A Bayesian Estimation of Child Labour in India

Jihye Kim [1] · Wendy Olsen [1] · Arkadiusz Wiśniowski [1]

Abstract

Child labour in India involves the largest number of children in any single country in the world. In 2011, 11.8 million children between the ages of 5 and 17 were main workers (those working more than 6 mo) according to the Indian Census. Our estimate of child labour using a combined-data approach is slightly higher than that: 13.2 million (11.4–15.2 million) for ages 5 to 17. There are various opinions on how best to measure the prevalence of child labour. In this study, we use the International Labour Organization (ILO)'s methodology to define hazardousness and combine it with the most recent United Nations Children's Fund (UNICEF)'s time thresholds for economic work and household chores. The specific aims of this study are to estimate the prevalence of child labour in the age group 5 to 17 and to suggest a combined-data approach using Bayesian inference to improve the accuracy of the child labour estimation. This study combines the National Sample Survey on Employment and Unemployment 2011/12 and the India Human Development Survey 2011/12 and compares the result with the reported figures for the incidence of child labour from the Indian Census. Our unique combined-data approach provides a way to improve accuracy, smooth the variations between ages and provide reliable estimates of the scale of child labour in India.

✉ Jihye Kim
  jihye.kim@postgrad.manchester.ac.uk

  Wendy Olsen
  wendy.olsen@manchester.ac.uk

  Arkadiusz Wiśniowski
  a.wisniowski@manchester.ac.uk

[1]   Social Statistics Department, School of Social Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Springer

1 Introduction

India has the largest number of working children of any country in the world, with a Census estimate of 12.66 million children aged from 5 to 14 holding this status in 2001, falling to 4.35 million in 2011 (according to the Ministry of Labour and Employment n.d.). According to the analysis of 'child workers' (ages 5–14), as per the 2011 Census of India, Uttar Pradesh state had the largest number of child workers (2.1 million, 4.1%) and Bihar had the second-largest number (1.1million, 3.9%; Samantroy et al. 2017, p.46). In terms of incidence, Nagaland and Himachal Pradesh have shown the highest proportion of child workers at 13.2% and 10.1% respectively (ibid.). However, the numbers vary according to datasets and definitions. For example, the proportion of child labourers reached 11.8% among children aged 5 to 14 in 2012, as calculated by UNICEF (n.d.),[1] which is roughly 29 million children in total. Increased household income has led to a reduction over time, but many industries and farms still employ child labourers. Unpaid household services (household chores) are ones of the most prevalent types of child labour and could be increasing, but how this work can be adequately measured as child labour is still debatable.

There has been extensive discussion of how to define child labour. The differentiation between child labour and child work and setting up time boundaries are the key issues. The ILO and UNICEF have moved toward an international agreement on definitions of child labour. However, their definition is still different in some aspects from the definition used at the national level in India. In 2016, the Indian government amended the Child Labour Act (the Amendment Act) to adopt a strict policy of banning children under the age of 14 from work. However, the Amendment Act excludes 'helping families or working in family enterprises' from the category of child labour. Also, there is no time limit for children's weekly working hours in the Amendment Act.

Besides being a matter of definition, there is a need to address the issue of how to measure the number of child labourers with accuracy. A measurement error is a departure from reality in the measurement provided (Groves et al. 2011, p.52). One possible measurement error is an intentional or unintentional misresponse: for example, parents' non-responses due to their increasing awareness of the illegality of child labour (Basu 1999); and children being involved in farm work but not being recognised as 'child labour' (Chaudhri et al. 2003; Chaudhri and Wilson 2000). In general, 'child labour' is a subset of all the forms labouring and child work undertaken by children under the age of 18 (variants are described by Dubey et al. 2017). Alternatively, a measurement error might arise owing to limited information, such as precise working hours, industrial categories and working conditions.

Furthermore, child labour is sparse, especially among younger age groups. There could be overrepresentation when we calculate this number using survey weights only. A model-based approach can provide more accurate information regarding the number of child labourers. The National Sample Survey (NSS) and the India Human Development Survey (IHDS) provide qualified data relating to child labour, with each dataset having different strengths. We have chosen to use the Indian Census 2011 as auxiliary information. Despite its large population coverage, it provides only the aggregate

---

[1] Available at https://www.unicef.org/infobycountry/india_statistics.html (Accessed 16 May 2019). The population of children aged 5 to 14 is estimated at about 260 million by the Indian Census, 2011.

number of child labourers based on two broad categories – children who are main (i.e. working more than 6 mo a year in economic activity[2]) or marginal (working less than 6 mo a year) workers.

The contribution of this paper is twofold. First, we investigate how the international definition of child labour affects the measured prevalence of child labour. We compare any differences in the measurement of child labour between the major stakeholders – the ILO and UNICEF – and the Indian government's version before justifying the use of the definition and criteria of child labour. Then we apply these criteria to the estimation of the number of child labourers in India by using a Bayesian hierarchical model. This hierarchical model has layers of parameters; some parameters are relatively innocuous feed into estimates of the key unknowns, such as the number of child labourers. This study is the first attempt to integrate two data sources by using Bayesian inference to estimate child labour disaggregated by age and state. Thus, we overcome the limitations of using a single dataset. Ultimately, this study aims to reveal whether the combined-data approach is efficient in addressing measurement errors regarding estimates of child labour.

This paper is organised as follows. Section 2 introduces the background of the research and discussions on the definition of child labour, which justify our own definition. Section 3 explains our methodology, including a review of Bayesian inference, and describes our models. Lastly, we summarise the results in Section 4 and provide key policy implications in Section 5. We find it informative to use the two sample datasets together alongside the Census.

2 Differentiation Separating Child Labour from Child Work

Earlier research on child labour defined child labour in an extensive way in order to put pressure on legislative interventions. Weiner (1991, p.3007) suggests that children who are not in school are potential child labourers. Economic studies on child labour have shown wide inclusion of various types of work. For example, Basu and Van (1998) define child labour as any economic activity of children, and Basu (1999) notes that this definition includes part-time workers. In a later study, domestic chores are also included in the work done by child labourers (Basu et al. 2010).

There have been ongoing efforts to distinguish child labour that is harmful to children's development from socially accepted child work. Since the early 2000s, the ILO definition of child labourers as 'economically active' children (Ashagrie 1993) has been widely recognised in research. Ray (2000, 2002), for example, uses wages to distinguish 'child labour' from 'child work'. Meanwhile, conceptualisations of child labour have gradually expanded and included unpaid and irregular work. For example, the issue of girls being heavily burdened with domestic chores is now more recognised (Das and Mukherjee 2007, 2011; Kambhampati and Rajan 2008).

---

[2] In the Indian Census data, economic activity is working as cultivators, agricultural labourers, household industry workers and other workers.

Despite a conceptual expansion of child labour, estimation of its extent has been methodologically limited due to inconsistent standards of calculation and lack of relevant datasets at the national level. The use of working hours to separate child labour from child work is agreed by most international agencies and many other stakeholders. The 20th ILO Conference of Labour Statisticians (ICLS) re-confirms the use of working hours as a threshold for child labour in both the System of National Account (SNA) production boundary and the general production boundary (ILO 2018a).[3] However, how many hours of domestic chores are considered harmful for children, and what periods (month, quarter or year) should be regarded as the "usual status", are still under discussion. Also, the ILO (2018a) has suggested that thresholds of hours are determined by national law, and in the absence of such law, by adult workers' normal working hours. There is some ambiguity because the maximum hours dictated by national law are different in each country.

Studies in the methodology of estimating the correct number of child labourers are rare. Levison and Langer (2010) use two datasets to count domestic servants in Latin America. They do not use models but a weighted count. Webbink et al. (2015) provide a Tobit model of working hours in paid work but do not critically differentiate child labour from child work. Giri and Singh (2016) attempt to count child labourers in India by integrating economic activities and domestic chores and including 'nowhere children'[4]; they do not make use of models but multiply the ratio of child labourers by the population. Moreover, considering all 'nowhere children' as potential child labourers might bring about misunderstandings of reality (Lieten 2002). Thus we argue that it is best to measure the number of actual child labourers and not the potential size of this group. This paper gives advice on how to do so.

2.1 Review of Definitions and Measurement of Child Labour

There has been an effort to define child labour by international agencies. According to the United Nations SNA, economic activity includes some types of family work but excludes household chores (ILO 2017, p.17). Some economic work, such as production of own-use goods, is within the Gross Domestic Product (GDP), while some forms of domestic work, such as unpaid household services, are outside of the GDP (Hirway and Jose 2011). However, the ILO agrees that unpaid household services are part of child labour if they are performed in hazardous conditions (ILO 2017). In this section, we describe the widely-used ILO and UNICEF definition and measurement of child labour and then compare these with the Indian government's current version. UNICEF has used the same definition of child labour as the ILO since 2008 after the 18th ILO Conference of Labour Statisticians (ICLS),[5] but they have applied slightly different methods in capturing child labour statistics.

---

[3] The SNA is a set of standards to measure economic activity, initiated by the United Nations Statistical Commission. The general production boundary covers all kinds of activities producing goods and services (ILO 2018a). Own-use production work of services, such as washing and preparing meals, is excluded from the SNA production boundary but included in the general production boundary (ILO 2018a).

[4] Chaudhri et al. (2003) and Chaudhri and Wilson (2000) introduce the concept of 'nowhere children' who are neither in school nor work and insist they should be counted as child labourers.

[5] Available at https://www.unicef.org/protection/57929_child_labour.html (Accessed 11 Jan. 2020)

2.1.1 Definitions of Child Labour

In the ILO definition, child labour means children in employment, excluding 'children who are in permitted light work and those above the minimum age' (ILO 2017, p.17). The ILO's first focus is on the hazardousness of work that children are engaged in (Omoike 2010). The worst forms of child labour are part of hazardous work. The worst forms of child labour are any work which, by its nature and circumstances, harms children's health, safety or morality, such as slavery, prostitution, and illicit activities.[6] The ILO (2017) also recognises the significance of hazardous unpaid household activities, but it does not provide an explicit method for estimating household chores. The ILO (2018a) has moved in the direction of allowing for hazardous forms of domestic work (notable in the Annex).

According to the ILO's (2017) minimum age standard,[7] the minimum age should not be less than the age of completion of compulsory schooling and in any case, no less than 15 years old (14 for developing countries). The ILO also shows a concern for children who are 16 or 17 years old. The ILO's Worst Forms of Child Labour Convention prohibits the worst forms of child labour for any children under the age of 18 (ibid.).

UNICEF's definition agrees conceptually with the ILO's (2017) version, but methodologically it shifts interest toward including child's domestic work. UNICEF emphasises the importance of domestic work performed by children, which is measured by different time boundaries for ages 5–11, 12–14 and 15–17 (Chaubey et al. 2007, p.2). As a result, the number of child labourers in UNICEF's standard shows significant growth when compared with the ILO (2017).

The Indian Government's Amendment Act, which came into effect in September 2016, defines child labour as any work of a child (up to the age of 14 years) except for helping their families, working in family enterprises after school hours or artistic work, and adolescents (15–17) working in hazardous industries. However, criticism has been raised because family work and family enterprises might allow for exceptions. Also, the Amendment Act considers only three types of work as hazardous – mining, working with inflammable substances or explosives, and working in a hazardous process. The Indian government's definition of child labour is narrower than the definition of international agencies, as it still allows many hazardous work activities and disregards the effects of long working hours.

2.1.2 Measurement of Child Labour

The table below summarises the ILO estimation procedure of child labour (ILO 2017). There are various child labour categories, such as children aged 5–11 in any work, children aged 12–14 who are in more than light work, and children aged 15–17 who are in hazardous work. Hazardousness is specified by industrial and occupational types, working conditions, long-hour work and hazardous unpaid household activities. Table 1.

---

[6] Worst Forms of Child Labour Convention initiated in 1999, No. 182
[7] Minimum Age Convention in 1973, No. 138

Table 1 Latest Standard of Child Labour by the ILO (2017)

Age groups	5–11	12–14	15–17
ILO Child Labour Criteria	- In hazardous industries/in hazardous occupations/long hours of work/in hazardous working conditions	- In hazardous industries/in hazardous occupations/long hours of work/in hazardous working conditions	- In hazardous industries/in hazardous occupations/long hours of work/in hazardous working conditions
	- Any involvement (at least 1 h or more) in economic activity[1]	- 14 h or more in economic activity	- 43 h or more in economic activity
	Hazardous unpaid household activities by children[2]		

Source: Summarised from the ILO conceptual framework (ILO 2017, p.56)

Notes: [1] Economic activity does not include household chores; [2] hazardous unpaid household activity (household chores) means general production boundary work (a) for long hours, (b) in an unhealthy environment, or (c) in dangerous locations (see ILO (2018a), the 18th ICLS resolution paragraphs 15(c), 36 and 37)

The ILO does not specify a limit of working hours for unpaid household activities. Children's involvement in unpaid household services for long hours or in an unhealthy environment and dangerous locations is child labour, though there are no specific criteria for the measurement of these activities (ILO 2016, pp.55–57).

We use UNICEF's most recent time boundaries for child labour in each age group (Table 2). According to the current database, UNICEF (2019)[8] categorises the criteria as follows: (a) children 5–11 years old who undertook at least 1 h of economic activity or at least 21 h of household chores per week; (b) children 12–14 years old who undertook at least 14 h of economic activity or at least 21 h of household chores per week; and (c) children 15–17 years old who undertook at least 43 h of economic activity per week. UNICEF's change to their standards of child labour reflects the concerns of the ILO (2013, 2016) that more than 20 h of household chores negatively affects children's education. Previously, the time threshold for household chores for ages 5–14 was 28 h (UNICEF 2017), but this has now been reduced to 21 h. Children 'working in hazardous working conditions' or children (15–17 years) who spend long hours doing household chores, which were criteria of child labour in the Multiple Indicator Cluster Surveys (MICS) from UNICEF (2017), are no longer included.

2.2 Conceptual Framework

Hereinafter, we define child labour as including children between 5 and 17 years of age who are engaged in any work that is harmful to their development as well as household chores that require considerable amounts of time (defined below). Our definition is along the same lines as the definition of the ILO. It includes any types of child labour that hampers children's physical, intellectual and mental development (Weiner 1991; Weston 2005). Thus we believe that children working in hazardous industries and occupations should be considered to be child labourers regardless of working hours.

---

[8] This update took place in October 2019 (https://data.unicef.org/topic/child-protection/child-labour/ accessed 11 Jan. 2020).

Table 2 The Latest Standard of Child Labour by the UNICEF 2019 Criteria

Age groups	5–11	12–14	15–17
UNICEF Child Labour Criteria	- At least 1 h of economic work or 21 h of unpaid household services per week	- At least 14 h of economic work or 21 h of unpaid household services per week	- At least 43 h of economic work per week

Source: Summarised from the UNICEF database (UNICEF 2019)

Time thresholds are applied differently for the age groups depending on the types of work. We follow the ILO's minimum ages: age 15 for basic work and age 18 for hazardous work.

Our measurement is composed of several steps. Firstly, borrowing knowledge about hazardous occupations and industries from the ILO criteria, we categorise any children involved in harmful areas of work as child labourers regardless of the amount of time spent working. The list of hazardous industries and occupations is shown in Appendix Fig. 9. We keep time thresholds for economic activity (43 h for ages 15–17, 14 h for ages 12–14, and 1 h for ages 5–11). Then we use UNICEF's weekly time thresholds for unpaid household services: at least 21 h a week of household chores for ages 5–14. In Table 3, we summarise our operationalisation for this study of the definition of child labour used by ILO and UNICEF.

3 Methodology

3.1 Review of Methods

We have devised a bespoke statistical model that is capable of estimating the number of child labourers using two different sources of data and reconciling the differences between them. We use Bayesian inference to produce the point and interval estimates of the number of child labourers with the accompanying measures of uncertainty in the

Table 3 Our Criteria for Child Labour Estimation

	Age Groups	5–11	12–14	15–17
Child Labour Criteria	Hazardousness[1]	- In hazardous industries or occupations	- In hazardous industries or occupations	- In hazardous industries or occupations
	Time thresholds[2]	- At least 1 h of economic work or - at least 21 h of unpaid household services per week	- At least14 hours of economic work or - at least 21 h of unpaid household services per week	- At least 43 h of economic work per week

Notes: [1] The ILO scheme is shown in the category of hazardousness; [2] the UNICEF scheme is shown in the category of time thresholds; weekly working hours as a usual (principal) status

form of posterior probability distributions. Bayesian inference combines the model for the observed data specified in the form of a likelihood function with the prior distributions for the unknown parameters.

There are many advantages to using the Bayesian approach in this piece of research. Firstly, the Bayesian approach quantifies the uncertainty of conditions that may not be observable, taking it as a prior distribution (Gelman et al. 2013, p.8). We test several priors and choose the best-performing one to produce our final results. Secondly, the Bayesian approach provides direct interpretations of the posterior probability distribution. While the frequentist approach uses confidence intervals which are the ranges of chances to include the true value of parameters (i.e. 95% confidence), the Bayesian approach uses predictive intervals, which specify the probabilities that the true values lie within them (Gelman 2013, p.95). The predictive intervals allow a clear interpretation of the estimated number of child labourers as well as other model parameters. Thirdly, it is relatively straightforward to fit models with many parameters together with a multi-layered probability structure (Gelman et al. 2013, p.4).

In this particular application, the Bayesian inferential framework presents us with an efficient way to combine different datasets. The uncertainty of various measurements is translated into model parameters. This is especially useful when multiple datasets provide partial (biased), imprecise or conflicting information about the unobserved (latent) quantity of interest. A statistical model that accounts and corrects for the inaccuracies and biases in the data can be used. The resulting estimates can thus be more precise as information from both datasets is used, compared with using each data source on its own.

3.2 Data

This research maximises the accuracy of measurements of child labourer numbers in India, combining both the NSS 2011 and the IHDS 2011, and using the Indian Census 2011 as auxiliary information.

The NSS Employment and Unemployment Survey is the most commonly used dataset on employment as it provides details of work types and industrial categories. The NSS 68th round (July 2011–June 2012) is a large sample survey (number of respondents for ages 5–17 = 122,630, representing 0.04% of the population of the same age group), covering all 35 states in India. A stratified multi-stage design is applied; the first stratum is based on urban-rural characteristics, and the second stratum is based on household wealth (National Sample Survey Office 2013).

The India Human Development Survey (IHDS) is a panel survey of two rounds – Wave 1 in 2004/05 and Wave 2 in 2011/12 – but we use only the second wave for this study to match the year with the NSS. The sample size of the IHDS is about half of the NSS's (number of respondents for ages 5–17 = 51,556, representing 0.02% of the population of the same age group). It covers 33 states in India, excluding Andaman and Nicobar Islands and Lakshadweep (Desai and Vanneman 2015, p.2). The variables missing from these two States are treated as missing items in our models. The samples in rural areas are drawn from participants in a previous survey by the National Council of Applied Economic Research and the samples in urban areas are selected by sampling proportional to population (ibid.).

The Indian Census 2011 (Ministry of Home Affairs 2011) does not include the specific data required to measure child labour according to our definition (Section 2.2), such as industries or working hours; it only gives the aggregate numbers of main and marginal workers defined by the period of work. We use this information as an auxiliary variable. Also, we obtained the population size by age and state from the Indian Census 2011. The date of the Census is close to the date of the other two surveys (2012 vs 2011).

3.3 Undercount Parameter

The IHDS has a clear limitation regarding working hours in domestic work, but it provides accurate working hours for economic activity, while the NSS has an approximation of work intensity, but covers all types of work. The IHDS includes working hours in a family business and household farming work but does not include working hours for household chores. Hence, some of the IHDS sample children who work long hours at home cannot be included as child labourers, which might cause a significant undercount. The descriptive analysis shows that there could be a systematic undercounting of child labourers in the IHDS figures compared to those from the NSS (see Section 3.5).

Considering the overall differences between the IHDS and the NSS data, we suggest using the combined-data model with a parameter that measures any systemic undercounting by the IHDS against the NSS numbers. A systematic over- or under-estimation in one dataset might be solved by applying an over- or under-count parameter (e.g. Wiśniowski 2017; Wiśniowski et al. 2013). The under- or over-counting parameter in this study informs us about the relationship between two datasets. Thus it is not possible to use it for a single-dataset model (see Appendix Table 5).

3.4 Matching Two Key Datasets

To combine the datasets, we have reviewed the variables and questionnaires and then matched the types of work (industries and occupations) and the time-use information. Firstly, the IHDS's industrial and occupational codes are matched with the NSS's. The NSS provides five-digit codes for industries of the National Industrial Classification (NIC) 2008 and three-digit codes for occupations of the National Classification of Occupation (NCO) 2004, which correspond to those used by the ILO (ILO 2016). The IHDS uses its own industrial codes and the NCO 1968.

Regarding time-use, the IHDS asks respondents how many hours a day they usually work, whereas the NSS asks for a daily time disposition of activity based on a one-week recall. The IHDS provides daily working hours (0 to 16 h)[9]; however, the NSS offers the intensity of time disposition for each activity (None = 0, half = 0.5, and full = 1.0) for the last 7 d. (Maximum points are 7.0 per activity.)[10] In this paper, we use time thresholds on a weekly basis; therefore, the daily hours of the IHDS need to be multiplied by seven, based on an assumption that children work every day in the week.

---

[9] 'How many hours did you work in a usual day?' (Question 7.8, IHDS 2011).
[10] 'Time disposition during the week ended on ....' (intensity of activity for 7 d, full mark: 1.0 and half-mark: 0.5) (Question 5.3, NSS 2011).

The NSS sets the maximum working hours at seven points in a week, which is converted to 70 points and regarded as equal to 43 h.

This study rigorously follows the concept of 'usual hours of work per week' suggested by the ILO (ILO 2018a).[11] We use a principal activity status provided by both datasets which indicates the activity on which a person spent a long time during a reference year, as a usual status.

3.5 Descriptive Analysis of the Data

Our criteria for measuring child labour provide time thresholds both for economic activity that is a type of work counted in GDP, including work in family-owned farming or business, and for unpaid household services. In terms of industries and occupations, construction, mining, and waged agricultural or fishery work are the areas with the highest number of child labourers. Children working in textile manufacturing and the service sector, such as street vendors, also appear in large numbers (see Appendix Fig. 8). Moreover, many child labourers are found to be unpaid household labourers, and most of these are female children (see Appendix Fig. 9).

Figure 1 provides a gross-weighted count of children who are considered to be labourers according to two different measurement methods – one from the ILO and the other from UNICEF – calculated by using the NSS and IHDS in 2011/12. In the ILO standard, the gross-weighted figure of child labourers according to the IHDS is 11.4 million, while the same figure according to the NSS is 13.7 million. Meanwhile, the UNICEF measurement excludes child labour in hazardous activities and adds unpaid household services for long hours to the category of child labour. Accordingly, it generates a huge discrepancy in the numbers of child labour calculated from two datasets: 7.3 million and 13.22 million from the IHDS and the NSS respectively. After applying the UNICEF standard, the IHDS has a significantly reduced number of child labourers because there are no records of children doing household chores and because of the exclusion of those in hazardous work. However, the NSS, even after excluding hazardous work from the definition, still shows many child labourers because of including children who spend excessive time in the recall week doing household chores.

Figure 2(b) shows the result of applying the newly-constructed integer weights to the counting of child labourers and multiplying it by the population from the Census using the NSS and IHDS 2011/12 under our measurement criteria. The gross-weighted number of child labourers shows an irregular pattern by age (Fig. 2(a)). However, using a relative weight reduces any sharp decrease between the ages. The relative weights are obtained by taking the sampling weights divided by the mean and rounding them up to the nearest integer greater than 0. The ranges of the new weights are from 1 to 26 for the IHDS, and from 1 to 39 for the NSS, which indicates the number of duplications of each value from the survey used to construct the estimate. Thus, in models, we prefer to use the relatively weighted number of child labour and to multiply
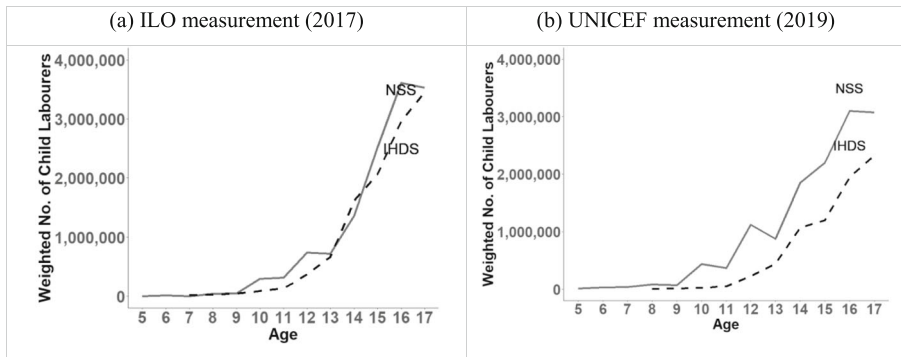
Fig. 1 Gross Weighted Number of Child Labourers in India Using Different Measurement Schemes. (**a**) ILO measurement (2017), (**b**) UNICEF measurement (2019)

it by the population from the Indian census, as it smooths variations by age. A relative weighted count of child labour in the combination of the true population suggests 14.8 million based on the NSS and 11.4 million when using the IHDS.

An obvious limitation of both datasets is that many children in early childhood are categorised as either "other" or "too young" in their principal activity status, so they are not included as child labourers. A lower incidence of child labour among younger children might be related to who is in those categories. An "other" category includes abandoned work such as begging, waste-picking, etc., but it is difficult to obtain a specific status for each child in that category.

Figures 3 and 4 compare different numbers of child workers and child labourers in India. The Indian Census suggests that the number of child workers between the ages of 5 and 17 was 23.8 million in 2011/12. Our estimation suggests there were 11.4–15.2 million child labourers in the same year (see Section 4). In both results, Uttar Pradesh has the largest number of child workers as well as child labourers because of its large population.
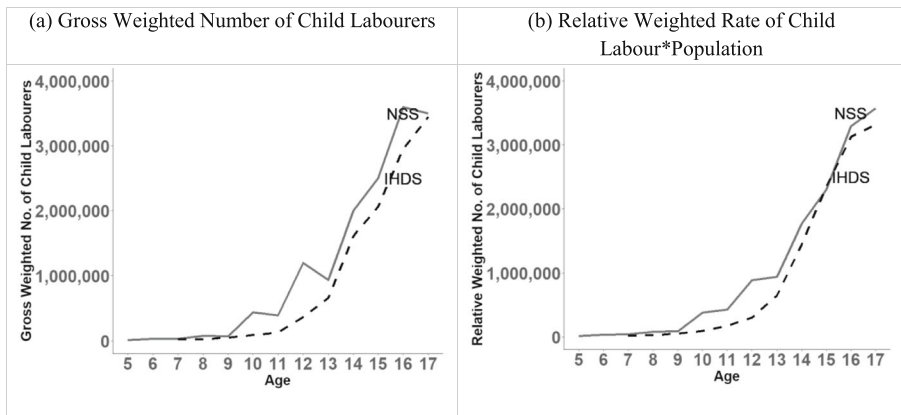

Fig. 2 Gross vs Relative Weighted Count of Child Labourers in India (Non-model, Authors' Definition). (**a**) Gross Weighted Number of Child Labourers, (**b**) Relative Weighted Rate of Child Labour*Population. Notes: weighted counting with the use of measurement criteria of this study (see Table 3)
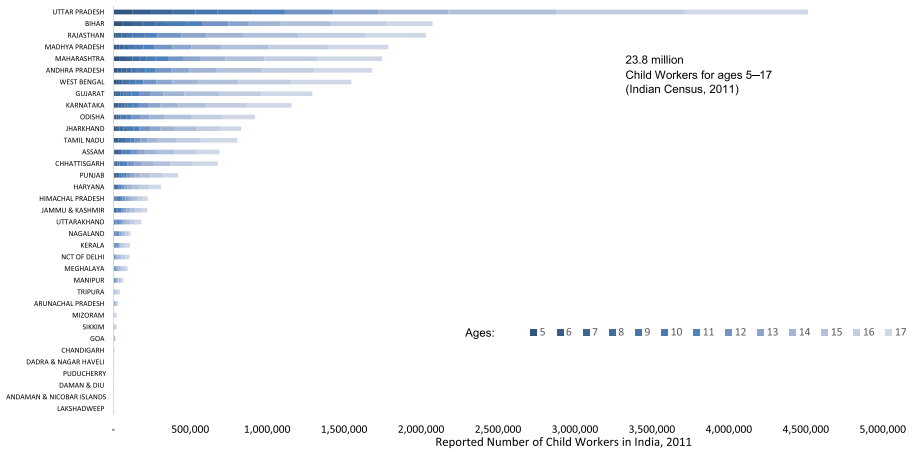
Fig. 3 Child Workers by Age and State from the Indian Census, 2011. Notes: Main workers (working for more than 6 mo) + marginal workers (working less than 6 mo)

This study uses the aggregated and weighted numbers of child labourers per age i per state j (13 ages * 35 states). In 2011/12, there were 35 states, which have shown different trends in the prevalence of child labour. This study estimates the number of child labourers by using a Bayesian hierarchical Poisson log-normal model and obtaining the posterior distributions for child labour by age and state. Based on these posterior distributions, we produce summaries such as medians and posterior predictive intervals as the point and interval estimates of the number of child labourers, respectively.

The specification of all models is presented in Appendix Table 6. By μ_{ij}, we denote a key parameter in child labour estimates: the true ratio of children in child labour to all children of age i and in state j. We use $n.a_{ij}$ and $n.b_{ij}$ to denote the sample sizes by age i and state j in the IHDS and the NSS, respectively, weighted to adjust regional
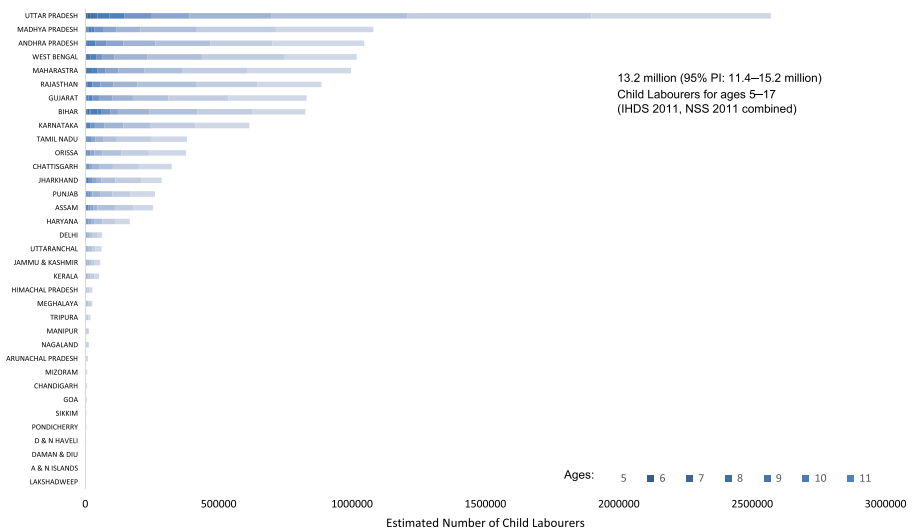

Fig. 4 Child Labour by Best Estimates, 2011/12

differences according to the Indian Census 2001 (Desai and Vanneman 2015; National Sample Survey Office 2013). Firstly, we estimate μ_{ij} separately for the NSS and the IHDS (Model 1 and Model 2). Then we combine the IHDS and NSS datasets in Model 3 (a Poisson model) and Models 4.1 to 4.3 (Poisson log-normal models). Thus, for each model, the expectation of the Poisson model is the outcome of the product of μ_{ij} and either $n.a_{ij}$ or $n.b_{ij}$. By using each model, we obtain a suitable expectation parameter for estimating any integer count (Gelman et al. 2013, pp.42–44). In Models 4.1–4.3, we use a discount parameter υ to capture a possible under-counting of child labourers in the IHDS data compared with the NSS.

In the models, $y.a_{ij}$ and $y.b_{ij}$ represent the observed counts of child labourers in age-state group ij in each survey. The $y.a_{ij}$ and $y.b_{ij}$ are assumed to be drawn from the Poisson distribution with the true ratio of children in child labour to all children (μ_{ij}) multiplied by the sample size ($n.a_{ij}$ from the IHDS or $n.b_{ij}$ from the NSS). Poisson distribution is a natural candidate for modelling counts of persons or, more precisely, counts of "events" where a child is identified as a labourer by our definition (see Section 2). By $\hat{y}_{ij}$ we denote the posterior median of the predicted distribution of the number of child labourers aged i and in state j. This value can be obtained by projecting the estimated rate of child labour, μ_{ij}, onto the population, N_{ij}. Then $\hat{y}_{i+}$, the sum of the $\hat{y}_{ij}$ for all states, shows the number of child labourers by each age.

The models include a few explanatory variables, such as age (x_i) and the log-ratio of main workers (z_{ij}), obtained from the Indian Census 2011 (defined by its narrow definition of work), which explain the true child labour rate μ_{ij} in a log link function. Parameter β_0 denotes the intercept; β_1 and β_2 are the coefficients of the covariates x_i and z_{ij}.

Lastly, we assume over-dispersion in addition to the Poisson variability (see Table 6, the last column). In rows 4 and 5 of the last column, $\lambda.a_{ij}$ and $\lambda.b_{ij}$, are assumed to be normally distributed to incorporate over-dispersion in each dataset. In row 7, λ_{ij} also allows for overdispersion to predict the true child labour rate (ψ_{ij}). Overdispersion parameters ($\lambda.a_{ij}$, $\lambda.b_{ij}$ and λ_{ij}) allow the mean to vary by observation and explain more variability (Lunn et al. 2012, p.227; see Table 5). In Models 4.1–4.3, the true rate of child labour is ψ_{ij}, which is an adjusted mean of the Poisson distribution. Overall, this method of modelling permits a more robust description of the uncertainty of the measured child labour from two data sources.

To obtain posterior distributions, we have used the Markov Chain Monte Carlo (MCMC) method as implemented in the R packages JAGS (Plummer 2003) and R2jags (R Core Team 2018). After discarding the first 40,000 simulation runs, we implemented 360,000 iterations and thinned them by eight, producing an effective total of 40,000 posterior samples.

3.6 Prior Distributions

The priors for the intercept and coefficients of age and the log-rate of main workers β_k and $k = 1, 2, 3$, respectively, are assumed to be normally distributed with mean 0 and a large variance (i.e. small precision, which is the inverse variance $\tau = 1/\sigma^2$). These non-informative priors allow data to play a dominant role in the inference (Gelman et al. 2013).

$$\beta_k \sim \text{Normal}\left(0, 10^{-6}\right), k = 1, 2, 3$$

where 10^{-6} denotes precision. In Models 4.1–4.3, we assume a vaguely informative prior. Gelman (2006) suggests three different priors: uniform, inverse gamma, and half Cauchy. We have tested these priors and compared the sensitivity of the results to various specifications by using the DIC (Deviance Informative Criterion; Spiegelhalter et al. 2014) score: a model with a uniform prior (Model 4.1), inverse gamma prior (Model 4.2) and a half Cauchy prior (Model 4.3). As a result, we have chosen to use an inverse gamma prior, as specified below. The effect of the choice of the prior distribution is explained in the next section.

$$\tau.a, \tau.b, \tau \sim \text{Gamma}\,(0.5, 0.5)$$

In the combined data model (Models 3 & 4.1–4.3), there is an under-counting parameter, υ, that controls any systematic undercounting of child labour in the IHDS compared to the NSS. The under-counting parameter, υ, is assumed to follow a uniform distribution $\upsilon \sim$ uniform (0, 1) which reflects our lack of knowledge about the undercount.

4 Results

4.1 Comparing Models

We have reviewed the DIC to compare the goodness of fit as well as the complexity of the models. DIC is the posterior mean of the deviance plus the effective number of parameters (pD). Comparing DIC of Model 1 and Model 2 is not feasible because the NSS and the IHDS have different sample sizes. Model 4.2 has the lowest DIC among the other models (3, 4.1, 4.2, and 4.3). Table 4.

Both Model 1 (with the IHDS 2011/12) and Model 2 (with the NSS 2011/12) show a narrow posterior uncertainty, as they assume that the variances are equal to the means. The two models generate different predictions for child labour: Model 1 estimates this

Table 4 Aggregate Child Labour Estimation, India, 2011/12

Models	DIC	Deviances	pD	No. of Child Labourers[1]			
				Age 5–17	sd	Age 5–14	sd
Model 1	1319.3	1316.9	2.4	11,700,959	238,241	2,618,101	99,144
Model 2	3951.6	3848.5	3.1	14,103,812	158,733	4,390,055	79,728
Model 3	5387.0	5384.4	2.7	13,484,868	132,013	3,929,058	63,166
Model 4.1 (Uniform Prior)	2919.8	2367.4	552.4	13,545,971	915,409	3,329,530	258,347
Model 4.2 (Inverse Gamma Prior)	2894.3	2358.9	535.4	13,194,114	963,859	3,222,680	292,084
Model 4.3 (Half Cauchy Prior)	2914.4	2368.1	546.3	13,570,367	901,155	3,341,185	252,466

Notes: [1] Median; pD = var.(deviance)/2

number at 11.7 million (median) and its 95% predictive interval (PI) is 11.2–12.2 million; Model 2 at 14.1 million (95% PI: 13.8–14.4 million).

The variance of Poisson distribution should be equal to the mean, which may not realistically capture the over-dispersion in the data. Figure 5 shows over-dispersion happening in Models 1 to 3. Several observations are outside of the predicted posterior distribution. Thus a simple Poisson model does not capture the variability of data. Once the data from both surveys are combined, a Poisson model (Model 3) estimates the number of child labourers at 13.5 million (95% PIs: 13.2–13.7 million) for ages 5 to 17, and 3.9 million (3.8–4.1 million) for ages 5 to 14. However, it still does not capture the observed variability very well. No observations lie within the 95% PIs.

A Poisson log-normal model (Model 4.1–4.3) uses the adjusted mean of the Poisson distribution (ψ_{ij}), which allows over-dispersion (larger variance) of each parameter. Models 4.1–4.3 reduce deviances compared with Model 3. The result indicates that Models 4.1–4.3 are superior to Model 3, since the DIC value is smaller for these than for Model 3 even after incorporating the large penalty of complexity.

The simulation of different priors implies that the best prior is an inverse gamma prior (Model 4.2), although the difference in DIC between the models is not large. In the next section, we based our predictions of child labour in Indian states using Model 4.2.

4.2 Result of Poisson Log-Normal Model

4.2.1 Parameter Estimation

The MCMC algorithm shows proper convergence in the posterior parameters of interest in Model 4.2. Figure 6 shows a histogram of MCMC samples taken from Model 4.2, median and 95% intervals of estimated parameters. The intercept (β_1), the coefficient of age (β_2), and the coefficient of the log ratio of main workers from the Indian Census (β_3) show stable convergence, resulting in statistically meaningful outcomes. The coefficient for age is positive, as the rate of child labour increases with age. The coefficient for the rate of main workers
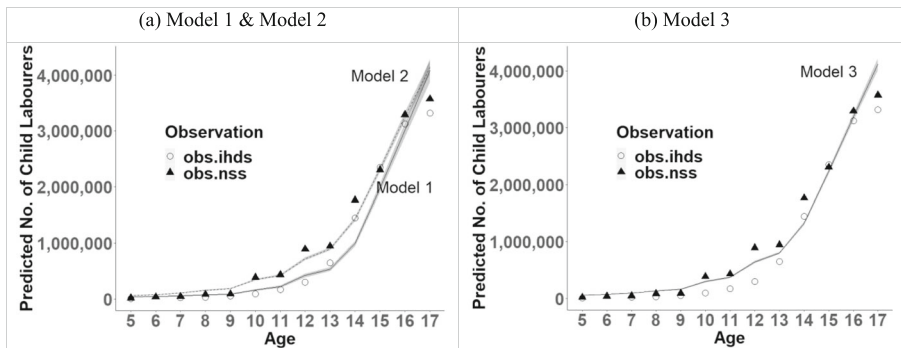

Fig. 5 Posterior Mean Numbers of Child Labourers in India, 2011/12, by Ages (Models 1–3). Notes: the shaded areas indicate 95% intervals
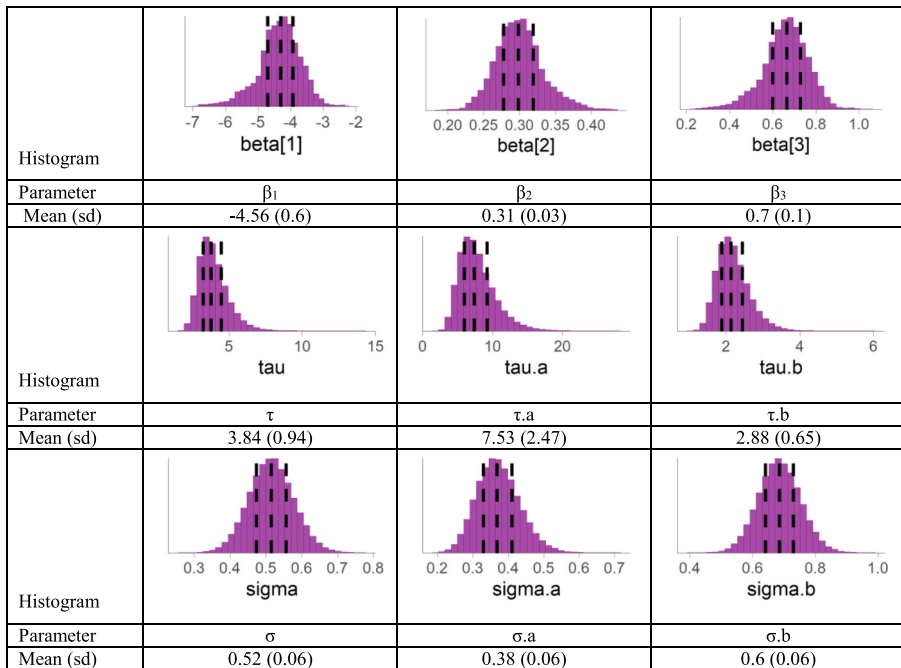
Histogram	beta[1]	beta[2]	beta[3]
Parameter	β_1	β_2	β_3
Mean (sd)	-4.56 (0.6)	0.31 (0.03)	0.7 (0.1)
Histogram	tau	tau.a	tau.b
Parameter	τ	$\tau.a$	$\tau.b$
Mean (sd)	3.84 (0.94)	7.53 (2.47)	2.88 (0.65)
Histogram	sigma	sigma.a	sigma.b
Parameter	σ	$\sigma.a$	$\sigma.b$
Mean (sd)	0.52 (0.06)	0.38 (0.06)	0.6 (0.06)

Fig. 6 Histogram of Parameter Estimates. Notes: Model 4.2; burnin = 40,000; iterations kept = 40,000 (2 chains); thin by 8; median and 95% lower and upper bounds

from the Census is also positive. Using the Indian Census as auxiliary information increases the mean rate of child labour and reduces the gap between ages. As a result, it contributes to smoothing the graph of child labour by ages.

The under-counting parameter (υ) indicates the posterior mean 0.81 (95% PI: 0.77–0.94), which shows that there is a slight undercount (around 2.3 million child labourers aged 5 to 17) in the IHDS. Under-counting of the IHDS is mostly caused by a lack of information on household chores. Accordingly, the combination model puts greater weight on the observations from the NSS.

According to the results of our final model (Model 4.2), the number of child labourers (ages 5–17) is estimated at 13.2 million (4% of the child population aged 5–17) in 2011/2012. The 95% PI for the number of child labourers ranges from 11.4 million to 15.2 million. That is, with a probability of 95%, the true number of child labourers lies within this interval. The estimate for ages 5–14 is around 3.2 million and 95% PIs are 2.7 to 3.8 million. Figure 7.

The number of child labourers is estimated to be higher than the figure proposed by the Indian Census for main workers aged 5 to 17 (working more than 6 mo in any economic activity). The number of child labourers surveyed by the Indian Census is 11.8 million for ages 5 to 17, which is smaller than our point estimate but lies within the 95% PI. The Indian Census figure of main workers for ages 5 to 14 (4.35 million) is larger than the forecast number of child labourers and lies outside the 95% PI. Our estimates do not adequately capture the child labourers under the age of 10, due to there being only a small
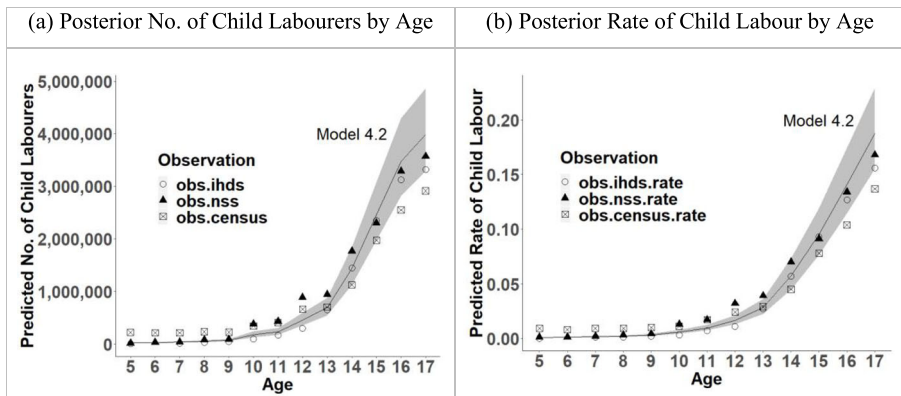
(a) Posterior No. of Child Labourers by Age	(b) Posterior Rate of Child Labour by Age

Fig. 7 Results of Bayesian Poisson Log-Normal Model Using a Combination of Datasets (India). (**a**) Posterior No. of Child Labourers by Age, (**b**) Posterior Rate of Child Labour by Age. Notes: the shaded areas indicate 95% intervals; the census data is presented in a graph for purposes of comparison

number of observed child labourers at the early ages (see Section 3.5). Child labourers who are under 10 might be underestimated because some children who work are categorised in "other" or "too young", so they are not included as child labourers.

4.2.2 Summary of Findings

Combining datasets with a Poisson log-normal model provides a reliable figure for child labour in India and incorporates uncertainty in models supported by the use of available observations. The use of an under-count parameter is a useful way to reduce any systematic error, which might be caused by the lack of information in one dataset compared to the other. In addition, the model shows the clearest age trend of child labour, as it has confirmed the effect of smoothing the variation between ages. The auxiliary variable from the Indian Census ratio of child labour introduces smoothing of the trend of child labour by age; it reduces the gaps between the single-year age groups.

   A large increase in the percentage of child labourers appears at age 14 when many more children are likely to be involved in labour compared to earlier ages. This trend is related to the education system in India. In Indian law, the Right of Children to Free and Compulsory Education Act 2009 defines education as free for children from 6 to 14 years, but some children become full-time workers before they move to secondary school. This finding supports the importance of secondary school education in preventing children from becoming full-time workers (Chaudhri et al. 2003; Charudhri and Wilson 2000).

5 Discussion and Policy Implications

The discrepancies between definitions and measurement of child labour have long been discussed. Our research has tried to reduce the gaps. We have relied

on the international definition of child labour and offered a novel solution to estimate it by applying current international criteria within one country, India. Through using the available datasets and accounting for their limitations, we provided authoritative estimates for 2011/12. This study includes a child's work in household chores as one of the main aspects of child labour if the child worked more than the specified time threshold for their age group. The recent proposition of the ILO is that not only is children's domestic work undertaken with the aim of creating manufactured goods or any product that competes in the market and counts toward GDP considered child labour, but also domestic work, i.e. services for other household members. This domestic work must also exceed certain time thresholds, relative to the age group, to be counted as child labour. This particular work on domestic tasks is outside GDP but inside the 'general production boundary', and is thus non-SNA work (ILO, 2018a).

Our estimation of child labour reflects the best and most recent knowledge regarding the differing prevalence of child labour across India's states. Given the definition of child labour using the hazard elements used by the ILO and the time criteria used by UNICEF, the probability distribution represents 13.2 million for child labourers aged 5 to 17 overall, which is larger when compared, for example, to the Indian Census, where this number was 11.8 million. As new datasets emerge, the method can be used further.

The study provides an accurate number of child labourers based on a definition consistent with the one that the international agencies offer. The focus is on any work that is harmful to children's development, including both economic work and unpaid household services that require considerable time. In addition, time boundaries play a key role in our measurement of child labour. The selected categories of working hours, based on the UNICEF guidelines, help with the capturing of child labour as a category representing work that is harmful to children. However, the most recently suggested UNICEF measures do not include children working in hazardous industries as child labourers and thus might underestimate child labour. Therefore, we suggest using both the concept of hazardous work and the concept of time thresholds to calculate child labour. With the data-combining methods, which are not data-pooling methods, achieving these estimates becomes a feasible calculation task.

There are a few further critical and practical points to make about using hourly thresholds for discerning child labourers. Working hours is a broad term in itself: for example, some datasets use daily working hours, and some use weekly working hours. In the Indian case, the IHDS 2011/12 has daily working hours and the NSS 2011/12 has a roughly weekly basis work intensity. However, seasonality is not handled well in either survey. The best suggestion for any further survey on working hours is to collect hourly information as specifically as possible within a limited time and budget. Daily working hours during a reference week, as well as a relevant period of actually working, perhaps during 2–3 seasons, is required. Also, for weighting, the NSS uses 2001 Census which may be out of date at the time of data collection in 2011/12.

The other obvious concern is the under-reporting of child labour. Although the IHDS provides better time information, it does not cover the domestic sector, so we

needed to use an undercount parameter. Furthermore, the observations used in this study fail to capture some of the child labour in early age groups, below the age of 10. Considering that a large number of children's labouring statuses are not reported or under-reported, relying only on working hours might lead to an underestimation of child labour.

A simplistic counting of child labour relying on one single dataset should be avoided. As a minimum, a clear method of calculation is required that is comparable to international standards and definitions. This piece of research makes a significant methodological contribution to child labour studies in several ways. Firstly, it has introduced the use of a relative weight and multiplied it with a true population, so that we can reduce the amount of error related to the population ratio of any survey.

Secondly, we demonstrate how a Bayesian hierarchical model can be used to combine different datasets to benefit from an increased sample of observed child labourers in two data sources, especially when datasets have different strengths. The combined-data approach can account for any potential systematic under- or over-counting of child labourers and provide more trustworthy estimates for an unknown parameter and the "true" estimated number of child labourers.

Third, we suggest using a Poisson log-normal model, which accounts for over-dispersion of counts. It provides an efficient way to incorporate uncertainty raised by the rare number of observations of child labour. The posterior probability distribution allows reliable estimation as it maximises the use of information using different datasets. In our case, a prediction is smoothed by using age as a covariate and by borrowing information from the Census data, where the less precise definition of child labour is used. Further research can be developed with other multi-dimensional variables to explain the prevalence of child labour.

We suggest the following implications for policymakers. Our results recognise that unpaid household services are non-ignorable aspects of child labour in India. The Indian Child Labour Amendment Act (2016) allows children to help family and work in a family business. However, if unpaid household service work exceeds the time thresholds of a reference age, it is regarded as hazardous according to both UNICEF (2019) and the ILO (2017). As a large number of child labourers are engaged in unpaid household services in India, there should be more support for children who spend long hours on housework and so are deprived of education. Secondly, the Amendment Act does not provide time limitations for child work. Setting up maximum working hours will be an important next step towards harmonising with the international standard (e.g. 40 h a week for ages 16–17 in the UK; 35 h a week for ages 15–17 in South Korea). Lastly, the profile of child labourers by age produced by our model shows a clear age trend for when children become labourers. It is found that children might become full-time workers after completing elementary school (ages 13–14) or before entering secondary school at the age of 14. Although further investigation of the relationship between education and child labour is needed, interventions are necessary for children at those ages.

Appendix

Hazardous Industries	
IHDS 2011/12	NSS 2011/12

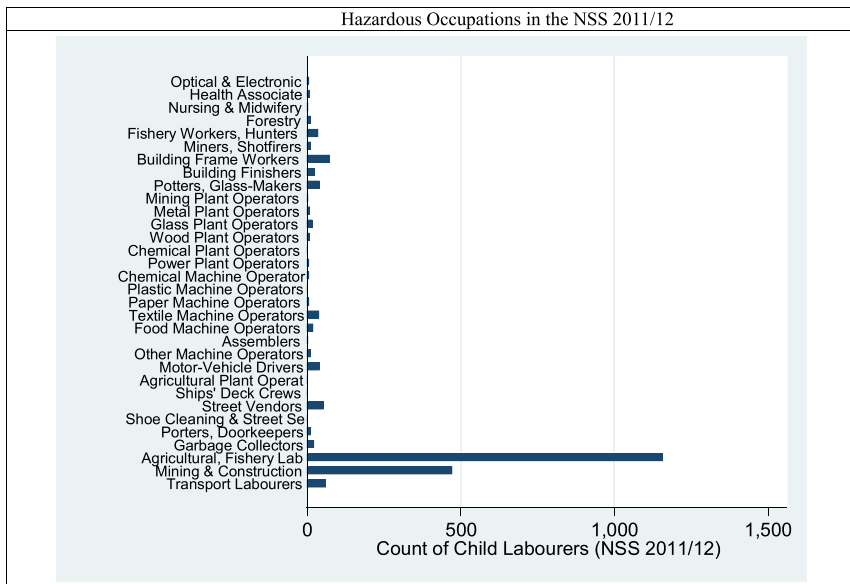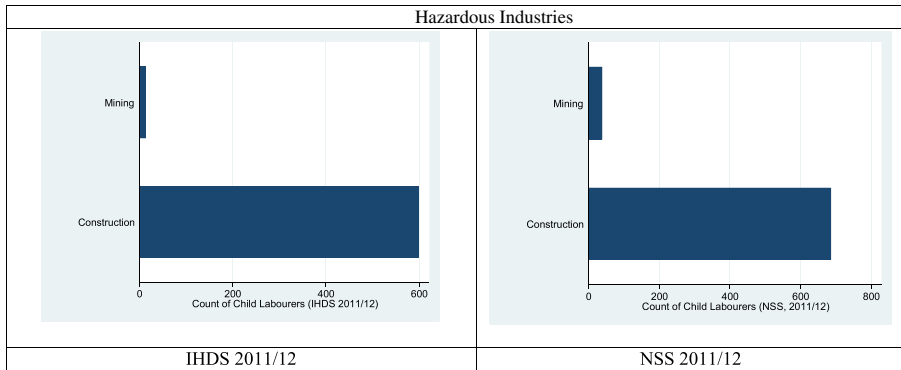Hazardous Occupations in the NSS 2011/12

Fig. 8 Count of Child Labourers in Hazardous Industries and Occupations from the IHDS and NSS, 2011/12. Notes: Weighted Count, Source: NSS 2011/12 & IHDS 2011/12
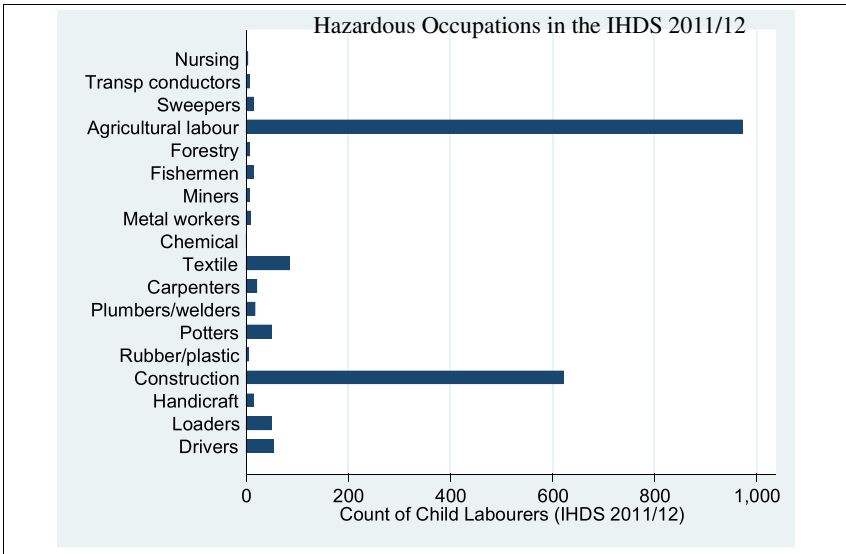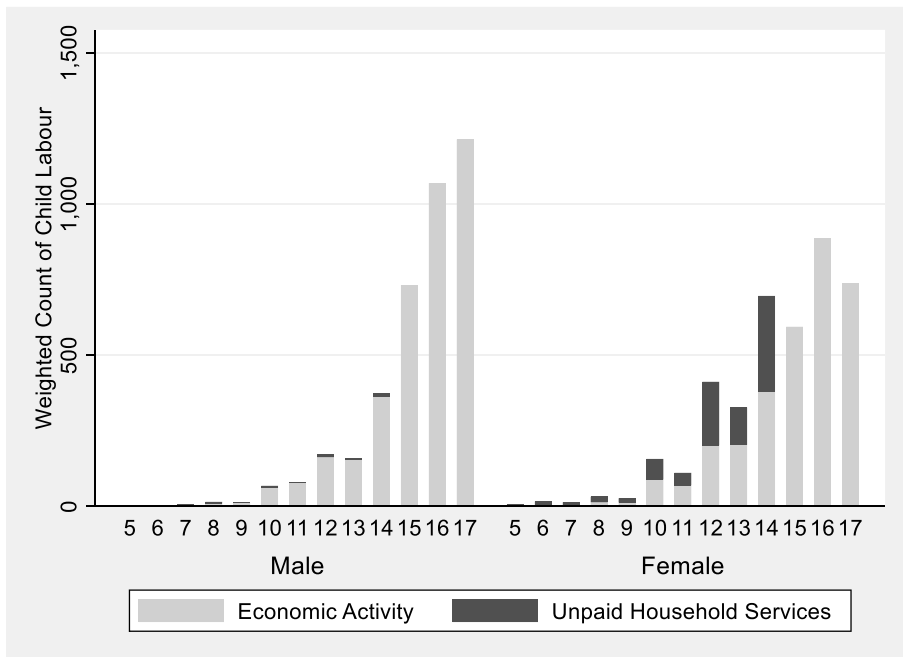
Fig. 8 (continued)

Fig. 9 Child Labour by Economic Activity vs Unpaid Household Services. Notes: Weighted count with our criteria measuring child labour; unpaid household services work is defined by principal activity status code 92 ("attended domestic duties only")

Table 5 Over-dispersion and undercount

Over-dispersion

The use of an overdispersion parameter does not affect the coefficient estimates. Instead, this well-established parameter increases the standard error of each parameter estimate. (In terms of standard usage, this would widen the confidence interval. In this paper it tends to widen the predictive interval.) The overdispersion parameter arises to compensate for a situation when a Poisson model assumption is not met. The assumption in question is that the "mean" equals the "variance" of the key random variable. Since the assumption does not seem to hold in the current application, the adjustment factor that is normally distributed around zero is inserted into the equation as a constant term. Whilst allowing a better fit and smoothing the results, with improved behaviour of the likelihood, this does not change the estimates at all for the percentage or number of "child labourers". An alternative method of accounting for over-dispersion to the one presented in this article is by assuming that counts of data follow a negative binomial distribution.

Undercount

We have used υ (read as "upsilon") as our notation for the undercount. The undercount is a measure of the broad under-representation of the dependent variable: in this case, the number of "child labourers" in one dataset versus another dataset. We only need an undercount parameter when two datasets are used in the same model. It has been used in a similar context when comparing data from different sources in, e.g. Wiśniowski 2017; Wiśniowski et al. 2013. The use of an undercount parameter is consistent with the optimal use of the information that we know to be true. Here, the Indian NSS dataset is considered a baseline for calculating this parameter. The IHDS dataset is thought to have an undercount because it does not have records of children's domestic work. Because the NSS does have such records, we consider there to be a broad bias downward in those who do excessive amounts of domestic work (above the child labour thresholds) in the NSS. In summary, predictions from our model ignore the undercount, and the level of "undercount" is an approximation of the broadly based bias due to lack of domestic-work measures in one survey vs the other survey.

We explain in a little more detail two innovations in the models of this paper

Table 6 Summary of the Models

Types of Models	Model 1 IHDS (Poisson)	Model 2 NSS (Poisson)	Model 3 Combining two datasets (Poisson)	Models 4.1–4.3 Combining two datasets (Poisson Log-Normal)
Likelihood	$y.a_{ij} \sim$ Poisson($\mu.a_{ij} * n.a_{ij}$)	–	$y.a_{ij} \sim$ Poisson($\mu_{ij} * n.a_{ij}$)	$y.a_{ij} \sim$ Poisson($\upsilon * \mu.a_{ij} * n.a_{ij}$)
	–	$y.b_{ij} \sim$ Poisson($\mu.b_{ij} * n.b_{ij}$)	$y.b_{ij} \sim$ Poisson($\mu_{ij} * n.b_{ij}$)	$y.b_{ij} \sim$ Poisson($\mu.b_{ij} * n.b_{ij}$)
Overdispersion	–	–	–	$\log(\mu.a_{ij}) = \log(\psi_{ij}) + \lambda.a_{ij}$
				$\lambda.a_{ij} \sim$ Normal $(0, \tau.a)$
	–	–	–	$\log(\mu.b_{ij}) = \log(\psi_{ij}) + \lambda.b_{ij}$
				$\lambda.b_{ij} \sim$ Normal$(0, \tau.b)$
Prediction	$\widehat{y}_{ij} \sim$ Poisson ($\mu.a_{ij}*N_{ij}$)	$\widehat{y}_{ij} \sim$ Poisson ($\mu.b_{ij}*N_{ij}$)	$\widehat{y}_{ij} \sim$ Poisson ($\mu_{ij}*N_{ij}$)	$\widehat{y}_{ij} \sim$ Poisson ($\psi_{ij}*N_{ij}$)
	$\widehat{y}_{i+} = \sum_i \widehat{y}_{ij}$	$\widehat{y}_{i+} = \sum_i \widehat{y}_{ij}$	$\widehat{y}_{i+} = \sum_i \widehat{y}_{ij}$	$\widehat{y}_{i+} = \sum_i \widehat{y}_{ij}$
Model for true child labour rate	$\log(\mu.a_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij})$	$\log(\mu.b_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij})$	$\log(\mu_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij})$	$\log(\psi_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij}) + \lambda_{ij}$
				$\lambda_{ij} \sim$ Normal$(0, \tau)$

Notes: i – age; j – state; τ is a precision (inverse variance)

Table 7 Summary of Child Labour Data in the IHDS, the NSS and the Indian Census

| | IHDS[1] | | | | NSS[2] | | | | Indian Census | | |
| | Relatively Weighting | | | Gross Weighted | Relatively Weighted | | | Gross Weighted | | | |
Age	No. of Child Labourers (y.a)*	No. of Children in Sample (n.a)	Ratio (y.a/n.a)* Population(N)	No. of Child Labourers	No. of Child Labourers (y.b)	No. of Children in Sample (n.b)	Ratio (y.b/n.b)* Population(N)	No. of Child Labourers	Population(N)	No. of Main Workers[3]	No. of Main and Marginal Workers[4]
5	1	4876	5342.12	4908	8	13,589	15,335	8303	26,048,171	223,354	430,785
6	–	4831	–	–	18	14,152	32,622	27,000	25,647,854	211,068	442,565
7	4	5457	18,193	22,419	22	13,027	41,917	32,000	24,820,355	214,041	491,150
8	5	4684	28,780	22,525	47	15,932	79,537	78,000	26,961,440	234,439	583,419
9	9	4196	50,230	48,085	40	10,833	86,471	69,000	23,418,444	225,906	585,719
10	17	5640	92,066	87,647	225	17,988	382,059	440,000	30,544,351	344,651	925,032
11	29	4196	170,944	129,918	193	11,144	428,360	390,000	24,733,883	409,365	1,003,678
12	70	6504	299,949	365,350	583	18,292	888,254	1,200,000	27,869,538	663,856	1,579,741
13	140	5240	648,541	660,536	485	12,502	941,679	940,000	24,273,967	699,458	1,636,946
14	324	5669	1,443,139	1,613,758	1071	15,342	1,762,695	2,000,000	25,250,481	1,127,109	2,449,628
15	436	4691	2,346,879	2,063,072	1326	14,548	2,301,494	2,500,000	25,250,481	1,975,126	3,865,154
16	612	4815	3,124,739	2,956,737	1956	14,617	3,289,798	3,600,000	24,584,341	2,552,054	4,738,080
17	737	4715	3,315,434	3,450,573	1951	11,595	3,568,955	3,500,000	21,210,681	2,911,827	5,047,586
Total	2384	65,514	11,544,238	11,425,529	7925	183,561	13,819,175	14,784,303	330,613,987	11,792,254	23,779,483

Notes: 1), 2) Our definition of child labour is applied; 3) main workers (working more than 6 mo); 4) marginal workers (working less than 6 mo); the definitions of child workers of the Indian Census do not correspond to the definitions used in IHDS or NSS

Table 8 Estimates of Child Labour by Bayesian Models

Age	Model 1 (IHDS) No. of Child Labourers[1]	95% PI LL[2]	HL[3]	Model 2 (NSS) No. of Child Labourers[1]	95% PI LL[2]	HL[3]	Model 3 (Combination, Poisson) No. of Child Labourers[1]	95% PI LL[2]	HL[3]	Model 4.2 (Combination, Poisson Log-Normal) No. of Child Labourers[1]	95% PI LL[2]	HL[3]
5	37,040	29,439	46,396	58,850	52,173	65,918	53,131	48,068	58,865	18,635	12,349	27,469
6	43,641	35,779	52,929	77,790	70,238	85,703	68,650	63,000	74,876	24,731	17,392	34,908
7	54,378	45,760	64,192	103,561	94,975	112,418	90,229	83,804	97,169	37,565	27,044	53,258
8	73,650	63,422	85,140	153,376	142,598	164,458	131,365	123,285	139,732	54,750	40,719	74,145
9	86,395	75,725	98,102	186,045	174,888	197,656	158,204	149,870	166,942	70,295	52,692	95,591
10	159,559	141,922	178,695	345,394	327,510	364,125	293,773	280,194	307,961	173,988	129,882	238,627
11	221,169	200,471	242,847	423,080	404,536	441,988	369,143	354,932	383,708	233,032	178,415	310,999
12	419,892	387,271	453,081	717,915	692,150	744,040	641,154	620,933	661,848	453,096	354,457	589,039
13	531,603	496,261	567,967	895,156	867,523	923,409	801,921	779,977	824,472	688,439	543,064	879,992
14	990,742	39,958	1,039,762	1,429,311	1,394,349	1,464,918	1,321,528	1,293,408	1,350,510	1,448,009	1,132,768	1,852,563
15	1,978,574	1,897,530	2,060,682	2,317,759	2,264,396	2,371,082	2,238,103	2,194,621	2,281,819	2,459,231	1,968,253	3,078,760
16	3,017,526	2,891,381	3,147,408	3,258,159	3,177,513	3,340,424	3,199,082	3,131,798	3,267,308	3,472,349	2,815,995	4,299,760
17	4,083,994	3,881,273	4,299,388	4,135,596	4,013,209	4,262,375	4,117,271	4,012,018	4,223,355	3,988,387	3,293,778	4,868,578
5–14	2,618,101	2,427,560	2,813,987	4,390,055	4,233,903	4,548,574	3,929,058	3,808,048	4,055,095	3,222,680	2,689,389	3,837,063
5–17	11,700,959	11,235,388	12,171,560	14,103,812	13,793,031	14,419,353	13,484,868	13,227,510	13,740,797	13,194,114	11,395,127	15,192,988

Notes: 1) median; 2) LL refers to low level (2.5%); 3) HL refers to high level (97.5%)

References

Ashagrie, K. (1993). Statistics on child labor. *Bulletin of Labour Statistics, 3*, 11–24.
Basu, K. (1999). Child labor: Cause, consequence, and cure, with remarks on international labor standards. *Journal of economic literature, xxxvii*(3), 1083–1119.
Basu, K., Das, S., & Dutta, B. (2010). Child labor and household wealth: Theory and empirical evidence of an inverted-U. *Journal of Development Economics, 91*(1), 8–14.
Basu, K., & Van, P. H. (1998). The economics of child labor. *The American Economic Review, 88*(3), 412–427.
Chaubey, J., Perisic, M., Perrault, N., Laryea-Adjei, G., & Khan, N. (2007). *Child labour, education and policy options*. New York: UNICEF.
Chaudhri, D., Nagar, A., Rahman, T., & Wilson, E. (2003). Determinants of child labour in Indian states: Some empirical explorations (1961-1991). *Journal of Quantitative Economics, 1*(1), 1–19.
Chaudhri, D. P., & Wilson, E. (2000). The challenge of child labour in rural India: A multi-dimensional problem in need of an orchestrated policy response.
Das, S., & Mukherjee, D. (2007). Role of women in schooling and child labour decision: The case of urban boys in India. *Social Indicators Research, 82*(3), 463–486.
Das, S., & Mukherjee, D. (2011). Measuring deprivation due to child work and child labour: A study for Indian children. *Child Indicator Research, 4*(3), 453–466. https://doi.org/10.1007/s12187-010-9097-8.
Desai, S., & Vanneman, R. (2015). India human development survey-II (IHDS-II), 2011-12. ICPSR36151-v2. Ann Arbor, MI: Inter-university consortium for political and social research [distributor], 31.
Dubey, A., Olsen, W., & Sen, K. (2017). The decline in the labour force participation of rural women in India: Taking a long-run view. *The Indian Journal of Labour Economics, 60*(4), 589–612.
Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis, 1*(3), 515–534.
Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.
Giri, A. K., & Singh, S. P. (2016). Child labour and 'nowhere' children in post-reforms India. *Indian Journal of Human Development, 10*(1), 97–110. https://doi.org/10.1177/0973703016654562.
Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology* (Vol. 561). John Wiley & Sons.
Hirway, I., & Jose, S. (2011). Understanding women's work using time-use statistics: The case of India. *Feminist Economics, 17*(4), 67–92.
ILO. (2013). Unpaid household services and child labour. In *Presented at the 19th international conference of labour statisticians*. Geneva: International Labour Office.
ILO. (2016). *Methodology: Global estimate of child labour, 2012–2016*. Geneva: International Labour Office.
ILO. (2017). *Global estimate of child labour: Result and trends, 2012–2016*. Geneva: International Labour Office.
ILO. (2018a). Resolution to amend the 18th ICLS resolution concerning statistics of child labour. ILO. https://www.ilo.org/wcmsp5/groups/public/%2D%2D-dgreports/%2D%2D-stat/documents/meetingdocument/wcms_648624.pdf. Accessed 11 January 2020.
ILO. (2018b). Hours of work. International Labour Office. https://www.ilo.org/ilostat-files/Documents/description_HRS_EN.pdf. Accessed 11 January 2020.
Kambhampati, U. S., & Rajan, R. (2008). The 'nowhere' children: Patriarchy and the role of girls in India's rural economy. *The Journal of Development Studies, 44*(9), 1309,1341.

Levison, D., & Langer, A. (2010). Counting child domestic servants in Latin America. *Population and Development Review, 36*(1), 125–149.

Lieten, G. K. (2002). Child labour in India: Disentangling essence and solutions. *Economic and Political Weekly, 37*(52), 5190–5195.

Lunn, D., Jackson, C., Best, N., Spiegelhalter, D., & Thomas, A. (2012). The BUGS book: A practical introduction to Bayesian analysis. Chapman and Hall/CRC.

Ministry of Home Affairs. (2011). Census Data, https://censusindia.gov.in/2011census/population_enumeration.html. Accessed 11 Jan 2020.

Ministry of Labour and Employment. (n.d.). Census Data on Child Labour. https://labour.gov.in/childlabour/census-data-child-labour. Accessed 11 January 2020.

National Sample Survey Office. (2013). *India - Employment and Unemployment July 2011–June 2012*. http://mail.mospi.gov.in/. Accessed 11 January 2020.

Omoike, E. (2010). Child domestic labour: Fostering in transition? Child Slavery Now*: A Contemporary Reader, 203.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, pp. 1–10). Vienna, Austria.

Core Team, R. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing https://www.R-project.org. Accessed 11 January 2020.

Ray, R. (2000). Child labor, child schooling, and their interaction with adult labor: Empirical evidence for Peru and Pakistan. *The World Bank Economic Review, 14*(2), 347–367.

Ray, R. (2002). The determinants of child labour and child schooling in Ghana. *Journal of African Economies, 11*(4), 561–590.

Samantroy, E., Sekar, H. R., & Pradhan, S. (2017). State of child Workers in India. V.V.Giri National Labour Institute.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 76*(3), 485–493.

UNICEF. (2017). Access the data - child labour. UNICEF. https://data.unicef.org/wp-content/uploads/2015/12/Child-labour-database_Nov-2017.xls. Accessed 11 January 2020.

UNICEF. (2019). Child Labour – Notes on the data. https://data.unicef.org/topic/child-protection/child-labour/. Accessed 11 January 2020.

UNICEF. (n.d.). India statistics. *UNICEF*. https://www.unicef.org/infobycountry/india_statistics.html. Accessed 16 May 2019.

Webbink, E., Smits, J., & Jong, E. (2015). Child labor in Africa and Asia: Household and context determinants of hours worked in paid labor by young children in 16 low-income countries. *The European Journal of Development Research, 27*(1), 84,98.

Weiner, M. (1991). *The child and the state in India : child labor and education policy in comparative perspective*. Princeton, N.J. ; Oxford: Princeton, N.J. ; Oxford : Princeton University Press.

Weston, B. H. (Ed.). (2005). *Child labor and human rights: Making children matter*. Boulder, Colorado: Lynne Rienner Publishers.

Wiśniowski, A. (2017). Combining Labour Force Survey data to estimate migration flows: the case of migration from Poland to the UK. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 180*(1), 185–202. https://doi.org/10.1111/rssa.12189.

Wiśniowski, A., Bijak, J., Christiansen, S., Forster, J. J., Keilman, N., Raymer, J., & Smith, P. W. (2013). Utilising expert opinion to improve the measurement of international migration in Europe. *Journal of Official Statistics, 29*(4), 583–607. https://doi.org/10.2478/jos-2013-0041.