# Revisiting Global Trends of Child Labour from 2000 to 2024

Fengqing Chao[1], Estelle Laurière[2], and Andrea Verhulst*[2]

[1]Computational Social Science Division, School of Humanities and Social Science, The Chinese University of Hong Kong, Shenzhen, Shenzhen, Guangdong, P.R. China

[2]INED – Institut National d'Etudes Démographiques, Aubervilliers, France

October 6, 2025

*Extended abstract for EPC 2026*

For EPC session chairs only, not for circulation

*Corresponding author; Email: andrea.verhulst@ined.fr

# 1 Introduction

Ending child labour in all its forms by 2025 is the aim of Target 8.7 of the Sustainable Development Goals (SDGs) [10]. Yet, according to a recent report of the International Labour Organization (ILO) and United Nations International Children's Emergency Fund (UNICEF), 138 million children between the ages of 5 and 17 were still engaged in child labour in 2024 [4].

The data used by ILO and UNICEF to produce the child labour estimates are well known to population scientists, as they are derived from specialized modules embedded in surveys such as the Multiple Indicator Cluster Surveys (MICS) and Demographic and Health Surveys (DHS). They also include stand-alone National Child Labour Surveys (NCLS) and other national socio-economic surveys.

Despite this, the reliability of global child labour estimates remains uncertain. The methodology used by ILO and UNICEF to produce global figures raises several questions [5]. The accompanying methodological report does not address the comparability of survey types and provides no assessment of statistical uncertainty. Furthermore, the estimates are based on a linear regression model with fixed effects, which heavily relies on the quality of socio-economic covariates and the number of observations available within each country. In contexts with limited or no child labour data, this approach may produce biased estimates. To address these limitations, we propose a Bayesian hierarchical time series model that relies more directly on the data and allows information to be shared across countries within the same region.

Moreover, global estimates published in the ILO–UNICEF report, as well as the SDG indicators, are grounded in the definition of child labour adopted at the 18th International Conference of Labour Statisticians (amended in 2018) [3]. This definition, based on international human rights and labour standards, identifies prohibited forms of labour and situates child labour as a subset of the broader category of economically active children. Under this framework, working more than one hour per week in economic activities is prohibited for those under 15, except for children aged 12–14 who may work up to 12 hours per week. For those aged 15–17, economic activity is permitted up to 43 hours per week.

In this paper, we aim to broaden the definition of child labour by leveraging available data. Our approach incorporates domestic labour ("household chores") and other unconventional indicators, applying more flexible thresholds for age and hours worked. To support this effort, we introduce the Child Labour (CL) Database, a publicly accessible tool that provides updated estimates of child labour with detailed disaggregation by age, sex, type of activity, and hours worked. In this way, the database addresses a long-standing gap identified in the literature [2, 6]. The CL Database will be made available through the new website of the French Institute for Demographic Studies (INED), expected in November 2025.

The objective of this paper is therefore to revisit existing global trends in child labour using advanced Bayesian methods, in conjunction with the CL Database, to document the multiple dimensions of this phenomenon with full research transparency.

# 2 Data

Preliminary results presented in this extended abstract are based on MICS surveys conducted between 2005 and 2023 (phases 3–6) and incorporated into the CL Database. The analysis focuses on conventional indicators of the proportion of children engaged in economic activities among the age groups 5–11 and 12–14. Altogether, these surveys cover 100 countries and subnational areas. For simplicity, we hereafter use the term "child labour proportion" to refer to the share of children engaged in economic activities.

Table 1 gives an overview of the observations by age group and MICS round. There are 173 and 169 observations available from 100 countries/areas in ages 5–11 and 12–14 respectively. The estimates were computed from individual-level data obtained from the IPUMS-MICS program [1].

## 2.1 Calculation of sampling errors

We used the Jackknife approach to calculate the sampling errors which reflect the uncertainties resulting from the MICS survey sampling design.

The jackknife sampling error is replaced by its corresponding bootstrap error if the bootstrap approach produce a bigger error than its jackknife counterpart. Most of such replacements are carried out to observations with small numbers of children.

| MICS round | Age 5–11 | Age 12–14 |
|---|---|---|
| 3 | 39 | 39 |
| 4 | 41 | 39 |
| 5 | 38 | 38 |
| 6 | 55 | 53 |
| **Total** | **173** | **169** |

Table 1: **Observation distribution by age group and MICS round.** MICS: Multiple Indicator Cluster Surveys.

# 3 Methods

In the following Section 3.1 and Section 3.2, we summarize the model to estimate child labour proportion for age groups 5–11 and 12–14, where same model was applied to each age group. The logit child labour proportion for a certain age group is modeled as the combination of the regional temporal pattern, and national natural fluctuation.

All notations refer are age-specific and we omit the age index for simplicity purpose. The outcome of interest is $R_{c,t}$, the child labour proportion for age groups 5–11 and 12–14 in country $c$ in year $t$. Throughout the report, $\mathcal{N}(\mu, \sigma^2)$ refers to a normal distribution with mean $\mu$ and variance $\sigma^2$ and $\mathcal{U}(a, b)$ refers to a continuous uniform distribution with lower and upper bounds at $a$ and $b$ respectively. The detailed explanations of the model and the joint density of posterior distribution are in the rest of this section.

## 3.1 Data model

The observations of the age-specific child labour proportion are indexed by $i \in \{1, \cdots, n\}$. $n$ refers to the number of observation. $r_i$ denotes the $i$-th observed age-specific child labour proportion, in country $c[i]$ in year $t[i]$. $R_{c,t}$ is the outcome of interest, the true age-specific child labour proportion, for country $c$ in year $t$. $r_i$ is rescaled by logit to ensure the model estimates fall between 0 and 1. We assume for $i \in \{1, \cdots, n\}$:

$$\text{logit}(r_i) = \text{logit}(R_{c,t}) + \delta_i, \quad (1)$$

$$\delta_i \sim \mathcal{N}(0, \sigma_i^2 + \omega^2), \quad (2)$$

$$\log\left(\frac{1}{\omega^2}\right) \sim \mathcal{PC}(\sigma_\omega, 0.01), \quad (3)$$

where $\delta_i$ is the error term for observations on the logit-scale. The error term $\delta_i$ follows a normal distribution with mean at zero. The variance of the distribution is the sum of a known sampling error variance $\sigma_i^2$ (Section 2.1) and unknown non-sampling error variance parameter $\omega^2$. The sampling error variance is pre-calculated to account for the uncertainty due to the sample design of the MICS surveys. The non-sampling error variance $\omega^2$ is a model parameter for uncertainties resulting from aspects that can be minimized but difficult to eliminate entirely, such as non-responses, data input errors, and recall bias.

## 3.2 Process model of child labour proportion

We assume the true underlying age-specific child labour proportion to be modeled as follows for country $c$ in year $t$:

$$\text{logit}(R_{a,c,t}) = W_{r[c],t} + P_{c,t}. \quad (4)$$

$W_{r[c],t}$ models the regional non-linear relation between the age-specific child labour proportion and time in region $r$ where country $c$ belongs to. $W_{r[c],t}$ is modeled with an random walk of order 2 (RW2) structure. For region $r \in \{1, \cdots, 10\}$, we assume:

$$\Delta^2 W_{r,t} = W_{r,t} - 2W_{r,t+1} + W_{r,t+2}, \quad (5)$$

$$\Delta^2 W_{r,t} \sim \mathcal{N}(0, 1/\tau_r^w), \text{ for } t \in \{2000, \cdots, 2024-2\}. \quad (6)$$

We assigned a Penalized Complex (PC) prior to the regional precision parameter $\tau_r^w$ for $r \in \{1, \cdots, 10\}$:

$$\tau_r^w \sim \mathcal{PC}(w, 0.01), \text{ for } r \in \{1, \cdots, 10\}. \quad (7)$$

where $w$ is the standard deviation of logit$(r_i)$. The PC prior is a vague prior. [9] documented the PC prior specification in detail.

$P_{c,t}$ accounts for the within country temporal fluctuations. We use a first-order random walk (RW1) model for $P_{c,t}$ and assign a PC prior to country-specific precision parameter $\tau_c^p$. For country $c \in \{1, \cdots, 100\}$, we have:

$$\Delta P_{c,t} = P_{c,t+1} - P_{c,t}, \quad (8)$$
$$\Delta P_{c,t} \sim \mathcal{N}(0, 1/\tau_c^p), \text{ for } t \in \{2000, \cdots, 2024-1\}, \quad (9)$$
$$\tau_c^p \sim \mathcal{PC}(1, 0.01), \text{ for } c \in \{1, \cdots, 100\}. \quad (10)$$

**Statistical computation** We use the Integrated Nested Laplace Approximation (INLA) for Bayesian inference ([7]). We use the R-package `R-INLA` ([8]) to implement the INLA estimation procedure.

# 4 Preliminary results

## 4.1 National results for age group 5–11

Figure 1 illustrates the child labour proportion for ages 5–11 estimates in Central African Republic, and Djibouti. There are three observations in Central African Republic and the model estimates are in line with the data in this country. Whereas in Djibouti, only one data point is available. The Bayesian hierarchical model enables us to produce estimates based on patterns from neighboring countries, meanwhile the level is informed by that single available data point.

## 4.2 Regional results for age group 5–11

Figure 2 presents the regional average trend for child labour proportion in Sub-Saharan Africa. This figure further explains how the Bayesian hierarchical modeling structure can improve the national estimate for countries lacking of data. As shown in Figure 2, the regional child labour average trend is well informed by 62 observations (36% of the total observations) from 37 countries/areas. The Bayesian hierarchical model follows the overall trend based on all data and weight observations according to their sampling errors in Sub-Saharan Africa. The resulting regional trend is non-linear overtime and is consistent with the pattern suggested by data.

While both countries Central African Republic and Djibouti are from Sub-Saharan Africa, their national estimates are influenced by this regional pattern. Since Djibouti only have one data point in 2006, the model imputation is largely based on this underlying regional pattern.

# 5 Discussion

Our preliminary results suggest an unexpected increase in child labour in sub-Saharan Africa, peaking between 2010 and 2015. To our knowledge, this trend has not been previously documented. By contrast, ILO–UNICEF estimates for the region indicate a peak around 2020, though their reports neither present the underlying observed data nor provide measures of statistical uncertainty.

In the final paper, we will examine whether this discrepancy stems from differences in modelling approaches or the types of surveys included in the analysis. We will also investigate whether the observed increase in child labour may be partially explained by changes in MICS questionnaires over time. In addition, we will expand the interpretation of country-specific results, conduct model validation exercises, and discuss the study's limitations.

We will explore the possibility to incorporate external covariates in our Bayesian model for the final paper and conduct sensitivity anslysis on whether the added covariates in the model can explain more uncertainties in the data.

The final results will present detailed country- and region-level comparisons with ILO–UNICEF estimates, drawing on a broader set of available data from 2000 to 2024. They will also incorporate unconventional indicators, with flexible age/hour thresholds and for non-economic activities, made available through the Child Labour Database.

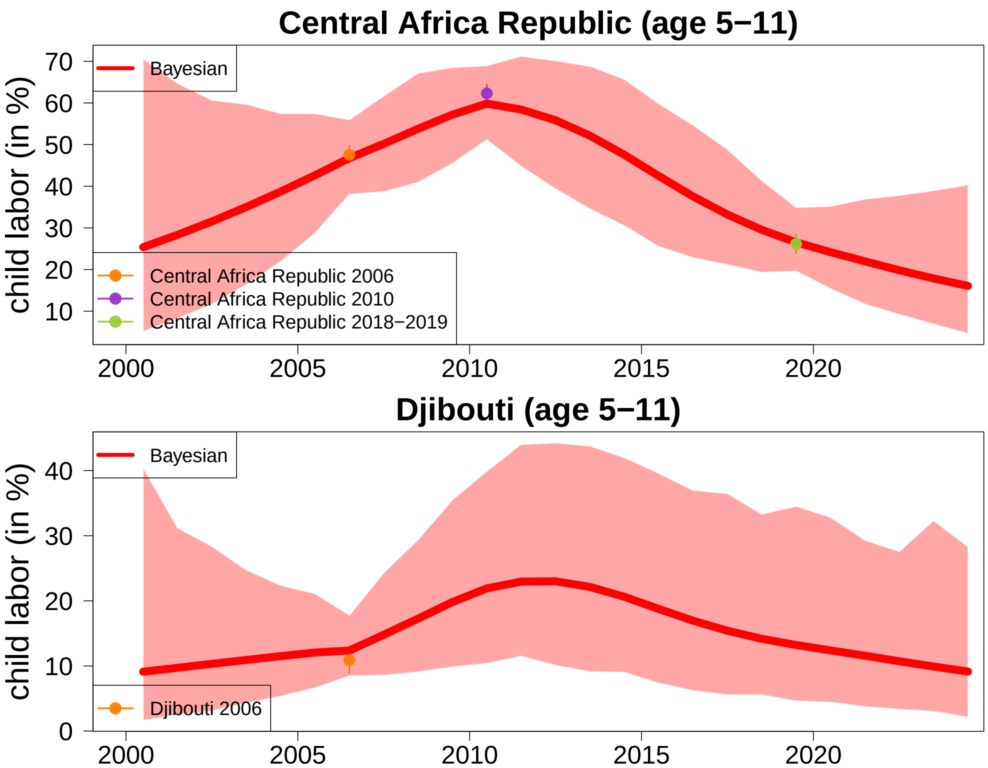


Figure 1: **Child labour proportion for ages 5–11 for selected countries.** Red curves are posterior medians. Red shades are 95% uncertainty intervals. Observations are in dots. The vertical bars around each dots refer to the corresponding pre-calculated sampling errors.

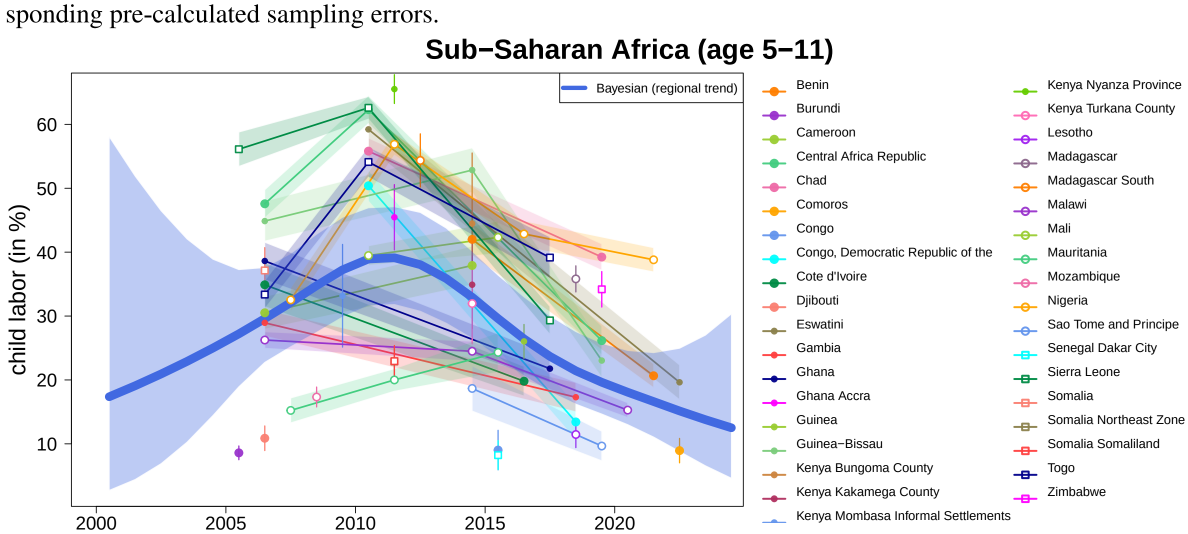


Figure 2: **Regional average trend for child labour proportion for ages 5–11 for Sub-Saharan Africa.** Blue curve is posterior medians. Blue shades are 95% uncertainty intervals. Observations from all countries within Sub-Saharan Africa are in dots while dot colors differentiate the countries. The vertical bars and shades around dots refer to the corresponding pre-calculated sampling errors.

# References


[1] Bolgrien, A., E. H. Boyle, M. Sobek, and M. King (2024). IPUMS MICS data harmonization code. version 1.2 [stata syntax]. IPUMS: Minneapolis, MN. https://doi.org/10.18128/D082.V1.2.

[2] Edmonds, E. V. (2007). Chapter 57 Child Labor. Volume 4 of *Handbook of Development Economics*, pp. 3607–3709. Elsevier. ISSN: 1573-4471.

[3] International Labour Organization (2018). Resolution to amend the 18th ICLS Resolution concerning statistics of child labour. Conference: 20th International Conference of Labour Statisticians (ICLS), 10-19 October, 2018.

[4] International Labour Organization and United Nations Children's Fund (2025a). *Child Labour: Global estimates 2024, trends and the road forward*. New York: ILO and UNICEF.

[5] International Labour Organization and United Nations Children's Fund (2025b). *Methodology of the 2024 ILO-UNICEF Global Estimates of Child Labour*. New York: ILO and UNICEF.

[6] Ortiz-Ospina, E. and M. Roser (2016). Child Labor. Published online at OurWorldinData.org: https://ourworldindata.org/child-labor.

[7] Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society B 71*, 319–392.

[8] Rue, H., S. Martino, F. Lindgren, D. Simpson, and A. Riebler (2013). R-inla: Approximate bayesian inference using integrated nested laplace approximations. *Trondheim, Norway*. Available at: http://www.r-inla.org.

[9] Simpson, D., H. Rue, A. Riebler, T. G. Martins, and S. H. Sørbye (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical science*, 1–28.

[10] UN (2015). *Transforming our world: The 2030 agenda for sustainable development (A/RES/70/1)*. Available at https://sustainabledevelopment.un.org/post2015/transformingourworld/publication. Accessed 14 Aug 2025.