Online supplementary methods

1. Multilevel logistic random-effects model

Multilevel model

A three level multilevel model is used to capture four sources of variation in the probability analysis: individual child and her mother and household characteristics, primary sampling unit, district-level factors, and random errors (Equations (1) and (2)):

$$Logit(childcare_{ijk}) = \beta_{0jk}X_0 + \beta_1 Mutuelles_{ijk} + \boldsymbol{\beta X_{ijk}} + \beta_2 PSU\text{-}mutuelles_{jk} \quad (1)$$
$$\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk} \quad (2)$$

where *Logit(childcare$_{ijk}$)* represents the probability of using medical care for the *ith* child in the *jth* PSU and kth district, X_0 is a constant, *Mutuelles$_{ijk}$* represents the *Mutuelles* status for *ith* child in the *jth* PSU and kth district. $\boldsymbol{X_{ijk}}$ is a vector that contains the child, maternal, household characteristics, and survey year for the *ith* child in the *jth* PSU at the kth district. "*PSU-mutuelles$_{jk}$*" captures ecological effects at the PSU level. $\boldsymbol{\beta}$ is a vector of coefficients for each of the $\boldsymbol{X_{ijk}}$. β_2 quantifies contextual effects at the district level. Finally, v_{0k} and u_{0jk} represent between-district random-variation and between-village/within-district random variation, respectively.

Intraclass correlation (ICC)

We reported the intraclass correlation at the district- and PSU-level for the main association between *Mutuelles* and medical care use. To estimate the ICC for our three level logistic models for PSUs within districts, we followed a latent-response formulation, as follows:[1]

$$\rho_k = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_u^2 + \pi^2/3}, \quad (3)$$
$$\rho_{jk} = \frac{\sigma_v^2 + \sigma_u^2}{\sigma_v^2 + \sigma_u^2 + \pi^2/3}, \quad (4)$$

where ρ_k represents the correlation of respondents in different PSUs within the same district, and ρ_{jk} the correlation of respondents in the same PSU within the same district; σ_v^2, and σ_u^2 are the variance estimates of random effects v_{0k} and u_{0jk}, respectively. In the latent-response approach, the residual variance or variance at the individual level is estimated by $\pi^2/3$. [1]

We also reported the correlation of PSUs within the same district, using the following expression:

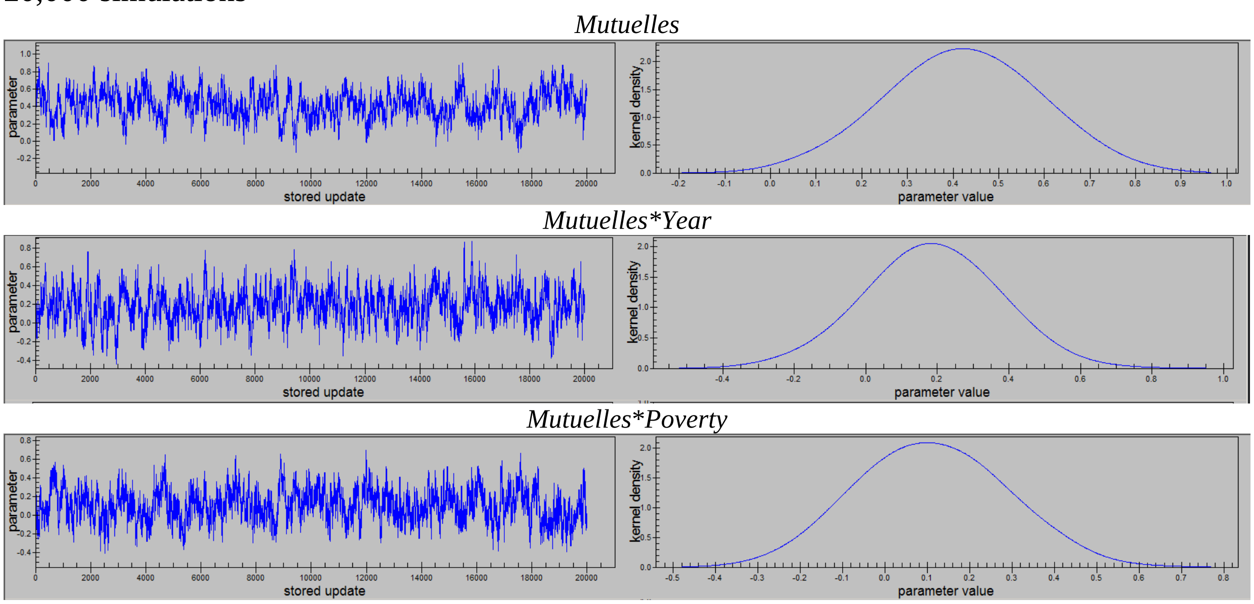
$$\rho_j = \frac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2}, \quad (5)$$

2. Model Optimization using Iterative Generalized Least Squares (IGLS) and Markov Chain Monte Carlo (MCMC)

We used a two-step procedure that combines a Likelihood approach (IGLS with Marginal Quasi-Likelihood –MQL) and the Bayesian MCMC method to optimize discrete outcomes multilevel models.[2] We applied IGLS/MQL in the first step to generate initial estimates and MCMC to get the final posterior estimates and random effects parameters in the second step. We applied this procedure because recent evidence demonstrates that Bayesian uninformative prior methods performed better than likelihood-based methods in calibrating point and interval estimates in a three-level random effects logistic regression model.[3] Details of the MCMC procedure are as follows.

We use diffuse priors to indicate no-belief of the true parameter. In MCMC, initial estimates are usually not correct because the Markov Chain has not stabilized. We then burin-in for 500 simulations and discard them to get the equilibrium distribution.[2,4] Afterwards, we applied 20,000 simulations while monitoring and plotting running cycles with estimates of covariates in cycle order to check for convergence of the model. We expected to see a pattern of “white noise” if the estimates converged.[2,4] We illustrate this monitoring process using 20,000 simulations in Figure S1, where plots of *Mutuelles*-related variables for Model 4 of the childcare analysis are displayed. The left-hand figures show sticky and trendy patterns, as well as trajectories with no-constant variance. Therefore, no sign of “white-noise” trajectories can be inferred from these patterns, indicating that 20,000 simulations were not enough to get stable trajectories for these covariates. The interested reader can obtain plotted figures for all models upon request.

Figure S1 Diagnostic plots for stunting analysis: *Mutuelles*-related variables in Model 4 after 20,000 simulations



We performed the Raftery-Lewis (RL) and Effective Sample Size (ESS) diagnostics for further validation. The former diagnostic provides an approximation of the number of simulations required for the estimation of a particular quantile of the posterior distribution.[5,6] Since we focus on the tail of the posterior distribution, we estimate the 2.5% and 97.5% quantiles using the default precision (tolerance of 0.005 and probability 0.95). The latter ESS diagnostic estimates the equivalent number of independent observations that resulted from 20,000 simulations, where higher ESS is preferable to get more efficient and reliable estimates, whereas low values may indicate highly correlated simulations.[2] These diagnostics are shown in Table S1, suggesting that 20,000 simulations were not enough to get the selected quantiles with the desired precision. For instance, the RL test for *Mutuelles* in Model 4 using 20,000 simulations shows that 69,304 simulations would be needed to estimate the 2.5% quantile and 51,978 simulations would be needed to estimate the 97.5% quantile. Furthermore, with 20,000 iterations, the ESS is only 162 for the *Mutuelles* estimate in Model 4.

The previous diagnostics suggested that more simulations were needed to get a stable posterior distribution. Therefore, we increased the number of simulations to 100,000 and performed the same diagnostics again. The RL was satisfied for all models and the ESS improved substantially (by more than four times for all models) compared to those obtained with 20,000 iterations. Figure S2 presents the posterior distribution of the estimates of the same variables featured in Figure S1 with 100,000 iterations. The left-hand side plots show clearer “white-noise” patterns than those displayed in Figure S1. Meanwhile, the right-hand side plots show the posterior distributions of the estimates, which look symmetric (the median and mean estimates are practically the same –not shown), suggesting that the predicted posterior data are evenly divided around the mean. We conclude from a visual inspection that it is very likely that the models have reached a stable equilibrium distribution with 100,000 simulations. To illustrate this, the point estimate in Model 4 for the *Mutuelles* effect for children in 0-23 month and living above the poverty line is around 0.42 and almost the entire distribution is situated to the right of zero, which indicates that the estimate is highly significant. All these diagnostics indicate that 100,000 simulations were enough to reach convergence of the posterior distribution.

Table S1 Diagnostics for the MCMC procedure for child medical care utilization analyses

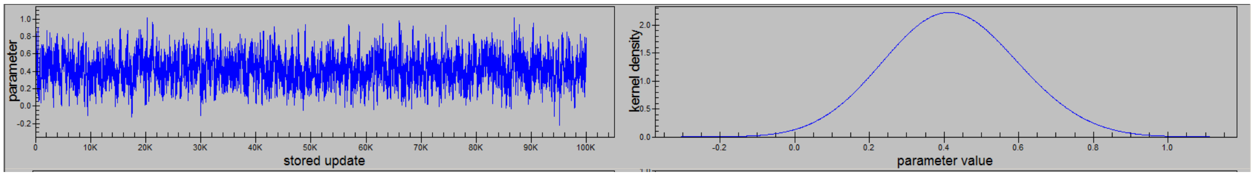
Model*	**Variable**	**Raftery-Lewis (RL)**		**Effective Sample Size (ESS)**	
Simulations		**20000**	**100000**	**20000**	**100000**
1	*Mutuelles*	(40236–40215)	(33884–27974)	615	2983
2	*Mutuelles*	(27776–48921)	(35176–38148)	414	1774
	*Mutuelles**year	(70168–75090)	(46278–46520)	281	1292
3	*Mutuelles*	(50230–55701)	(41124–45348)	195	1147
	*Mutuelles**poverty	(114920–70376)	(41376–38662)	208	1189
4	*Mutuelles*	(69304–51978)	(45650–51702)	162	912
	*Mutuelles**year	(43134–67275)	(49400–43148)	270	1338
	*Mutuelles**poverty	(31674–37989)	(37910–35346)	185	1090

*: For Model 1 to Model 4, please refer to Table 1.

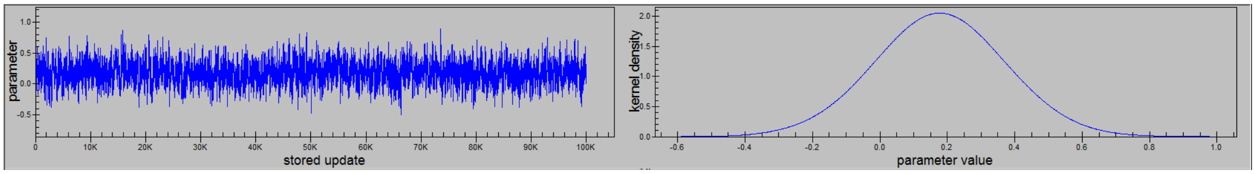
Odds ratios and credible intervals (CrI) are reported. For credible intervals, an estimate will lie within the interval with a probability of 95%.[4,7]

Figure S2 Diagnostic plots for stunting analysis: *Mutuelles-related* variables in Model 4 after 100,000 simulations

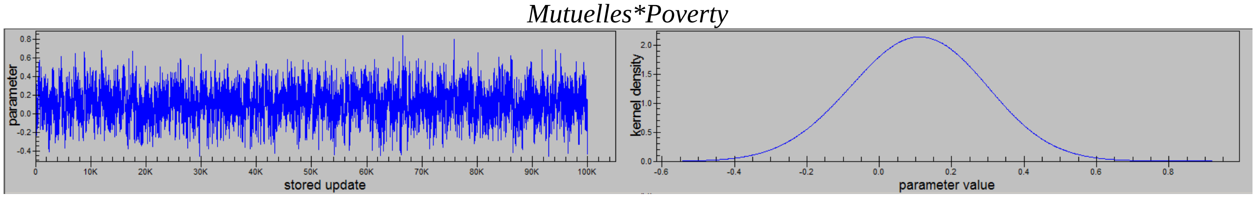
Mutuelles



*Mutuelles*Year*



*Mutuelles*Poverty*



Supplementary references

1. Goldstein, H. (2003). Multilevel statistical models. London: Arnold, 3rd edition.
2. Rasbash J, Steele F, Browne WJ, Harvey G. A User's Guide to MLwiN (version 2.26). Centre for Multilevel Modelling, University of Bristol; 2012.
3. Browne WJ, Drapper D. A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian Analysis*. 2006;1(3): 473-514.
4. Browne WJ. MCMC estimation in MLwiN (version 2.26). Centre for Multilevel Modelling, University of Bristol; 2012.
5. Raftery AE, Lewis SM. How many iterations in the Gibbs sampler? In: Bernardo JM, Berger JO, Dawid AP, editors. *Bayesian Statistics* 4. Oxford: Oxford University Press. 1992;763-773.
6. Raftery AE, Lewis SM. Implementing MCMC. In: Gilks WR, Spiegelhalter DJ, Richardson S, editors. Markov Chain Monte Carlo in Practice. London: Chapman and Hall. 1996; 115-130.
7. Gelman A, Carlin JB, Stern HS, et al. Bayesian Data Analysis. CRC Press, FL; 2014.