg (2003a). The Nonparametric Identification of Treatment Effects in Duration Models, *Econometrica*, **71**, 1491-1517.

[2] Abbring, J. H. and G. J. van den Berg (2003b). Analyzing the Effect of Dynamically Assigned Treatments Using Duration Models, Binary Treatment Models, and Panel Data Models, *IZA Discussion Paper*, No. **831**.

[3] Aassve, A., A. M. Kedir, and H. T. Woldegebrie (2006). State Dependence and Causal Feedback of Poverty and Fertility in Ethiopia *Working Paper*, No. **30**, Institute for Social and Economic Research, Colchester, University of Essex.

[4] Athey, S. and G.W. Imbens (2004) Discrete chioce models with multiple unobserved choice characteristics, *International Economic Review*, **48**, **4**, 1159 – 1192.

[5] Baccini M., F. Mealli, C. Frangakis, F. Li, Y. Yu, S.R. Cook, and D.B. Rubin (2004). Multiple Imputation in the Anthrax Vaccine Research Program, *Chance* forthcoming.

[6] Becker, G. (1960). An economic analysis of fertility, in G. S. Becker (ed.) *Demographic and Economic Change in Developed Countries*. Princeton: Princeton University Press.

[7] Becker, G. and Lewis, H. G. (1973). Interaction between quantity and quality of children, *Journal of Political Economy*, **81**, S279-288.

[8] Birdsall, N. M. and C. C. Griffin (1988). Fertility and Poverty in Developing Countries, *Journal of Policy Modeling*, **10**, 29-55.

[9] Chamberlain, G. and K. Hirano (1999). Predictive Distributions based on Longitudinal Earnings Data, *Annales d'Économie et de Statistique*, **55-56**, 211 – 243.

[10] Chib, S. (2003). Markov Chain Monte Carlo Methods: Computation and Inference, in J. Heckman and E. Leamer, eds., *Handbook of Econometrics*, North Holland, Amsterdam.

[11] Deaton, A. and Zaidi, S. (2002). Guidelines for Constructing Consumption Aggregates for Welfare Analysis, Living Standards Measurement Study Working Paper No. 135, The World Bank.

[12] Gelman, A., J. Carlin, H. Stern anf D. Rubin (2004) *Bayesian Data Analysis*, Chapman and Hall, London.

[13] Geman, S. and D. Geman (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, 609 – 628.

[14] Gelfand, A. and A. Smith (1990). Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, **85**, 398 – 409.

[15] Geweke, J. (1997). Posterior Simulators in Econometrics in D. Kreps and K. Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications*, Cambridge University Press, Cambridge (UK).

[16] Gilks, W., S. Richardson and D. Spiegelhalter (1996) *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London.

[17] Hall, R.E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence, *Journal of Political Economy*, **86**, 971-987.

[18] Hasting, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika***57**, 97 – 109.

[19] Hirano, K. (2002). Semiparametric Bayesian Inference in Autoregressive Panel Data Models *Econometrica*, **70**, 781 – 799.

[20] Kim, J., H. Engelhardt, A. Prskawetz and A. Aassve. (2005). Does Fertility Decrease the Welfare of Households? An Analysis of Poverty Dynamics and Fertility in Indonesia, *Working Paper*, No. **6**, Vienna Institute of Demography.

[21] Kim, J., and A. Prskawetz (2006). External Shocks, Household Consumption and Fertility in Indonesia, *Working Paper*, No. **4**, Vienna Institute of Demography.

[22] Koop, G. (2003). *Bayesian Econometrics*, Wiley, England.

[23] Lancaster, A. (1990). *The Econometric Analysis of Transition Data*, Econometric Society Monograph.

[24] Lancaster, A. (2004). *An Introduction to Modern Bayesian Econometrics*, Blackwell Publishing Ltd.

[25] Lanjouw, P. and M. Ravallion, (1995). Poverty and Household Size, *Economic Journal*, **105**, 1415 - 1434.

[26] Little, R.J.A. and D.B. Rubin (1987). *Statistical Analysing with Missing Data*, New York, Wiley

[27] Livi-Bacci, M. and G. de Santis (1998). *Population and Poverty in the Developing World*, Clarendon Press, Oxford.

[28] Mattei, A (2007). Estimating and Using Propensity score in Presence of Missing Background Data. An Application to Assess the Impact of Childbearing on Wellbeing, *Statistical Methods & Applications*, forthcoming.

[29] McNicoll, G. (1997). Population and poverty: A review and restatement, *Policy Research Division Working Paper*, No. **105**. New York: Population Council.

[30] Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, **21**, 1087 – 1091.

[31] Moav, O. (2005) Cheap Children and the Persistence of Poverty. *Economic Journal* **115**, 88-110.

[32] Rubin, B.D. (1976), Inference and missing data, *Biometrika*, **63**, 581 – 592.

[33] Rubin, B.D. (1988) Using the SIR algorithm to simulate posterior distributions, in J.M. Bernardo, M.H. DeGroot, D.V. Lindley, A.F.M. Smith, eds., *Bayesian statistics*, **3**, Oxford University Press, Oxford, 395 – 402.

[34] Tanner, M. and W. Wong, (1987). The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association* **82**, 52840.

Appendix 1: The MCMC algorithm

We construct a general state space Markov chain with unique invariant equilibrium distribution given by the joint distribution of the model parameters $\boldsymbol{\theta}$, the hyper-parameter σ_u, the individual-random effects $\{u_i\}_{i=1}^N$, and the missing consumption in $\{C_{i0}\}_{i=1}^N$ and $\{C_{it}\}_{t=1}^{T_i}$, $i = 1, \ldots, N$. The Markov chain algorithm uses Gibbs Sampling (Geman and Geman, 1984; and Gelfand and Smith, 1990) and Metropolis–Hastings algorithm (Metropolis et al., 1953; Hastings, 1970) with parameter and data augmentation method of Tanner and Wong (1987). Our approach relies critically on viewing the missing consumption as well as the individual specific parameters as unobserved random variables to be imputed given the observed variables. This appendix sketches the algorithm.

Let $\left(\{C_{i0}^j\}_{i=1}^N; \{C_{it}^j\}_{t=1}^{T_i}, i = 1, \ldots, N; \sigma_u^{2j}; \boldsymbol{u}^j; \boldsymbol{\theta}^j\right)$ denote the state of the chain at time j. The state of the chain at time $j+1$ follows from applying the following steps. We first draw missing consumption in $\{C_{i0}\}_{i=1}^N$ and $\{C_{it}\}_{t=1}^{T_i}$, $i = 1, \ldots, N$. We cannot draw directly from the appropriate conditional distributions; however, we can easily calculate the (complete-data) posterior density up to a normalizing constant at any consumption value, so we can use Metropolis-Hastings steps. For $i = 1, \ldots, N$, let

$$
\begin{aligned}
\mu_{i0} &= \boldsymbol{w}_i \boldsymbol{\tau} + \alpha u_i \\
\mu_{it} &= \rho C_{it-1} + \boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{h}_{it-1}\boldsymbol{\gamma} + u_i, \quad t = 1, \ldots, T_i \\
\pi_{it} &= \frac{\exp(\lambda C_{it-3} + \boldsymbol{x}_{it}\boldsymbol{\delta} + \boldsymbol{f}_{it-1}\boldsymbol{\psi} + \eta u_i)}{1 + \exp(\lambda C_{it-3} + \boldsymbol{x}_{it}\boldsymbol{\delta} + \boldsymbol{f}_{it-1}\boldsymbol{\psi} + \eta u_i)}, \quad t = 3, \ldots, T_i.
\end{aligned}
$$

Then, the (complete-data) posterior density of C_{i0}, consumption in period 0 for woman i, is

$$
\begin{aligned}
& p\left(C_{i0} | \boldsymbol{W}_i, \{C_{it}\}_{t=1}^{T_i}; \sigma_u, u_i, \boldsymbol{\theta}\right) \propto \\
& \quad \frac{1}{\sigma_\zeta}\phi\left(\frac{C_{i0} - \mu_{i0}}{\sigma_\zeta}\right)\left[\frac{1}{\sigma_\varepsilon}\phi\left(\frac{C_{i1} - \mu_{i1}}{\sigma_\varepsilon}\right)\right]^{\mathbb{1}(T_i \geq 1)}\left[\pi_{i3}^{\Delta N_{i3}}\left(1 - \pi_{i3}\right)^{1 - \Delta N_{i3}}\right]^{\mathbb{1}(T_i \geq 3)}.
\end{aligned} \tag{8}
$$

To draw C_{i0}, we sample a candidate value C_{i0}^{cand} from a normal jumping density with parameter values given by

$$\mu_{C_{i0}} = \sigma_{C_0}^2 \left(\frac{\boldsymbol{w}_i \boldsymbol{\tau} + \alpha u_i}{\sigma_\zeta^2} + \rho \frac{C_{t+1} - (\boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{h}_{it-1}\boldsymbol{\gamma} + u_i)}{\sigma_\varepsilon^2} \right) \qquad \text{and} \qquad \sigma_{C_0}^2 = \frac{1}{\frac{1}{\sigma_\zeta^2} + \frac{\rho^2}{\sigma_\varepsilon^2}},$$

which is the probability density function resulting from the product of the first two terms on the right hand of equation (8). The candidate draw is accepted with probability

$$\min \left\{ \frac{p\left(C_{i0}^{\text{Cand}} | \boldsymbol{W}_i, \{C_{it}^j\}_{t=1}^{T_i}; \sigma_u^{2j}, u_i^j, \boldsymbol{\theta}^j\right)}{p\left(C_{i0}^j | \boldsymbol{W}_i, \{C_{it}^j\}_{t=1}^{T_i}; \sigma_u^{2j}, u_i^j, \boldsymbol{\theta}^j\right)} \cdot \frac{\phi\left(\frac{C_{i0}^j - \mu_{C_{i0}}}{\sigma_{C_0}}\right)}{\phi\left(\frac{C_{i0}^{\text{cand}} - \mu_{C_{i0}}}{\sigma_{C_0}}\right)}, 1 \right\}.$$

The drawing of consumption in period $t > 0$ is done in the analogous way. The (complete-data) posterior density of C_{it}, consumption in period t for woman i, is

$$p\left(C_{it} | \boldsymbol{W}_i, C_{i0}, \{C_{is}\}_{s=1, s\neq t}^{T_i}; \sigma_u^2, u_i, \boldsymbol{\theta}\right) \propto \tag{9}$$
$$\frac{1}{\sigma_\varepsilon} \phi\left(\frac{C_{it} - \mu_{it}}{\sigma_\varepsilon}\right) \left[\frac{1}{\sigma_\varepsilon} \phi\left(\frac{C_{it+1} - \mu_{it+1}}{\sigma_\varepsilon}\right)\right]^{\mathbb{1}(t \leq T_i - 1)} \left[\pi_{it+3}^{\Delta N_{it+3}} \left(1 - \pi_{it+3}\right)^{1 - \Delta N_{it+3}}\right]^{\mathbb{1}(t \leq T_i - 3)}$$

For $t > T_i - 3$, we can directly draw C_{it} from its conditional posterior distribution using Gibbs sampler. For $t \leq T_i - 3$, we use Metropolis–Hastings steps, drawing a candidate value C_{it}^{cand} from the normal jumping density with parameter values given by

$$\mu_{C_{it}} = \frac{\rho C_{it-1} + \boldsymbol{x}_{it}\boldsymbol{\beta} + \boldsymbol{h}_{it-1}\boldsymbol{\gamma} + u_i + \rho\left[C_{it+1} - (\boldsymbol{x}_{it+1}\boldsymbol{\beta} + \boldsymbol{h}_{it}\boldsymbol{\gamma} + u_i)\right]}{1 + \rho^2} \qquad \text{and} \qquad \sigma_{C_t}^2 = \frac{\sigma_\varepsilon^2}{1 + \rho^2},$$

and accepting the candidate draw with probability

$$\min \left\{ \frac{p\left(C_{it}^{\text{Cand}} | \boldsymbol{W}_i, \{C_{is}^{j+1}\}_{s=1}^{t-1}, \{C_{is}^j\}_{s=t+1}^{T_i}; \sigma_u^{2j}, u_i^j, \boldsymbol{\theta}^j\right)}{p\left(C_{it}^j | \boldsymbol{W}_i, \{C_{is}^{j+1}\}_{s=1}^{t-1}, \{C_{is}^j\}_{s=t+1}^{T_i}; \sigma_u^{2j}, u_i^j, \boldsymbol{\theta}^j\right)} \cdot \frac{\phi\left(\frac{C_{it}^j - \mu_{C_{it}}}{\sigma_{C_t}}\right)}{\phi\left(\frac{C_{it}^{\text{cand}} - \mu_{C_{it}}}{\sigma_{C_t}}\right)}, 1 \right\}.$$

The next step consists of drawing the individual random effects, u_i. We first draw the hierarchical parameter σ_u^2 from its conditional posterior distribution: an Inverse-χ^2 with parameters

$$\bar{n}_u = \underline{n}_u + N \qquad \text{and} \qquad \bar{\sigma}_u^2 = \frac{\underline{n}_u \, \underline{\sigma}_u^2 + \sum_{i=1}^N u_i^2}{\bar{n}_u},$$

where $\underline{n}_u = 1$ and $\underline{\sigma}_u^2 = 10$ are the parameters of the Inverse-χ^2 prior distribution we use for σ_u^2.

Throughout this appendix, we use bars under parameters (e.g., $\underline{n}_u$) to denote parameters of a prior density, and bars over parameters (e.g., $\bar{n}_u$) to denotes parameters of a posterior density (Koop, 2003).

Then, we draw the individual random effects, u_i, by using Metropolis–Hastings steps. As jumping distribution we use a normal density with parameter values given by

$$\mu_{u_i} = \varrho_{u_i}^2 \left[\frac{\alpha^2 \left(C_{i0} - \boldsymbol{w}_i \boldsymbol{\tau} \right)}{\sigma_\zeta^2} + \frac{\sum_{t=1}^{T_i} \left(C_{it} - \left(\rho C_{it-1} + \boldsymbol{x}_{it} \boldsymbol{\beta} + \boldsymbol{h}_{it-1} \boldsymbol{\gamma} \right) \right)}{\sigma_\varepsilon^2} \right]$$

and

$$\varrho_{u_i}^2 = \frac{1}{\sigma_u^2 + \alpha^2 \sigma_\zeta^2 + T_i \sigma_\varepsilon^2}.$$

The conditional posterior density for each u_i is independent of $u_{i'}$ for $i \neq i'$, and is given by

$$p\left(u_i | \boldsymbol{W}_i, C_{i0}, \{C_{it}\}_{t=1}^{T_i}; \sigma_u, \boldsymbol{\theta}\right) \propto \frac{1}{\sigma_u} \phi\left(\frac{u_i}{\sigma_u}\right) \frac{1}{\sigma_\zeta} \phi\left(\frac{C_{i0} - \mu_{i0}}{\sigma_\zeta}\right) \prod_{t=1}^{T_i} \frac{1}{\sigma_\varepsilon} \phi\left(\frac{C_{it} - \mu_{it}}{\sigma_\varepsilon}\right) \left[\pi_{it}^{\Delta N_{it}} \left(1 - \pi_{it}\right)^{1 - \Delta N_{it}}\right]^{\mathbb{1}(t \geq 3)}.$$

Therefore, a candidate draw, u_i^{cand}, is accepted with probability

$$\min \left\{ \frac{p\left(u_i^{\text{Cand}} | \boldsymbol{W}_i, C_{i0}^{j+1}, \{C_{it}^{j+1}\}_{t=1}^{T_i}; \sigma_u^{j+1}, \boldsymbol{\theta}^j\right)}{p\left(u_i^j | \boldsymbol{W}_i, C_{i0}^{j+1}, \{C_{it}^{j+1}\}_{t=1}^{T_i}; \sigma_u^{j+1}, \boldsymbol{\theta}^j\right)} \cdot \frac{\phi\left(\frac{u_i^j - \mu_{u_i}}{\varrho_{u_i}}\right)}{\phi\left(\frac{u_i^{\text{cand}} - \mu_{u_i}}{\varrho_{u_i}}\right)}, 1 \right\}.$$

Finally, we draw for the following subvectors of $\boldsymbol{\theta}$ in sequence, conditional on all others: α, $\boldsymbol{\tau}$, σ_ζ, ρ, $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, σ_ε, $(\lambda, \boldsymbol{\delta}, \boldsymbol{\psi}, \eta)$.

The drawing of the coefficients α, $\boldsymbol{\tau}$, σ_ζ^2, ρ, $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, and σ_ε^2 is from their conditional posterior densities, and uses the Gibbs sampler algorithm. For the regression parameters α, $\boldsymbol{\tau}$, ρ, and $(\boldsymbol{\beta}, \boldsymbol{\gamma})$, we use (multivariate) normal priors densities, which lead to (multivariate) normal posterior densities, after multiplying times the complete-data likelihood function in Section 3.

The posterior distribution of α is normal with parameter values given by

$$\bar{\mu}_\alpha = \bar{\sigma}^2_\alpha \left[\frac{\underline{\mu}_\alpha}{\underline{\sigma}^2_\alpha} + \frac{\sum_{i=1}^N u_i \left(C_{i0} - \boldsymbol{w}_i \boldsymbol{\tau}\right)}{\sigma^2_\zeta} \right] \qquad \text{and} \qquad \bar{\sigma}^2_\alpha = \frac{1}{\frac{1}{\underline{\sigma}^2_\alpha} + \frac{\sum_{i=1}^N u_i}{\sigma^2_\zeta}},$$

where we set the prior parameters $\underline{\mu}_\alpha$ and $\underline{\sigma}^2_\alpha$, to 0 and 20, respectively.

The prior distributions on $\boldsymbol{\tau}$ and $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ are multivariate normals centered at zero, $\underline{\boldsymbol{\mu}}_{\boldsymbol{\tau}} = \mathbf{0}$ and $\underline{\boldsymbol{\mu}}_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \mathbf{0}$. As prior covariance matrixes, $\underline{\boldsymbol{V}}_{\boldsymbol{\tau}}$ and $\underline{\boldsymbol{V}}_{\boldsymbol{\beta},\boldsymbol{\gamma}}$, we use scalar matrixes with diagonal elements equal to 20. The mean vector and the covariance matrix for the posterior multivariate normale distribution of $\boldsymbol{\tau}$ are

$$\bar{\boldsymbol{\mu}}_{\boldsymbol{\tau}} = \bar{\boldsymbol{V}}_{\boldsymbol{\tau}} \left[\underline{\boldsymbol{V}}^{-1}_{\boldsymbol{\tau}} \underline{\boldsymbol{\mu}}_{\boldsymbol{\tau}} + \sigma^{-2}_\zeta \sum_{i=1}^N \left(C_{i0} - \alpha u_i\right) \boldsymbol{w}_i \right] \qquad \text{and} \qquad \bar{\boldsymbol{V}}_{\boldsymbol{\tau}} = \left[\underline{\boldsymbol{V}}^{-1}_{\boldsymbol{\tau}} + \sigma^{-2}_\zeta \sum_{i=1}^N \boldsymbol{w}_i \boldsymbol{w}'_i \right]^{-1},$$

whereas the parameter values of the normal multivariate posterior density of $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ are given by

$$\bar{\boldsymbol{\mu}}_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \bar{\boldsymbol{V}}_{\boldsymbol{\beta},\boldsymbol{\gamma}} \left[\underline{\boldsymbol{V}}^{-1}_{\boldsymbol{\beta},\boldsymbol{\gamma}} \underline{\boldsymbol{\mu}}_{\boldsymbol{\beta},\boldsymbol{\gamma}} + \sigma^{-2}_\varepsilon \sum_{i=1}^N \sum_{t=1}^{T_i} \left(C_{it} - \rho C_{it-1} - u_i\right) \left(\boldsymbol{x}_i, \boldsymbol{h}_{i,t-1}\right) \right]$$

and

$$\bar{\boldsymbol{V}}_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \left[\underline{\boldsymbol{V}}^{-1}_{\boldsymbol{\beta},\boldsymbol{\gamma}} + \sigma^{-2}_\varepsilon \sum_{i=1}^N \sum_{t=1}^{T_i} (\boldsymbol{x}_i, \boldsymbol{h}_{i,t-1})(\boldsymbol{x}_i, \boldsymbol{h}_{i,t-1})' \right]^{-1}.$$

Finally, the mean and the variance of the posterior conditional normal density for the autoregressive parameter ρ are

$$\bar{\mu}_\rho = \bar{\sigma}^2_\rho \left[\frac{\underline{\mu}_\rho}{\underline{\sigma}^2_\rho} + \frac{1}{\sigma^2_\varepsilon} \sum_{i=1}^N \sum_{t=1}^{T_i} C_{it-1} \left(C_{it} - \boldsymbol{x}_i \boldsymbol{\beta} - \boldsymbol{h}_{i,t-1} \boldsymbol{\gamma} - u_i\right) \right] \qquad \text{and} \qquad \bar{\sigma}^2_\rho = \frac{1}{\frac{1}{\underline{\sigma}^2_\rho} + \frac{\sum_{i=1}^N \sum_{t=1}^{T_i} C^2_{it-1}}{\sigma^2_\varepsilon}},$$

where we set the prior mean, $\underline{\mu}_\rho$, at 1, and the prior variance, $\underline{\sigma}^2_\rho$, at 0.25.

For the variance parameters σ^2_ζ and σ^2_ε we use Inverse-χ^2 with $\underline{n}_\zeta = \underline{n}_\varepsilon = 1$ degree of freedom and scale parameter $\underline{\sigma}^2_\zeta = \underline{\sigma}^2_\varepsilon = 10$ to describe their prior distributions. Therefore, the posterior distributions are Inverse-χ^2 with parameter values given by

$$\bar{n}_\zeta = \underline{n}_\zeta + N \qquad \text{and} \qquad \bar{\sigma}^2_\zeta = \frac{\underline{n}_\zeta\, \underline{\sigma}^2_\zeta + \sum_{i=1}^{N} \left(C_{i0} - \boldsymbol{w}_i \boldsymbol{\tau} - \alpha u_i\right)^2}{\bar{n}_\zeta}$$

and

$$\bar{n}_\varepsilon = \underline{n}_\varepsilon + N \qquad \text{and} \qquad \bar{\sigma}^2_\varepsilon = \frac{\underline{n}_\varepsilon\, \underline{\sigma}^2_\varepsilon + \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left(C_{it} - \rho C_{it-1} - \boldsymbol{x}_i \boldsymbol{\tau} - \boldsymbol{h}_{it-1} \boldsymbol{\gamma} - u_i\right)^2}{\bar{n}_\varepsilon},$$

respectively.

It remains to sample from the posterior distributions of the parameters of the birth process, λ, $\boldsymbol{\delta}$, $\boldsymbol{\psi}$, and η, for which we assume (improper) uniform priors. A draw from their joint posterior distribution is approximated by using the Sampling-Importance-Resampling (SIR) algorithm of Rubin (1988), which is based on the importance sampling of the Monte Carlo integration. The samples of the importance sampling are resampled with probabilities equal to the importance weights. Specifically, we first simulate a pool of candidate values as a large number (e.g., 1000) of draws from a normal distribution centered at the maximum likelihood estimate of $(\lambda, \boldsymbol{\delta}, \boldsymbol{\psi}, \eta)$, and with covariance matrix set to the inverse of the observed Fisher information. For each draw, we calculate the importance ratio of the actual posterior density to the approximate normal density. Finally, we sample one of those draws with probability proportional to the importance ratios. Note that, the maximum likelihood estimate of the parameters is included in the pool of candidates, in order to avoid that the final draw is forced to be extreme merely because there is no candidate with a high importance ratio in the pool of candidates[6].

[6]We implement the SIR method by using routines written by Rubin et al., 2004

Appendix 2: Additional tables

Table A1 Summary statistics of the time-variant exogenous covariates (n = 632)

Variable	Wave 1993 Mean	Wave 1993 sd	Wave 1997 Mean	Wave 1997 sd	Wave 2000 Mean	Wave 2000 sd
Urban	0.540	0.499	0.519	0.500	0.511	0.500
Access to safe drinking water						
PAM/PUMP water	0.475	0.500	0.546	0.498	0.560	0.497
In the village there are ...						
Systems of sewage channels/gutters	0.563	0.496	0.616	0.487	0.559	0.497
Factories or cottage industries	0.785	0.411	0.823	0.382	0.859	0.348
Medicinal posts or midwives	0.391	0.488	0.647	0.478	0.570	0.496
Season						
January – March	0.003	0.056	0.016	0.125	0.223	0.417
April – June	0.348	0.477	0.454	0.498	0.634	0.482
July – September	0.611	0.488	0.500	0.500	0.142	0.350
October – December	0.038	0.191	0.030	0.171	0.000	0.000
The farm business is owned by the Household	0.318	0.466	0.318	0.466	0.318	0.466
Household size	4.524	1.886	4.759	1.647	4.837	1.520
Age	25.426	3.465	29.402	3.478	32.229	3.481
Number of children	1.972	1.254	2.560	1.326	2.869	1.355
Log real consumption per head	5.916	0.778	6.105	0.643	6.204	0.656

Table A2 Summary statistics of the time-invariant exogenous covariates (n = 632)

Variable	Mean	sd
Cohort		
Before 1979	0.028	0.166
Between 1980 and 1984	0.277	0.448
Between 1985 and 1989	0.399	0.490
Between 1990 and 1993	0.296	0.457
Province		
North Sumatra	0.070	0.255
West Sumatra	0.046	0.209
South Sumatra	0.073	0.260
Lampung	0.038	0.191
Jakarta	0.093	0.291
West Java	0.177	0.382
Central Java	0.112	0.316
Yogyakarta	0.047	0.213
East Java	0.136	0.343
Bali	0.052	0.223
West Nusa Tenggara	0.062	0.241
South Kalimantan	0.044	0.206
South Sulawesi	0.049	0.216
Muslim	0.897	0.304