A Bayesian Hierarchical Model to Estimate and Predict Child Labour in India

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities

2021

Jihye Kim

School of Social Science/ Social Statistics

Table of Contents

Word count: 71,300 (excluding bibliography, appendices and annexes)

List of Tables

List of Figures

List of Abbreviations

AC	Assembly Constituencies
CRC	Convention on the Rights of the Child
CRY	Child Rights and You
DHS	Demographics and Health Surveys
DIC	Deviance Information Criterion
FA	Factor Analysis
FSU	First-Stage Units
GAD	Gender and Development
GDP	Gross Domestic Product
GOI	Government of India
GPNs	Global Production Networks
HDPI	Human Development Profile of India
HH	Household
HL	High-Level
ICLS	International Labour Organization Conference of Labour Statisticians
IHDS	India Human Development Survey
ILO	International Labour Organization
IWI	International Wealth Index
LFP	Labour Force Participation
LL	Low-Level
LOOIC	Leave-One-Out Cross-Validation Information Criterion
LPPD	Log Pointwise Posterior Predicted Density
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
MLT	Multipliers
MPCE	Average Monthly per Capita Expenditure
MSE	Mean Squared Error
N/A	Not Applicable
NCO	National Classification of Occupation
NCPCR	National Commission for Protection of Child Rights
NGO	Non-Governmental Organisation
NIC	National Industrial Classification
NPEGEL	National Programme for Education of Girls at Elementary Level
NSC	Combined Number of Sub-Samples
NSS	National Sample Survey
NSS	Number of Sub-Samples (p.89)
NSS EUS	National Sample Survey Employment and Unemployment Surveys
NSSO	National Sample Survey Office
NWNAS	Non-Workers and Not Attending School
OBC	Other Backward Classes
PC	Parliamentary Constituencies
PCA	Principal Component Analysis
PI	Prediction Interval
PLFS	Periodic Labour Force Surveys
PPP	Posterior Predictive P-Value
PS	Polling Stations
REML	Restricted Maximum Likelihood
S.D.	Standard Deviation
SC	Scheduled Castes

SDG	Sustainable Development Goals
SNA	System of National Accounts
SPP	Sampled Posterior P-Value
SRMR	Standardised Root Mean Squared Residual
SSA	Sarva Shiksha Abhiyan (‘Education for all’)
ST	Scheduled Tribes
UN	United Nations
UNCRC	UN Convention on the Rights of the Child
UNICEF	United Nations Children's Fund
UP	Uttar Pradesh
UT	Union Territory
WAIC	Widely Applicable Information Criterion
WID	Women in Development
WVS	World Value Survey
ZI	Zero-Inflated

The University of Manchester
Jihye Kim
Thesis submitted for the degree of Doctor of Philosophy in the Faculty of Humanities

A Bayesian Hierarchical Model to Estimate and Predict Child Labour in India

Abstract

Child labour – work identified as harmful to children aged 5–17 years – is not yet eradicated in India. Consequently, there is a strong demand for the number of child labourers to be measured with more accuracy. Furthermore, understanding the diverse routes to and causes of child labour is essential for the improvement of its estimation and prediction. In India, child labour is a combined result of social structure, institutions, and norms. Notably, the gendered aspects of child labour add complications: largely based on gender roles and attitudes, and norms in India. Accordingly, this study aims to offer a new way to gauge the number of child labourers in India by using advanced modelling and revealing the relationship between its socioeconomic and cultural attributes. This study mainly uses the Indian Human Development Survey 2011/12 and the National Sample Survey on Employment and Unemployment 2011/12 for the analysis of the child-labour problem in India.

This thesis is composed of three journal articles that address key conceptual and methodological questions regarding child labour. The first article investigates how to improve the estimation of child labour using a Bayesian hierarchical model within a combined data approach. The second examines the interactive effects of social group, class and gender on child-labour participation and child labourers' working hours. The third article reveals the impact of gender norms on child-labour risks in terms of household occupation and state. All of these are based on the gender and development approach, in which social and gender relations explain agents' decisions regarding child labour. Gender is constructed within social and cultural circumstances which decide whether and where boys and girls work.

In the first article, a Bayesian statistical method was used for analysis with a combination of the two representative datasets in India from 2011/12. A combined-data approach enables us to accumulate knowledge from multiple data sources, as it incorporates diverse priors and assumptions as a probability distribution. Second, a Bayesian hurdle Poisson model provided a consistent and efficient way to estimate the prevalence and magnitude of child labour and to reveal its gender-based trends. The third article suggested a child-labour risk model that could be incorporated together with the norms and socioeconomic variables used in the estimation of child labour, thereby providing a way to predict risks of child labour according to households' socioeconomic backgrounds and locations.

The main findings include a more accurate estimation of the number of child labourers across the whole of India. The best estimate for child labourers aged 5–17 years old in 2011/12 is 13.2 million (95-percent predictive intervals are 11.4–15.2 million). The suggested Bayesian model can be used to project the potential numbers of child labourers in the near future, through applying more accurate population data and priors. This study confirms the importance of including unpaid household services in the definition of child labour as a way to make girls' labour more visible. Moreover, this research has revealed new evidence of the relationship of gender, class, and social groups with child labour, such as the low levels of participation in the work force among female children in contrast with their long working hours as informal labourers. Lastly, the findings demonstrate that group-based norms and institutions, along with household structural positions, play clear roles in decisions regarding child labour. Norms supporting women in working outside the home are found to help reduce child labour. This research accounts for diverse aspects of child labour – social structure, institutions/norms and gender – advocating therefore for an integrated and long-term approach towards the child-labour problem. As the findings imply, the distribution of labour across children according to gender should be taken into account in dealing with child labour in India.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

My sincere thanks go to my supervisors, Professor Wendy Olsen and Dr Arkadiusz Wiśniowski, for their full support, guidance, and great supervisions. I appreciate Professor Wendy Olsen's inspiration, encouragement, and sharing so much knowledge and passion. I thank Dr Arkadiusz Wiśniowski for advising me with kindness and patience and for providing me with many valuable lessons. It was a privilege for me to have these wonderful supervisors. I would also like to thank Professor Yaojun Li and Professor Supriya Garikipati for examining the thesis and giving me insightful feedbacks. Thanks to Professor Amaresh Dubey for advice, support and constructive comments.

Colleagues and friends in India and the UK made my PhD full of joyful memories. I would like to thank my PhD colleagues, who we spent a difficult time together with. Thanks to Indian friends who helped me when I visited India in 2017 and 2018. I am particularly indebted to Abinash Lahkar and Dr Santosh Kumar Singh. My sincere thanks to my old friends and friends from the Alliance Manchester for their consistent support and friendship. I am grateful to Oksu Lee for being a great mentor.

I want to thank my husband, Sunbin Kim, for his love, companionship, and encouragement. I am in debt to my parents and parents-in-law. Their love and prayers made me complete this PhD. Thanks to my grandparents for their love and concern for me. I will remember them in my mind.

Chapter 1. Introduction

The position of children in the family, community and society is not simple. Their status is decided by the diverse relations surrounding them, and those relations are not always equal and sometimes exploitable. Child labour is a clear violation of children's human rights that arises from structural inequalities. In this sense, it is a global, structural problem based on social oppression of children. The recent emphasis on children's agency and children's right to work does not diminish its structural properties. Supporting children's agency includes not only respecting their decisions but also protecting and safeguarding them. The complexity of child labour – existing between individual decisions and structural pressure – makes it difficult to reach a simple solution; however, such a solution requires more profound analysis of how child labour is built upon the social structure that individuals belong to by looking at the diverse social relations of children.

Acknowledging the role of structural factors in child labour, this thesis contends that institutional and cultural variations are also associated with child labour. There are, in fact, enormous differences between countries, regions and sectors in terms of child-labour practices. Because children's participation in the labour force is formed through social construction, such as gender, ethnicity or social groups, its characteristics are varied, changing over time and different locations. For this reason, looking at one country's specific example, e.g., India in 2011/12 in this research, could give insight into the causes of child labour. Furthermore, gender roles and expectations are deeply entrenched in children's working lives. Nevertheless, any gendered pattern of child labour is obscure, despite the significant implications of revealing underlying structures and institutions. While it is known that girls' labour market participation is less than boys' in India, girls' huge contribution to production has not been discussed enough.

This study, in particular, is interested in revealing the magnitude and the causes of child labour in India in recent decades. Despite the economic growth in the country, child labour is deeply embedded in Indian society. The Indian labour market is not free from social prejudice and cultural bias against women and children. A long-lasting concern is that the child-labour problem is coupled with gender inequality. Gender-based discrimination and women's low level of participation in the labour force raises concerns for the status of children in India. Gender relations are not only decided by materialistic relations but also by cultural constructions. Accordingly, this study will need to balance the structural and constructional aspects of child labour while looking at gendered trends.

This research will draw great interest from policymakers, international agencies and non-governmental organisations (NGOs) who have active roles in addressing the child labour problem in India. International organisations, such as UNICEF and the ILO, lead the global discussion, make standards, and build a consensus on child labour. This research will help them define child labour with more precision by showing the way to predict child labour using an advanced model based on India's example. This research could be used by many non-governmental organisations such as Save the Children, Childline India Foundation and CRY (Child Rights and You), who have active interests in eliminating child labour in India. The main theme of this research, the gender roles of children, could help meet the needs and interests of these NGOs in protecting children with a more strategic approach.

In analysing determinants of child labour in India, this thesis is based on on-going theoretical discussions. Because of the complexity of the position of a child, more sensitive analysis and perspectives from more than one discipline are required in order to reveal the causes of child labour. With the growth in literature concerning child labour, its causes at the household level and its relationship with globalisation have been primarily revealed. On the one hand, child labour has been attributed to the family's economic conditions (such as lack of resources). On the other hand, it has been linked to the global structure of the production network. Comparatively recently, child labour as a socialisation process or emphasising children's autonomy has also been discussed. Although all of those explanations are plausible, explaining the variations in the status of children requires a more comprehensive theory. Further discussions are available in Chapter 2. This research found that the gender and development approach, which interlinks social and gender relations with regard to social inequality, is a suitable choice to investigate what causes and sustains child labour in India (see Sections 2.5, 2.6 and 7.2).

This study emphasises the importance of finding evidence that confirms children's socioeconomic status in a specific location and age group. It employs a quantitative analysis method to examine the combined influences of social structure and social construction on child labour. Moreover, Bayesian statistics can help to identify the less well-understood figures and characteristics of the problem (see Chapter 3. Methodological Review). The limits of research into child labour stem from considerable uncertainty regarding its inconsistent definitions, its illegal characteristics, and the limited available data. We need an innovative statistical method to overcome the enormous uncertainty related to those unaddressed issues. A data-combining approach can provide a way to integrate information for estimating numbers and assessing variables without merging different data sources (Sub-Section 3.2.2). In a Bayesian approach, combining multiple data sources is conceptually justifiable and methodically backed.

In this first chapter, I introduce the status of child labour in India through analysing reported figures and trends, and observations I made at local level through a scoping visit (the summary of the scoping visit is in Section 1.2 and full details are in Annex A). The chapter then further explains the research gaps that this study will attempt to address. These gaps then form the definition of the aim and scope of this research (Section 1.4). This provides a conceptual framework for analysing child labour (Section 1.5). The thesis structure is outlined at the end of this chapter (Section 1.6) and describes the linkage between articles and the authors' individual and shared responsibilities over them.

1.1. The Child Labour Situation in India

Child labour is an ongoing and prevalent issue in India, the nation with the largest number of child labourers in the world. The ILO (2017) reported that in 2016 the global number of child labourers stood at 151 million among children from 5 to 17 years old and that among them 62 million were in Asia. In India, estimates of child-labour numbers in 2012 ranged from 4.5 million to 27 million, according to different studies. The 2011 census figure for main and marginal workers aged 5–17 is around 23.7 million, a number which includes not only the child labourers but also child workers. In India, the main areas of child labour are the primary sectors, including farming, mining, and agricultural activities. Child labour also occurs in the service sectors such as construction, trade and hospitality services; some child labourers are also seen in less organised or domestic manufacturing production.

In fact, there has been a significant decline in the number of child labourers in India in the period 1990–2018. In 1991 Weiner (1991, p.155) found that the estimates of child-labour numbers ranged from 13.6 million to 44 million. This number was consistently reduced by 2011/12, a phenomenon largely due to India's economic growth. A more recent study reports that child labour numbers show a clear downward trend, decreasing from 2.7 million in 2011/12 to 1.1 million in 2017/18. (National Sample Survey 2011/12 and Periodic Labour Force Survey 2017/18 are compared using the usual principal status of children aged 0–14, Bhandari and Dubey, 2019.) Nevertheless, to date, child labour is a persistent concern in India, implying that economic growth might not be a panacea for child labour.

The Indian Census provides comparable figures for children in labour over several past decades. Table 1.1 shows the transformation in the number of main and marginal workers between 1981 and 2011. 2021 census data has not been released yet when this analysis is carried out. Since the 1981 Census, main (worked for more than six months) and marginal workers (worked for less than six months) have been separately recorded. In general, the numbers of Indian child workers aged 5–14 have gradually reduced over the last three decades. Moreover, estimates of main workers have shown a consistent decrease during those periods. However, there were fluctuations in the numbers of marginal workers.

Table 1.1. Changes in the number of working children in India (main and marginal workers)

	1981	1991	2001	2011
Total workers	13.6	11.3	12.7	10.1
Main workers	11.2*	9.1*	6.9**	5.8**
Marginal workers	2.4*	2.2*	5.8**	4.3**

Sources: Central Statistics Office, 2018; *Poddar et al. (n.d.) ;** Giri, 2017
Notes: in a million; ages 5–14

While the Indian census provides the aggregate number of child labourers based on the large population, there are a few limitations regarding the measurement. Firstly, it does not provide information on working hours of children that are the key criteria to separate child labour from child work (see Table 3.3. Comparing datasets - IHDS, NSS and Indian Census). Secondly, in the Census data, only four types of industries are provided at an aggregate level. According to the ILO’s standard, however, detailed information on children’s industries or occupations is required to classify hazardous jobs to children (ILO, 2017). Indian Census relies on the number of months spent on work as the key criterion of child labour – e.g. main workers mean children who work more than six months a year in economic activity and marginal workers working less than six months a year (Office of the Registrar General & Census Commissioner, 2011). This study uses detailed household surveys to estimate the number of child labourers so that the internationally used criteria of child labour can be applied.

Although child labour shows a general downward trend in India, questions could be asked regarding the sudden increase in the number of marginal workers between 1991 and 2001, which increases overall during that period. This can be explained by the growth of the informal sector. Since the Indian market reforms in 1991, the country has undergone tremendous economic change and the number of informal jobs has grown tremendously (Das and Das, 2009). Many industries were transformed into home-based work, and the number of marginal workers increased as a result (Ministry of Labour, 2001). The change in the child population is another potential reason. Chaudhri and Wilson (2000) pointed out that some states such as Uttar Pradesh and Rajasthan saw increases in their child populations between 1970 and 1990, which could be a leading factor in the increases in child labour in those states. This fact indicates that we should consider not only the number of child labourers but also the child-labour rate.

Most child labour occurs in the agricultural sector, but there is also some diversity in industries. In 2011/12, 63 percent of child labour occurred in the agricultural sector (calculated using the IHDS 2011/12). Previously, children employed in the agricultural sector represented about 75 percent of full-time child labourers in 1993/94 (National Sample Survey 50th round; Chaudhri and Wilson, 2000). However, over two decades, the concentration of child labour in the agricultural sector had been reduced to a large extent. This change implies that a certain portion of child labourers had moved over to the urban and informal sectors since the industrial structural reform in India during the 1990s.

The declining trend of women's participation in the labour force may also play a role in the incidence and prevalence of child labour (Table 1.2). Female labour-force participation increased between 1994 and 2005 and then showed a significant drop between 2005, 2012 and 2018. Andres et al. (2017) indicate that the addition of the female labour force happened in rural areas before the decrease was caused by the withdrawal of those rural workers. Women's intensive involvement in reproductive labour is a possible cause of the falling rate of their labour force participation (Rao, 2018; Bhandari and Dubey, 2019). The withdrawal of women from the labour market is a critical factor regarding child labour as it deepens the gender division of labour and allows households to accept more patriarchal norms (Abraham, 2013).

Table 1.2. Trends of labour force participation[1] (%), all-India, between 1994 and 2018

	1994	2000	2005	2012	2018
Male	84.1	82.8	82.7	79.1	74.6
Female	30.8	30.7	32.2	23.3	20.8
Total	58.1	57.2	57.8	51.6	48.0

Source: ILO STAT database (available at https://www.ilo.org/shinyapps/bulkexplorer39/?lang=en&segment=indicator&id=EAP_DWAP_SEX_AGE_RT_A, accessed 1 Aug 2020)
Notes: Labour force participation indicates the proportion of 'total' age population who actively participate in the labour market (ILO, 2020).

Compared to the general figure, children are more intensively concentrated in the agricultural sector (63 percent in IHDS 2011/12). Note that employment in the agricultural sector accounts for about 50 percent of general employment in 2010 (ILO, 2020). A study tracing child-labour-related products revealed that paddy, wheat, other crops, construction, cotton, and tobacco were the commodities most closely involved with child labourers in India, in 2011/12 (Gómez-Paredes et al., 2016). If large supply chains are considered, construction involved the largest number of child labourers, cotton was the most child-labour-intensive commodity, and tobacco was mainly produced by girls (ibid.).

[1] The LFP rate is the ratio of the people with self-employment combined with employment to the whole age-group population.

In 2011/12, based on the IHDS, the percentage split of boy and girl child labour does not show any critical difference in the primary sectors (Table 1.3: Boys=33% and girls=30% of all child labourers). The second largest section of child labour was construction, which involved 13 percent of child labourers, mostly boys. The service sector is more likely to engage boy labourers in construction, wholesale, and retail sectors: three sectors which together represent 15 percent of the total child labour. Meanwhile, the manufacturing sector takes 10 percent of the total child labour, including the 3 percent employed in home-based manufacturing. Small-scale domestic manufacturing is quite common in India, in which case children help production undertaken at home. Interestingly, an almost equal number of boys and girls are found in domestic manufacturing production in the IHDS 2011/12.

Table 1.3. Industrial composition of child labour in 2011/12

Industry	Male		Female		Total	
	Weighted count	% of total child labourers	Weighted count	% of total child labourers	Weighted count	% of total child labourers
Agriculture, forestry & fishing	1,528	33%	1,415	30%	2,943	63%
Mining & quarrying	12	0%	2	0%	14	0%
Manufacturing	213	5%	98	2%	311	7%
Home-based manufacturing	66	1%	66	1%	132	3%
Construction	473	10%	114	2%	587	13%
Wholesale & retail	213	5%	83	2%	296	6%
Restaurants & hotel	19	0%	3	0%	22	0%
Other services	277	6%	64	1%	341	7%
Total	**2,801**	**60%**	**1,845**	**40%**	**4,646**	**100%**

Source: IHDS 2011/12
Notes: ages 5–17; survey weights are applied; market labour only (children aged 5–11 who worked over 1 hour a week and children aged 12–17 who worked more than 38.5 hours a week)

These simple descriptive comparisons suggest that child labour mainly happens in the informal sector, including farming, informal manufacturing and services using cheap labour. As mentioned earlier, economic growth has reduced the scale of child labour, but as the informal economy is growing, child labour might be sustained. Another clear trend among child labourers in India is the gendered division of work. Boys are involved in both agricultural and non-agricultural industries, and girls are involved in the agricultural sector as marginal farmers or agricultural labourers but less likely to work in services or manufacturing. There is a role played by gender norms which assign different jobs and work environments to boys and girls. In general, in 2011/12, 60 percent of child labourers were males, and the other 40 percent were females (Table 1.3). As the participation of girls in the labour market is culturally limited, they are largely found in the agricultural sector. Furthermore, many girls work in domestic areas without being recognised as labourers. Given boys and girls have different economic roles and work in different industries, disregarding gender can misrepresent the realities that girls and boys experience. A close examination of the industrial structure is necessary to avoid any bias caused by the different gender roles for children in India.

This study recognises that gender inequality is closely associated with child-labour risks. In the informal sector, in particular, female labour and child labour are significantly related. On the one hand, female workers bring children to the workplace, in which women and child labour are integrated. Although an early reported case, the match factories of Savakasi saw both women and children constitute a large part of the industry (Narayan, 1988, p.158). On the other hand, women’s low level of economic participation can be a risk factor for children. If female adults are prevented from participating in the labour market, the economic burden might increase the need for child labour. For example, in northern India, where female seclusion is common, children’s participation in the labour market is more prevalent. This study focuses on the latter relationship whereby women’s lower economic position might increase child labour risks.

Before proceeding, this study focused on child labour in 2011/12. IHDS 2011/12 and NSS 2011/12 were the most recent available datasets at the starting point of this study. It should be noted that in 2011/12, which is the focus year of this study, some types of work were allowed for children in India according to the Child Labour Prohibition and Regulation Act, 1986 (Child Labour Act 1986, GOI, 1986). The Child Labour Act 1986 prohibits children under the age of 15 from being involved in 18 specified hazardous occupations and 65 processes. Child labour cases occurring in sectors other than those restricted occupations were not illegal. In addition, children aged 15 years or more were not restricted by the Act. It is only since 2016 that child labour became entirely illegal (among children under the age of 14) in India. In 2005, the National Commission for Protection of Child Rights (NCPCR) was established, an organisation responsible for ensuring that all laws, policies, and programmes are in accordance with suggestions from the UN Convention on the Rights of the Child (UNCRC; Central Statistics Office, 2018). The Child Labour Amendment Act was then established in 2016, which bans any employment of a child under the age of 14 and makes it illegal for adolescents aged 14–17 from being employed in hazardous occupations and processes.

1.2. Review of the Scoping Visit

This section provides a general overview of the situation regarding child labour in India based on a scoping visit (during 7–26 January 2018). The aim of the visit was to understand current trends in child labour in India and to obtain wide-ranging opinions from diverse experts and stakeholders. This field visit was not intended to intervene with children but performed to have advice from local experts and professionals. I used the University of Manchester Ethical Decision Tool for this visit, and ethical approval was not required. While this study mainly uses the quantitative research method, some reflections of the scoping visit are presented in this sub-section to provide a general understanding of child labour in India. A detailed summary of the scoping visit is provided in Annex A. As this visit was made in the early stages of the research, the interview was not structured but dealt with broader issues of child labour and was later summarised by themes. I used English for the interviews, but the local languages are Hindi, with some Bhojpuri and Urdu.

In this scoping visit, discussion with stakeholders (excluding children) was not the only method employed as informal observations were also used. The following summaries are based on both interviews and observations. Interviewees included people from the government research institute, non-governmental organisations (NGOs) and academics researching or working in the field of child labour in India. Topics ranged from general views on child labour to more specific areas such as education, working areas and working hours of child workers and labourers. In addition, I visited NGO projects in rural and urban areas. The project visits provided me with a clear view of family socioeconomic conditions and community environments. In this subsection, I clarify a few topics regarding child labour based on the findings of the visit, which ultimately helped decide the scope of this study.

- **Informal and formal/urban and rural child labour**

Child labourers are prevalent in the informal sector, including agricultural labour, construction, services work, and small-unit production or home-based manufacturing. Moreover, urban and rural areas have different types of informal child labour. In rural areas, child labourers are involved in agricultural jobs, family farms and domestic chores. The interviewee from the school explained that children help family with farming on a regular or seasonal basis; therefore, they are absent from schools when the busy season approaches. Fetching water, taking care of their siblings, doing laundry, and washing are common tasks for children in villages. The same interviewee said that many children are involved in domestic work after school, but people do not consider children helping families child labourers. Children working in restaurants were often observed in the town and cities.

It is difficult for outsiders to discover children working in the formal sector, such as firms or factories. The manager of the carpet company said that employers are well aware that child labour is illegal. On one occasion, in Uttar Pradesh, several carpet companies were criticised for recruiting children. To avoid further consumer complaints, the carpet company that I visited now uses a monitoring system to screen the local production units through video cameras. Public pressure against child labour is intense, especially involving commodities exported abroad. There is also growing legal pressure against child labour. Furthermore, one local professional said that industrial structural changes had made child labour less popular in the carpet industry; as the industry has shrunk, there is no longer demand for child labour. Production has become mechanised, while high-skilled jobs cannot be taken on by children. Consequently, there has been a sharp decrease in child labour over recent decades in the carpet industry, and it is hard to see children becoming involved in that industry any longer: at least, in factories. One interviewee in the community told me that child labour no longer exists in the carpet industry, but the children might be employed in restaurants, motorbike shops, and tea shops.

The situation in big cities, such as Delhi, is different from that of villages or towns. Children are involved in more diverse industries, including construction, household subcontracting and other services. Construction is one of the acknowledged sources of child labour as many construction workers are migrants from rural areas and bring their children to work. The worst forms of child labour are easily found in urban areas in which a considerable number of children work as street vendors; some of them work with parents, and some others work on their own, possibly connected to gangs. Begging is found in busy downtown areas and rag-picking mostly occurs in the slums.

Child labourers are also employed in small-scale units or home-based manufacturing in urban areas, many of them involved in the garment industry (in Delhi), which are undertaken at the household level. Child labour occurs near the biggest factories that distribute orders to small units. Often, the whole community in an urban slum is involved in subcontracting. One NGO staff explained that child labourers in the garment industry are mostly from other states, referring to Bihar, Uttar Pradesh, West Bengal, and Assam.

The types of work and working conditions that child labourers engage in are diverse. Their work is mostly less skilled, less well-paid and undertaken in small-scale units or at household level. Labour-intensive industries are the areas posing the greatest child labour risks. Thus, it is difficult for child labourers to develop their skills or abilities in order to move onto better-paid work.

- **Child labour and education**

Child labour happens more in urban slums and poor rural areas where children have less access to schools. Many interviewees agreed that child labour should be examined in terms of its close relationship with education. One interviewee said that working children show a high drop-out rate in educational settings; if they are girls and live in rural areas, drop-out rates increase even more sharply. People are concerned about the quality of education, which can be the reason for the persistence of child labour. Lack of quality schools (people complain about the quality of public schools) and the financial demands of secondary education contribute to a high drop-out rate; thus, children frequently drop out before completing school.

Ages 12 and 13 are the ages, many interviewees believe, at which children may drop out of school and become full-time child labourers. Elementary school finishes at grade 8 (ages 13–14)[2] and secondary school starts at grade 9 (14–15 years old). Thus, children may decide whether they will continue to study in secondary school, move to a training school, or become full-time workers before they turn age 14. One interviewee from research institute said that drop-out could be related to the board exam for higher education, which is held at grade 10 (ages 15–16). Some students do not have the motivation to take that exam nor enough support to prepare for it, so they drop out.

The situation is better for children who go to private schools, such as the school run by the NGO I visited in Uttar Pradesh. The school covers grades 1 to 12, but not all students continue to attend school until grade 12. The school project was initiated by one of the biggest companies working to remove child labour from the carpet industry. Most of their students belong to carpet-weaving families, but they are not labourers. They help their parents after school, sometimes working on the farm, and so, during harvest season, children's attendance is low. Drop-out often occurs primarily among girls aged 12 or 13 years. Teachers acknowledged that the conditions of government schools are much worse. Many rural students drop out of school because of inadequate school facilities and poor-quality education.

On the other side, urban child labourers often drop out or temporarily stay out of school because of their family situation or work duties. Although elementary education is free and schools are accessible, child labourers often cannot continue their schooling. The NGO staff said that migrants' children are required to submit certificates from the original school to the new one, and during this process, delays or disruptions occur. Meanwhile, families need the help of their children because of large workloads, and it becomes difficult for children to go back to school.

[2] India's elementary school system is slightly different from the primary and middle school system: primary school is for grade 1-5 and middle school is for grade 6-8.

- **Working hours of child labourers**

The NGO staff explained that children who are involved in the readymade garment industry usually spend 3–4 hours a day helping family work, even on Saturdays and Sundays. If children go to school, they work less than that, about 1–2 hours per day. Older children aged from 15 to 18 work at an *addas*, a small unit of production, so they can earn income. This refers to a household-based unit of the garment industry, a small space where collective work is made. Finishing tasks such as stone pasting and thread cutting is undertaken at household level, while intricate embroidery is undertaken at *addas* (Save the Children, 2015). Those children work for 6–10 hours per day (ibid., p.28).

In villages, young children are given light work, so they work for about 1–3 hours a day after school, according to the school officer. Older children work for longer than that. As children grow up, they perform not only household chores or household production but also their own work (being hired out). Older boys go to towns to find a job, such as working in the services sector, and girls move to different cities to work as domestic service workers. Working hours are longer for full-time child labourers, although it differs by industry or occupation.

- **Social-cultural factors in child labour**

Child labour is a multifaceted matter involving gender, caste, religion, social class and migration. In the garment sector in Delhi, the majority of child labourers are from Muslim families and also many belong to OBCs, SCs and STs. Thus, this type of child labour is a mixed result of caste and religion. The NGO staff explained that the child labourers' parents have migrated from Uttar Pradesh, Bihar, Jharkhand and West Bengal before settling down in the community in Delhi; therefore, the child-labour problem is mixed with migration. The VV Giri National Labour Institute has observed that migration is a crucial factor in increasing incidences of child labour. According to their research in the state of Nagaland, a large number of migrant children become domestic workers, and families migrate from rural to urban areas because of cultivation difficulties and unemployment. Migration occurs between states or within them. According to a CRY report, between the years 2001 and 2010, there was a sharp increase in the number of child labourers aged 5 to 9 years old due to increased migration from rural areas to urban cities.

Furthermore, the patterns of child labour among boys and among girls are not the same. It is acknowledged that girls are more likely to be found in unpaid work. The garment industry, for example, is carried at household level and many child labourers are girls who stay at home and help their parents. Save the Children reported 69 percent of home-based child labourers were girls in their project sites. In contrast, in *addas*, which is production outside the home, all child labourers are boys. In the case of the garment industry, boys tend to work more independently, working in factories or taking occupations that differ from those of their families. Girls help family production or become domestic workers after dropping out of school.

To summarise, this short field visit provides a picture of the areas of work in which child labourers are involved and how stakeholders react towards child labour. Generalising child labour should be avoided. Children's gender, region, and family socioeconomic status allow different working conditions, hours and sectors. Nevertheless, many child labourers are observed in both urban and rural areas. Boy child labourers are more visible, and they can work independently. Some girls are also found working in urban circumstances; they can work outdoors with their parents or their brothers, but they are less likely to work on their own. In a rural context, girls' contribution to farming and household production is enormous. In addition, formal employment of children (such as in a factory) is forbidden because of child protection policies and the awareness of the illegality of child labour. Nevertheless, household production and unpaid household services are the exceptions, and helping with household work is not considered 'labour', which is commonly accepted as light work (in fact, it is often not light). Considering all these factors, a careful approach to children's diverse economic roles both within and outside the home is a necessary step to better understand child labour.

1.3. Research Gaps in Child Labour Studies

One of the questions in this research asks what leads a child to become a 'child worker' or 'child labourer'. Despite ongoing international discussions regarding child labour, how to define child work and child labour remains unclear. Furthermore, there are continuing academic discussions about whether child labour should be entirely banned. Although in general, scholars agree that child labour is harmful to children, children's situations differ because of social and cultural contexts.

On the one hand, scholars support a global ban on child labour, focusing on children's universal rights (Weiner, 1991; Weston, 2005). In this approach, the focus is on 'child labour' rather than on 'child work'. The key to this approach is to identify and separate child labour from acceptable light work. Regarding the protection of children's rights, international interventions are required, such as setting the minimum age of employment at a global level. In this approach, child labour is considered a human rights problem; it violates children's rights as human beings as well as their right to be educated (Weston, 2005, p.5). Sen (2005) recognises human rights as human capabilities composed of basic needs for life, health, emotional needs, and political control. In this consideration, child labour is a risk to human rights and human capabilities. Weiner (1991) emphasises the role of the political system and inappropriate education that stratify the social order, insisting children from lower classes have to earn money instead of studying. Thus, a child-labour problem represents a political failure in providing education (ibid.).

On the other hand, some scholars insist that the drive to prohibit child labour is based on a presumption that all child labour is harmful and ignores the fact that many children want to continue schooling and earn money (White, 1996, 1999; Bourdillon, 2006; Bourdillon et al., 2009; Woodhead, 1999). Bourdillon (2006) argues that the harms resulting from child labour are arbitrary: some other activities, such as sport, are harmful too. Participating in work can be a subjective decision taken by children, rather than a position forced on them by adults (White, 1996). One empirical study, for example, shows the explicitly autonomous behaviour of migrant child workers in rural Karnataka, India (Iversen, 2002). Domestic work for people outside families follows diverse patterns in different countries, and although it inflicts psychosocial harm on child domestic workers, a total ban on this type of work is difficult because economic reward and educational access are given to children in return (Gamlin et al., 2015). In this approach, the socialising effect of child labour and active roles of children are prioritised. Therefore, children voluntarily helping their parents is not considered harmful to children.

While children's agency should be respected, I believe that the child-labour issue should be looked at from the perspective of children's universal rights which state the child should be protected from any harmful activities. Advocating children's agency often fails to address children's urgent needs, and children's rights as human beings and their specific needs due to their special status as children require stricter safeguards and protection (Abernethie, 1998). That the debate on the eradication of child labour is ongoing can partly be attributed to a lack of evidence regarding the harms or benefits of child labour. Therefore, there should be more research in order to provide evidence of the causes and results of child labour in different societies. Abebe and Bessell (2011) assert that analysis of child labour requires a combination of socio-cultural and political perspectives. Child labourers are not homogeneous, and norms and ideas, as well as structural power, form part of the diversity within it (Bessell, 2011). I agree with their points and would like to emphasise the importance of interlinking norms and structures regarding the incidence and perseverance of child labour. This study will reveal how child labour is structurally grounded and how it is shaped differently by social constructions such as gender and social norms, using the Indian example. In this regard, Olsen and Morgan (2015) have a clear position whereby structures are not the norms but operate with agents who have norms and are located within the structure.

Liebel (2004) emphases that child labour should not be limited to a Westernised concept of childhood but understood in all its diversity including different social and economic factors, as well as children's subjective willingness. Liebel (2004, p.77) believes the ILO and UNICEF restrictions on child labour are too rigid. Children's work can help 'children's social recognition, independence, rights, participation and personality development'. He puts culture at the front of the discussion, showing that children in early years take independent economic roles and learn skills from working processes. Following that, he adds a review of social structure – for example, low-caste children are more exploited in India – and he argues that there is a tension between 'culture' and 'society'. However, his analysis of child labour has a number of limitations. It tends to overlook the fact that culture and society are not precisely separated but interlinked or interwoven. For example, the caste system is not only a social structure but also a culture. His examples of 'culture' are focused on tribal children, but culture is more than ethnicity. Furthermore, gender, which is closely related to culture and structure, is rather overlooked.

Balagopalan (2014) argues that a 'multiple childhoods' approach supports the cultural diversity of child labour but victimises childhood in the Global South, excluding some childhoods such as those of middle-class children. In historical thinking, postcolonialism should be understood together with the role of culture in the current sophisticated practices of child labour (ibid., p.131). Experiences of street children in Calcutta, India are based on the devaluation of education for the working classes in the early twentieth century and putting economic growth before children's welfare during postcolonial development (ibid.). I fully agree with the point that we should avoid childhood essentialism, such as treating child labour as a natural choice, but postcolonialism need not undermine the cultural diversity of childhood. Instead, it can show how child labour is originated from the social structure and appears in different forms within a society or community.

Moreover, previous research has exclusively focused on family decision-making without paying much attention to social and cultural effects. In the famous study by Basu and Van (1998), the influence of family wealth on child labour incidence is analysed. They tend to undermine the interactive relationship between family wealth and other socio-cultural factors. On a macro-level, economic growth is believed to reduce child labour, but its impact is unclear. Moreover, children in lower social groups benefit less from economic development. Chaudhri et al. (2003) insist that economic growth leads to more children becoming employed because of higher demand for child labour; therefore, the offset of child labour is only possible when the economy reaches a sustainable level. Kambhampati and Rajan (2006, 2008) have also found economic growth actually increases child labour; demand for child workers increases in rural areas in particular as there is greater demand for family farming.

Overall, we need an integrated approach that takes into account household socioeconomic status and cultural aspects. Economic and cultural factors cannot be separated, as they are mutually constituted (Kabeer, 1994, p.134). In many studies on child labour, social norms are ignored. Social norms and institutions can greatly influence children, parents, and their communities. Agarwal (2016, p.148) points out that bargaining power increases when there are more resources and is confirmed by social norms and institutions. The economic motivation of child labour might be counteracted if parents were aware of their children's rights. If people knew of the illegality of child labour or if social norms valued children's education, the incidence of child labour might be reduced.

Moreover, gender is intertwined with culture, such as caste and religion, as well as structure. For example, the association of gender and ethnicity leads to a particular pattern of child labour in some countries. In Bolivia, for example, indigenous girls are more likely to be child labourers (Zapata et al., 2011). In many studies of child labour, gender relations are classified in a simple or fixed way, e.g., boys in paid work and girls in unpaid household work. I believe that girls' economic roles are more diverse than that, and looking at gendered labour is important for children.

The last concern is an insufficient methodological development in studying child labour. Although child labour in India has been a concern for a long time, few studies have been made that estimate the number of child labourers and predict risks employing an advanced statistical methodology. A sizeable statistical error due to insufficient coverage of the child population, an undercount due to parents' increasing awareness of the illegality of child labour, and the unavailability of suitable datasets are all related problems. Furthermore, inconsistent survey design and reference periods between datasets make it more difficult to use time-use measurement (ILO, 2018c). This research aims to estimate the prevalence and causes of child labour across India, with a precise measurement of time spent on work by child labourers. It predicts the accurate number of child labourers with an improved statistical model, reflecting gender and social relations. The research applies a Bayesian analysis for the estimation of child labour and offers an innovative data combination method to obtain accurate estimates of child labour. Bayesian analysis has been growing in popularity within the social sciences, but it is still new in the study of child labour. It can bring advantages by overcoming the shortage of datasets and insufficient information in order to estimate the number of child labourers. A Bayesian approach is thoroughly reviewed in Chapter 3.

1.4. Aims and Scope

As discussed in previous sections, several concerns have been raised about child labour. The main limitations of the child-labour study are the generalisations regarding child labour, lack of socioeconomic and cultural considerations, insensitivity to gender roles, and insufficient methodological innovation. The scoping visit helped to understand the types, locations and sectors of child labour in India. Situations vary depending on children's age, gender, household socioeconomic status and cultural background. All this information suggests that we should avoid assuming similarity across child labourers' conditions but examine in detail the different characteristics of child labourers. Socioeconomic structure sits at the root of child labour, and gender, norms and values are factors that should be clearly understood.

The scope of this study includes quantifying child labour in India using an advanced methodology, revealing the causes of child labour in association with gender and social relations, and predicting child labour risks using multiple variables. Firstly, this study carefully defines child labour and estimates the number of child labourers by applying a Bayesian hierarchical model. The second part of the study reveals what impact gender, caste and class have on the incidence and severity of child labour. Third, this study predicts child labour risks on the basis of social norms and gender, applying accurately constructed time-thresholds of child labour. Finally, based on the findings, the key institutional interventions that are needed to help reduce the child labour problem in India are discussed.

This study aims to provide the most accurate measurement of child labour in India and suggest policy implications to reduce it, based on a clear understanding of gender and the socioeconomic and cultural determinants of child labour. Theoretically, it employs the gender and development approach (GAD) as it can contribute to explaining how gender and social determinants of child labour are interwoven. Methodologically, this study contributes to developing the measurement of child labour and allows us to predict it with accuracy based on innovative and advanced estimation methods. It defines child labour by international standards and uses consistent criteria to calculate its scale. The model includes carefully selected variables that represent social and gender dimensions. Finally, this study provides the best risk model for predicting the rate of child labour among the population by state.

The thesis has three key questions. The first asks how a fair number of child labourers in India can be calculated by applying a Bayesian data combination method. Secondly, it asks how the social and gender-based characteristics that are associated with the incidence and working hours of child labour might be figured out. The last question discusses how to obtain the best child labour risk model in order to predict child-labour rates in key states using norm indicators and applying cutting-edge guidelines from key international and national stakeholders.

The research questions are as follows:

I. **How does a Bayesian hierarchical model with a combined data approach improve the estimation of child labour in India?**

- How does including domestic chores in the measurement criteria of child labour change its estimation?
- How does combining multiple datasets and using a Bayesian hierarchical model improve predictions regarding child labour?

II. **How do social group, class and gender influence the labour-force participation and labour hours of children?**

- Does an interaction exist between social group, class and gender regarding children's participation in labour and their labour hours?
- What difference is found between informal and formal sector jobs involving children, based on their social group, class and gender?

III. **How do gender relations and norms improve the prediction of a child labour risk by household occupation and state in India?**

- How can we measure a time threshold that shows at what point working becomes harmful to children's development?
- What differences are found in comparing predicted risks of child labour by state?
- What are the implications for gender relations and norms regarding interventions to reduce child labour?

1.5. Conceptual Framework

Figure 1.1 describes the conceptual framework of this study that interconnects gender and social relations. Firstly, this study accounts for structural effects on child labour. More specifically, it looks at whether a household's class status influences a child's participation in child labour. Household socioeconomic class is more complex than poverty; structural and relational inequality are involved. Children from landless labour households might engage in the worst forms of rural work such as *kuulie* work (casual paid work) or bonded labour. Landless rural workers are often low-paid, exploited by landowners, and accordingly constrained by economic poverty and social exclusion (Olsen and Neff, 2007). Meanwhile, farmers' households involve a large number of child labourers, too; while men and boys are involved in non-agricultural activities to diversify the household's livelihoods, women and girls are restricted to agricultural activities. Class can be proxied by family wealth as well as the amount of land a family owns. Land ownership typically represents class status in a rural part of India; in an urban area of India, class relations are not necessarily decided by land ownership. In this case, asset ownership can be used as a proxy for class to observe what effects household class can produce in terms of child labour.

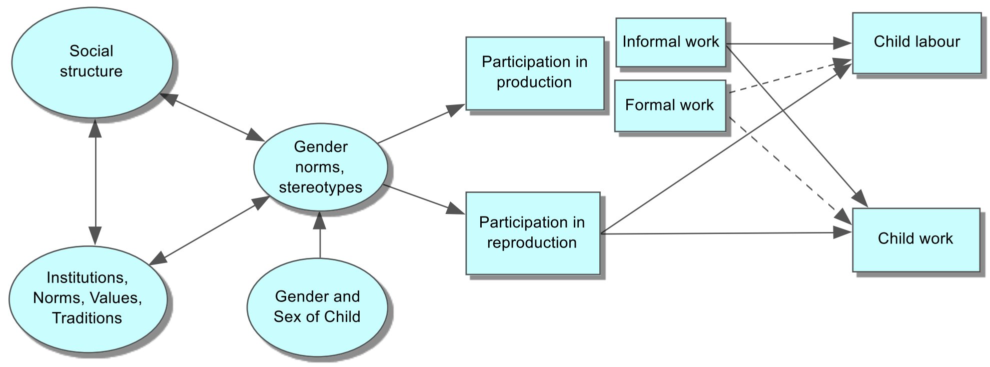


Figure 1.1. Conceptual framework

Notes: The dashed lines indicate that the path is possible but less frequent. Child labour harms children's development, which should be prohibited, while child work is acceptable and does not harm children. Not all production work is harmful to children, and so some roles in production within the limited hours are not child labour but child work. Meanwhile, child labour occurs more in the informal sector than in the formal sector. Working in reproduction, caring work, is often considered as children's work, but depending on the amount of time a child works, the effects can be severe. Therefore, some reproductive work falls within the boundaries of child labour. In India, girls are more likely to perform duties in reproduction.

Secondly, this study views the combined effect of structure and institutions (norms, values, and traditions) so that it extends the boundaries of relations. Many social phenomena in India are rooted in the caste system, which is an institution as well as a structure. Caste determines not only the social division of labour but also the sexual division of labour (Desai and Krishnaraj, 2002, p.303). Caste is related to social honour, thereby putting pressure on girls to stay at home. The upper caste social norm immediately withdraws girls from economic activities to practice purdah. At the same time, lower-caste girls are also alienated from outside work when their families follow upper-caste social norms. Girls from scheduled castes and scheduled tribes, especially in villages, do more outside work alongside doing domestic chores. In this study, norms and beliefs have important roles in explaining child labour. They are constructed among agents and compose institutions; they are not entirely the same as structure but are combined with it.

Thirdly, this study reveals a gendered pattern of child labour in India. This study includes gender-related norms, values, and roles as well as sex and gender identities, implying that gender norms and values are constructed based on biological sex and the notion of gender. Work segregation of labour by children's gender is one example that shows how culture forms different gender roles for children. Furthermore, girls are called to pick up their absent mother's work because of culturally entrenched gender norms. Boys and girls start with similar tasks on family farms, usually within their mother's sphere, and as they grow, their agricultural duties are differentiated according to their adult gender roles (De Lange, 2009, p.4). In spite of the significant prevalence of girl child labour in the areas of reproduction and household production, the economic participation of girls is less visible. Thus, looking at reproductive and productive work and informal and formal sectors separately is an important step towards revealing gendered patterns in child labour.

1.6. Thesis Structure

This research uses a journal format, and the thesis is composed of three main journal articles and other related chapters. In this section, I provide the outline for the following chapters and an overview of shared responsibilities in three journal articles.

Chapter 2 contains the literature review for the thesis, drawing on the gender and development (GAD) approach. To explain the need for this approach in a study of child labour, I review related theories, such as intra-household decision-making theory and the structural approach, including Marxist feminism. Then, I apply the GAD approach in order to better comprehend child labour. Next, gender and social relations – class and social groups – are interchanged in terms of child labour in more detail. Child labour is studied employing an extended view, including productive and reproductive, and informal and formal work. The different gender roles of children are examined.

Chapter 3 introduces Bayesian methods, covering the general advantages of a Bayesian approach and a specific methodology for the study. The focus is given to Bayesian explanations of the data-combining approach and the use of a Bayesian hierarchical model. Furthermore, how to deal with a large number of zeros to select priors for parameters and to use a posterior predictive check are also discussed.

In Chapter 4, I present a published paper titled *Bayesian Estimation of Child Labour in India*. This paper elaborates on the definition of child labour and provides a Bayesian estimation of the number of child labourers by combining two datasets – the National Sample Survey and the Indian Human Development Survey in 2011/12. Child labour reflects both children's economic activity and unpaid household services (domestic chores). A Bayesian combined-data method contributes to improving accuracy in its estimation.

In Chapter 5, the second journal article is presented, entitled *Girl Children's Labour Participation, 'Child Labour' and Decent Work: Results from India*. It reveals the effect of social determinants such as caste, class and gender on child labour. This chapter also provides an analysis of labour hours of children showing that girls work long hours in the informal sector, and boys work in both the informal and formal sectors.

In Chapter 6, the last paper, entitled *Predicting Child Labour Risks by Gender and Norms in India*, provides a concrete view of predicting the number of child labourers in India. It estimates the time thresholds after which the rate of children dropping out school sharply increases. The child labour risk model, in combination with norm indicators, helps to predict the prevalence of child labour as accurately as possible, revealing how gender and social norms affect the incidence of child labour.

Finally, Chapter 7 provides a discussion based on findings through the analysis of three main chapters. This chapter has three key points: how the GAD approach can be extended to include children in the discussion; the further use of a Bayesian method to measure child labour; and, finally, some policy implications of this study. At the end of the chapter, the implications and limitations of the study are discussed, and the future direction of the study is suggested.

Table 1.4 shows the three main chapters of this thesis. Chapters 4 to 6 are journal articles dealing with different topics within the scope of this thesis. Specific methodologies and literature reviews are included in each chapter. In terms of the publication of the three papers, my supervisors, Wendy Olsen and Arkadiusz Wiśniowski, and I shared co-authorship, taking equal responsibility for their contents. As the lead author of all three articles, I reviewed the literature, prepared data and modelling, analysed the results, and wrote all the manuscripts. Wendy Olsen contributed through introducing key theories that pertain to child labour and the socioeconomic backgrounds of India, and developing and analysing the models. Arkadiusz Wiśniowski has contributed by suggesting the Bayesian statistical method, developing and checking models, and analysing the results. During the whole process, we communicated closely and reviewed the papers together.

The first paper has already been published in *Child Indicators Research*, and the second and third papers are in the process of submission. The first paper was presented at the Royal Statistical Society 2018 International Conference. The second paper was presented at the International Association for Feminist Economics 2019 Annual Conference and the Stochastic Modelling Techniques and Data Analysis 2020 International Conference.

Table 1.4. Titles of articles and selected journals

No.	Format	Chapter No.	Title of Article	Journals
1.	Journal Article	Chapter 4	Bayesian Estimation of Child Labour in India	*Child Indicators Research* (Published)
2.	Journal Article	Chapter 5	Girl Children's Labour Participation, 'Child Labour' and Decent Work: Results from India	TBC
3.	Journal Article	Chapter 6	Predicting Child Labour Risks by Gender and Norms in India	*Work, Employment and Society* (Under review)

Chapter 2. Research Context

This chapter aims to compare different theoretical views on the causes of child labour. Firstly, I critically review the two most widely acknowledged theories on child labour – neoclassical economics and the structuralist approach. Both theories provide different insights into the causes of child labour: the former emphasises insufficient family wealth, and the latter explains child labour in terms of unequal power distribution in the society. I introduce the GAD approach as a way to understand the interlinkage of gender and social relations around child labour. Child labour is a complicated matter that involves household socioeconomic status, gender division of labour, and social norms. The GAD approach helps explain the complex causes of child labour through comprehensive accounts of social, economic, and cultural dynamics.

This chapter unfolds in several sections. In Section 2.1, I briefly introduce how children and child labourers can be theoretically defined using the concept of socially constructed or socially structured children. In Sections 2.2, 2.3 and 2.4, I analyse the selected theories' strengths and weaknesses in explaining child labour. Section 2.5 outlines the gender and development (GAD) approach in general, and Section 2.6 discusses how to apply the GAD approach in a study on child labour including consideration of class, social groups, and gender. Section 2.7 elaborates norms and institutions from the viewpoint of GAD. Finally, in Section 2.8, I clarify the key concepts used in this research: definitions of child labour, children's working hours, and informal and formal work.

2.1. Children and Child Labourers

Defining children and child labourers is not a simple task and few works in the literature provide meaningful theoretical definitions of a child or children. James and James (2004) explicitly explain the concepts of children, childhood and the child: 'Childhood is the structural site that is occupied by "children" as a *collectivity* ... within this collective and institutional space of "childhood", as a member of the category "children", [wherein] any individual "child" comes to exercise his or her unique agency' (James and James, 2004, p.14). Childhood is a natural phase of life for all people (ibid., p.14), and various children have experienced 'childhood' differently in any given society (ibid., p.22). Each (individual) within the children who compose the collectivity of children acts as a social agency, and it is essential to understand the uniqueness and commonality of his/her childhood (ibid., p.27).

Historically, the concept of childhood did not exist in the medieval period (Aries, 1961 in Wyness, 2006). The modern images associated with childhood such as dependency, happiness and the right to protection came from more recent cultural and historical backgrounds (James, Jenks and Prout, 1998, p.62). Wyness (2006) finds that the modern concept of childhood refers to a 'lack of work' (p.9), which means child labour is regarded as problematic (p.11). Although social constructionism has become a dominant approach in the study of childhood, childhood is now being considered as an unequivocal status tied to 'what adults do and think' (ibid., p.26).

The discussion on children's agency is comparatively recent in sociological or historical studies of child labour (Goose and Honeyman, 2013). For example, diversity in child labour was already found in nineteenth-century Britain. Goose and Honeyman (2013) find that child labour makes different childhood experiences according to local or regional characteristics. Child employment in factories in Britain during the Industrial Revolution is a well-known image of child labour, but in fact, children were also engaged in more diverse sectors – in agriculture and services – and also as apprentices (ibid., p.5).

The structuralist approach brings a different understanding of childhood and children when compared to social constructionism. It locates children in political, cultural and economic realms, as a subordinate group within society: they have limited access to political power, behave under the frames built by adults, and are less visible in terms of their economic positions (Wyness, 2006, pp.29–32). The social structure where children are subject to dependence explains why children's groups experience a different childhood over and across time (Wyness, 2006, p.35).

The theoretical discussion between social constructed and socially structured childhood provides meaningful insight into how child labour is viewed in society. James and James (2004) offer a conceptual mapping of theories regarding children (Figure 2.1). Social structural childhood views that children, as components of social structure, interact with other social groups in society and focuses on the commonality of childhood at a national or global level (James and James, 2004, p.60). In seeing a child as a social actor, children's problems are understood as concerns shared by all children, not by an individual child (ibid., pp.58–59). On the other hand, socially constructed childhood focuses on the diversity, and local particularities of childhood, which views the concept of 'the child' or 'childhood' is constructed in the local context (ibid. p.60). James and James (2004, pp.22–27) take the view that childhood is neither the result of structure nor the outcome of the discourse constructing childhood, but it is a product of children's relations with other social components and the participation in the institutions. Children's agency and their activities as social agents play a role in constructing childhood (ibid.).

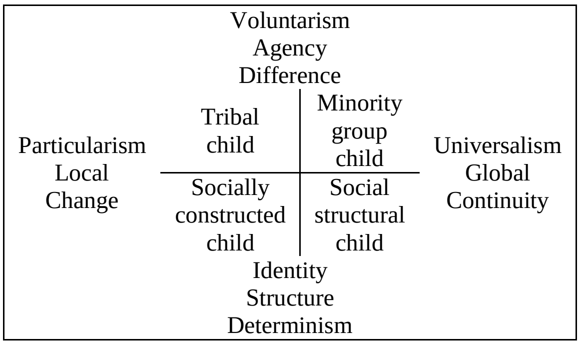


Figure 2.1. Four theoretical focuses on children: 1) tribal children, 2) minority group children, 3) socially constructed children, and 4) social structural children
Source: James and James, 2004, p.58

James and James (2004) imply that social structures determine childhoods, but they are only deterministic in some idealised sense. What is more common is a child has a range of influences as they grow up, including structural and agency influences. Social structure is not universalism since there are interactive effects of culture, tradition and institutions. Therefore, the social-structural element highlighted here by the authors is not kept clear of the other elements shown. Structure as determinism and universalism is limited in explaining the diversity and power imbalance in people's relations. Power of who has the power or capacity to influence how patriarchy works and thus how children's gender is constructed does not appear in the James and James diagram. They do discuss these matters, however. By looking closely at relations that surround children, we notice the reasons for child labour in a more realistic and precise way. Further emphasis on a relational structure is placed in Sub-Section 2.7.2. The diagram helps to identify areas of social theory (Figure 2.1). In this sense, it is very useful; therefore, I have copied the diagram verbatim. Moreover, it is useful in that it can be applied in many geographic contexts.

It is also limited by the absence of gender and intersectionality from the diagram (and in some ways from their whole book). Gender is an important component in the shaping of childhood. Gender itself is a social construction; moreover, it constitutes structure and structural relations. Children experience different gender hierarchy in society and, accordingly, gender constructs the norms of children. Thus, gender connects social structures with social construction. The limitation of the lack of gender in James and James (2004) can be overcome here in this new child-labour project.

This research recognises children and child labourers are strongly affected by social structure. It means that child labour is not an individual-specific phenomenon but a socially-oriented one. Therefore, this study seeks to understand the social structure that children and their families belong to. However, this approach towards social structure does not downplay the agency and diversity of child labourers being influenced by their cultural and historical roots. Although this study uses social structure or labour relations to explain child labour, it considers the diversity of child labour associated with institutions and culture to the same degree. Accordingly, 'child' in this study is located between a 'socially constructed child' and a 'social structural child'.

In short, this study supports the view that children and child labourers are socially constructed within the given social structure. Cultural differences and changes over time happen, but the view of socially structural children does not necessarily contradict that. Social construction and social structure interact with one another and thereby produce variations in the statuses of children. This study reinstates the idea of child labour as a result of social structure; so, it supports the use of universal law against child labour as a way to eliminate child exploitation. At the same time, this study recognises the cultural diversity of child labour, and therefore it requires a broader understanding of traditions, social norms, and gender roles. In the next two sections, I review the literature of two theories – a conventional intra-household decision-making theory and a structuralist approach – and explain why we require the GAD approach to better describe child labour in India.

2.2. Limits of the Intra-Household Decision-Making Model

The neoclassical model pays attention to the economic role of child labour within the family. In early explanations, the family, as one unit, decides about child work or child labour based on economic motivations. Here, neoclassical theory underscores the households' choice based on thinking that child labour is a reasonable choice for a household. Nardinelli (1980, p.740) reiterates the role of child labour in the Industrial Revolution, which confronts the traditional condemnation of the exploitation of children. It is not legislative pressure but the growth in real income that leads to changes in child labour in nineteenth-century Britain (ibid.). His main argument is that child labour is not a matter of morality but a choice for the household.

In these earlier studies, a household is considered a single group, and the choices of individual household members are disregarded. The unitary model assumes that households base their preferences on rationality (Ellis, 2000); however, different household members have different preferences, and they might be cooperative or non-cooperative in the decision-making. This collective household model does not consider children's subjective choices and ignores the reality that women and children have less power in the decision-making process.

Later, intra-household decision-making incorporates morality into the approach. Parental altruism (the 'luxury axiom') is the best-known concept in child labour research, which was introduced in a child-labour study by Basu and Van (1998). They explain that parents send their children to work only when household income is insufficient, and thus low adult wages cause child labour. This approach to analysing child labour takes family dynamics into account. There are multiple equilibria – one with high adult wages and no child labour, and the other with low adult wages and child labour (ibid.). The way to reach a better equilibrium is improving remuneration for adults because the banning of child labour might produce unexpected outcomes (ibid., p.425).

Many other researchers have studied child labour based on the 'luxury axiom'. From this point, the coverage of household wealth becomes diverse, including factors like land ownership, access to the credit market, and non-income shocks. For example, Ray (2000) adds empirical evidence of a relationship between child labour and household wealth with cases from Peru and Pakistan; rising wages for adult females increase children's participation in the labour market in Pakistan. Dessy and Pallage (2005) insist that harmful forms of child labour have economic roles, such as human capital accumulation. They use a model of parental decision-making regarding child education and child work and insist that the worst forms of child labour emerge when children are better remunerated. They suggest that poverty alleviation through interventions such as food for education programmes could be a better solution than entirely removing the child-labour option through a banning policy (ibid., pp.85-86). Basu et al. (2010) reveal an inverted-U shape relationship between landholding and child labour. They show that more land ownership increases children's working hours until a point, beyond which working hours are reduced (Basu et al., 2010).

Lima et al. (2015, pp.80–81) prove that the relationship between land and child labour is stronger for 'non-altruistic families' where children work many hours per day. Households with limited access to credit show higher trends of sending children to work (Alcaraz et al., 2012; Dumas, 2012). Non-income shocks, such as agricultural shocks, are also associated with a higher incidence of child labour (Soares et al., 2012; Bandara et al., 2015).

Gender is sometimes referred to in a family decision model. Also, Basu et al. (2010) point out an imperfect labour market as a significant cause of the gender bias in child labour. Kumar (2013) suggests that son-preference in Bangladesh can be a reason for girls in child labour, and the problem is also the result of family poverty. A recent study has revealed that being a child labourer increases the probability of vulnerable employment in adulthood, especially among girls in Tanzania (Burrone and Giannelli, 2020).

Despite its explanation in terms of a relationship between family wealth and child labour, there are several shortcomings in the neo-classical approach to child labour. Firstly, following this approach, the family decision-making process excludes children; therefore, child labour exists to meet the preferences of adults not of the children themselves (Elson, 1982, p.482). Secondly, the neo-classical economic approach does not adequately explain the relationships outside families (Elson, 1982, p.483). Furthermore, this approach does not take into account different power relations. Young (1993) pointed out money is unequally allocated within households; thus, poverty alleviation should directly support children (or the most vulnerable). Girls are physically impaired as they are least supported by resources (Drèze and Sen, 1995; Young, 1997). Although the neoclassical economic approach has included diverse aspects leading to child labour such as land ownership or financial access, it overlooks the social and cultural roots of child labour. The effects of family wealth on child labour are over-simplified, although it is closely associated with class and culture. Furthermore, gender is not a priority despite its critical impact on child labour, and economic motivation does not fully explain the underlying causes of the gender gap in child labour. Lastly, this approach is based on an assumption that there is human capital accumulation from child labour (for example, Dessy and Pallage, 2005). However, child labourers face high risks to become low-paid workers as they lose opportunities for education and preparation for the future.

2.3. Limits of Structuralist Views on Child Labour

Marxist economists have looked at power relations in capitalist society with the understanding that children are physically and psychologically weaker than adults (Elson, 1982, p.485). According to Marxist economics, children's work is not a matter of family preference but the capitalist social system (ibid., p.485). She explains child labour is a result of age hierarchy wherein children are subordinated to adults, just like women are subordinated to men under the gender hierarchy.

The Industrial Revolution in Britain provides empirical evidence to support how the capitalist social system increased children's engagement in factory work. Although children’s family work existed before the Industrial Revolution, the introduction of the factory system in the late nineteenth century led to a critical change in the position of children; they were absorbed into the labour market and took a subsidiary role in mechanised production (Kirby, 2003, pp.71–72). Lavalette (1994, p.9) considers child labour as a structural phenomenon in capitalist societies: ideological change, such as the development of the ideology of childhood, and political change that legitimated child labour affected the recognition of child labour in Britain. The same author (2000) focuses on the emergence of children’s jobs, which are severely marginalised in the labour market but receive less attention, in association with the establishment of the working class and the concepts of childhood and development of child-related social policy. Even in modern Britain, child exploitation has continued in the form of menial jobs (delivering, selling and services; ibid.)

Humphries (2010) provides a unique approach to using historical evidence from Britain in the nineteenth century. She uses both class and demographic changes to explain child labour in Britain. Child labour in the Industrial Revolution was not only caused by increased demand for child labour but also because of changing family structures, such as the earlier independence of children (ibid.). A considerable increase in fertility was another reason for the spread of child labour during the Industrial Revolution (ibid., pp.38-39).

The phenomenon of child labour is one piece of evidence supporting the idea of ‘social categorisation’ (Tilly, 1998). Tilly (1998) explains that categorical inequality is institutionalised through exploitation, opportunity hoarding, enumeration and adaptation; for instance, child labour rises in numbers as categorical inequality between work with firms and non-firms, formal and informal labour, which is coupled with age and gender-based inequality. Following him, Mosse (2010) discusses how poverty is a result of social categorisation, referring to the case of Indian Adivasi migrant workers, including child workers, who were exploited due to their lack of power. In both studies, the exploitation of children is a relational and structural problem. It is not only class-based struggle but also based on other relations such as gender and social groups.

The growth of the global market has threatened the positions of women and children in the informal sector. Phillips et al. (2014) point out the exploitation of children in developing countries is perpetuated through the segmentation of production processes in global production networks (GPNs). In the new social and power relations formed by GPNs, child labour becomes the lowest tier of production and its form is usually home-based work; therefore, it is difficult for it to be captured by the law or institutions (ibid.). Child labour in the garment sector in Delhi is one example that shows how home-based production uses child labourers and global value chains make the situation worse (Bhaskaran et al., 2014). These studies also accept that exploitation is closely related to gender, caste, and race (Phillips, 2013; Bhaskaran et al., 2014). In addition, Nieuwenhuys (2007) focuses on the global order of child labour, finding that child labour is concentrated in the Global South, especially in the area of reproduction, and she refers to children that are trapped in the 'global womb' (ibid.).

In global production networks that are based on sub-contracting and domestic production, controlling or auditing is never easy to achieve. The difficulty in auditing unethical labour in the global value chains has been pointed out by many authors (Barrientos and Smith, 2007; LeBaron et al., 2017). Barrientos and Smith (2007) analysed the impact of the use of Ethical Trading Initiative codes and revealed that it has a limited impact on improving labour standards. LeBaron et al. (2017) suggest that the disproportionate power relation between companies and auditors makes it impossible for auditors to reach the end producers, and they have access only to limited and partial information.

Structuralist research attempts to make children visible in society (Scott, 2000). However, there are a few limitations to be overcome in order to satisfy the diverse aspects of child labour, such as children's agency and their autonomy. The focus of the Marxist approach has mainly been placed on social relations. As Marxist economists and feminists emphasise the relationship between production and capitalism, gender relations are downplayed (Walby, 1990, p.38). Moreover, this approach does not capture children's complex and diverse characteristics beyond 'social class' (Wyness, 2006, p.49). Despite their efforts to take non-class-based groups into account, culture (or categories) is often considered less important in a relational approach (Tilly, 1998, p.21). Without taking gender relations into account, it is not feasible to explain the differences in society between male and female child labour.

Furthermore, in this approach, there is a tendency to view children as a passive and minor group in society. Wyness (2006, p.236) argues that by seeing children as a subordinate group, children come to have less power in society and are not accepted as independent social agents. Finally, in Marxist economics, children are treated as a minority group in the same way women are (Oakley, 2002). However, it is still questionable that children and women can be allocated to the same position in a parallel way. Women's struggles cannot be the same as what children experience in society. Particularly, any exploitative relations against boys and girls are distinct from relations that male and female adults have. There is more need to shed light on the gender hierarchy among children themselves, which is interlinked with adult-child relationships.

2.4. Family Structure and Child Labour

Family structure, in particular the sibling effect, is one of the important aspects that can affect the incidence of child labour. For the neo-classical economists, the sibling effect on child labour is in line with parental preferences regarding gender and sibling order, e.g., a preference for eldest sons. The empirical studies have found the number of siblings has a substantial impact on child labour (Patrinos and Psacharopoulos, 1997; Dessy, 2000; Ravallion and Wodon, 2000; Ray, 2000). Several studies suggest that a high fertility rate is a determinant of child labour (Grootaert and Kanbur, 1995; Basu and Van, 1998). Basu and Van (1998) insist that when fertility increases, more children will work at a lower wage. Its implication is straightforward – the policy needs to transform society from a high-fertility economy into a low-fertility one (ibid., p.425). Some research also insists that older sisters are more likely to work than their younger brothers (Parish and Willis, 1993; Edmonds, 2006; Dammert, 2010); however, the interaction between gender and sibling order is not consistent, differing according to the culture. Edmonds (2006) notes that the sibling effect is related to the presence of household production. He insists that it is important to take into account not only parents' preferences but also the different responsibilities in the domestic work done by sons and daughters.

On the other hand, Marxist economists view that the number of siblings is associated with social class. In this view, child labour is a cause of high reproduction rates, not a consequence of it (White, 1982). It is the conditions and relations within the work that affect households' decisions of their reproductive strategies (ibid.). For the proletariat, having children is a way to earn income when adults' wages are not enough (Brezis and Young, 2003). As adults' wages are low, children's wages are more necessary, which result in a fertility increase (ibid.). Furthermore, as siblings increase, income is diluted; hence, they need more income from children (Brezis and Young, 2016).

Overall, both views indicate that the number/order of siblings or a higher fertility rate could have a positive relationship with child labour. They provide a different explanation on this: one explains that having more children is a way to obtain more human resources, while the other sees an increased fertility and child labour as results of class struggles. It is difficult to tell which approach is more relevant to the Indian case. The link between the sibling effect and child labour should be looked at in a broader social system and cultural background. Given its complicated characteristics, this study does not include the sibling effect yet in the models.

2.5. The Gender and Development Approach

The gender and development (GAD) approach is an interdisciplinary theory, placing gender in the middle of the process of development and discerning gender inequality rooted in social structures. It is the best approach to explain the gender relations of children. In this section, I will explain the GAD approach by focusing on the concept of gender, the establishment of structural gender relations, and social construction of gender inequality.

The GAD approach brought ‘gender’ to the discourse of development, discussing it as a social construction that explains social inequality. Ostergaard (1992, p.6) defines gender as ‘the qualitative and independent character of women’s and men’s position in society’. It ‘distinguishes the biologically founded, sexual differences between women and men from the culturally determined differences between the roles given to or undertaken by women and men in a given society’ (Ostergaard, 1992, p.7). Moreover, GAD is based on ‘the political motivated assumption’ that believes ‘woman’ is a socially constructed category (Baden and Goetz, 1998, p.34). Momsen (2010, p.13) clarifies that the GAD approach is ‘based on the concept of gender (the socially acquired ideas of masculinity and femininity) and gender relations (the socially constructed pattern of relations between men and women)’ and analyses ‘how development reshapes these power relations’. Together, in the GAD approach, gender concepts and gender relations start from accepting the biological differences between males and females and are based on how this difference is culturally elaborated and established. The constructed relations regarding gender are structured as relations of power and hierarchy as development proceeds; thus, gender relations are not entirely constructional nor structural.

The GAD approach does not limit its view to social relations but gender relations. Gender relations are beyond social relations of production, but they entail the ‘relations of everyday life’ (Kabeer, 1994, p.65). The constructions of gender are everywhere and bounded with culture; therefore, gender relations cover much broader areas of relations. While Marxist feminists place a clear focus on class relations of gender, the GAD approach analyses broad interests in social and gender relations. Ostergaard (1992, pp.6–7) explains that the relations of power and dominance constitute gender relations. The Women in Development approach, for example, is rooted in modernisation theory; therefore, it does not recognise the social inequality found in dependency theory or a Marxist approach (Rathgeber, 1990). The GAD approach is able to bridge 'the gap left by the modernization theorists by linking the relations of production to the relations of reproduction and taking into account all aspects of women's lives' (ibid., pp.493-494).

In addition, gender relations in everyday lives mean taking diverse gender roles into account. Moser (1993) defines the triple role of women – reproductive work, productive work, and community managing work. Women's role in reproductive work is often invisible, and their community roles are also less recognised. Momsen (2010, pp.188–192) provides an empirical example of women's roles in rural areas in developing countries due to both productive and household care work; in West India, for example, female farmers work much longer hours than men, especially in the busy season of the agriculture. As women have double and even triple the burden of work, they might suffer from long working hours and low wages.

How development affords different outcomes to males and females is a fundamental question of GAD. This question originated from Women in Development (WID), an organisation that started in the 1970s as an approach to integrate women in the development process (Kabeer, 1994). The scholars of WID locate the causes of women's problems in their needs rather than the unequal access to resources that originates from the unequal power balance between males and females (ibid., p.7). For example, Boserup (1970) has explained that industrial segregation is caused by women's preference to stay in less risky types of work, such as domestic services and home-based industries. In her explanation, the gender division of work is regarded as a reasonable choice or a lack of relevant skills rather than a symptom of men's and women's unequal access to power. Contrarily, the GAD approach, originating in the Marxist feminist tradition, finds fundamental causes in the unequal social relation of power and its reproduction through gender. That said, women's integration into the bottom of the structure originates from capital accumulation as well as gender relations (Kabeer, 1994, p.63).

Boserup (1970, p.212) explains that the integration of women into productive working areas could be achieved by providing better education to women. However, this is a 'treating cancer with bandaid' as it neglects the fundamental causes of gender inequality, such as the gap in capital accumulation and male power in gender relations (Benería and Sen, 1981; Kabeer, 1994, p.31). Furthermore, WID undermines the roles of women's reproductive work. It disregards the importance of women's participation in reproduction and emphasises women's similarities with men and their participation in income-generating activities (Kabeer, 1994, pp.28–30). WID overlooks how women have made a role in the traditional sectors (White, 1997).

Benería (2003) gives a clear view of how the globalisation and development process has affected women. She reveals the economic basis of women's subordination and unequal power relations by focusing on gender that is culturally and materialistically constructed. Economic restructuring followed by a prevalence of subcontracting and informal work has shifted women's and men's positions (Benería, 2003, pp.95–96). Globalisation has accelerated the feminisation of low-wage production and low-skill industrialising areas, particularly the service sector or low-paid manufacturing sectors (Benería, 2003, p.163). In spite of an increase in paid employment, many women remain at the bottom of the social structure, because of an increase in part-time and precarious work and the concentration on home-based work (ibid., p.164). She finds that the cultural construction of gender is essential in explaining those challenges of women in development and globalisation (ibid., p.29).

Furthermore, the GAD approach has another root in feminist anthropology, which considers intra- and inter-cultural differences (Pearson and Jackson, 1998, p.6). The social construction of gender is explained in many ways. Kabeer (1994) emphasises the role of 'agency' and the different experiences of women in different locations. Agency indicates individuals who can make choices and act accordingly (Burr, 1995, pp.146-147). Here, culture has a role in confirming gender roles and gender status. Nieuwenhuys (1994) analyses boys' and girls' economic duties in Kerala, India, and avoids universalisation and gender neutrality. Child labour is exploitation against children through families and kinship within a capitalistic structure in which boys and girls perform different economic roles (ibid.). The GAD approach is interested in 'people's experience of gender differentiation' (El-Bushra, 2000, p.61). Thus, geographical and ethnographic perspectives are essential in GAD. Benería (2003, p.29) points out the major challenge for feminists is 'neglecting the important lesson learned from the critical perspectives represented by the "turn to culture" in the disciplines'. In globalisation, culture can bring a clearer view of gender relations (ibid.). Mohanty (2002, p.72) insists that 'women are constituted as women through the complex interaction between class, culture, religion and other ideological institutions and framework'. Thus, women are 'cultural and ideological composite' (ibid., p.62). She explains that gender studies are different from cross-cultural approaches that assume women are a single group. Moreover, a 'cross cultural singular' notion of patriarchy or male dominance should be avoided (ibid.).

Culture is related to all aspects of the quality of life, such as work, health, education, and even life itself (Nussbaum and Glover, 1995). Respecting culture means understanding human beings and achieving justice and the development of human capabilities (ibid.). Okin (1994) emphasises that women's different experiences should not weaken the concept of gender inequality. These discussions on 'gender and culture' imply that gender studies do not need to take either universalist or cultural relativist positions. Gender itself reflects cultural diversity as well as social inequality. Culture in the GAD context is defined as 'the distinctive patterns of ideas, beliefs, and norms which characterise the way of life and relations of a society or group within a society' (Reeves and Baden, 2000, p.4). Culture involves people's attitudes, values and norms in relation to gender position. In this study, I particularly focus on the role of norms in identifying gender differentiation in child labour.

Some studies believe that cultures should be mainstreamed in gender discussion. For example, Chua et al. (2000) criticise feminist theories (as well as the GAD approach) like the multi-ethnic or multi-racial studies of gender relations, as they mis-specify Third World women as victims. They theorise 'Women, Culture and Development' beyond the GAD model. They insist 'culture' should be viewed as 'the relationship between production and reproduction in women's lives', and thus, that women's agency should become more explicit. Nevertheless, I do not take that view that 'culture' should be at the forefront of the discussion because social relations should not be reduced to culture. Although it is important that this approach recognises the role of culture in identifying women's problems, social and gender relations are rooted in structure, which should not be belittled.

In summary, the unequal distribution of power and resources via gender generates structural inequality in gender relations. In a male-dominated society, women's main roles in reproduction marginalise women's positions. With the growth of global production, unorganised and informal jobs are given to women, and the feminisation of low-skilled industries becomes clearer in developing countries. All unequal relations are interwoven with socially constructed concepts of gender, gender norms, expectations, stereotypes, and gender roles. Thus, it is crucial to understand how one society has its cultural peculiarities of gender and gender relations in revealing the structural inequality.

2.6. GAD and Child Labour

Although gender relations are essential when considering child labour, they have not been appropriately discussed. The GAD approach does not fully focus on children, but its implications for women and girls, as well as for men and boys, provide meaningful insights for the study of child labour. There are a few considerations on how gender relations affect the incidence of child labour in India.

- Boys and girls have different positions in society. In ‘traditional’ families, sons are preferred over daughters. Hence, girls might occupy the lowest position in the gender and age hierarchy and accordingly the lowest jobs (with longer working hours and lower wages).
- Gender roles and expectations toward children can be compounded by their household status. According to families’ socioeconomic positions (such as class and caste) and economic needs, some households can differently decide their sons’ and daughters’ labour force participation.
- Adult female and male relations can also be influential in participation in the children’s labour market. When mothers are not able to work outside the home, children might take more economic responsibilities.
- Gender relations are formed in the reflection of social and gender norms. According to norms, some types of labour are allowed for boys or girls. Most clearly, girls’ outside activities are restricted by cultural norms in India. Whether agents accept or reject these norms might affect their decisions regarding child labour.

Taking these insights forward from the GAD approach, I reveal that structure (class, social groups) and institutions/norms are compounded in the decisions about child labour in India. In the following subsections, I introduce more details of how class, social groups and gender relate to child labour.

2.6.1. Class on child labour

Class heavily affects childhood experiences. Marx classified class into wage-labourers, capitalists and landlords, based on the ownership of the means of production. On top of this classification, Weber (1978 cited in Breen, 2005) specifies class in terms of ‘economic order’ and ‘social order’; class forms one’s economic conditions, while social status, such as caste, forms one’s social position. Weber’s view provides a firm ground for explaining child exploitation by class and culture. However, in his view, where we can locate gender in social relations is still a question.

The previous literature helps this study to define social class and understand the class structure in India. Bardhan (1982) defines social class based on land ownership and the means of production. Originating from Roemer (1982), five classes – a capitalist landlord, a rich farmer, a family farmer, a poor peasant, and landless labourer – are constructed in terms of status of self-employment and the process of hiring and being hired (Bardhan, 1982). Athreya et al. (1987) likewise define social class by land relations and labour relations. Further, they use a surplus, after considering income, cost, and wage for hiring labour, to separate class structure, which leads them to define social class as surplus appropriators, middle peasants, poor peasants and agricultural labourers. Although these studies are based on the agricultural sector, in India, the agrarian and non-agrarian sectors are rather 'integrated and mutually influence each other' (Athreya, 1990, p.154).

While the aforementioned studies do not take a gender approach, some scholars consider both social and gender relations. Mies (1980a) contends that subsistence reproduction is a base of capital accumulation.

> *The fight against sexism must, therefore, become an integral part of any strategy aimed at eliminating all class rule. To define women's emancipation only as a tactical goal in the process of class struggle misses the main point.* (Mies, 1980a, p.12)

Harriss-White (2003, pp.43–45) evokes interest in 'the intermediate social class' composed of small land-owners, rich and middle peasants, rural merchants and small-scale manufacturers and retailers. They are not necessarily rich but constitute a large population (ibid.). The intermediate social class might relate to gender-based work segregation of child labourers. For example, family firms involve sons as family labourers, while daughters are not allowed to work there (Harris-White, 2003, pp.113–115).

It seems that the parents' class affects children's work. Children experience different childhoods by class, especially in some cultures where children spend much of their time in domestic and paid work (Blasco, 2005). In India, children of farmers with low wealth and of casual labourers are at greatest risk of child labour. In peasant classes, children tend to work as family farmers, especially when families own land but do not have enough wealth to hire employees. Children's work on the family farm enables their parents to do other work or earn more income. Among casual labourer households, children work as non-agricultural or agricultural labourers. They might help to farm as well if those households have marginal land. Each class shows distinctive expectations for boys and girls. For example, in Africa, more girls than boys from lower classes tend to work in agricultural tasks as part of their household duties (De Lange, 2009).

This study applies class-gender relations in both the agricultural and non-agricultural sectors, though a comparatively small number of child workers or child labourers are found in the non-agricultural sector. The feminisation of agriculture in India has been supported by much research (Da Corta and Venkateshwarlu, 1999; Olsen and Mehta, 2006a; Garikipati, 2009). It is observed that men move into self-employment or workers in the non-agricultural sector, while women are kept in the agricultural sector as casual labourers. A similar trend is found among girls in India who are more likely to take responsibility for agriculture.

Class status can be proxied by land ownership and household assets or consumption. In Chapter 5, a household class is categorised based on the occupational status of household heads. With the IHDS 2011/12, I have used occupational codes (WS4 and NF1) that are based on National Occupation Classification (NCO) 2004. Firstly, farmers who have spent most of their time in farming (using the variable FM38) are categorised by the size of owned land. Occupation codes WS4 and NF1, 60–62 are also categorised as farmers. Second, waged workers who are paid daily (using WS9=1) are classified as agricultural and non-agricultural labourers. They are the key groups involved with child labour. Then, workers with occupational codes (WS4, NF1) 0–29 are categorised as professionals, which include managers and government officials. Finally, both salaried workers who are paid monthly or annually (WS9=2 or 3) and self-employed workers and do not belong with professionals, farmers or manual labourers are categorised as workers. Therefore, ‘workers’ in this study are a heterogeneous group including lower and higher occupational groups, clericals, other unclassified salaried workers, and business class (Iversen et al., 2017; Vaid and Heath, 2010; Vaid, 2012). The details on how to classify them are explained in Sub-Section 5.4.2.

The industrial composition of child labourers by household occupational classes and gender is diverse, which shows a clear difference in the sectors of work performed by boys and girls (Table 2.1). Agricultural labour is one of the worst forms of work, and boys and girls from the working classes are greatly involved in that labour. Family farmers are also commonly found among boys and girls. Meanwhile, boy child labourers are found in diverse sectors, especially as non-agricultural sector labourers. In farmer-families, a high percentage of male children are found as non-agricultural labourers and workers, while girl children are concentrated in farm work. This accords with the previous findings of Swaminathan (1998); many of the boys were employed in trade, commerce and manufacturing, while a small proportion of girls were employed in services and household manufacturing in urban India. In addition to that, there is a strong tendency for girls to be concentrated within the agriculture sector – traditional gender roles among farming households or son-preference families leads to an increase in girls’ participation in agricultural activities.

Table 2.1. Industrial composition of child labour by gender, occupations and household class

a. **Male** child labourers as percentage of total male children in each household class

	Ag. labourers	Non-ag. labourers	Workers					Marginal farmers	Small farmers	Middle farmers	Large farmers	Professionals	Others
			Lower salariats	Higher salariats	Clericals	Other salariats	Business						
(a) % of non-child labourers	**89.6**	**92.6**	**94.5**	**98.1**	**98.5**	**92.5**	**94.5**	**92.2**	**90.7**	**93.8**	**94.6**	**97.1**	**94.3**
(b) % of child labourers by occupations													
Agricultural labourers	5.1	0.5	0.5	0.2	0.0	0.3	0.1	0.9	0.7	0.5	0.0	0.0	0.6
Non-ag. Labourers	2.4	3.8	2.0	0.8	0.3	2.0	1.0	1.5	1.4	0.8	0.3	0.2	3.0
Workers	1.2	1.2	1.9	0.6	0.4	2.3	3.3	0.7	0.9	0.1	0.0	0.2	1.0
Family farmers	1.5	1.8	1.1	0.3	0.7	1.4	0.9	4.4	6.3	4.9	5.1	1.3	0.9
Others	0.2	0.0	0.0	0.0	0.1	1.4	0.2	0.2	0.0	0.0	0.0	1.2	0.2
Subtotal of (b)	**10.4**	**7.4**	**5.5**	**1.9**	**1.5**	**7.5**	**5.5**	**7.8**	**9.3**	**6.2**	**5.4**	**2.9**	**5.7**
Total: (a) + (b)	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

b. **Female** child labourers as percentage of total female children in each household class

	Ag. labourers	Non-ag. labourers	Workers					Marginal farmers	Small farmers	Middle farmers	Large farmers	Professionals	Others
			Lower salariats	Higher salariats	Clericals	Other salariats	Business						
(a) % of non-child labourers	**92.5**	**94.7**	**97.9**	**98**	**98**	**96**	**95.9**	**93.4**	**95.8**	**95.7**	**91.1**	**99.1**	**97.2**
(b) % of child labourers by occupations													
Agricultural labourers	5.3	1.3	0.7	0.2	0.3	1.8	0.6	1.2	0.8	1.0	0.3	0.2	1.1
Non-ag. Labourers	0.6	1.2	0.3	0.4	0.2	0.4	0.4	0.8	0.3	0.2	0.3	0.1	0.8
Workers	0.2	0.4	0.8	0.4	0.2	0.4	2.3	0.1	0.1	0.0	0.0	0.1	0.4
Family farmers	1.3	2.3	0.3	0.9	1.3	1.4	0.5	4.2	2.9	2.8	8.3	0.4	0.4
Others	0.1	0.1	0.0	0.1	0.0	0.0	0.3	0.3	0.1	0.3	0.0	0.1	0.1
Subtotal of (b)	**7.5**	**5.3**	**2.1**	**2**	**2**	**4**	**4.1**	**6.6**	**4.2**	**4.3**	**8.9**	**0.9**	**2.8**
Total: (a) + (b)	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: IHDS 2011/12, *Notes:* Child labour is defined by the same definition applied in Chapters 5 and 6; household class is defined by household head's job with the longest hours usually worked and size of land owned; each cell indicates percentage of male or female child labourers of the total number of the same sex children in each household class; the groups that have high risks compared to other groups are coloured (i.e., having more than 2 percent).

2.6.2. Social groups on child labour

Caste has a close relation to the incidence of child labour. Historically, Varna classifies four occupational orders (Brahmin – priest; Kshatriay – warrior; Vaishya – trading and artisan; and Sudra – agriculture), and Jati is a local, endogamous and hereditary group associated with one or more traditional occupations (Srinivas, 1984). Nowadays, the caste system does not limit people's occupations, but it influences their lifestyles, norms and values.

Cultural groups comprise the forward caste, other backward classes (OBC), scheduled tribes (ST or Adivasi), and scheduled castes (SC or Dalit). The scheduled caste (SC) was previously known as the untouchables, based on Hindu practices, and they make up 17 percent of the population (excluding Muslim and Christian converts; Mosse, 2018). On the other hand, OBC was defined, within a political context, as the community or caste with low socioeconomic status but not as low as the untouchables (Deshpande and Ramachandran, 2014). Majumdar and Madan (1967, in Hasnain 2016) defines a tribe as a social group with territorial affiliations, endogamous, hereditary, united in language, and following tribal traditions, beliefs and custom. They constitute a significant proportion (around 9 percent) of the Indian population (Joshi, 2010). The reality of scheduled tribes is that they are highly heterogeneous; however, there are concerns regarding low economic status, low population growth, education (illiteracy), and the use of resources (Chaudhuri and Chaudhuri, 2005).

Diverse child deprivations (poverty, mortality, education) of ST/SC children by state have been observed, e.g. severe child deprivation in Jharkhand, Rajasthan and Bihar (Ekbrand and Nandy, 2017). Indigenous (Adivasi) children are deprived of educational opportunities, and gender gaps in education are extended among older aged children because of tribal girls' active economic roles (Joshi, 2010). Adivasi girls are largely involved in the agricultural sector but also show comparably high participation in the non-agricultural sector. There are roles played by egalitarian social norms.

Sharma (1985) pointed out that caste and class overlap in rural India: large landlords are mostly from the upper caste, agricultural labourers are predominantly from the scheduled caste, and peasants from the middle agricultural castes. Upper-caste and lower-caste women experience different socioeconomic constraints under control of purdah (Sharma, 1985). Lower-caste women work freely on their own but have a dual burden of market work and domestic chores while upper-caste women are secluded in the domestic sphere. Both castes' women are responsible for household chores (Sharma, 1985, pp.75–76).

Liddle and Joshi (1989) emphasise the similarity of the gender restrictions in the upper or lower castes to those in the class system. The same authors (1989, pp.106–108) suggest that upper-caste (Brahmin) women have dual burdens from caste and class. In the caste system, women are under pressure to marry and prohibited from employment (ibid., p.106). Under the class structure, women's employment, though not entirely banned, is restricted, so women are confined to a limited variety of occupations (ibid.).

It should also be noted that caste and class are not the same. Caste is a tradition, but class is determined by the socioeconomic status of individuals or households. Owing to the similarities between caste and class, it is not easy to distinguish how they affect child labour. Mohanty (2002) asserts that castes form the class relations in Indian society: for example, land ownership is affected by caste status. On her study of Bihar, she emphasises that caste is 'a fundamental basis of social inequality' because caste determines whether people have access to social and material resources. Raju and Bagchi (1993) have found that the role of class and caste are mixed in the occupational division of labour by sex. They found that middle- and upper-class households are more likely to undertake agricultural management, but landless employees do all the physical work (Raju and Bagchi, 1993, p.101). There are variations by states: for example, in Tamil Nadu and Kerala, low-caste women usually transplant paddy, and some of higher caste women might also join the same work, although it is considered shameful. In Bengal, weeding is regarded as a job for lower-caste women, and harvesting is acceptable for the higher caste women, but no higher caste women would transplant paddy (ibid., pp.105-106). Overall, it should be acknowledged that the influences of class and caste on child labour can overlap; thus, the relationship between caste and class should be more carefully examined in a study of child labour in India.

Purdah, as a custom for Muslim and Hindu women in India, has greatly affected women's lives; the urban middle classes consider purdah as backwardness, but its values and attitudes remain (Mies, 1980b, p.68). The Hindu joint-family excludes girls from the right of inheritance and leaves them stereotyped gender-roles. Under purdah, women are forbidden from any economic activity (ibid., p.185). However, lower social groups show variations of purdah: women can work in farm or market, make bangles, or work as domestic workers (Mandelbaum, 1993, p.34). As families become more prosperous, purdah seclusion increases, but the education of girls is allowed (ibid.). Chen (1995) shows an example of how tradition limits women's positions in Bangladesh and North India. It secludes women and restricts their participation in work; female heads of households are not allowed to work outside the home because of kinship pressure, but in reality they have to cultivate their farms (ibid.). Meanwhile, middle-class women experience another variation of purdah, as they take up outside activities but still keep purdah values (ibid.). Tribal women in Jharkhand do not recognise that they have low status, as they have a prominent position in families and societies; however, inequality against women there still exists (Kelkar and Nathan, 1991).

Land relations are strongly related to women's caste membership (Agrawal, 2016). Women's power is controlled because Hindu caste social norms, region, and ethnics limit their land ownership. Notably, Hindu caste women are excluded from the inheritance of land, which generates more gender inequality. Among some religious or tribal groups such as Muslims or tribal communities in the North Eastern states, customary practices limit women's access to land, too. Agrawal (2016, p.141) emphasises that the limitation on the power of women represented by limited land ownership is instituted by social norms interacting with social structure (class, caste and race).

The literature has demonstrated that social groups divide labour among children. Being in a lower caste is found as a significant social factor that increases child labour, bringing a different pattern for rural and urban boys and girls (Das and Mukherjee, 2011; Das, 2012). Das and Mukherjee (2011) find that tribal girls in rural areas participate less in work, compared with other social group girls. It might be because their model combines child work and child labour. (The definitions of these terms are presented later on.) To the extent of severe labour, i.e., child labour, girls from the lower social group – scheduled tribes and the scheduled caste – might be more commonly engaged than those from other social groups. The caste system has a significant effect on workforce participation of girls (Kambhampati and Rajan, 2008). In some cases, the caste effect overweighs the gender effect on children's work participation (Majumdar, 2001).

2.6.3. Gender roles and child labour

Gender stereotypes and expectations produce segregation between the roles of female and male child labourers. In general, girls are more likely to be involved in domestic work while boys are more involved in waged work (Zapata et al., 2011; Koissy-Kpein, 2012). Benería (1979) explains that there is an underestimation of women’s domestic labour and their roles in rural or farm work because they are extensions of domestic work. Likewise, girls spend time completing household duties; therefore, they are often invisible and not counted. However, girls’ roles are not limited to domestic work, and sometimes they work outside in certain sectors. Girls work longer hours than boys with each increase in age in the Philippines; a similar trend is found in both market and domestic work in Ghana (Murray, 2004). Robson (2004) shows a great example of how to understand the different economic roles of girls and boys. She focuses on the spatial and gendered practice of child labour in rural Nigeria, where Hausa girls spend considerable working hours in reproduction and trading. They replace secluded female adults who are affected by seclusion norms. The findings of that study show how material gender relations are ‘socially constructed and rooted in structures’ (ibid., p.195).

In India, females’ roles in production are valued less than those of males. The feminisation of the agricultural sector is an example of restricted women’s roles in the primary sector. Similarly, girl child labourers are confined to reproduction or informal jobs, such as farming and home-based manufacturing. Nieuwenhuys (1994) explains that Indian rural child labour mostly occurs through families and kinship; while boys are employed and remunerated, girls’ work is carried out at home with other domestic duties and so they are less valued. Nieuwenhuys (1994) underlines that children’s work in peasant classes has not been adequately investigated, although it demonstrates diversity according to class, gender and age. In the Indian state of Kerala, while boys are hired as labourers, girls primarily work at home, in subsistence forms of production (ibid.). Girls’ work, such as peeling prawns or making coir yarn, are part of domestic duties, and therefore they are undervalued and poorly remunerated (ibid.). Among poor households, the incomes of adult males or boys are used for providing food for a family (ibid., pp. 25-26). In contrast, women have less-rewarding tasks at home, and girls’ roles are to assist their mothers (ibid.). This contradiction of child labour in Kerala occurs ‘on the basis of gender and seniority, with very young children and girls at the bottom and adult males at the top’ (ibid., p.199).

Another key gender construct is a possible exchange between child labour and women’s labour participation. Many studies believe that women’s work participation increases the burden on children (Weiner et al., 2006; Ilahi, 2001). In Pakistan, among the households where women are employed, the incidence of child labour increases (Ray, 2000). Kambhampati and Rajan (2008) insist that girls acting as substitutes for household chores when their mothers work outside the home is much more common than boys doing the same. Their finding is based on the relationship between women's labour force participation and child labour in domestic work. Meanwhile, this research hypothesises that supportive norms regarding women's employment might reduce child labour (see Chapter 6). This is based on the idea that norms can change individual behaviours, and it may take a long time to bring about social change. Hence, the relationship between norms and social structure requires in-depth understanding and long-term perspective.

Given that child labour is a result of the decisions of agents that are influenced by norms and values, it is important to understand which norms and institutions are reflected in the agents' decisions. Prior studies have rarely found the relationship between norms and child labour. Burra (2001) discusses how parents' unequal gender norms could devalue girls' education and increase their economic participation. Olsen and Morgan (2010) explain the existence of child labourer norms in terms of being good or honourable rural workers. Girl child bonded labour is an example that shows the integration of norms, institutions and structures; in many cases, girl children are bonded to labour usually because of their father's debts and their cultural obligations (Olsen and Morgan, 2015, pp.193–195; Venkateswarlu, 2007).

In short, child labour in India is the result of the combination of diverse norms within a given structure. Therefore, it is vital to investigate the influence of norms on child labour, in terms of gender and social structure. Chapter 5 reveals that social group-based norms limit girls' public work while girls and boys both work for many hours. In Chapter 6, two gender norms – seclusion norm ('practising purdah') and norms on women's work – are tested on the incidence of child labour. This research hypothesises that supportive norms regarding women's employment can reduce child-labour risks while seclusion increases them. Revealing the norms of children or parents is a difficult task, but it can provide meaningful implications that address the issue of child labour. In Chapter 7, based on the implications of findings, a few prestigious institutions are recommended, such as girls' education and women's empowerment, and setting a strict maximum working hours limit. Moreover, changing gender norms and institutionalised gender roles may require collective activities, such as grassroots movements.

2.7. Norms and Institutions and the GAD Approach

This research understands that child labour occurs within a social structure wherein girls and boys are differently affected by institutional attributions, such as norms, values and traditions. Sociological concepts greatly help to extend discussions on the norms and institutions in the context of GAD. In the next two sub-sections, I elaborate on how norms and institutions should be defined in the study of child labour, in accordance with the GAD approach.

2.7.1. Norms

In the GAD approach, norms are the keys to interpret agent behaviours in social and gender relations. Women's and gender studies have found that changing norms and improving women's agency are crucial stages in achieving gender equality (Drèze and Sen, 1995). Bourdieu's concept of '*doxa*' provides a basis for many gender studies that are interested in how individuals make decisions regarding gender practice. 'Doxa' is a self-evident part of the social world (Bourdieu, 1977, p.164). It is distinguished from the discourse, such as orthodoxy and heterodoxy that are debatable and disputable, as it is granted and undisputed, and therefore it reproduces structures (ibid.). It is a 'self-evident and natural order which goes without saying and therefore goes unquestioned' (ibid., p.166).

A social norm is, broadly speaking, a part of 'doxa' that is a natural and self-evident element of the social order (Bourdieu, 1977, pp.167–170 citied in Agarwal, 1997, p.159). Admittedly, social norms are less restricted forms than doxa, as they are subject to some argument, albeit a limited range of them (ibid.). Agarwal (1997) focuses on the role of norms in determining power (and the necessity of contestation). The author pointed out that social norms are everywhere: deciding categories of persons, gender division of labour, whether women should work outside the home, household decision-making, and how society shares resources. At the same time, norms can weaken the women's positions by discouraging them from working outside the home, restricting their bargaining power, and also, setting cultural expectations for gendered behaviours (ibid.). Control over property and key institutions, as well as group solidarity, make it possible for people to challenge norms (ibid., p.21).

Kabeer (1999, p.441) employs the concept of 'doxa' as a taken-for-granted, naturalised way of perceiving needs and interests. It means traditions and beliefs, which exist beyond discourse (ibid.). There is a role for 'critical consciousness' and 'being and doing' to move from naturalised 'doxa' to 'discourse' (ibid.). Thus, only when individuals are available with a 'critical consciousness' can the social order be revealed (ibid.). She emphasises that taken-for-granted rules, norms and customs are keys to forming gender relations, and women's adherence to social norms helps the reproduction of gender discrimination in South Asia.

Another concept from which norms are derived is 'habitus' (Bourdieu, 1977, 1990). Bourdieu casts light on the relationship between social structure, habitus and practices in his explanation of the theory of practice. Habitus is a 'system of dispositions' that are constructed through an individual's cognition and perceptions given the class of conditions (Bourdieu, 1990, p.52). Thus, it is strongly related to individuals' past experiences within a structure and a product of history (ibid., p.54). Although it is individually constructed (individual habitus), its basis in social structure comprises a group habitus (ibid., p.60). Habitus, as a perceived structure, produces practices, thoughts, and perceptions, establishes common sense, and reproduces regularities (ibid., pp.54–56). Habitus is a 'subjective but not individual system', as it generates perceptions and actions common to all members of groups and 'preconditions of objectification' (Bourdieu, 1977, p. 86).

Social practices rest upon social norms which are connected with Bourdieu's concept of habitus (Olsen, 2009). According to Bourdieu's (1990) concept, agents are likely to conform to norms, and so agents' confrontations with norms and, thereby, change of norms (habitus), are not highlighted. However, Olsen (2009) emphasises that the habitual and normed nature of social practices are subject to change due to agents' complex decision-making and 'strategy'. 'Strategy' is an ontological concept that goes beyond practices and explains the divergence of norms more carefully as part of the process of agents' decision-making in relation to the world (ibid.).

Overall, gendered practices are driven by agents' norms and beliefs. Identifying norms behind decisions helps reveal underlying gender-related problems. Challenging norms, such as moving from 'doxa' to 'critical consciousness', is necessary in order to change child labour practices (Kabeer, 1999). Norms as a habitus are built on people's relationships with others and their practices in a social structure. It is important to understand that changes in norms are possible and natural. Norms can explain the reasons why child labour is allowed and why boys and girls are involved in different work in a similar society.

2.7.2. Institutions

Norms and institutions are not entirely divided. Norms are expectations regarding people's behaviours, and institutions are more structured and settled social forms. Parsons (1952, p.39) defines an institution as 'a higher order unit of social structure' that is 'made up of a plurality of interdependent role-patterns'. He uses examples of property and marriage as the integration of actions and expectations, considering norms the bricks that compose social institutions. Bourdieu (1977, p.167) describes the relation between 'doxa' and 'institutions' in his explanation of 'symbolic capital'. 'Doxa' is reinforced by people's practices and institutions which constitute the collective thought of members of the group; therefore, 'instituted discourses' reduplicate 'the self-evidence of the world' (ibid.). Scott (2011) focuses on showing how norms, which are expectations and assumptions about normality, form a pattern, are clustered in society, and construct people's social positions, referring to this system as 'institutional structure'.

López and Scott (2000) explain that institutions 'are carried in the minds of individuals, but they have virtual objectivity that put them beyond particular individuals' (ibid., p.23). Therefore, institutions are objective but subjective as well in terms of individuals' social interactions (ibid., p.24). They admit that institutions are built from norms because those normative patterns are standardised, regulated, and shape people into social systems (ibid, p.25). Likewise, in the GAD approach, institutions are the ways through which gender practices are guided and regularised. Kabeer (1994, xiii) explains that 'gender and development focus on the construction and reinforcement of gender inequalities through the rules, procedures and practices of the key institutions.'

López and Scott's (2000) concept of 'institutional structure' provides insights for this study in terms of culture: 'Social institutions are cultural phenomena, and this preamble on culture has been necessary to bring out the fact that the institutional patterns that comprise a social structure have the same virtual existence as all other cultural phenomena' (López and Scott, 2000, p.22). Institutions related to child labour, such as family, patriarchy, regulations, laws and gender roles, are deeply associated with culture. By looking at institutions, we might comprehend how child labour is culturally constructed and accepted.

The process through which individuals are guided by institutions is driven by power. Revealing the relationships between power (structure) and institutions in child labour is a clear task for this study, though they are not explicitly distinguishable. Garikipati and Olsen (2008, p.329) state that 'agents behave according to their internal composition and their external relations'. Olsen (2009) emphasises that agents are neither entirely structural nor entirely individual. According to the GAD approach, a power that influences individuals through institutions is dependent on social and gender relations. Olsen and Morgan (2015) introduce an understanding of the conceptual differences between norms, institutions and structure: 'institutions are sets of norms which are predominantly followed in a particular concrete historical place in society. Structures are not simply the norms but rather the whole ensemble of parts into whole' (ibid., pp.188-189). Institutions are reflective of both internal norms and structure; therefore, revealing the roles of institutions is key to understanding child labour in a particular society.

Furthermore, there are multiple explanations of how structure affects agency. Elder-Vass (2007, 2008) emphasises the existence of 'causal power' originating from a social structure whereby a norm group conforms to the normative standard. Agents belong to multiple normative groups and different powers are involved (ibid.). However, in his emphasis on 'social structure as whole' (Elder-Vass, 2007, p.465), divergent decisions made by individuals and their relations with others are dismissed. I emphasise here that once power is given to individuals, it becomes internal; thus, the causal mechanism is not always evident. In addition, the GAD approach focuses on the relational structure rather than on structure as decisive power; agents reflect the (relational) structure in diversified ways through different institutions and norms.

Lastly, I want to point out that institutions exist not only at the local level but also at a global level. National laws or international regulations, education and working hours policies are the institutions most closely related to child labour. Some scholars insist that global regulations, such as minimum age laws, are unnecessary because child labour is a construction in the cultural context (White, 1996; Bourdillon, 2006). In contrast, James and James (2004) provide insights into how international conventions, such as the UNCRC on children, can reconcile the construction of childhood within societies. Children's universal rights can be satisfied for individual children in the cultural context (ibid., p.83). Similarly, education policy and laws are imposed in the context of culture by conventions ('primary means'; ibid., p.85). Thus, international level institutions can be influential over institutions at a local level as well as an individual level. This study supports the view that international, national and local institutions are equally important as they complement each other.

2.8. Key Definitions and Criteria

In this section, firstly, how to define child labour in this study is clarified based on the key stakeholders' definitions. Then, the time-allocation of child labourers is discussed, in order to find a suitable time threshold for child labour in India. Lastly, child labour is separated into formal and informal work. Describing the sectors of child labour in detail is important because gender roles are not the same in each sector of work.

2.8.1. Child labour and child work

This study defines child labour as work that is harmful to children's development, including domestic and non-domestic work that requires considerable time. Child labour results in physical or mental harm for children and therefore should be banned. In contrast, child work is work permitted provided hours worked are less than the time threshold. Child work does not lead to exploitation. Nevertheless, there could be risks of child work; child workers are often unpaid and the work might also impact children's learning outcomes and reduce their leisure time.

While this study uses a conservative definition of child labour using types of occupations and the number of working hours, many studies use broader definitions of child labour which extend the boundaries of limited forms of labour. For example, Kak (2004) includes any children who are categorised as 'others' rather than 'students' in a survey designed to define child labour. In this way, children who leave education early are included in the measurement of child labour, based on the assumption that children who are not in school possibly engage in work, especially in rural areas. Giri and Singh (2016) also use a broad definition of child labour, referring to 'nowhere children' who are neither in school nor at work; this definition adds to the numbers in child labour to a large extent compared to the nationally reported figure of child labour.

2.8.1.1. Definition of child labour by the UN

International policy interventions have been made to cut off global child labour by monitoring global production and pushing governments to follow the guidelines and regulations. The universalisation of the law has brought about a significant impact on the repositioning of children with the Convention on the Rights of the Child (CRC) being adopted in 1989. The CRC clarifies that every child has the right of harmonious development. Children have a right to be protected from economic exploitation that is harmful to their physical, mental and social development. The UN has set international targets (Sustainable Development Goals, SDG) to stop child labour by 2030[3].

[3] SDG 8.7: 'Take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour, including recruitment and use of child soldiers, and by 2025 end child labour in all its forms.'

In SDG 8.7, child labour is regarded as similar to other types of illegal labour such as forced labour, modern slavery and child soldiers. The protection of children, indicated in SDG 16.2[4], is the basis of the prevention of child labour. The indicator of child labour used in the SDGs is straightforward (Indicator 8.7.1: proportion and number of children aged 5–17 years engaged in child labour, by sex and age). The UN focuses not only on children but also on adolescents (ages 15–17). The International Labour Organization (ILO) and the United Nations Children's Fund (UNICEF) are the two most important agencies regarding international strategies for child-labour elimination. They suggest more detailed indicators than the SDG reporting figures should be used (and they also compile the data), which are listed below:

> Indicator 1: Proportion and number of children aged 5-17 years engaged in economic activities at or above age-specific hourly thresholds (SNA production boundary basis);
>
> Indicator 2: Proportion and number of children aged 5-17 years engaged in economic activities and household chores at or above age-specific hourly thresholds (general production boundary basis)[5].

These indicators include both 'economic activities' and 'household chores' as types of child labour; 'household chores' are referred to as 'unpaid household services' and included in the general production boundary but excluded from the United Nations System of National Accounts (SNA) production boundary. While discussions are ongoing over whether household chores should be included by the SNA, the ILO and UNICEF agree that they consider them part of child labour, as specified in SDG 8.7.1 – Indicator 2.

2.8.1.2. Definition of child labour by the ILO

The ILO provides the most detailed way of defining child labour, dividing it by hazardousness of work – hazardous working conditions or working hours – and by separate age groups. The ILO distinguishes a 'child in labour' from a 'child in employment' or a 'child in light work' (Edmonds, 2009). It does not provide an explicit definition of child labour but bases it on two conventions – the Minimum Age Convention of 1973 (No. 138) and the Worst Forms of Child Labour Convention initiated in 1999 (No. 182). Convention No. 138 clarifies that the minimum age for labour shall not be less than 18 years in any case. The worst forms of child labour are slavery, prostitution, illicit activities and hazardous

[4] SDG 16.2: 'End abuse, exploitation, trafficking and all forms of violence and torture against children.'
[5] Available on https://unstats.un.org/sdgs/metadata/files/Metadata-08-07-01.pdf, accessed 22 Jan 2020

work, and it is defined as any work that harms children's health, safety or morality. With the advent of the ILO's Worst Forms of Child Labour Convention (No. 182), the ILO's child labour policy intends to screen for the most severe forms of child labour and abolish them (van Daalen and Hanson, 2019). Since 2003, the ILO has not emphasised the improvement of working conditions of children under the minimum age but focused more on the eradication of all forms of child labour (ibid.).

The ILO estimation of child labour is based on the above conventions. The 18th International Conference of Labour Statisticians in 2008 provides clear categories and methodologies with which to measure child labour. The hazardousness of work lies at the centre of the ILO definition of child labour. Although the UN SDGs do not use hazardous industrial sectors as criteria of child labour, this study will stick with the hazardousness concept of child labour from the ILO. That is because using working hours only for calculating child labour could result in underestimation.

2.8.1.3. Definition of child labour by the Indian government

The Indian Child Labour Prohibition Act 1986 defines children as people below the age of 15 and bans them from physically dangerous occupations, such as transport, construction and some hazardous activities including bidi-making, carpet-weaving, and manufacturing cement, clothes and matches, etc. The list of occupations prohibited for children has been expanded with a further 18 specified occupations and 65 processes (activities). The act regulates hours of work for children; children shall not work more than three hours before a break for at least one hour, and the total hours (including break) shall not be more than six hours a day. The 1986 act did not ban children from working in domestic settings, agriculture and services, despite the fact these sectors represent a large part of child labour in India.

The Indian government's Amendment Act, which came into effect from September 2016, bans children (up to the age of 13 years) from work in any occupation 'except helping families or working in family enterprises after school hours and child artist' and bans adolescents (aged 14–17) from working in hazardous industries. However, criticism has been raised because family work and family enterprises might allow for exceptions. Hazardous industries include only three types – mining, work under inflammable and explosive conditions, and work in hazardous process[6].

[6] The Factory Act, 1948.

The Indian government's Amendment Act was established just before the Indian government ratified the ILO Child Labour Conventions in 2017. Although they are improvements following the efforts to eliminate child labour, further progress is needed. In some way, the Amendment Act does not reach international standards. For example, the definition of child labour from the Indian government is narrower than the international definition, as it allows domestic work for long hours and disregards many types of hazardous work being done by adolescents.

2.8.2. Child labourer working hours

Analysis of children's working hours plays a major role in defining child labour, and time thresholds are important criteria that differentiate child labour from child work. Nonetheless, there is ambiguity regarding the number of hours that should be used as the threshold. Moreover, child labourer's working hours represent the severity of their work. It is vaguely known that children work long hours but they are not precisely measured. The ILO's Hours of Work (Industry) Convention 1919 limits working hours to 8 hours a day or 48 hours a week, and these are legally binding restrictions. However, it does not specify the working hours of children separately. Most commonly, the ILO's (2017) time criteria for child labour are used: any economic activity $\geq$ 1 hour for ages 5–11; any economic activity $\geq$ 14 hours for ages 12–14; and any economic activity $\geq$ 43 hours for ages 15–17. UNICEF (2019) suggests categorising 21 hours of unpaid household services for ages 5–14 as another category of child labour. These thresholds reflect studies on the effects of different working hours; however, more research is required in order to reveal the relationship between working hours and possible harm to children in India.

A lack of time-use data is another problem when measuring a time threshold in India. Although the large datasets, such as the National Sample Survey (NSS) or the Indian Human Development Survey (IHDS), provide the weekly or daily numbers of hours children have spent in work, they have certain limitations, such as the exclusion of domestic working hours (e.g., the IHDS 2011/12) and less accuracy in the reported number of hours (e.g., the NSS 2011/12). A comparably small amount of data provides more specific information on time allocation, and this might be derived from a far less number of samples.

A handful of studies have provided information on Indian children's working hours. Analysing the NSS 1993–1994 time use survey, Hirway (2002) concludes that boys and girls aged 6–14 spent 24.3 hours and 18.6 hours a week, respectively, on SNA work. The weekly hours are then disaggregated for boys: 32.7 hours for manufacturing; 26.2 hours for construction; and 21.5 and 20.1 hours for animal grazing and farming (ibid.). Girls' working hours are recorded as 27.6 hours for manufacturing, 22.3 hours for construction, 20.8 hours for farming and 18 hours for animal husbandry, respectively (ibid.). Bhat and Rather (2009) show that child labourers in the handicrafts sector in Kashmir work 6–8 hours every day.

Regarding measuring time, Guarcello et al. (2004) provide the most relevant time-measurement model for child labourers in association with physical harm. They measure the probability of injury or ill health, using weekly hours as an independent variable, and compare the marginal effect on the number of harms in each additional weekly hour (ibid.). They found in Bangladesh that the probability of ill health increases sharply from approximately 30 hours per week (ibid.). Guarcello et al. (2007) regress child school attendance with ‘effective working hours’, which are calculated using hours spent on market and non-market activity. Each country shows different trends; for example, in Guatemala, an increase in work time from 20 hours to 40 hours reduces school attendance from 70 percent to 50 percent (ibid.). A steeper negative relationship is also found between working over 30 hours and school attendance in Cambodia (Dorman, 2008). A similar regression method was used across 44 countries with unpaid household services and the probability of attending school, but because of the large variation between countries, it only provided a general confirmation that 20 weekly hours is a suitable threshold for children performing unpaid household services (ILO, 2013).

Those time-measuring methods demonstrate a general trend of risk associated with children’s long working hours, either in a linear or non-linear relationship, but they do not provide accurate time points based on statistical tests. The study provides a time threshold model to estimate the critical point that might be considered the time threshold after which educational harm sharply increases. The ILO’s time thresholds are commonly applied and modified in many studies without considering country-specific conditions. This study avoids assuming harmful working hours but provides the best way to estimate a time threshold with real data (Chapter 6). The result of the findings shows that the international and national level of time thresholds are not consistent. Some child labour is not considered hazardous if the ILO’s time threshold is applied, such as 43-plus weekly working hours for ages 15–17. However, this study suggests using a time threshold established by evidence to account for the situation in India. As a result, it suggests the use of 38.5 hours of work per week for children aged from 14 to 17 years[7], as the time threshold for child labour in India (see Chapter 6).

[7] Child labour among children under age 14 is already illegal regardless of working hours according to the Child Labour (Prohibition and Regulation) Amendment Act, 2016.

2.8.3. Informal and formal work

There has been a tendency to underestimate the difference between children's informal and formal employment. The sectoral patterns of child labour are shaped by gender. Samantroy and Sekar (2016) found, using Census 2011, that the proportion of marginal workers among girls was 1.5 times greater than the main workers, and the gap does not exist among boys. For example, girls are more likely to work in agricultural activities that are highly casual and seasonal in nature (ibid.). This study focuses on the concentration of girl child labourers in the informal sector and their invisibility in the labour market. Separating informal and formal work is an important step towards including all groups of children who participate in diverse work. It is also important because the formal and informal sectors have varied work situations and difficulties. In Chapter 5, I divide industrial categories into formal and informal sectors in order to understand the pattern of female child labour more clearly.

In Breman (1996)'s definition, which focuses on the non-agrarian sectors, the informal sector includes self-employment and casual labour, which comprises 90 percent of all workers in India (ibid.). The amount of capital involved is commonly used as a standard method to separate the formal and informal sectors. With globalisation, the informal sector has been growing in India; it is also typical that many women are found in the informal sector. Informal sector workers usually have less capital and weak collective power, and therefore they suffer from low wages, job insecurity, no protection and sometimes debt bondage. Although Breman (1996) does not directly mention child labourers, many nuances are applied to children. Parents who work in the informal sector, such as urban migrated workers, might also involve their children in the informal sector.

Breman (1999) provides a good summary of the characteristics of informal labour in India: 1) the simplicity in production because of limited capital investment; 2) less protection from legal regulations; 3) difficulty forming unions; 4) a large number of home-workers who are women and children; 5) waged labour, such as contracting or sub-contracting, that involves middle men; 6) a high proportion of female and child labour; and 7) seasonality, in particular related to diversification in the rural economy. Olsen and Mehta (2006a) point out that in both rural and urban areas, many people work in informal enterprises, which are not reported as proper employment. Olsen and Neff (2007) draw attention to rural tenants, who tend to spend a considerable amount of time in unpaid domestic work, tenancy work and informal jobs.

Evidently, the majority of child labourers is found in the informal sector, in which Chapter 5 shows a clear trend of child labour. Although boys' labour force participation in the informal sector is much higher than that of girls, there is no clear difference in working hours. In fact, it is found that female child labourers' working hours are longer than those of male child labourers in the informal sector.

Children's participation in the formal sector is limited but exists. The probability of boys and girls in the participating labour force in the formal sector does not exceed 1 percent of all children (based on our own calculation using the IHDS 2011/12). Nevertheless, some adolescent boys are employed in agriculture, construction, textiles, and sales, etc. and are categorised as formal labourers (IHDS 2011/12). Regular farm employment, which is defined as farm-hand by Breman (1996), can include children, but it is different from casual farm labour. The tea plantation sector is another example of the traditional formal sector, in which child employment is classified into two types: permanent employment or temporary work (Bhowmik, 2002, p.149). Different types of exploitation happen within the formal sector because some permanent child labourers do not have a chance to attend school; in contrast, child labourers in temporary status suffer from lower wages.

Chapter 3. Methodological Review

A Bayesian approach is employed as the methodological basis for the empirical parts of this study. As using a Bayesian inferential framework is very rare in the areas of measuring child labour, and using the Bayesian method is almost new, I review the possibility of using this method in this chapter. I provide an overview of the Bayesian approach and the methodological considerations applied, starting with a brief introduction to the Bayesian approach and its advantages. The second section considers several methodological concerns raised in the three papers in this study, including model selection, data combination, model checking, and the selection of prior distributions. Following this, the chapter discusses key datasets that are used in this study – the NSS 2011/12 and the IHDS 2011/12.

3.1. Review of the Bayesian Method

3.1.1. Introduction to a Bayesian approach

A Bayesian approach is often compared with a frequentist approach. The fundamental difference between a Bayesian approach and a frequentist approach lies in how the nature of probability is regarded (Kaplan, 2014). While in a frequentist approach probability for any parameter relies on long-run frequencies, in a Bayesian approach the probabilities used as a means of presenting uncertainty and it thus has an epistemic status (ibid., p.284). The frequentist approach considers most parameters to be a scalar, whose value can be known in large samples or other asymptotic conditions. Gelman et al. (2013, p.11) explain that a Bayesian approach accesses whether any two parameter estimates are the same or not from previous knowledge or experience, and he proposes that estimates are based on given knowledge baskets. Thus, the Bayesian approach transforms the uncertainty of making assumptions into a matter of uncertainty in the information available, then thence into a complex probability distribution (ibid.). Therefore, it is certainly important to understand how uncertainty becomes altered into probabilities in a Bayesian approach. Priors are a special kind of parameter not found in the frequentist approach. The use of priors in a model usually is the key to the Bayesian approach.

The objectiveness or subjectiveness of a Bayesian approach has been discussed often, and I want to briefly introduce those discussions. Some consider Bayesian statistics a subjective approach (Gelman et al., 2013). The observer or scientist is the subject who perceives the information about a situation. When employing a frequentist approach, probability is based on the long-run frequency of events and uncertainty is somewhat fixed and objective (Jackman, 2009). Conclusions are considered independent of the observer. By contrast, in a Bayesian approach probability is based on a degree of belief and beliefs about types of things, and therefore uncertainty is considered more random and subjective (ibid.). Furthermore in a Bayesian approach if the evidence is held by an observer then there may be disagreement about the conclusions, so that even if there is 'objective' reality of the world to which the data refer, there is at the same time a 'subjective' element in the scientist's knowledge about that world. Theoretically, 'rationality' is the basis of the belief driving people's choices when they act (Bernardo and Smith, 2000); thus, belief in this context is not illogical but rather a 'logical process of decision making' in situations of uncertainty (ibid., p.15). Furthermore, some scholars insist that 'objective priors' can be applied (Berger, 2006). Kaplan (2014) highlights that the use of the 'evidence-based subjective' Bayesian approach, which uses priors based on evidence, can make objective use of previous knowledge. In this sense I accept that in a Bayesian approach there is some overlap of the 'subjective' role of the observers and the 'objective' nature of the entities to which our knowledge refers (Olsen, 2011, p.73–75)

Before discussing a Bayesian approach in the particular context of this study, I explain basic Bayesian statistics. Bayes' theorem is a simple formula explaining the probability of event A given that event B has already occurred (Eq. 3.1)[8]. It provides a straightforward way to understand the concept of Bayesian inference. In a Bayesian approach, the most interest lies in finding a posterior probability of key parameter θ given observation y. According to Bayes' theorem, a 'posterior probability' becomes proportional to the product of the probability of the observation given a prior assumption, $p(y|\theta)$, which is called the 'likelihood', and the prior probability (simply called a prior), $p(\theta)$ (Kaplan, 2014, p.8). Here, $p(y|\theta)$ is obtained from data, and $p(\theta)$ is based on our knowledge and beliefs.

$$p(\theta|y) \propto p(y|\theta)p(\theta) \qquad (3.1.)$$

In this study, the core aim lies to find correctly the number of child labourers, via its posterior distribution, $p(\theta|y)$. To calculate this, we need a likelihood, $p(y|\theta)$, that is a function of the data conditional on a prior distribution and some knowledge of the prior distribution, $p(\theta)$. In a frequentist approach, a likelihood is enough to make an inference, but for a Bayesian approach, some assumptions about the prior distribution are essential. A Bayesian model uses prior probabilities to describe current knowledge so that we can have

[8] $p(A \mid B)= p(B \mid A)\cdot p(A)/p(B)$

new posterior probabilities (Christensen, 2011, p.18). Assumptions made about a prior can transform the outcome, and this is addressed in Sub-section 3.2.5.

Furthermore, a Bayesian approach uses different inference procedures from those used in a frequentist approach. Bayesian inference means a process drawing a statistical conclusion about a probability of a parameter θ or unobserved data $\tilde{y}$ based on observed data y (Gelman et al., 2013, p.6). Thus, the primary tasks of Bayesian inference are to develop the model and perform computations to estimate 'posterior predictive distribution' of $p(\theta|y)$ using a prior distribution and observed data (ibid., p.7). A posterior predictive distribution means we predict values that are potentially observable but not yet observed, using a future sample and a posterior distribution (Cowles, 2013). Practically, we can calculate all relevant probabilities using Bayesian inference packages such as JAGS or Stan (see Annex C). Markov Chain Monte Carlo (MCMC) is the most common method of drawing values of θ from distributions and correcting those to approximate the posterior distribution, $p(\theta|y)$ (ibid., p.275). In Bayesian inference, credible intervals are used, rather than confidence intervals. A credible interval indicates that the area the parameter of interest falls into, which has a simpler interpretation than a classical confidence interval does (Lynch, 2007). Lynch (2007) refers to it as a 'credible interval' or 'posterior probability interval' while a (credible) interval of posterior predictive distribution for y is often called a 'predictive interval' (PI).

3.1.2. Advantages of a Bayesian approach

The Bayesian method is a useful way to estimate an unknown number of child labourers as it converts uncertainty into the matter of parameters. Child labour estimates have a large zone of uncertainty because of the complexity of defining child labour and the limited data sources, such as the true rate of child labour and over- or undercount of child labour cases. We need a flexible but accurate method to transform the various uncertainties into probabilities. In a Bayesian approach, uncertainty can be defined as a probability (Gelman et al., 2013, p.11). Bayesian statistics can incorporate uncertainty related to the measurement of child labour in a model by transforming it into a parameter and using a prior for the parameter. For example, we can use an 'undercount' parameter in a model to represent our partial information. A uniform distribution between 0 and 1 can be set as a prior of an undercount parameter as we know that one dataset does not have any records of children's domestic work, while the other has. In this way, the use of priors helps combine previous knowledge to obtain the best measurement result.

Bijak and Bryant (2016) describe the key advantages of a Bayesian approach as forecasting uncertainty and combining information such as data and expert knowledge. First, forecasting uncertainty is a crucial part of demographic studies, and a Bayesian approach incorporates all types of statistical uncertainty into a joint probabilistic forecasting model through prior distributions (ibid., pp.7-8). A Bayesian forecasting process is evaluated by calibration[9] of the probability distribution and of the precision of the error distribution, which provides an efficient way to interpret the outcome (ibid.). Secondly, a Bayesian hierarchical model uses both a direct estimate (likelihood) and a prediction from a model (prior), which allows a balance between robustness and sensitivity (ibid., p.7). For example, when the number of observations is small and erratic, the posterior distribution can be more strongly affected by a prior than by a likelihood.

Kaplan (2014) summarises the Bayesian advantages in a few categories: a marginal likelihood, $p(y|\theta)$, is directly compared through Bayesian factors, which is compared to a frequentist approach that needs nested models for a likelihood ratio chi-square test. Furthermore, it has flexibility in handling complex data structures, especially in hierarchical models – it does not matter if parameters are fixed or random, because any parameters are all considered uncertain (random) in a Bayesian framework; therefore, it is easy to fit any additional information into a model (pp.288–290).

Based on the literature review, I have three reasons for employing a Bayesian approach in this study of child labour. First of all, it provides a mathematical and practical way of combining data. This is a special feature which is original in the models of Chapters 4 to 6 in this study. Multiple datasets can inform various parameters in a model; those parameters may reflect various aspects of the problem under study (e.g. refer to different subpopulations). Exchangeability justifies the combination of multiple data sources. Under hyperprior (or hyper-parameter) characteristics of the population, different aggregates of datasets ($y_1, y_2,\ldots,y_n$) become exchangeable, which means they are considered random samples of the population. Methodologically, data integration is possible through using priors or model parameters in a Bayesian approach.

Secondly, a Bayesian approach can help solve the problem of rare events in the study of child labour. This aspect is well known, and it includes an ability to handle small sub-groups within multilevel models. It is not easy to measure the effect of regressors if the number of events is small, and also the estimation of a rare event can involve bias. Unlike a restricted maximum likelihood (REML) method, the Markov Chain Monte Carlo (MCMC) method performs efficiently with small sample sizes because it does not require either an adjusted degree of freedom (for the Kenward–Roger adjustment) or finite-sample

[9] Bayesian calibration means 'the activity of adjusting the unknown rate parameters until the outputs of the model fit the observed data' (Kennedy and O'Hagan, 2001, p.426).

corrections (for the REML; McNeish, 2016; Baldwin and Fellingham, 2013). However, if probabilities are small, the prior information has more effect on the posterior probability (Christensen et al., 2011, p.100). Thus, comparing priors and a careful selection of prior distributions are essential for measuring events where a small number of cases are available as data.

When multilevel models are estimated using maximum likelihood methods, the advantages of the multilevel model include shrinkage and efficiency of estimates in the small-sample sub-groups (Jackman, 2009). In this study, multilevel models are estimated with multiple rectangular datasets, which has rarely been carried out before – and never in the child-labour context. Such an estimate cannot be obtained using maximum-likelihood estimation methods, i.e. frequentist methods.

Third, Bayesian methods allow the accumulation of previous knowledge about child labour and maximise their use in a model, which is made possible by using informative priors. It provides a way to incorporate (prior) results from previous research into current findings (Lynch, 2007, p.71). Although this study uses noninformative and weakly informative priors, it is possible to derive a prior probability and use it to improve estimates of child labour in future.

3.2. Analytical Methods

In this section, I step forward to introduce methodological concepts of models and priors. This study has two basic frameworks – a Bayesian hierarchical model and a data-combined approach. Based on them, I choose a model for counts with many zeros, such as the number of child labourers and the number of working hours. Lastly, different priors are reviewed for application in the selected models. I suggest a few priors that are suitable for use with estimates of child labour.

3.2.1. A Bayesian hierarchical model

A Bayesian hierarchical model is the base model of this study: it assumes a hyperparameter at a higher level that affects a lower-level parameter. In the Bayesian hierarchical model, the observed outcome is modelled with parameters that are reliant on hyperparameters (Gelman et al., 2013). Gelman et al. (2013, p.101) explain that a hierarchical model is useful for large datasets that require a large number of parameters; a non-hierarchical model tends to overfit data, bringing an inferior prediction, but a hierarchical model reduces the chance of an overfitting problem by using enough parameters (ibid.). Overfitting refers to models that fit the existing data well but does not predict new data well (ibid., see Figure 3.2(c)).

‘Exchangeability’ is a key concept for understanding the Bayesian hierarchical model. It is a weaker assumption than independence, $p(A|B) = p(A)$; if two events are independent, they are exchangeable, but exchangeability does not necessarily mean independence (Kaplan, 2014, p.17). Under exchangeability, the joint distribution of data is irrelevant to the order of observations. Goldstein and Wooff (2007, pp.178–179) explain exchangeability as follows: our belief about the unobservable probability is not affected by any permutation and invariant of the probability for obtaining the subset of any data. Exchangeability is not a term that applies only to a hierarchical model, but it also justifies the use of a Bayesian method. Bernardo (1996) and Bernardo and Smith (2000) provide a clear explanation of the basic idea of exchangeability and how it allows the combination of a likelihood and a prior. If we have a joint distribution of y_i for a set of i={1, 2, 3,…, n}, under exchangeability, all possible permutations, π, of the {1, 2, 3,…, n} (regardless of the order of them), will have the same probability as below[10]:

$$p(x_1, x_2, \ldots, x_n) = p(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}). \qquad (3.2)$$

Exchangeability is important because it is a strong condition of the General Representation Theorem (De Finetti, 1974). According to the General Representation Theorem, if $\{x_1, x_2, \ldots\}$ is infinitely exchangeable and $\theta \in \Theta$ (Θ is the limit of the function of x_i), the joint probability of the subset $p(x_1, x_2, \ldots, x_n)$ takes the below form[11]:

$$p(x_1, x_2, \ldots, x_n) = \int_\Theta \prod_{i=1}^{n} p(x_i|\theta)p(\theta)d(\theta). \qquad (3.3)$$

The above formula is a simplified version of de Finetti’s theorem (Bernardo, 1996), but the meaning is the same. If observations are exchangeable, any subset of them is a random sample from some model ($p(x_i|\theta)$) and there exists a prior distribution ($p(\theta)$) (ibid.). Here, θ might not belong to the subset but is available information and also affects the model ($p(x_i|\theta)$). Thus, exchangeability justifies the procedure of combining a likelihood and a prior (Bernardo and Smith, 2000, p.174), thereby advocating the use of a Bayesian approach. Furthermore, it justifies the integration of different samples under exchangeability, which I am going to explain further in the next section (Sub-Section 3.2.2). While infinite exchangeability is not feasible in reality, we could consider ‘partial exchangeability’ (Bernardo and Smith, 2000, p.170). Kaplan (2014) calls this ‘conditional exchangeability’, whereby exchangeability is allowed for each group in a hierarchical model that has a similar condition, as long as groups rely on hyperparameters.

[10] Bernardo, 1996, p.2
[11] ibid.

We can assume different priors for each group; in such a case, exchangeability is still allowed but conditional to different priors for each group. Kaplan (2014, pp.184-185) explains that observations are clustered into second-level groups; they are exchangeable, conditioned by our knowledge of groups. More specifically, if groups are not different from one another, we apply the same prior for all individuals, but if not, we use 'conditional exchangeability' for observations within groups (ibid., p.17, pp.184–185). If exchangeability is allowed between groups, then they are considered random samples from a population (Bernardo, 1996).

Here is the notation for a Bayesian hierarchical model (Bernardo, 1996). The individual samples, $\{x_1, x_2, \ldots, x_n\}$, are exchangeable under the condition of the parameter, θ_i (Eq. 3.4); the group parameters $\{\theta_1, \ldots, \theta_n\}$ are also exchangeable under the condition of Φ (Eq. 3.5). The Φ is called a hyper-prior and is characteristic of a population (Eq. 3.6). Φ not only affects group parameters but also affects the posterior distribution of individuals. If there is no information for the parameter, Φ, such as population mean or variance, it can be derived from the posterior distributions of individuals (ibid.). That is possible because individual samples are drawn from the same population.

$$p(x_1, x_2, \ldots, x_n | \theta_1, \ldots, \theta_n) = \prod_{i=1}^{n} p(x_i|\theta_i), \quad (3.4)$$

$$p(\theta_1, \ldots, \theta_n \mid \Phi) = \prod_{i=1}^{n} p(\theta_i|\Phi), \quad (3.5)$$

$$p(\Phi).^{12} \quad (3.6)$$

In addition, in a hierarchical model, all the information (individual and group level) is integrated, so that shrinkage occurs. A posterior mean is shifted away from the group mean (θ_i in the above equation), and shrunk toward the grand mean (Φ) of a distribution, which is referred to as shrinkage (Jackman, 2009). In a Bayesian sense, shrinkage is caused when the posterior estimate is moved from the sample mean towards the prior mean (Zhao et al., 2010). When the number of samples is small so that variances in individual data are large, shrinkage increases; on the other hand, when group variance is large, shrinkage decreases because the distribution relies more on the information of each group and is less affected by the information from other groups (Jackman, 2009).

3.2.2. A data-combining approach

This study uses a data combination of multiple data sources (in Chapters 4 and 6). Bayesian data combination has been tried in several areas. For example, in epidemiology, 'multiparameter evidence synthesis' is used to synthesise information, such as combining

[12] ibid., p.5

multiple data sources, aggregate and individual data, and different models and methods, all of which became possible using Bayesian inference (Ades et al., 2008). In a demographic study, Wiśniowski (2016, 2017) has estimated the size of migration using a Bayesian data combination that combines two different data sources – migration data from a sending country and data from a receiving country.

Again, exchangeability justifies the combining of information. For the purpose of 'borrowing strength' (Jackman, 2009, p.307), we could consider using the different sources of data. A data combination is possible given two related conditions: one is hierarchical modelling, and the other is exchangeability of group means (Jackman, 2009). A hierarchical structure of the model allows a hyperprior (or a hyperparameter) that is conditioned by priors from each group (e.g., $p(\theta_1, \ldots, \theta_n | \Phi)$). Group parameters are exchangeable, given a hyperprior; therefore, they are regarded as random samples of the population. Here, group information can come from different data sources. The information from the first group is reflected in those hyper-parameters, and under exchangeability these hyper-parameters affect the posterior density of the remaining groups (Jackman, 2009, p.307). The other way around is also feasible.

In a Bayesian context, as long as there is a prior or hyperprior, this data-combining method is justifiable. Moreover, data combination is possible through the use of diverse types of prior. For example, historical data can be used as a prior for a model; this is called a 'power prior' (Ibrahim et al., 2000; 2015). Recently, Sakshaug et al. (2019) have tried to combine small probability samples that are accurate but expensive and those that are less expensive but also less accurate and potentially biased non-probability samples. They have been attempting to obtain informative priors from non-probability samples and then apply them to small probability samples, which was found to be efficient in reducing errors (ibid.). Another possible method is to use a weight parameter (shrinkage parameter) that originates from a different source. For example, Kaizer et al. (2018) name this type of model a 'multisource exchangeability model', in which non-exchangeable supplementary data is combined as a weight prior. A data-combining approach makes it feasible to combine individual- and regional-level data from multiple data sources (Olsen et al., 2020).

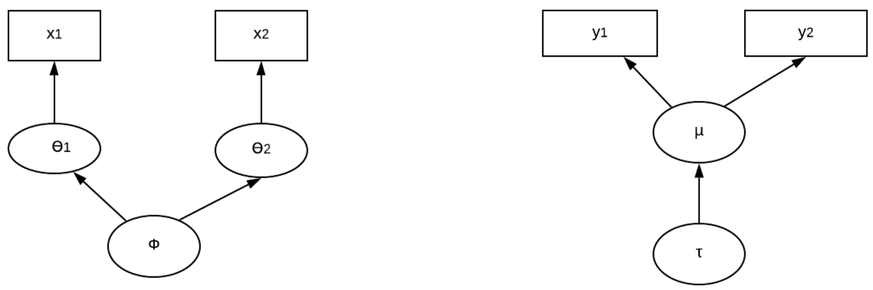
Figure 3.1(b) shows a simplified framework of the data combination in this study. This study combines different data sources via the common parameter of the interest (μ) (Eq. 3.7). y1 and y2 are sets of grouped data obtained from various sources. Both are based on a common parameter (μ) based on a hyperparameter (τ) (Eq. 3.8). y1 and y2 are exchangeable, conditioned on the common parameter, μ, and also affected by the hyperparameter, τ (Eq. 3.9). This hierarchical structure is slightly different from the Bernardo's (1996) model

(Figure 3.1(a)) because the structure (Figure 3.1(b)) does not involve individuals. In Figure 3.1(b), each group mean (y1 and y2) replaces the position of individual data (x1 and x2), and the common parameter, μ, becomes a second-level mean. τ is a hyperparameter. The primary purpose of this study is to find the correct number of child labourers derived from μ; therefore, this framework generates a probability distribution of a true mean based on two data sources. The model of this study can be simply written as below:

$$p(y_1, y_2 | \mu) = \prod_{i=1}^{2} p(y_i | \mu), \quad (3.7)$$

$$p(\mu | \tau), \quad (3.8)$$

$$p(\tau). \quad (3.9)$$



(a) Barnardo's (1996) model (b) The model of Chapter 4

Figure 3.1. Data-combining model structures

In Chapters 4 and 6, observations are at group level, and they are assumed to be conditionally exchangeable as they are from the same population. Although individual-level responses are not directly used in those chapters, individuals are assumed to be conditionally exchangeable, within groups such as states, ages or occupational groups. In short, data combination is justified based on exchangeability and hierarchical modelling, which are key components of this study.

3.2.3. Modelling with excess zeros

There is an excessive number of zeros due to the underreporting of children's work participation or working hours, and this study paid some attention to dealing with them. Much of the literature suggests a careful use of models when many zeros appear (Humphreys, 2013; Cameron and Trivedi, 2005; Winkelmann, 2003). The solutions to deal with zero inflation include use of a Tobit model, Heckman model, or a hurdle Poisson model. In this sub-section, I compare the three possible models dealing with excessive zeros and explain the reasons I chose the Poisson hurdle model for Chapter 5. Before explaining those models, I first describe a Poisson model, which is a basic step towards a Poisson hurdle model.

3.2.3.1. A Poisson model

The response of children's work participation is used as a key variable in this study and has a binary outcome of 0 and 1. Therefore, the count of child labourers is discrete and positive numbers, including zero counts. A Poisson model is used for count data, which has λ (rate parameter) as a key parameter (Hilbe, 2011). A Poisson distribution has non-negative integers as y-values; therefore, it includes the possibility of y=0 (ibid.). Another possible model for the discrete count of child labour is a binomial model (Gelman and Hill, 2007, p.112). However, a Poisson model is more suitable in this study because the probability of child labour (p) is small while the total size of n is quite large. There are many non-child labour cases, and we need a distribution to capture the probability that is lower than 0.05. When n is large but p is small, a Poisson model is a better choice than a binomial model (Pollard, 1997).

The number of working hours can be considered a discrete variable, too. In both IHDS 2011/12 and NSS 2011/12, the number of hours or intensity of work was recorded as discrete numbers. In Chapter 5, the count of working hours is positive numbers, including zero counts (before the truncation), so I used Poisson distribution that uses an exponential function for the linear predictor. The binomial model is another possible model to measure the number of hours spent on work. Similarly to the above, children's working hours include many zero cases, and for low probability with the large size of n, a Poisson model can be a suitable choice (ibid.).

Overall, a Poisson distribution can be applied to both children's work participation and working hours. It uses an assumption that the mean and variance of a model should be identical. However, the presence of over- or under-dispersion often violates this assumption. Inclusion of outliers, important explanatory variables being missing, and an insufficient number of datasets can be reasons for dispersion (Hilbe, 2011, p.39) and the inclusion of a large number of zeros in the data can result in overdispersion. More details about overdispersion are dealt with in a combined data model (see Chapter 4). How to address it using an adjustment parameter (i.e., Poisson log-normal model) is explained there.

3.2.3.2. A Tobit model

A Tobit model (Tobin, 1958) is another possible model with which to analyse working hours when there are many zeros in the data. Webbink et al. (2015) have applied a Tobit model to analyse working hours in paid work done by children. A Tobit model uses a normal regression, censoring figures below a threshold, usually zero (Cameron and Trivedi, 2005). The standard Tobit model can be denoted as below[13]:

$$y^* = x_i\beta + \varepsilon, \qquad (3.10)$$

$$y = \max(0, y^*). \qquad (3.11)$$

y_i^* are treated as zero if $y^* \leq 0$, although x_i is observed. A Tobit model is one censored regression. Therefore, unlike a two-stage model, we cannot measure the effect of children's labour force participation and working hours separately. A critical comparison leads us to prefer using a hurdle model rather than a Tobit model. Stewart (2013) proves that estimates from a Tobit model show more bias than a hurdle model through his simulation study of daily time-use. A Tobit model results in misreading as it treats a participation decision as the same as an amount decision (Humphreys et al., 2010). Overall, when there is a clear different indication of participation and quantity in any case, a hurdle model might bring about a better result as it distinguishes two different aspects of work – work participation and working hours.

3.2.3.3. A Heckman (sample selection) model

Unlike a Tobit model, a Heckman model (Heckman, 1979) has two steps. It introduces two variables – y_1 (participation) and y_2 (frequencies), which are explained by x_1 and x_2, and the correction of the selection bias (ibid., p.550). A Heckman model can be simplified as below[14]:

$$y_2 = x_2\beta_2 + \rho\sigma_2\lambda(x_1\beta_1) + u_i. \qquad (3.12)$$

In a Heckman model, positive value y_2 is regressed on a truncated explanatory variable ($x_2\beta_2$), the inverse Mills ratio ($\lambda(x_1\beta_1)$) derived from a first-step probit regression using outcome y_1, predictor x_1, and u_i (error term; Cameron and Trivedi, 2005, p.550). For example, y_1 can be children's labour force participation and x_1 can be socioeconomic variables. ρ indicates a correlation between the two errors in the first and second steps, and σ_2 is the standard deviation of u_i (ibid.).

One difference between a Heckman model and a hurdle model is the inclusion of the inverse Mills ratio. In applying the inverse Mills ratio (IMR), collinearity often occurs, and when the inverse Mills ratio is an approximated linear, such as with $x_1=x_2$, a multilinearity problem can be severe (Cameron and Trivedi, 2005). Puhani (2000) explains that there is

[13] Winkelmann and Boes, 2006, p.212
[14] Cameron and Trivedi, 2005, p.550

possible collinearity with most observations unless the probability of selection is higher than 97.5. Including a variable in x_2, which is a good predictor of y_2 while it does not appear in x_1 nor is associated with y_1, can resolve an identification problem (Little and Rubin, 1987).

Identification of a sample section model is a critical issue in both frequentist and Bayesian approaches. However, only a few Bayesian studies consider this during the sample selection (Li, 1998; Chib et al., 2009). Van Hasselt (2014) explains that when two sets of covariates (one set for a sample selection, and the other for a second stage) are identical, many causal effects become non-identifiable. The solution involves including in the selection stage at least one variable (called an instrument) that is not included in the linear regression in the selection stage, which allows identification of the effect of 'unobserved confounding' so that all the causal effects are recovered (Li and Tobias, 2014).

A Heckman sample selection model can be applied in order to explain the different determinants of labour participation and working hours. However, there are a few reasons that this study prefers a hurdle Poisson model to a Heckman sample selection model. Firstly, a Heckman model does not link the outcome of y_1 (sample selection) with y_2 (frequencies). It is different from a hurdle model in that outcome is directly affected by the probability of being observed. If we want to look at a joint outcome of work participation and working hours, a hurdle model is a more suitable choice. Secondly, the impact of adding the inverse Mills ratio (IMR) in the model is not so immense in the case of this study. In Chapter 5, I have checked the effect of IMR on Model 1, Model 2 and Model 3, using maximum likelihood estimation. The difference due to including the IMR in the models was only marginal, bearing that the results of the maximum likelihood estimation and Bayesian methods are not directly comparable. Lastly, there is a technical pitfall. A Heckman model is mostly used with linear regression; therefore, if we apply it to Poisson regression, extra coding is required[15].

3.2.3.4. A hurdle Poisson model

Cragg's model (1971), which is most commonly referred as a hurdle model, shares a similarity with the Heckman model; it does not use the inverse Mills ratio so that it does not correct selectivity but it generates a joint outcome of participation and frequencies. In the first stage of a hurdle model, decision y (0 or 1) is regressed on explanatory variables. In the second stage, if y>0, a truncated normal distribution (or truncated Poisson in this study) is

[15] For example, in STATA, a Heckman Poisson model has been recently included, and in R, the generalised linear model (e.g. glmer in lme4 package) does not include yet the option to calculate the inverse Mills ratio.

applied, but it is linked with the probability of crossing the hurdle. The function of y given x is specified below (Cameron and Trivedi, 2005, p.545):

$$f(y|x) = Pr(d=0|x) \qquad \text{if } y=0, \qquad (3.13)$$

$$Pr(d=1|x)*f(y|d=1, x) \qquad \text{if } y>0. \qquad (3.14)$$

The first part (1) is a probit or logit model, and the second part is a truncated normal regression with positive values. Regressors appear the same (x) in both parts to simplify the model, but they do not need to be identical (ibid.). A hurdle model can be applied to count data. Mullahy (1986) first discussed a Poisson hurdle model for it that uses a binomial probability model[16] and then a truncated-at-zero-count data model[17] with positive figures after the hurdle is crossed. Here are the notations of a Poisson hurdle model (Neelon et al., 2013).

$$Pr(y_i=0|x_1) = 1 – p, \quad 0 \leq p \leq 1, \qquad (3.15)$$

$$Pr(y_i=k|x_1, x_2) = p* \frac{\mu^k e^{-\mu}}{k!(1-e^{-\mu})}, \ k=1, \ldots, \infty, \ 0 < \mu < \infty. \qquad (3.16)$$

The first part (1) is a probit or logit model, and the second part is a truncated normal regression with positive values. Regressors appear the same (x) in both parts to simplify the model, but they do not need to be identical (ibid.). Like a Heckman selection model, having the same set of regressors in the first and second parts makes identification of the parameters difficult (Newman et al., 2003). As a solution, we can consider using exclusion restrictions that are variables applied in one equation but excluded from the other.

In the study of child labour (Chapter 5), a probit function is used to calculate the probability of zeros (attending work). This probability, p, is linked to the log-likelihood of a Poisson distribution. I apply different regressors for two steps: on the first stage, x_1 is used; and on the second stage, x_2 is used. x_1 is reflected in the probability (p), which also affects the second stage. p is quite small because of an excess of zeros, which might lead to over-dispersion. The expected value of a Poisson model decreases because of p, which introduces a difference between mean and variances, and therefore results in overdispersion in the case of $0< p <1$ (Winkelman, 2003, p.140).

[16] Mullahy's (1985) examples suggest using a Poisson or logit model for the binomial probabilities.
[17] A truncated Poisson or truncated geometric model is introduced with a positive count (Mullahy, 1985).

3.2.4. Model checking

In this sub-section, I introduce methods for checking the fit of a model and comparing models. Model checking is a vital process; much of the literature suggests diverse methods for evaluating the appropriateness of a model using deviance or residuals. It can also be made in a qualitative way such as looking at the features of replicated data (Gabry et al., 2019). Gabry et al. (2019) show that the effect of the prior (without data) can be checked through visualising the prior predictive distribution. More discussion on priors is provided in Sub-Section 3.2.5. By contrast, if a model estimate is based on extensive relevant information, the posterior result does not have much difference from the direct estimates. Regression usually sits between these two extremes.

3.2.4.1. Posterior predictive check

A posterior predictive check compares observed data with simulated values from the joint probability distribution (Gelman et al., 2013, p.143). Here I introduce a Bayesian r-squared and a Bayesian p-value, as well as a graphical way to do posterior predictive checks.

More recently, Gelman et al. (2018) suggest using a Bayesian r-squared. The classical r-squared statistic has two critical problems in a Bayesian context: 1) it does not reflect uncertainty in coefficients and so cannot show the overfitting problem, and 2) because of using a prior, sometimes fitted variance is bigger than observed variance, yielding an r-squared greater than one (ibid.). We can use the predicted values of simulated draws, $y_n^{\text{pred}_s}$, instead of point predictions $\hat{y}_n$, together with the residual variance ($\text{var}_{\text{res}}^{s}$): both are conditioned on parameters. Thus, a Bayesian r-squared is the ratio of the variance of predicted values to the sum of the variance of predicted values and expected errors, which is always between 0 and 1. However, in a Bayesian r-squared, the denominator is not fixed; therefore, comparing R^2 between models is not possible. Furthermore, it does not yet overcome overfitting. I used the Bayesian r-squared in Chapter 6.

Table 3.1. Comparison of classical and Bayesian r-squared

	Classical R^2	Bayesian R^2
Definition	Variance of the predicted values divided by the variance of the data	Variance of the predicted values divided by the variance of predicted values plus the expected variance of the errors
Formulas	$R^2 = \frac{V_{n=1}^{N}\hat{y}_n}{V_{n=1}^{N}\, y_n}$	$\text{Bayesian } R^2 = \frac{V_{n=1}^{N} y_n^{\text{pred}_s}}{V_{n=1}^{N} y_n^{\text{pred}_s} + \text{var}_{\text{res}}^{s}}$

Notes: Summary from Gelman et al., 2018; V stands for $V_{n=1}^{N} z_n = \frac{1}{N-1}\sum_{n=1}^{N}(z_n - \bar{z})^2$ for any vector z; $y_n^{\text{pred}_s}$ stands for the predicted values of simulated draws; $\text{var}_{\text{res}}^{s}$ stands for residual variance.

A Bayesian p-value provides a test score for the fit of the data to the model. It is similar to the significance test but not precisely the same (Christensen, 2011, p.59). A Bayesian posterior p-value means there is a probability that a future observation is more extreme than the data, being obtained by average p-values (Gelman, 2005). A Bayesian p-value can be categorised into three by type of measurement – the posterior predictive p-value (ppp), the sampled posterior p-value (spp) and the calibrated posterior predictive p-value (Zhang, 2014). The ppp value uses the same data twice – first, to obtain the posterior distribution and, second, to calculate the probability (ibid.). Conversely, the spp value compares observations and multiple replications drawn from 'a single random parameter value', so it can avoid double use of data (ibid.).

The posterior predictive p-value, which is written as 'Pr(T(y.rep) > T(y) |y)'[18], has 'a distribution that is uniform, if the model is true' (Gelman et al., 2013, p.151). P-values close to 0 or 1 are regarded as extreme; therefore, a Bayesian predictive p-value suggests a good fit at around 0.5. Usually, a Bayesian predictive p-value between 0.05 and 0.95 is considered an adequate fit (Gelman et al., 2013; Neelon, 2013). However, there are some concerns regarding the use of a Bayesian p-value. Bayarri and Berger (2000, p.1128) point out that if 'the null model has unknown parameters, p values are not uniquely defined'. Furthermore, for most cases, 'exact uniformity' for the p-value is not achieved (ibid.). Gelman (2013, p.151) explains that the p-value test shows a 'statistical significance' but does not demonstrate a 'practical significance', e.g., overfitting. Accordingly, when applying a Bayesian p-value, cross-validation is required, which is 'fitting the model to training data and then evaluating this predictive accuracy on a holdout set' (Gelman et al., 2013, p.170).

The predictive p-values can take different discrepancy measures (Gelman et al., 1996). Discrepancy measures can be residuals (such as a Pearson's chi-square test score[19]), deviance residuals or a likelihood ratio test ($2\sum y\log(y/E(y|\theta))$) (Lunn et al., 2012, pp.143–156). Another simple method of model checking is using a Bayesian mid p-value. A mid p-value can be used for discrete data and is the probability if y.pred>y.obs plus half of the probability if y.pred=y.obs[20] (Lunn et al., 2012, p.149). In this test, 0 shows an extreme case when predicted numbers are always bigger than observations, and 1 indicates the other way around. A mid p-value around 0.5 is considered desirable.

Many studies use a posterior predictive check using the same samples. It is possible, to use the same data in estimation and in predictions, but it is preferable not to. Out-of-

---

[18] T(y) refers to test statistics.

[19] Pearson's chi-square test can be calculated by $X^2 = \sum(y-E(y|\theta))^2/E(y|\theta)$ (Christensen, 2011, p.59).

[20] pr(y.pred>y.obs) $+\frac{1}{2}$pr(y.pred=y.obs) (Lunn et al., 2012, p.149)

sample test methodology provides a way to minimise the effect of the double use of the data. A Bayesian p-value that is used in this research is the posterior predictive p-value (ppp). As mentioned earlier, using the ppp requires a little caution – I would use it with the cross-validation method and check it with other statistical indices. A splitting dataset, such as a 90 percent training sample and 10 percent out-of-sample method, is commonly used for cross validation. Lunn et al. (2012, p.149) suggest using a subset of data; about 10 percent of an entire dataset (y_c) is left out and compared with the predicted outcome, y_c^{pred}, given the information from the other 90 percent sample of data (y_f). Thus, a Bayesian p-value is written as p-value = $pr(T(y_c^{pred}) \leq T(y_c)|y_f)$ (ibid.).

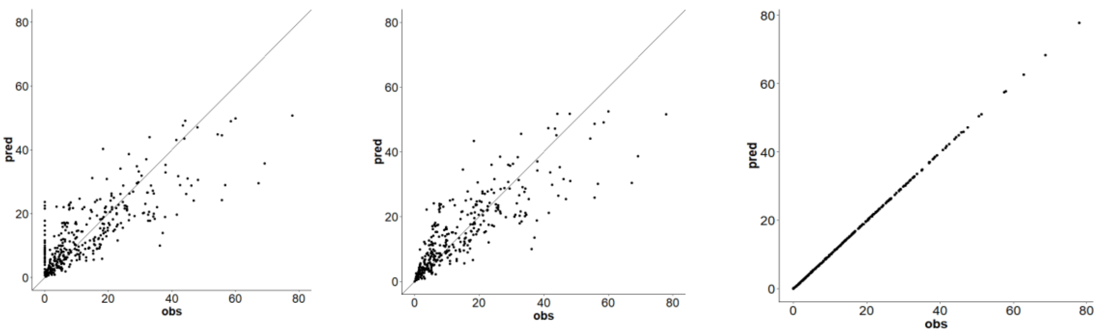
Here I suggest the example of a posterior predictive check of the data and models used in Chapter 5. First of all, a Bayesian posterior predictive p-value helps to check whether there are any extremes in the models. For test purposes, three different models – a hurdle Poisson, hurdle and zero-inflated (ZI) Poisson and hurdle log-normal Poisson – for child-labour informal hours are compared using the same data. For model checking, a 10-percent out-of-sample prediction method is applied. In Table 3.2, the hurdle Poisson model or hurdle and zero-inflated Poisson model show that around 50 percent of simulations show a Bayesian mid p-value between 0.05–0.95, indicating there is not a significant discrepancy between prediction and observation. On the other hand, hurdle log-normal Poisson model indicates all the acceptable ranges of a Bayesian mid p-value in entire simulations. If overfitting occurs, any Bayesian p-value falls within acceptable ranges; it cannot prove whether there is an overfitting problem. It shows that using a Bayesian p-value without cross-validation could result in misunderstanding. Thus, it should be remembered that a Bayesian p-value is efficient but not sufficient to check a model. The following chapters use a Bayesian r-squared and various information criterion (such as the DIC and WAIC) together with a Bayesian p-value.

Table 3.2. Results from the cross-validation study

	Hurdle Poisson Model	Hurdle and Zero Inflated Poisson Model	Hurdle Log-Normal Poisson Model
Simulation 1	**0.36**	**0.07**	**0.70**
Simulation 2	**0.27**	**0.56**	**0.12**
Simulation 3	1.00	**0.77**	**0.15**
Simulation 4	**0.15**	0.99	**0.95**
Simulation 5	0.03	**0.93**	**0.82**
Simulation 6	1.00	**0.89**	**0.81**
Simulation 7	1.00	0.99	**0.82**
Simulation 8	1.00	0.03	**0.39**
Simulation 9	**0.84**	0.99	**0.63**
Simulation 10	1.00	0.02	**0.50**

Notes: the mid p-value, pr(y.pred>y.obs) +$\frac{1}{2}$pr(y.pred=y.obs), is applied (Lunn et al., 2012, p.149). P-value between 0.05 and 0.95 is considered adequate.

Next, I introduce how a posterior predictive check can be made graphically with the same data and models. In Figure 3.2, the plots comparing observations and predictions present a clear difference between three models. In a hurdle Poisson, simulations are generally appropriate, but there is a slight overestimation of zeros. In hurdle Poisson or hurdle and ZI Poisson models, there is more uncertainty as to the number of working hours. By contrast, it is obvious that overfitting occurs in a hurdle Poisson log-normal, which is caused by skewed residuals. Compared to the two other models, a Hurdle and zero-inflated Poisson, which incorporates a zero-inflation component within a hurdle Poisson, shows superior performance in predicting values. Meanwhile, despite the improvement in predictions in the hurdle and ZI Poisson model, the model becomes too complex. The impact of adding a zero-inflation component in the model is not substantial. Accordingly, I decided to use a hurdle Poisson model in Chapter 5.



(a) Hurdle Poisson (b) Hurdle ZI Poisson (c) Hurdle Log-Normal Poisson

Figure 3.2. Posterior predictive check using scatter plots
Notes: multivariate normal priors are used; 10-percent out-of-sample test is applied (N=361); a 45-degree line is drawn in (a) and (b).

3.2.4.2. Comparing models

There are many information criteria used for model comparison purposes. The deviance information criterion (DIC) provides a way to compare a relative fit of models. Following Dempster's (1997) suggestion of using the posterior distribution of likelihood for the significance test, Spiegelhalter et al. (1998) define Bayesian deviance, $D(\theta)$, as $-2\log\{p(y|\theta)\} + 2\log\{f(y)\}$ where $p(y|\theta)$ indicates a likelihood and $f(y)$ means a function of the data. $D(\theta)$ depends on log-likelihood because $f(y)$ does not affect the model (ibid., p.2). The posterior mean of the deviance ($\overline{D}$ or $\overline{D(\theta)}$) is used as a Bayesian measure of fit (ibid.); when deviance is lower, the log-likelihood is higher so that low deviance means that discrepancies between model predictions and the data are smaller.

Spiegelhalter et al. (1998, 2002) proposed the use of the deviance information criterion (DIC) for model comparisons. It can be written as follows:

$$p_D = \overline{D(\theta)} - D(\bar{\theta})^{21}, \qquad (3.17)$$

$$DIC = \overline{D(\theta)} + p_D^{22}. \qquad (3.18)$$

They define p_D as 'the difference between the posterior mean of the deviance and the deviance at the posterior means', and add it to the posterior mean deviance, $\overline{D(\theta)}$ (Spiegelhalter et al., 2002, p.584). p_D is a penalty on model complexity and required because a complex model usually has small deviance. Thus, the DIC combines the goodness of fit and complexity of a model; therefore, it allows comparison of different models.

Additionally, WAIC (widely applicable information criterion) and LOOIC (leave-one-out cross-validation information criterion) share a similar purpose with DIC. WAIC uses log pointwise posterior predictive density (lppd) and the sum of the variance of the log predictive density over samples (Gelman et al. 2013, p.193, Eq.3.20). We can calculate lppd using draws from the posterior simulations, θ^s (ibid., pp.168–169, Eq. 3.19). In a Bayesian context, WAIC is superior to DIC because it evaluates the predictions that are actually made (ibid., p.174). Furthermore, LOOIC provides an efficient cross-validation method to evaluate the predictions based on simulations. It uses lppd and corrects a bias that might arise from predictions based on n-1 data points (ibid., Eq. 3.21). In most cases, WAIC and LOOIC have similar values. These two indices can be written as below[23]:

$$\text{lppd} = \sum_{i=1}^{n} \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^s)\right), \qquad (3.19)$$

$$\text{WAIC} = -2\text{lppd} + 2\sum_{i=1}^{n} \text{Var}_{s=1}\left(\log p(y_i|\theta^s)\right), \qquad (3.20)$$

$$\text{LOOIC} = -2\text{lppd} + \left(\text{lppd} - \sum_{i=1}^{n} \log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^{is})\right)\right). \qquad (3.21)$$

Residuals are slightly different from deviances as they explain discrepancies between observations and expected values (Lunn et al., 2013). The most common form of discrepancy test using residuals is the mean squared error (MSE), which is the easiest to calculate – $MSE(y) = \sum (y - E(y|\theta))^2 / (n-1)$. It directly informs how close predictions are to observations. With this, we can compare different models with the same data.

3.2.5. Priors

[21] Spiegelhalter et al., 2002, p.587
[22] Spiegelhalter et al., 2002, p.603
[23] Gelman et al., 2013, pp.173–176; s stands for simulations (s= 1,…,S); n is the number of data inputs (i=1,…,n); θ^s means the posterior simulations, and θ^{is} is n different inferences of posterior simulations.

Priors play a pivotal role in Bayesian inference. Given the knowledge or beliefs regarding parameters, different priors can be established. If there is no proper knowledge of parameters, we might consider using non-informative or weakly informative priors. In this sub-section, I explain the rationale behind the priors used in this study.

Gelman (2006) and Gelman et al. (2013) provide three different priors for a variance of a hierarchical model (two-level normal model): a uniform prior, a half-Cauchy prior, and an inverse-gamma prior. In general, as a noninformative prior, uniform distribution such as uniform (0, 100) or half-normal distribution with mean 0 can be used (Gelman, 2006). As a weakly informative prior, a half-Cauchy distribution with a scale factor of 25 can be a good starting point (ibid.). With a large number of groups, a noninformative prior is enough; otherwise, weakly informative priors are required to estimate low variances between groups. On the other hand, an inverse-gamma prior is a 'conditionally conjugate' prior for a normal distribution. For convenience, gamma distribution is usually applied for precision (τ) of a normal distribution (precision means inverse of variance, $1/\sigma^2$), i.e., $\tau \sim$ gamma(a, b) means $\sigma^2 \sim$ inverse gamma(a, b) where a is a shape and b is a rate parameter. A conditionally conjugate prior means the posterior distribution conditioned on all the other parameters has the same distribution as the prior (Gelman et al., 2013, p.130).

This study uses a Poisson log-normal distribution in Chapter 4. The Poisson mean parameter is assumed to follow a normal distribution, and again, a normal distribution has a gamma distribution on precision (or inverse-gamma distribution on variance) as a conjugate prior. For example, a normal distribution, $N(\mu, \sigma^2)$, with a gamma prior on σ^{-2}, such as gamma(a,b), has a Student-t distribution as a posterior distribution (Lunn et al., 2012, pp.46–47). In many other studies using a Poisson hierarchical model, a gamma distribution is considered a prior. Wiśniowski (2017) uses a vaguely informative gamma prior to precision (inverse of the variance), which assumes τ= gamma(10^{-6}, 10^{-6}); Ntzoufras (2009, p. 245) uses $\tau \sim$ gamma(10^{-4}, 10^{-4}). As Gelman (2006) pointed out, we need to set a gamma prior (a, b) more carefully. Comparing with other priors is a necessary process to evaluate the appropriateness of the use of a gamma prior. It is also helpful to compare possible sets of (a, b).

Priors are also required for the parameters of linear predictors. Gelman et al. (2008) suggest an assumption of independence of priors for coefficients and a weekly informative Student-t prior distribution for each coefficient in logistic or generalised linear models. Further, they consider a Cauchy prior on coefficients as it can give flexibility in coefficient values. Recently, Gelman (2020) proposes using a normal prior, N(0, 2.5), as a default prior for coefficients in logistic or linear models. Also, Ghosh et al. (2018) suggest the use of normal or Student-t priors with the coefficients of a Bayesian logistic or probit regression. In Chapter 5, a probit regression against probability (p), which is at the first stage of a Poisson hurdle model, requires a prior for coefficients. I have compared the results of normal and Cauchy priors for coefficients on the first stage of the Poisson hurdle model, using the DIC, WAIC and LOOIC (see Sub-Section 5.4.3).

3.3. Data Sources

3.3.1. NSS EUS 68th Round, 2011/12

In this study, the National Sample Survey India – Employment and Unemployment surveys (NSS EUS) 2011/2012 was mainly used. That was the most recent dataset when this study began. The NSS EUS 2011/2012 comprises a representative dataset about employment with a large sample size. The NSS EUS was undertaken every five years from 1972/73 – with some periodic exceptions – until the last NSS EUS dataset in 2011/12 (68th round). Some years later, the Indian government launched the NSS Periodic Labour Force Survey (PLFS) 2017/2018, which employed a different sampling method than the NSS EUS; therefore, direct comparison of the NSS EUS and PLFS data requires caution. As this study is focused on the year 2011/12, in order to compare and combine with other available datasets such as the IHDS 2011/12 and the Indian Census 2011, this study utilises the NSS EUS 2011/12.

The NSS EUS 2011/12 has a large sample size: 456,999 in total with 122,630 samples for ages 5–17, covering all 35 states in India. The detailed sampling procedure is provided by the Government of India (n.d.). Here, I summarise the process. First of all, a stratified multi-stage design is applied; the first-stage units (FSU) are villages for the rural areas, which are obtained from the 2001 census, and the Urban Frame Survey blocks (2007–2012) for the urban areas. Any urban areas where the population is more than one million (according to the 2001 census) are given additional units. According to these procedures, a total number of 14,772 FSUs (8,432 villages and 6,340 urban blocks) are allocated.

Among them, different numbers of FSUs are given to each state and union territory (UT) in proportion to population as per the 2001 census; double weight is applied for the urban sector, and a minimum of eight FSUs are allocated to each rural and urban area. In rural areas, FSUs are selected by probability proportional to the size of the population with replacement; in urban areas, a random sampling without replacement is used.

Then, hamlet groups are formed in rural areas, and sub-blocks are built in urban areas. Rural FSUs do not have hamlet groups if the population is less than 600, and urban FSUs do not have sub-blocks if the population is less than 1,200. Lastly, in terms of wealth thresholds, a number of samples are given to each FSU. Annex Table B.1 indicates different thresholds of wealth in rural and urban areas. The average monthly per capita expenditure (MPCE) is collected from the NSS EUS 66th round.

The NSS uses a sophisticated way of sample selection, and therefore calculating weights is not simple. As explained, some FSUs have double samples via hamlet groups or sub-blocks, in which weights should be halved. Thus, if the number of sub-samples within the FSU (called the NSS) is the same as the combined number of sub-samples (called the NSC), the weight is the same as a multiplier; otherwise, it is divided by two. In the dataset, MLT includes two decimal points and should be divided by 100. Thus, weight proportional to the population is written as

$$\begin{aligned} \text{weight} &= \text{MLT}/100 \text{ if NSS=NSC} \\ &= \text{MLT}/200 \quad \text{otherwise.} \end{aligned}$$

The use of a simple weighted calculation of survey data is not free from the sampling problems within sub-strata (Chaudhri and Wilson, 2000, p.13). Because of the sensitivity to the child population in this study, using a gross weight might involve exaggeration of selected sample characteristics. Thus, in this study, I prefer to use the relative weights, which are the survey weights divided by the mean (see more details in Chapter 4). This method can reduce bias caused by extremely large survey weights.

3.3.2. IHDS-II 2011/12

The India Human Development Survey (IHDS) 2011/12 was selected because it was the most recent dataset when this study was conducted, and it dealt with the same year as the NSS 2011/12. The IHDS is a panel dataset with two waves – wave 1 in 2004/05 and wave 2 in 2011/12 – and this study uses only the second survey. The sample size of the IHDS 2011/12 is about 215,754 individuals (41,554 households) and covers 33 states in India, excluding Andaman and the Nicobar Islands and Lakshadweep (Desai and Vanneman, 2018).

The sampling procedure is more straightforward than that of the NSS. Half of the rural samples include previous interviewees of the Human Development Profile of India (HDPI) conducted by the National Council of Applied Economic Research (Desai and Vanneman, 2018). The other half is collected by stratified random sampling. In each district covered by the HDPI, two villages are randomly selected, and among each village 20 households are randomly interviewed (University of Maryland and the National Council of

Applied Economic Research, 2020)[24]. Meanwhile, urban samples are stratified and selected by probability proportional to population size at state level. Among the list of Census 2001 enumeration blocks (150–200 households), urban PSUs are randomly chosen, and 15 households are selected from those PSUs (ibid.).

The IHDS 2011/12 provides weights adjusted to population size from Census 2011; however, there should be caution in using these weights. As the IHDS website reveals, the urban samples are drawn at state level; therefore, weights are the same for every sample in each state. In such cases, it is difficult to use the weights for urban areas to represent the district population. Conversely, rural samples are taken from PSUs at each district which could represent the district population.

The IHDS has provided working hours in waged work, family businesses or farming, except for domestic chores. Waged work is categorised into agricultural (occupational codes 60 – 67), and non-agricultural activities, and family businesses and family farming are defined by IHDS questionnaire criteria. Children participate in several types of work at the same time. Although the IHDS does not provide domestic working hours, the primary activity status of children implies that some child labourers undertake both household chores and market labour. The IHDS Question no 2.7 (primary activity status) is used to recognise children's usual status, which indicates the activity on which a person spent a long time during the reference period, and, in fact, some child labourers indicate that their usual status is household worker/wife. More details about a usual status approach are explained in Sub-Section 3.3.3.2.

3.3.3. Comparisons of the NSS EUS 2011/12 and the IHDS 2011/12

3.3.3.1. Time use

Regarding working hours, the NSS 2011/12 asked for the time-intensity rating for each of the seven days preceding the survey (each broken into two halves). Over 95 percent of children and adults who participated in work ticked all 14 half-day slots in their returns. Thus, the time-intensity measures of the NSS are approximated. Compared to the NSS, the IHDS allows more specific information on time-use than the NSS can. The IHDS asks for time spent in a day on a usual basis (0–16 hours a day) as well as the number of total days

[24] Available at https://ihds.umd.edu/sites/default/files/publications/papers/technical%20paper%201.pdf (Accessed 16 May 2019)

per year that individuals work. This method makes it feasible to obtain more accurate thresholds for working hours.

Table 3.3. Comparing datasets - IHDS, NSS and Indian Census

	IHDS 2011/12	NSS 2011/12	Indian Census 2011
Unit	Individual	Individual	N/A
No. of samples (ages 5-17)	51,551	122,630	N/A
Industry	Own category (0-99)	5 digits of NIC 2008	4 types of industry
Occupation	Own category (0-99)	3 digits of NCO 2004	N/A
Working hours	Daily hours	Daily intensity for 7 days (None=0, half=0.5 and full=1.0)	N/A
Working days	Total days of work per year	N/A	N/A
Usual status	Primary activity status	Principal activity (relatively long time), subsidiary activity (relatively less time), and current weekly activity status	Main (more than 6 months) and marginal (less than 6 months) workers

Sources: Desai and Vanneman, 2018; Ministry of Home Affairs, 2011; National Sample Survey Office, 2013

Table 3.4. indicates how the work intensity recorded on the NSS become comparable with the working hours of the IHDS in Chapter 4. In the previous study, an intensity of 1.0 was considered equal to 4 hours (Gómez-Paredes et al., 2016). In this study, the work intensity of the NSS is multiplied by the ratio of the maximum working hours of time thresholds to the maximum working intensity (i.e., 43/7).

Table 3.4. Converting the NSS weekly intensity of work to weekly hours

	(1) Work intensity	(2) Adjusted working hours
Weekly categories	7	43
	6	37
	5	31
	4	25
	3	18
	2	12
	1	6
	0	0

Note: each cell in row (2) = each cell in row (1) * (43 / 7)

However, the IHDS does not detail how many days children work on average every week (which is provided in the NSS); therefore, the study assumes children work equally for seven days. In this sense, the NSS's one-week recall is better than the day-recall method of the IHDS. Nonetheless, the one-week recall method might overestimate working hours if interviewees record the same hours for a full week when they may not be full-time workers and they may not work after some seasons (Das et al., 2015).

Furthermore, while the IHDS provides the total working days in a year, seasonality is not clear. Recently, the Periodic Labour Force Survey (PLFS) 2017/2018, which follows the NSS 2011/12, has started to consider the seasonality of work. It provides hours actually worked for seven days through surveying four times a year. Hence, it addresses the seasonality issue and also overcomes the limits of time-intensity measurements.

The NSS 2011/12 includes children's work participation and work intensity in domestic chores. However, the IHDS does not provide information for domestic work hours, despite it being a considerable part of children's daily time-use in India. Considering the limitation and strength of each dataset – the IHDS and the NSS – I suggest a combination of the two of them as a way to reduce measurement error in estimating the number of child labourers (see Chapter 4).

3.3.3.2. Status of children

Many studies have used a 'usual status' of work approach when calculating child labourer numbers (Mukherjee, 2012; Kak, 2004). The IHDS 2011/12 and NSS 2011/12 use this usual status approach to ask individuals which activities they spend a longer time performing than others. For the NSS, a usual status (called principal activity status) means a status that people spend a relatively longer period within the 365 days[25]. The IHDS uses a 'primary activity status' but does not provide a specific definition for it. In the IHDS, primary activity status does not include missing items because children are assigned to the 'other' category if their current status is not identified. Among 'other' categorised children, some indicate that they are currently attending a school; in such cases, they are likely to be students.

The NSS 2011/12 usual status of children is quite specific. It even separates unpaid household services (code 92, 'attended domestic duties only') and household own-use production (code 93, 'attended domestic duties and was also engaged in free collection of goods (vegetables, roots, firewood, cattle feed, etc.), sewing, tailoring, weaving, etc. for household use'). These two categories are the key components in the current discussion on child labour (ILO). The former category is outside the System of National Accounts (SNA) production boundary, but the latter is within it.

For the purpose of data combination, this study applies a conservative way of defining children's status in Chapter 4. It is important to note that the NSS 2011/12 does not provide a total number of days or information on seasonality; therefore, it requires a

[25] Available at http://mail.mospi.gov.in/index.php/catalog/143/datafile/F5/V209 (Accessed 22 May 2018)

principal activity status of children to categorise their ‘usual status’. Although the IHDS 2011/12 provides the relevant information, such as total working days, we should also use the children’s primary activity status to define child labourers in order to make a fair comparison between the NSS and the IHDS. Thus, it involves the primary activity status of the IHDS and principal activity status of the NSS being applied to the ‘usual status’ of children. In this ‘primary (or principal) activity status’ approach, child labour indicates work takes place most days of a year.

In Chapters 5 and 6, child labour is not limited to working as principal activity. These two chapters use only the IHDS so that I can employ a more flexible way of defining child labour, in which a child’s status means any work that continues for 30 days or more per year. In this approach, children who combine working and schooling can be included as labourers according to the time spent in work. At the same time, since it excludes any cases of children working for less than 30 days a year, it can take into account seasonality of work, although this method is not perfect.

3.3.3.3. Missingness

In the NSS 2011/12, daily intensities of children’s activities are recorded as above zero, in any case. If someone has no intensity recorded in their work activities, he or she has an intensity in other activities such as schooling. If children do not work, those children are categorised as having ‘zero’ work intensity. The IHDS 2011 provides separate records of working hours greater than zero (1 to 16 hours) for farming, household enterprise work, and waged work, which shows that a huge proportion of children have absences in the records of their working hours. In a similar way, those children with missing data in the IHDS 2011/12 are assumed to have zero work hours. In both cases, children with no hourly information about work (or work intensity) are treated as if they do not participate in work at all (working hours=0). Imputation method is not applicable in such cases, as there is no clarity as to whether missing items are caused by non-response. Moreover, those non-zero-hour cases compose a large proportion of the children; therefore, data imputation could cause significant bias. A more reasonable solution to dealing with the absence in working hours is to introduce an ‘undercount’ parameter in a model.

In a Bayesian inferential framework, missing values in a dependent variable can be easily produced through a model. Any missing values are treated as the same as parameters through a Bayesian approach, and they have posterior predictive distributions based on likelihood and priors. Lunn et al. (2012) use a missing weight in a model, which can introduce the missing data mechanisms by a specific assumption on missing values. In

Chapter 4, data for two states are missing, and the Bayesian model efficiently imputes values for the missing data (the numbers of child labour cases in the two states). I do not apply different assumptions for those missing values in two states as they are based on a small size population and almost ignorable.

Meanwhile, Bayesian models do not produce automatic imputations for missing values in independent variables. If necessary, one can consider a linking prior between observed and unobserved values of covariates (Lunn et al., 2012). In Chapters 5 and 6, I introduce many explanatory variables. Among them, imputations are typically required for the asset index. The asset index is estimated through several items that households own, and some of the items have missing values. I chose the multiple imputation method, which is based on estimations of maximum likelihood and similar in a way to the Bayesian imputation. The multiple imputation method repeats generating each set of a completed dataset using a model and combines the results of imputations to adjust for missing values (Rubin, 1996). One option to combine the repeated results is to average imputed values. In a Bayesian context, multiple imputation is an average of repeated imputations drawn from the posterior predictive distribution of missing data (ibid.). In order to impute the asset index, I used a logit regression with maximum likelihood estimation (MLE) with all the items as variables, repeating imputations five times.

Besides, missingness is not observed in most of the demographic variables, such as age, sex, and urban or rural. Social groups and class variables have the category ‘others’ for those who do not have the relevant information. Migration and female household headship are dummy variables where children are given 1 if they meet a particular condition or 0 if not. None of those variables requires imputation.

3.3.4. Other supplementary data

3.3.4.1. Census 2011 India

Census 2011 India is the national census survey, covering 28 states and 7 union territories (referred to as 35 states). According to the data, the population of children aged 5–17 in 2011 is reported as 330 million. It also shows evident population characteristics. Among all children aged 5–17, 174 million were boys, and 158 million were girls. In geographical terms, 241 million children reside in rural areas, while 90 million children reside in urban areas. There is a gap in population sizes between age groups: for example, age 10 has the largest population among children from ages 5 to 17.

In Census 2011, ‘child workers’ means children who are involved in economic activities that include cultivation, agricultural labour, workers in household industries and others. It provides the number of main and marginal workers. The ‘main workers’ are those working at an economic activity for more than six months of a year, and ‘marginal workers’ are those working less than six months. Among children aged 5–17, 11.8 million were found as main workers, and 23.8 million are found as child workers (main or marginal workers). The Indian census reports that child workers aged 5 to 14 in 2011 number 10.1 million, which is 4 percent of all children in the same age group (Central Statistics Office, 2018, pp.53-54). However, Census 2011 does not detail the specific criteria that are required to measure child labour in this study. Therefore, it serves as auxiliary information to increase the precision of the estimates as it covers almost the entire population.

3.3.4.2. The World Value Survey Wave 6 – India 2012

The World Value Survey includes several questions about people’s norms and attitudes. In India, the survey was undertaken in 2012. It surveyed 19,444 individuals in 17 states. The samples are drawn via several steps: 1) firstly, 320 parliamentary constituencies (PCs) are chosen among 28 states and one union territory (Delhi) via random sampling; 2) after that, three assembly constituencies (ACs) in each PC and then two polling stations (PSs) in each AC (so, 320*3*2=1,920 PSs) are chosen; and 3) in all PSs, 8–13 respondents are surveyed.

The World Value Survey India 2012 does not provide a sampling weight. We could consider creating weight to be adjusted to the population at the state level; however, as the sampling process is not certain, applying weight could involve more errors. Thus, this study uses the unweighted mean at the state level from the World Value Survey India.

This survey provides useful information on people’s values regarding work, politics, tradition (religion) and gender, etc. However, it does not have a variable for child labour, and therefore direct analysis of the relationship between child labour and norms is not possible. In this regard, a combined-data approach is useful as it links different datasets that have different information regarding child labour. In the WVS 2012, there are diverse norm variables that can explain the variance in the incidence of child labour at state level. In particular, social and cultural norms are found closely related to the incidence of child labour (Chapter 6).

Chapter 4. A Bayesian Estimation of Child Labour in India

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 six months) 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.

Keywords: Child labour, Bayesian estimation, Combining data, India, Time threshold

4.1. 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.a). 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 percent) and Bihar had the second-largest number (1.1million, 3.9 percent; Samantroy et al. 2017, p.48). In terms of incidence, Nagaland and Himachal Pradesh have shown the highest proportion of child workers at 13.2 percent and 10.1 percent, respectively (ibid.). However, the numbers vary according to datasets and definitions. For example, the proportion of child labourers reached 11.8 percent among children aged 5 to 14 in 2012, as calculated by UNICEF (n.d.)[26], which is roughly 29 million children in total. Increased household income has led to

[26] 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.

a reduction over time, but many industries and farms still employ child labourers. Unpaid household services (household chores) are some 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. Moreover, 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 of labour 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 number of child labourers based on two broad categories – children who are main (i.e., working more than six months a year in economic activity[27]) or marginal (working less than six months 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 and 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.

4.2. 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

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

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).

Despite this 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, 2018b)[28]. However, how many hours of domestic chores are considered harmful for children, and which periods (month, quarter or year) should be regarded as the 'usual status is still under discussion'. Moreover, the ILO (2018b) 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'[29]; 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.

[28] 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, 2018b). 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, 2018b).

[29] 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.

4.2.1. 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 definitions and measurements 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)[30], but they have applied slightly different methods in capturing child-labour statistics.

4.2.1.1. 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[31]. 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 (2018b) 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[32], 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

[30] Available at https://www.unicef.org/protection/57929_child_labour.html (Accessed 11 Jan. 2020)
[31] Worst Forms of Child Labour Convention initiated in 1999, No. 182
[32] Minimum Age Convention in 1973, No. 138

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 (under the age of 14 years) except for helping their families, working in family enterprises after school hours or artistic work, and adolescents (14–17) 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.

4.2.1.2. 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 4.1. Latest standard of child labour by the ILO (2017) criteria

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 hour or more) in economic activity1)	- 14 hours or more in economic activity	- 43 hrs 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 (2018b), the 18th ICLS resolution paragraphs 15(c), 36 and 37).

The ILO does not specify a limit to 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 4.2). According to the current database, UNICEF (2019)[33] categorises the criteria as follows: (a) children 5–11 years old who undertook at least one hour of economic activity or at least 21 hours of household chores per week; (b) children 12–14 years old who undertook at least 14 hours of economic activity or at least 21 hours of household chores per week; and (c) children 15–17 years old who undertook at least 43 hours 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 hours of household chores negatively affects children's education. Previously, the time threshold for household chores for ages 5–14 was 28 hours (UNICEF, 2017), but this has now been reduced to 21 hours. Children 'working in hazardous working conditions' or children (15–17 years) who spend long hours doing household chores, which were criteria for child labour in the Multiple Indicator Cluster Surveys (MICS) from UNICEF (2017), are no longer included.

Table 4.2. 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 hour of economic work or 21 hours of unpaid household services per week	- At least 14 hours of economic work or 21 hours of unpaid household services per week	- At least 43 hours of economic work per week

Source: Summarised from the UNICEF database (UNICEF, 2019)

4.2.2. 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 child labourers regardless of working hours. 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

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

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 Figure A.1. We keep time thresholds for economic activity (43 hours for ages 15–17, 14 hours for ages 12–14, and 1 hour for ages 5–11). Then we use UNICEF's weekly time thresholds for unpaid household services: at least 21 hours a week of household chores for ages 5–14. In Table 4.3, we summarise our operationalisation for this study of the definition of child labour used by ILO and UNICEF.

Table 4.3. 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 hour of economic work or - at least 21 hours of unpaid household services per week	- At least14 hours of economic work or - at least 21 hours of unpaid household services per week	- At least 43 hours 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

4.3. Methodology

4.3.1. 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 form of posterior probability distributions. Bayesian inference combines the model for the observed data specified in the form of a likelihood function with 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 percent confidence), the Bayesian approach uses predictive intervals, which specify the probabilities that the true values lie within them (Gelman et al., 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.

4.3.2. 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 percent 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= 51,551, representing 0.02 percent of the population of the same age group). It covers 33 states in India, excluding Andaman and Nicobar Islands and Lakshadweep (Desai and Vanneman, 2018). 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).

4.3.3. 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 only 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 undercount parameter (e.g. Wiśniowski, 2017; Wiśniowski et al., 2013). The over- or undercounting 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 A.1).

4.3.4. 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 hrs)[34]; however, the NSS offers the intensity of time disposition for each activity (None=0, half=0.5, and full=1.0) for the last

[34] 'How many hours did you work in a usual day?' (Question 7.8, IHDS 2011).

seven days. (Maximum points are 7.0 per activity.)[35] 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. 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 hours.

This study rigorously follows the concept of 'usual hours of work per week' suggested by the ILO (ILO, 2018a)[36]. 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.

4.3.5. Descriptive analysis of the data

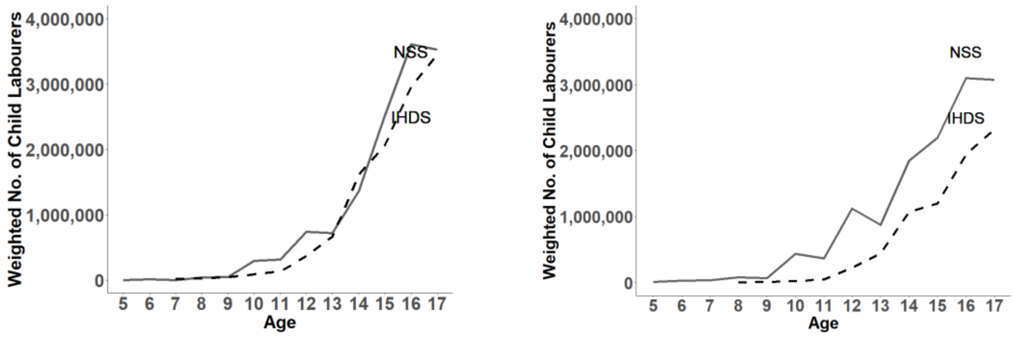
Our criteria for measuring child labour provide time thresholds both for economic activity that is a type of work counted in GDP, including work in family-owned farming or business, and for unpaid household services. In terms of industries and occupations, construction, mining, and waged agricultural 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 Figure A.1). Moreover, many child labourers are found to be unpaid household labourers, and most of these are female children (see Appendix Figure A.2).

Figure 4.1 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 performing child labour calculated from the 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

[35] 'Time disposition during the week ended on ....' (intensity of activity for seven days, full mark: 1.0 and half-mark: 0.5) (Question 5.3, NSS 2011).

[36] The ILO defines hours usually worked as 'hours (actually) worked in a job per a short reference period such as one week, over a long observation period of a month, quarter, season or year' (ILO, 2018a).

many child labourers because it includes children who spend excessive time in the recall week doing household chores.

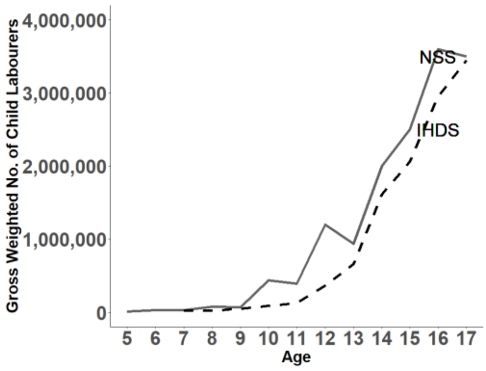


(a) ILO measurement (2017) (b) UNICEF measurement (2019)

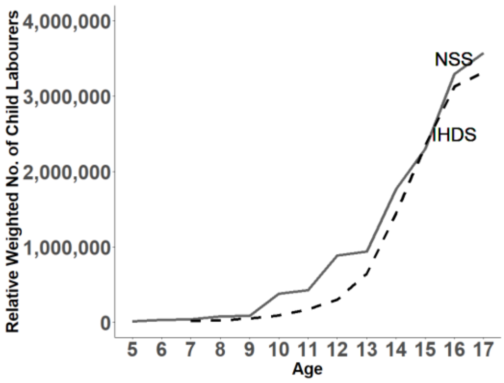
Figure 4.1. Gross weighted number of child labourers in India using different measurement schemes

Source: NSS 2011/12; IHDS 2011/12

Figure 4.2(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 (Figure 4.2(a)). However, using a relative weight reduces any sharp decrease between 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 for child labour and to multiply 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.



(a) Gross weighted number of child labourers



(b) Relative weighted rate of child labour*population

Figure 4.2. Gross vs relative weighted count of child labourers in India (non-model, authors' definition of child labour)

Source: NSS 2011/12; IHDS 2011/12
Notes: weighted counting with the use of measurement criteria of this study (see Table 4.3)

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.

Figure 4.3. shows the descriptive information from the Indian Census. The Indian Ceuns 2011 is useful in capturing the number of child workers. The census data provides the number of child workers who are categorised by the months of work involved in a reference year. Here, the term 'child work' is a broad category, referring to both main (working for more than six months a year) and marginal workers (working less than six months). It suggests that the number of child workers between the ages of 5 and 17 was 23.8 million in 2011/12.

Comparably, Figure 4.4. indicates the number of child labourers defined by this study. Our estimation suggests there were 11.4–15.2 million child labourers in the same year (see Section 4.4). In both results, Uttar Pradesh has the largest number of child workers as well as child labourers because of its large population.

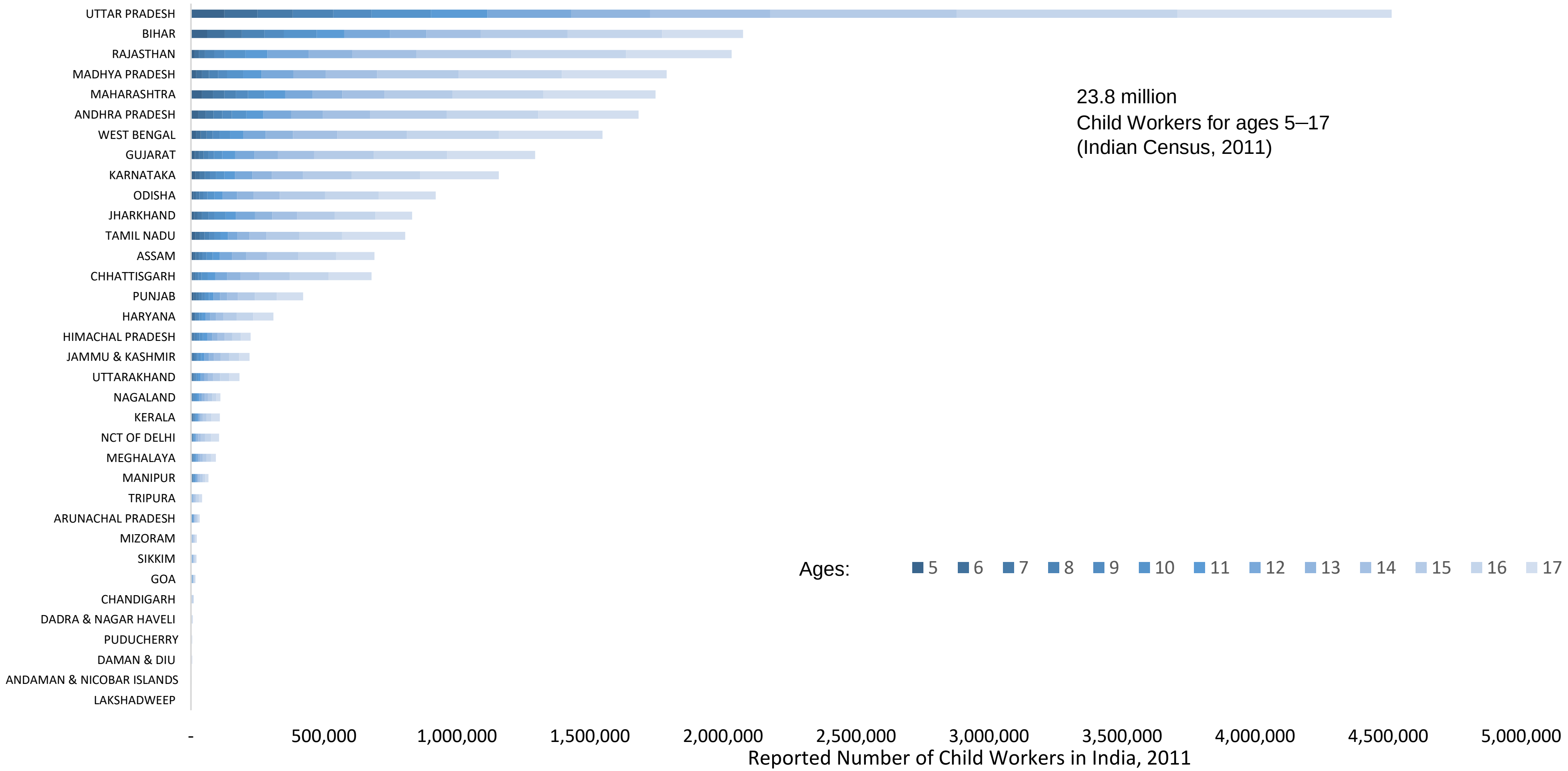


Figure 4.3. Child workers by age and state from the Indian Census, 2011

Source: Ministry of Home Affairs, 2011
Notes: Main workers (working for more than six months a year) + marginal workers (working less than six months)

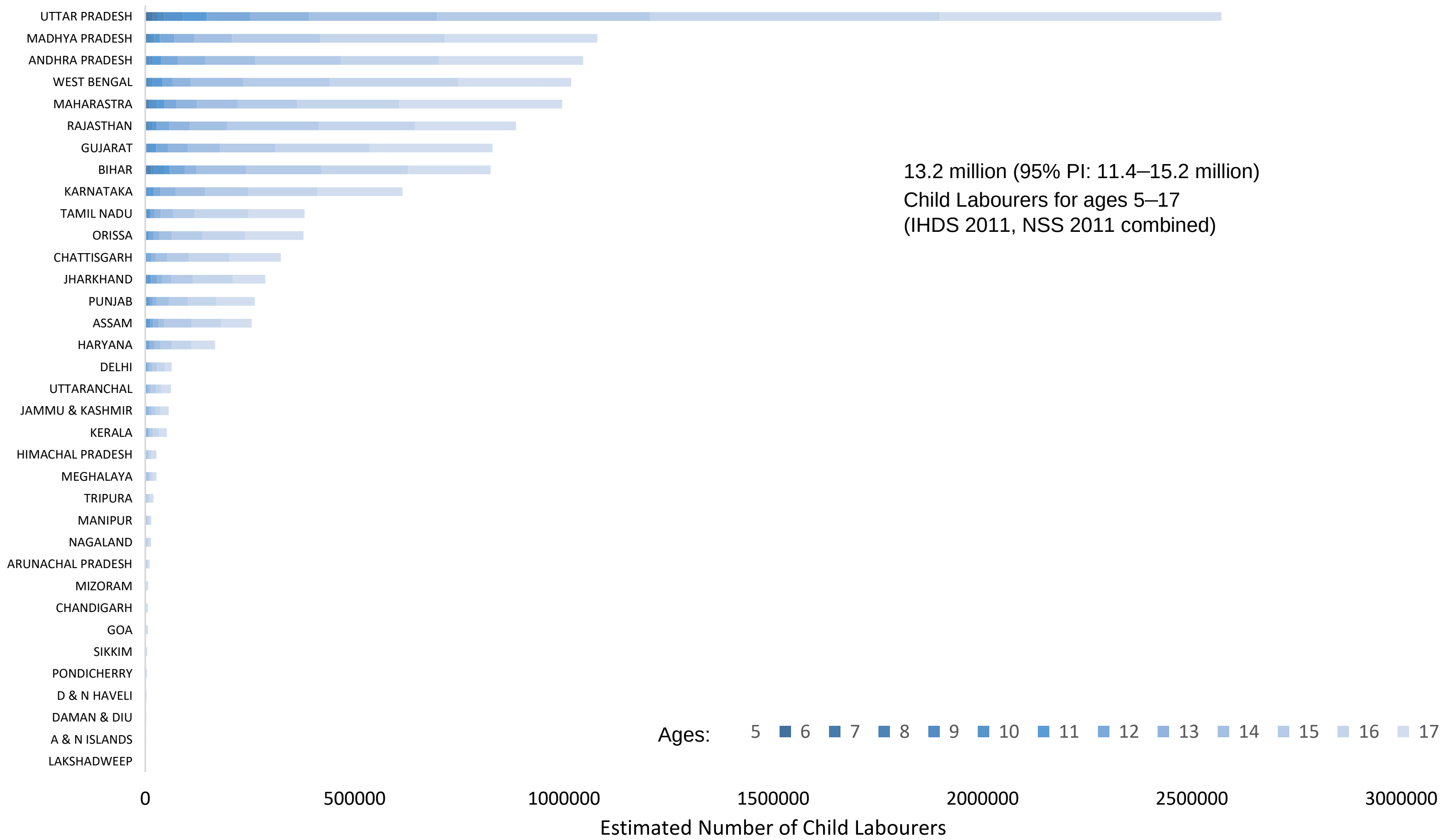


Figure 4.4. Child labour numbers by best estimates, 2011/12

4.3.6. Model framework

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 A.2. 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 for regional differences according to the Indian Census 2001 (National Sample Survey Office, 2013; Desai and Vanneman, 2018). 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 undercounting 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. $y.a_{ij}$ 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 4.2). 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 group.

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 Appendix Table A.2, 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 Appendix Table A.1). 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 the 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.

4.3.7. 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).

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

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

4.4. Results

4.4.1. Comparing models

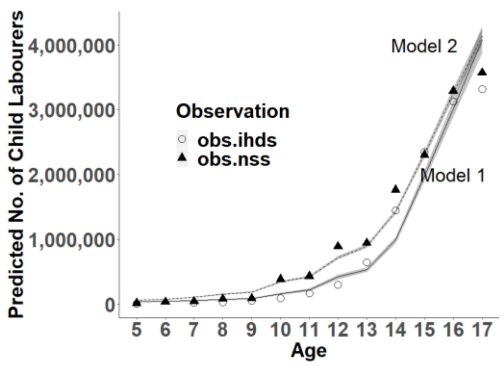
We have reviewed the DIC to compare the goodness of fit as well as the complexity of the models. DIC is the posterior mean of the deviance plus the effective number of parameters (pD). Comparing the DIC of Model 1 and Model 2 is not feasible because the NSS and the IHDS have different sample sizes. Model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

Table 4.4. Aggregate child labour estimation, India, 2011/12

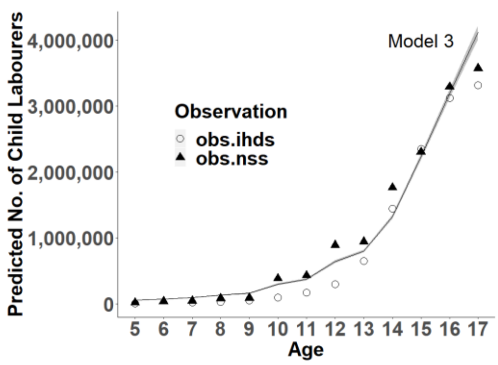
Models	DIC	Deviances	pD	No. of Child Labourers[1]			
				Age 5–17	sd	Age 5–14	sd
Model 1	1,319.3	1,316.9	2.4	11,700,959	238,241	2,618,101	99,144
Model 2	3,951.6	3,848.5	3.1	14,103,812	158,733	4,390,055	79,728
Model 3	5,387.0	5,384.4	2.7	13,484,868	132,013	3,929,058	63,166
Model 4.1 (Uniform Prior)	2,919.8	2,367.4	552.4	13,545,971	915,409	3,329,530	258,347
Model 4.2 (Inverse Gamma Prior)	2,894.3	2,358.9	535.4	13,194,114	963,859	3,222,680	292,084
Model 4.3 (Half Cauchy Prior)	2,914.4	2,368.1	546.3	13,570,367	901,155	3,341,185	252,466

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

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



(a) Model 1 & Model 2



(b) Model 3

Figure 4.5. Posterior mean numbers of child labourers in India, 2011/12, by ages (Models 1–3)

Notes: the shaded areas indicate 95-percent intervals

The variance of Poisson distribution should be equal to the mean, which may not realistically capture the over-dispersion in the data. Figure 4.5. 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-percent 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-percent 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.

4.4.2. Result of Poisson log-normal model

4.4.2.1. Parameter estimation

The MCMC algorithm shows proper convergence in the posterior parameters of interest in Model 4.2. Figure 4.6 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 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.

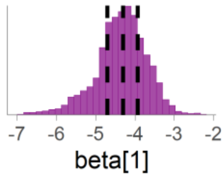
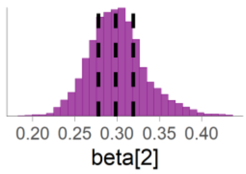
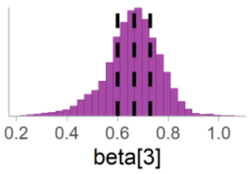
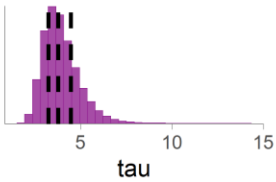
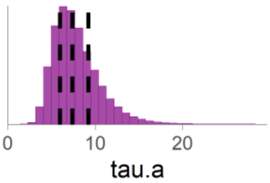
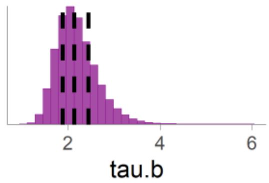
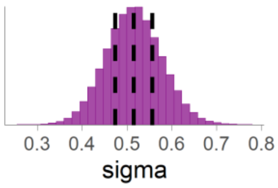
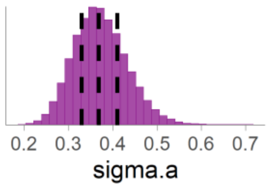
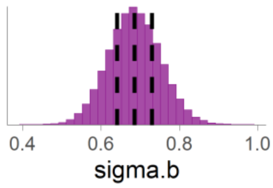
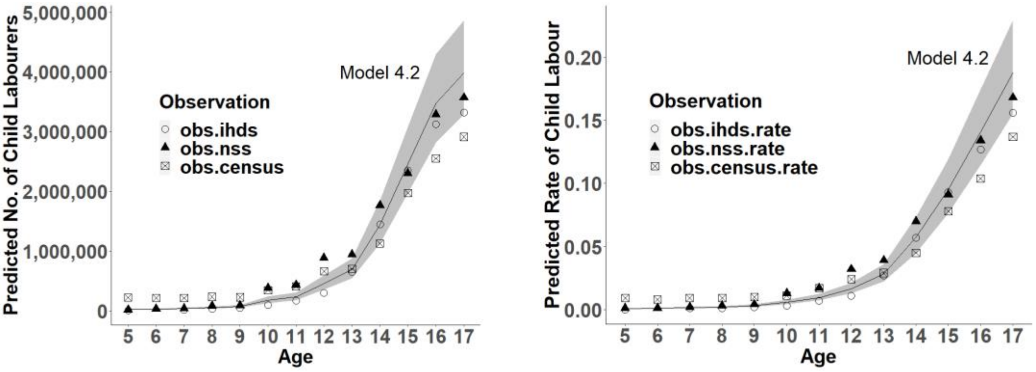
<table>
<tr><td>Histogram</td><td></td><td></td><td></td></tr>
<tr><td>Parameter</td><td>β_1</td><td>β_2</td><td>β_3</td></tr>
<tr><td>Mean (sd)</td><td>-4.56 (0.6)</td><td>0.31 (0.03)</td><td>0.7 (0.1)</td></tr>
<tr><td>Histogram</td><td></td><td></td><td></td></tr>
<tr><td>Parameter</td><td>τ</td><td>τ.a</td><td>τ.b</td></tr>
<tr><td>Mean (sd)</td><td>3.84 (0.94)</td><td>7.53 (2.47)</td><td>2.88 (0.65)</td></tr>
<tr><td>Histogram</td><td></td><td></td><td></td></tr>
<tr><td>Parameter</td><td>σ</td><td>σ.a</td><td>σ.b</td></tr>
<tr><td>Mean (sd)</td><td>0.52 (0.06)</td><td>0.38 (0.06)</td><td>0.6 (0.06)</td></tr>
</table>

Figure 4.6. 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

The undercounting parameter (υ) indicates the posterior mean 0.81 (95-percent 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. Undercounting 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 percent of the child population aged 5–17) in 2011/2012. The 95-percent PI for the number of child labourers ranges from 11.4 million to 15.2 million. That is, with a probability of 95 percent, the true number of child labourers lies within this interval. The estimate for ages 5–14 is around 3.2 million and 95-percent PIs are 2.7 to 3.8 million, as seen in Figure 4.7.



(a) Posterior No. of child labourers by age (b) Posterior rate of child labour by age

Figure 4.7. Results of Bayesian Poisson log-normal model using a combination of datasets (India)

Notes: the shaded areas indicate 95-percent intervals; the census data is presented in a graph for purposes of comparison.

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 six months 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-percent 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-percent PI. Our estimates do not adequately capture the child labourers under the age of 10, due to there being only a small 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’ and are, therefore, not included as child labourers.

4.4.2.2. 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. A Poisson log-normal model allows over-dispersion, widening predictive intervals (see Appendix Table A.1). The use of an undercount 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).

4.5. 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 domestic work too, 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, 2018b).

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 labour. 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 for 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 two or three seasons, is required. Moreover, 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 underreporting of child labour. Although the IHDS provides better time information, it does not cover the domestic sector; therefore, we need 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 are underreported, 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 over- or undercounting 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 hours a week for ages 16–17 in the UK; 35 hours 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.

Chapter 5. Girl Children's Labour Participation, 'Child Labour' and Decent Work: Results from India

Abstract

The intensive labour participation of girl children often constitutes damaging forms of 'child labour', defined clearly in this paper. In India, studies of gender differences in child labour are rare. This study aims to reveal socioeconomic causes – social group, class and gender – of children's labour force participation and labour hours in India. Here, we apply the gender and development approach (GAD) which stresses class differences in gendered work experiences and stereotypes. The paper covers all of India, providing a comprehensive analysis of female labour market participation for ages 5 to 17. We suggest a method to estimate female child labour population size as well as their working hours while exploring the interaction between gender, social group (notably, India's scheduled tribes) and class. Data from India's Human Development Survey covers labour participation and hours worked in specific task groups. In India, institutionalised barriers to working in public, such as upper-caste social norms and Sanskritisation, tend to exclude girls from participation in well-paid activities. Social class also strongly affects the working hours of girls. Girls are heavily involved in production, especially in the informal sector, whilst there is important socioeconomic diversity among female children.

Keywords: Child labour, Class, Gender and Development, India, Institutionalised norms, South Asia

5.1. Introduction

Indian child labour is characterised by featuring more boys than girls. Girls' lesser involvement in economic activities outside the home is reflective of caste norms, as well as a patriarchal social system in which girls work within a family boundary. Across the world, girls are much less visible in the labour market than boys, with more girls than boys working as labourers in only a few countries (Webbink et al., 2013). In India, although girls are again less visible in the labour market, their work contributes significantly to generating economic value (Nieuwenhuys, 1994). Moreover, the gender division of labour and a complex understanding of gender relations regarding child labour could exacerbate the discrepancy between labour participation and intensity of work performed by girls.

Previous studies on child labour found it to be the result of a lack of economic resources and focused on revealing the relationship between family wealth and child labour (Basu and Van, 1998; Lima, Mesquita, and Wanamaker, 2015). However, economic motivation is an insufficient explanation for a gendered pattern of child labour. Gender roles are based on socioeconomic and cultural dynamics, which are often missing from the economic approach. Some traditions, such as patriarchy, tend to increase children’s participation in housework or family businesses (Webbink et al., 2012). In India, it is known that girls spend fewer hours on economic activities than boys, but this result is reversed according to extended System of National Accounts (SNA) research (Hirway, 2002). Moreover, son preference affects the numbers of hours spent on domestic activities by each child gender in India (Lin and Adserà, 2013). In this sense, son preference might also affect hours spent in productive areas.

This research aims to reveal different causes for the incidence and magnitude of girl child labour participation in India based on socioeconomic and institutional aspects. We allow for explanatory factors of diverse kinds. We focus on girls’ labour-force participation and hours worked in economic activities, about which less is known compared with girls’ contribution to domestic chores. A lower incidence of girls in the labour market might be related to caste-gender relations, in which girls’ activities are restricted because of traditional gender norms. Furthermore, this study contributes to understanding the intensive workloads of girl child labourers. Girls’ labour hours might differ because of the household’s socioeconomic status. Gender norms, such as son preference within certain classes, and double burden placed on girls as family workers and waged labourers can be risk factors leading to an overall lifetime detriment for girls.

The paper is structured as follows. Firstly, we review diverse perspectives on child labour and explain the significance of the gender and development approach (GAD) in explaining a gendered pattern of child labour. Then, our conceptual framework is presented, based on the GAD approach. An overview of gender differences in child labour follows. This study uses a Bayesian Poisson hurdle model explaining the incidence and magnitude of child labour by caste, class and gender. Lastly, we highlight core findings and provide implications based on them.

5.2. Literature Review

5.2.1. Review of previous research into ‘gender and child labour’

Analysis of family decision-making regarding child labour helps to explain girls’ long working hours as a result of parental preferences. Edmonds (2006) found that in Nepal older girls work more and for longer than boys. Indeed, in general, son preference increases both the incidence and magnitude of child labour among girls (Edmonds and Pavcnik, 2005). However, realities vary according to family and socioeconomic backgrounds. For example, in India, girls are much less involved in market-based activities than boys because of norms and culture. The transformational effect of norms and values by class and social group results in divergent outcomes.

Webbink et al. (2013) analyse the number of working hours by children in paid work using a mixed-effects model of parent's education, mother's work status and family wealth. Their research contributes to the consideration of economic, demographic, and cultural factors and their effects on child labour, simultaneously. Moreover, the livelihood approach has generated interest in gender inequality, integrating the diverse components of livelihoods: culture, environment, and social factors (Ellis, 2000). However, in those studies, class relations are not taken into account, although they can have a significant impact upon children's working lives.

On the other hand, a standard structuralist approach explains child labour as an exploitation of labour based on class struggles and also caused by the power relations shaped by global production networks. Child labour is a type of unfree labour, which is an exploitation caused by a lack of power. Venkateshwarlu and Da Corta (2001, p.2) explain that it is more likely a recent change whereby women and girls have become unfree labourers in agriculture to replace the men and boys who used to work as bonded labourers. Many other studies discuss the problem of the feminisation of agricultural labour in India (Da Corta and Venkateshwarlu, 1999; Garikipati, 2009; Olsen and Mehta, 2006b). There has been a clear trend that many women and girls in India work as marginal farmers or agricultural labourers while girls are highly involved in subsistence and informal work (Hirway and Jose, 2011). Phillips et al. (2014) explain that informal and home-based production in global production networks drives child labour, within which women's and children's labour are strongly connected. 'Adverse incorporation' reinforces the division of work by gender; girls often take on their mother's work, such as taking care of siblings, while boys are responsible for earning income, often through household enterprises (Phillips, 2013). Overall, this structuralist approach provides a firm ground on which girl child labour can be explained as exploitation and associated with women's struggles within the class structure.

5.2.2. The view on child labour from the GAD approach

The GAD approach provides a better explanation for the complicated relationships surrounding gender and child labour, whereby gender is a social construct related to other social components such as class and culture (Momsen, 2010). It has made a significant contribution to understanding how women and men are differently affected by economic change due to sociocultural practices, such as women's subordination to men or the expectation of reproductive work (Momsen, 2010, pp.47- 49).

At the same time, gender is interwoven with other social inequalities such as class and race and requires analysis through a holistic framework (Kabeer, 1994, p.65). Social structure plays a key role in generating child labour, while, at the same time, its effect on child labour is integrated with gender and other constructed identities. For example, household wealth tends to increase the amount of child labour, but the result might be reversed in societies with non-traditional or diversified livelihoods (Kabeer, 2001).

Norms and values have a critical role in the gendered division of child labour. How parents consider girls' activities outside the home affects the girls' employment status. A gender relation is not fixed but flexible and is decided according to social norms. It should be emphasised that norms are changeable over time and in terms of location by agents who are affected by social structure and institutions. Diverse patriarchal bargains may occur across cultural and religious boundaries (Kandiyoti, 1988). On the other hand, Olsen and Zhang (2015) argue that deviances from norms are a regular and critical part of agency in the labour markets. For example, in a rural area of Nigeria, girls are more involved in domestic work, but at the same time, they are involved in hawking, replacing their mothers in the role, because of the general seclusion of women there due to the dominant religious norms (Robson, 2004, pp.204-205). Social norms on whether it is acceptable for girls to work outside can shift according to the social and cultural backgrounds of children and their families. Actions also may sometimes dis-conform to social norms.

Child labour is a complex outcome with multiple components such as class, social groups and gender. Attempting to explain child labour while relying on one academic discipline could bring about misleading conclusions. Agarwal (2016, p.142) insists that the bargaining model of intra-household dynamics does not fully cover the complexity of gender relations, including the roles of social norms and perceptions. Child labour studies need a deeper understanding of interlinkages between family, structure, and institutions and norms. There should be more specific consideration of children's participation in the labour force and their working hours based on those interrelations. Norms affect economic expectations for girls in combination with their socioeconomic status. Thus, we need to make a closer examination of both the socioeconomic and institutional factors influencing child labour. This paper is unusual in offering such an analysis for a huge and diverse country like India.

5.2.3. Effects of 'social group and gender' on work participation

Caste-related social norms are the key values linking gender inequality and social inequality in India. In a country where purdah, the general seclusion of women, is valued, daughters are more likely to stay at home and sons more likely to participate in activities outside the home: 'Caste ideology also affects whether women work at all, what work they can do, how far from home they may move' (Harriss-White and Gooptu, 2001, p.100). Olsen and Mehta (2006b) point out that social norms such as Sanskritisation, Brahmanical gender norms, and purdah lead many families to value women staying at home. Protection or seclusion of girls has a social and cultural value, which inhibits parents from sending their daughters to school (Lewis and Lockheed, 2007). Moreover, the family honour system restricts girls' activities outside the home (Kambhampati and Rajan, 2008).

Much research finds that social group is the main factor deciding girls' participation in child labour. It has been acknowledged that the Brahminical social norms do (or did) not let girls take part in any economic activity (Mies, 1980b, p.185). Neither are girls from lower social groups allowed to work outside the home when their families follow caste norms (Srinivas, 1956). Srinivas's (1956) theory of Sanskritisation supports the idea that a low-caste household with high wealth will assimilate upper-caste social norms. Hence, daughters rarely work outside the home and even less on the family farm (ibid.). Kambhampati and Rajan (2006) show that both Indian boys and girls of scheduled castes and tribes are more likely to participate in work outside the home, although the positive association of low caste with paid work is stronger for boys. The less control caste norms have over behaviours, the more girls might be involved in labour outside the home. In India, this particular pattern is well known for adults (Olsen and Mehta, 2006b).

5.2.4. Effect of 'class and gender' on working hours

Gender- and class-based exploitation can maintain an unequal distribution of working hours. There could be several reasons for increasing girls' working hours. Firstly, girls perform multiple jobs and take on the greater burden of unpaid non-SNA (the System of National Accounts) [37] productive work. The gender division of child work increases girls' domestic work burden, which requires a considerable amount of time (Koissy-Kpein, 2012; Robson, 2004). Burra (2005) notes that girls spend a significant number of working hours in both SNA and extended SNA activities. Girls experience greater involvement in non-economic

[37] The UN uses the System of National Accounts (SNA) to categorise and compare the economic activities of each country. Own-use production work on services is not included in the SNA production boundary but it is in the SNA general production boundary.

labour, which can associate them with activities featuring low productivity, low earnings, and long hours.

Secondly, girls, especially those from lower classes, spend a long time on agricultural activities. The feminisation of agriculture is a significant matter, not only for adult women but also for girls. Sharma (1985) has observed that low-class females have become concentrated in subsistence agriculture following the ‘green revolution’. This study will add empirical evidence to previous studies about the feminisation of the agricultural sector.

Thirdly, class is closely related to son preference, which might also affect the working hours of girls and boys. Previous literature explains that middle-class social norms exist in India; for example, middle-class families prefer to invest in their sons (Agarwal, 2016). Although there are enough resources, middle-class girls have less educational opportunities because more resources are spent on the son’s education (Bose, 2012). Therefore, lower-middle and middle-class families might expect their girls to do more work than their boys, if necessary. This contrasts with lower-class families, who need both sons and daughters to help the family economy.

There are two ways to investigate the class relations in India. Firstly, one can do it by looking at land and labour relations. For example, Bardhan (1982) describes the classes as follows: landless labourer, poor peasant, family farmer and rich farmer. Athreya et al. (1987) identify classes by their surplus and categorise them as rural proletariat, poor peasants, middle peasants and surplus appropriators. Olsen (1996) and Olsen and Mehta (2006b) extend class analysis by combining agricultural and non-agricultural activities. Secondly, asset conditions can serve as a proxy for class, which might give different results from a traditional definition of class – land and labour relations. In previous studies, wealth shows an inverted U-shape relationship with the incidence of child labour; more wealth decreases child labour and then subsequently increases it (Bhalotra and Heady, 2003; Kambhampati and Rajan, 2006).

5.2.5. Conceptual framework

Figure 5.1 shows the conceptual framework of this study. One of its hypotheses is that gender relations interact with other social relations to influence children’s labour participation and the number of hours they work. The literature review suggests that it is necessary to investigate the effects of social group dynamics on girls’ labour-force participation and class on girls’ working hours, respectively. However, to separate these two effects entirely is impossible. Social groups and class largely overlap in India in terms of the distribution of resources and power over males and females (Sharma, 1985). This study therefore controls the effects of social group and class on both labour-force participation and working hours. However, we focus on revealing the effects of social groups and gender on labour-force participation and the effects of class and gender on labour hours.

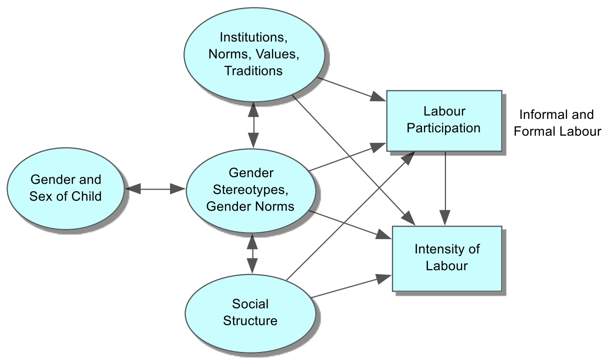


Figure 5.1. Conceptual framework of a gender, structure and agency model of child labour

Another assumption underlining this study is that girl child labourers are more likely to be involved in the informal sector because of gender segregation. Our classification of the formal and informal sectors accords with traditional economists' classifications of Indian labour (see Appendix Table B.1). The informal sector usually refers to less organised, less productive and small-scale sectors, in which peasants and landless labourers accommodate themselves (Breman, 1976).

5.3. The Context of Child Labour in India

5.3.1. Defining child labour

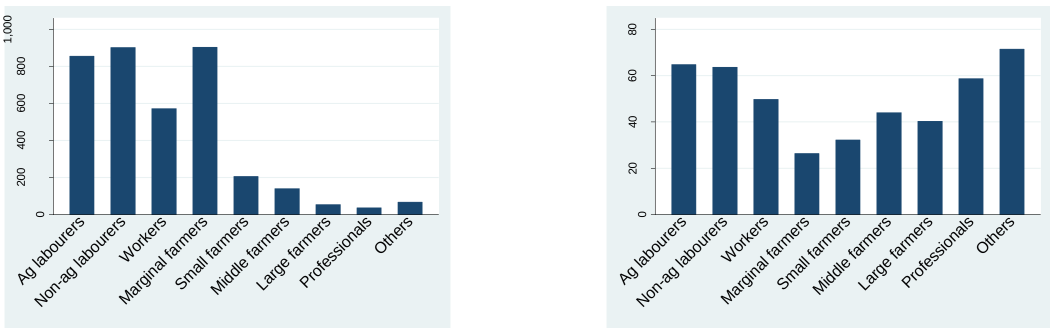
This study rigorously separates child labour from allowed child work. Child labour is harmful to children's development, while child work is not necessarily so. This study uses the ILO criteria to identify child labour, taking into account the definitions of hazardous industrial and occupational sectors. Hazardous industries include construction and mining; hazardous occupations include machinery workers, agricultural and fishery labourers, street vendors, and construction labourers (ILO, 2017, p.55). Children working less than 30 days a year are excluded from the categories of child labour so that the seasonality factor is discounted. In terms of a time threshold, this study estimates the number of working hours after which harm (being out of school) sharply increases and uses it as a threshold. According to the estimates, children aged 5 to 11 years who are involved in any form of employment (> 0 hours) and children aged 12 to 17 years working for 38.5 hours or more per week are categorised as child labourers. Details of how to measure time thresholds for these age groups are included in Appendix Table B.2. In using this method, we remove the reliance on conventional time thresholds.

5.3.2. Data sources and trends in India

The IHDS 2011/12 is a representative dataset on child labour in India. It provides detailed information such as industrial/occupational codes and working hours for three economic activities – waged work, farming, and family business. This study uses the total sums of hours worked in those categories. The total number of samples of children aged 5–11 is 51,551. Among them, 2,926 cases are found to be child labourers.

It is shown that about 5.7 percent of children aged 5–17 are engaged in child labour; of those, 61 percent are boys, and 39 percent are girls. These figures include only economic activity. Including domestic work within the category of child labour may greatly increase the percentage of girl child labourers. Note that the Indian census reported the number of main workers (those who work more than six months) among children aged 5 to 17 in 2011 at 11.8 million, which is 3.6 percent of all children in that age group (own calculation). Among them, 67 percent of child workers are male, and 32 percent are female.

The majority of child labourers belong to the lower classes, such as agricultural and non-agricultural labourers, and marginal farmers (Figure 5.2(a)). Descriptive figures show that children of agricultural and non-agricultural labourers have the highest working hours (Figure 5.2(b)). There are a few exceptional child labour cases in large farmers or professional groups, which makes an overrepresentation of the samples in terms of labour hours. Therefore, we need to estimate labour hours using a joint model of labour participation and labour hours. More details of household occupational groups are provided in Sub-section 5.4.2.



(a) Count of child labour by household occupations

(b) Weekly working hours(*) by household occupations

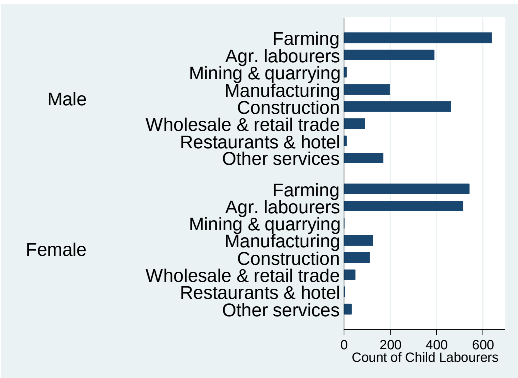
Figure 5.2. Count and working hours of child labourers by the household head's occupation.

Notes: IHDS, 2011/12; ages 5-17; survey-design weights (sampling weights divided by the mean and rounded up to the nearest integer greater than 0) were used; (*) weekly working hours without zero cases

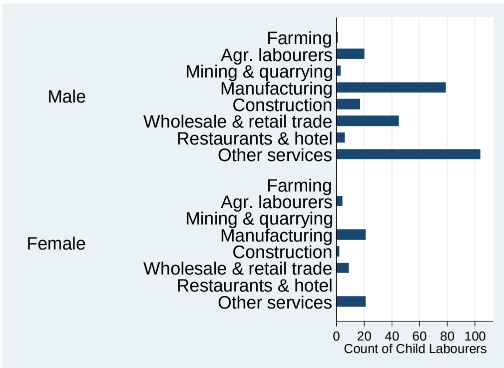
5.3.3. Gender division of child labour

Child labourers' working hours are categorised into two areas – informal and formal labour. The method used to define them is presented in Appendix Table B.1. Informal jobs include casual labourers, family workers in informal family businesses, and farming. We include farming as an informal labour category, as children work mostly on subsistence farms. In addition, child labour occurs in the formal sector too. Formal jobs include salaried workers or workers in family firms (with more than five employees). Firms are expected to have formal employment, but many employees work without a contract or are involved in outsourced work (Srivastava, 2012).

The informal sector is still the largest industry in India, and limited capital, less use of technology, low wages, and no contracts are its main characteristics (Parry et al., 1999). Child labour is found predominantly in the informal sector, such as agriculture, construction and low-skilled manufacturing. Figure 5.3 shows that girl child labourers are concentrated in the agricultural sector, but boy child labourers are spread across other industries, such as construction, manufacturing and services. There have been a few cases observed of girls in home-based manufacturing. Dual roles are also examined, especially among girls: 21 percent of girl child labourers indicate their primary status as household wife/household worker; and 23 percent of girl child labourers do both family work (mostly farming) and waged labour at the same time.



(a) Informal labour



(b) Formal labour

Figure 5.3. Industrial segregation of child labour by informal and formal labour

Notes: IHDS, 2011/12; ages 5-17; survey-design weights were used to obtain counts.

5.4. Methodology

5.4.1. Models

Our model is adapted from a Bayesian Poisson hurdle model (Neelon, Ghosh, and Loebs, 2013), which originated from the Mullahy (1986) Poisson hurdle model. There are several advantages of using a hurdle model in this study. It is a useful tool when regressing both labour participation and working hours on covariates, at the same time. It also addresses the zero inflation observed in the working hours of children[38]. Specifically, setting up a zero-inflated Poisson model is not necessary when using this model. Furthermore, a Bayesian approach provides advantages in dealing with complex models. The use of posterior predictive intervals provides more realistic assessment of uncertainty around children's labour-force participation and labour hours, which can be compared between girls and boys and between sectors.

The dependent variables are the labour hours of economic activity in each relevant sector, so hours spent on domestic chores are excluded. The first model uses total labour hours of child labourers as a dependent variable (Model 1). The second and third models use child labourers' labour hours in the informal (Model 2) and formal sectors (Model 3), respectively. Each model has two stages: (i) the decision on labour-force participation, and (ii) the number of working hours. The first part uses a probit regression that is based on a cumulative normal distribution function (denoted by Φ) of child labour participation, and the second part uses a truncated Poisson distribution[39]. The second stage of Model 1 is a truncated Poisson regression with data on non-zero cases only. By contrast, the second stages of Models 2 and 3 have many zero values and we apply the non-zero-truncated Poisson distribution.

The probability of labour-force participation (**p**) is regressed on $\mathbf{X_1}$, a vector of covariates (independent variables). The probability of **y** (labour hours) being zero is **1–p** (Eq.1). After the hurdle is crossed, the second part explains working hours, using a truncated Poisson regression and linking it with the probability of crossing the hurdle (**p**). Eq.2 explains the probability of **y=h**, given **y** exceeds the thresholds. The expected value is $\mathbf{E(y=h|\ X_1, X_2) = p*\mu/(1-e^{-\mu})}$. The probability of crossing the hurdle (**p**) directly affects the

[38] Alternative model is a 'Poisson regression with sample selection' derived from Terza (1998) and Cameron and Trivedi (2013) (https://www.stata.com/manuals/rheckpoisson.pdf, accessed 11 July 2021). To obtain the joint outcome of labour participation and working hours of child labourers, we select a Hurdle Possion model.

[39] The first part of a hurdle model can be either logit or probit. We choose a probit regression using the function 'Phi_approx' (https://mc-stan.org/docs/2_21/stan-users-guide/logistic-probit-regression-section.html, accessed 18 Aug 2020), which allows fast approximation to the model.

likelihood, when using a Poisson distribution. The vector $\mathbf{X_1}$ affects the $\mathbf{p}$, and $\mathbf{X_2}$ is a vector of covariates only affecting labour hours, $\mathbf{y}$. Models have a key parameter, $\boldsymbol{\mu}$, which is not a mean of labour hours but the 'original parameter' affecting the mean of labour hours (Geyer, 2006). Parameter $\boldsymbol{\mu}$ is assumed to be greater than zero hours for ages 5 to 11, and larger than 38.5 hours per week for ages 12 to 17, based on our close analysis of the data (Appendix Table B.2). In the models, maximum hours are set as 112 hours a week (16 hours*7 days). It should also be noted that any involvement in hazardous industries or occupations is child labour regardless of working hours or ages. The posterior predictive means of working hours are estimated by using the same set of independent variables in three models ($\mathbf{X_2}$) so that they are comparable between the models.

$$\Pr(y=0|X_1) = 1 - p,\ 0 < p < 1 \tag{1}$$

$$\Pr(y=h|\ X_1, X_2) = p \cdot \frac{\mu^k e^{-\mu}}{k!(1-e^{-\mu})}\ ,\ h=1, \ldots, 112,\ \begin{bmatrix} 1\ hr\ for\ ages\ 5-11 \\ 38.5\ hr\ for\ ages\ 12-17 \\ 1\ hr\ for\ hazardous\ work \end{bmatrix} \leq\ \mu\ \leq 112 \tag{2}$$

This first regression explains the probability of labour-force participation (now called, $\mathbf{p_{ij}}$) on the interactions between social groups and gender, and other demographic and social variables ($\mathbf{X_{1ij}}$). Subscript ***i*** indicates an individual level and ***j*** a hierarchical level representing 33 Indian states[40]. In the second stage, the regression explains the log of labour hours by class and gender, controlling for other characteristics ($\mathbf{X_{2ij}}$). To avoid the identification problem, two sets of covariates – one for sample selection and the other for a second stage – are set distinctively (van Hasselt, 2014; Newman et al., 2003). The first-stage regression includes the interactions between social groups and sex but excludes the interactions between class and sex. The second stage is the opposite, as it consists of the interaction between class and sex but excludes the others. In this way, we keep the model parsimonious. The first regression also includes two instrumental variables – female-headed households, and migration – which are not included in the second stage. While it is not easy to find instruments that affect one stage but not the other, a recent study proves that migration status is more influential over the decision of having a child work than the working hours (Ahmed and Ray, 2014).

The random intercepts, $\mathbf{u_j}$, include state-level intercepts in the first stage. In the first stage, a log of weights is applied as an offset; therefore, it adjusts estimates according to the proportion to the population. The weights are survey weights divided by the mean of weight and rounded up to the nearest integer above zero. Adding weights in the second stage would

[40] Two states - Lakshadweep and A & N Islands - are not covered by the IHDS 2011/12.

create a statistical error because we would truncate working hours between 0 and 112, so we avoid it.

$$Pr(y_{ij}=1) = p_{ij} = \Phi(\alpha_0 + \sum_{k=1}^{K} \alpha_k \cdot X_{1ijk} + \log(w_{ij}) + u_j), \quad (3)$$

X_{1ij} = {female, age, age^2, rural or urban, social group, social group* female[41], class + femheadedhh, migration},

$$Log(\mu_{ij}) = \beta_0 + \sum_{l=1}^{L} \beta_l \cdot X_{2ijl}, \quad (4)$$

X_{2ij} = {female, age, age^2, rural or urban, social group, class, class*female}.

All of the statistical analysis in this study is performed by using Bayesian inference, implemented in Rstan (R Core Team, 2018), i.e. the R interface to the program Stan (Carpenter et al., 2017)[42], and applies a cross-validation method to compare the sensitivity of models and check predictions.

5.4.2. Variables

Social group and class are included in both stages. The reference group for social groups is the so-called ‘forward’ social group[43]. Forward caste is a combined group of Brahmins and forward/general social castes defined by the IHDS 2011/12. The other groups are other ‘backward’ castes (OBC), ‘scheduled’ castes (SC or Dalit) and ‘scheduled’ tribes (ST or Adivasi).

We approximate class by two representative indicators – household occupational groups and the asset index. In line with previous studies, household is classified into to a few different groups using the primary activity of the head of the household. Details are in Appendix Table B.3. Being classified as agricultural labourers and non-agricultural labourers indicates that the head of household is hired and paid daily in each sector, respectively. Farmers are categorised into four groups according to the amount of owned land: marginal farmers that hold less than 1 hectare of land; small farmers that own between 1 and 2 hectares of land; middle farmers that own 2 to 5 hectares; and large farmers who own more than 5 hectares. Workers exclude casual or manual labourers but include salaried

[41] * means an interaction.
[42] Because of the large sample size (N=51,551), running two stages together takes too long. Thus, we separate the two stages and link them later. Estimated probability (‘p’) in the first part is later applied to the truncated Poisson (the second stage). Both stages are run in Rstan.
[43] The naming of social groups in India is made problematic by controversies around castes, caste-related laws, the Constitution and the presence of multiple religious groups. Here, religious group and caste group definitions are merged, as is commonly done in Indian social science; however, we grant no approval upon the status rankings known as ‘backward’ or ‘forward’ social groups. In India these are widely used, nominal group labels in the census and in law. Intriguingly, they do not merely group ‘castes’ together but also cut across a wide range of religions.

workers who are paid monthly or annually as well as self-employed workers. Workers include diverse occupational groups. For example, Iversen et al. (2017) used the categorisation of lower and higher status occupations and clericals. Vaid (2012) separated business class from salaried professionals. Lastly, professionals include professional workers as well as managers and government officials (Iversen et al., 2017). Others are the group that is not identified as one of the above-mentioned groups.

Table 5.1 describes the class and caste relations regarding the incidence of boy and girl child labour. In this table, the group 'workers' are separated into several occupational groups – lower- and higher-status occupations, clerical and other salaried workers, and business class (Iversen et al., 2017; Vaid and Heath, 2010; Vaid, 2012). To avoid a small sample size, in the regression, they are combined as workers. In general, ST and SC children are at great risk of being child labourers, especially if households belong to agricultural labourers or farmers' groups. Forward caste is heterogeneous as it includes not only Brahmins but also forward and general castes (Desai and Vanneman, 2018). A moderate percentage of child labourers are also found in the forward caste.

Table 5.1. The incidence of child labour and labour hours by class and social groups

Household occupational class		Boys					Girls				
		Forward caste	OBC	SC	ST	Others	Forward caste	OBC	SC	ST	Others
Non-ag labourers		6%	6%	9%	10%	2%	3%	5%	5%	9%	-
Ag labourers		7%	9%	12%	13%	3%	3%	8%	8%	11%	-
Workers	Lower salariats	1%	2%	2%	1%	-	1%	3%	2%	2%	-
	Higher salariats	1%	2%	2%	4%	-	2%	2%	2%	1%	-
	Clericals	4%	6%	5%	9%	-	2%	2%	2%	4%	-
	Other salariats	3%	9%	7%	11%	-	3%	7%	0%	5%	-
	Business (self-emp.)	5%	7%	4%	4%	2%	3%	5%	3%	7%	2%
Marginal farmers		7%	7%	11%	10%	11%	4%	7%	8%	8%	-
Small farmers		6%	10%	7%	16%	-	2%	3%	6%	10%	-
Middle farmers		5%	6%	15%	7%	15%	1%	5%	6%	15%	-
Large farmers		6%	5%	-	9%	14%	6%	10%	-	25%	-
Professionals/managers		3%	4%	3%	-	-	1%	1%	2%	-	-
Others		4%	5%	5%	4%	6%	1%	3%	2%	5%	7%

Source: IHDS 2011/12 (Desai and Vanneman, 2018)
Notes: % of child labour relative to the corresponding group; Top 10 items of each gender were highlighted; forward caste include Brahmins and forward/general caste; weighted counts are presented.

Asset index groups are also used as proxies for household socioeconomic class (see Appendix Table B.4)[44]. The final asset index score indicates a result consistent with the economic status of states (see Appendix Figure B.1). Using this asset index, we have created five quintiles as proxies of socioeconomic status. Female-headed households and migrated

[44] The IHDS provides its asset index too. However, its index is a simple sum of the number of asset items, so we have decided to build our own index.

households are controlled for in the first stage. They are under increasing economic pressure and lay progressively more burden on both male and female children. Migration, in this study, indicates whether households have any members left to find short-term work at least one month, such as harvest, brick kiln and construction, during the last five years (Desai and Vanneman, 2018).

Table 5.2. Description of variables

Variables	Mean	S.D.	Min	Max
Child labour rate	0.057	0.23	0.00	1.00
Weekly hours in any work by child labourers (without zeros)	50.05	0.44	7	112
Weekly hours in the informal sector (within child labourers)	43.33	0.48	0	112
Weekly hours in the formal sector (within child labourers)	5.06	0.27	0	84
Female (ref.: male)	0.48	0.50	0	1
Age	11	3.70	5	17
Urban (ref.: rural)	0.26	0.44	0	1
Female headed household	0.12	0.32	0	1
Migration	0.05	0.22	0	1
Social group (ref.: forward caste)				
Other backward caste (OBC)	0.43	0.49	0	1
Scheduled caste (SC or Dalit)	0.23	0.42	0	1
Scheduled tribes (ST or Adivasi)	0.09	0.28	0	1
Others	0.01	0.11	0	1
Household occupational group (ref.: Non-agr. labourers)				
Agr. labourers	0.13	0.34	0	1
Workers	0.23	0.42	0	1
Marginal farmers (<1 hec.)	0.13	0.34	0	1
Small farmers (1-2 hec.)	0.05	0.21	0	1
Middle farmers (2-5 hec.)	0.04	0.20	0	1
Large farmers (5+ hec.)	0.01	0.10	0	1
Professionals/managers	0.05	0.22	0	1
Others	0.10	0.30	0	1
Asset index category (ref.:3rd quintile)				
1st quintile	0.32	0.47	0	1
2nd quintile	0.20	0.40	0	1
4th quintile	0.18	0.38	0	1
5th quintile	0.12	0.32	0	1

Source: IHDS 2011/12 (Desai and Vanneman, 2018)
Notes: Ages 5–17; N=51,551; weighted mean and S.D.

5.4.3. Prior distributions

In the first stage, two prior distributions – normal (0, 1) and a Cauchy prior with mean 0 and scale, 2.5[45] – on coefficients are tested, with and without state-level intercept $\mathbf{u_j}$ (Table 5.3). According to the selected information criteria – DIC, WAIC, and LOOIC – Neither a normal

[45] Gelman et al. (2008) suggest a Cauchy prior with a scale of 2.5 as a default choice for coefficients of a logit regression, and then applying a Student-t prior distribution with mean 0 and standard deviation of 0.5.

nor Cauchy prior make a significant difference. Including second-level covariates to reduce all three values justifies the use of a two-level hierarchical estimation. The use of a Cauchy prior in a hierarchical model yields non-convergence for the states that do not have child labourers and requires a longer time to implement. Thus, in the first stage, we decide to use normal priors on both first- and second-level parameters. In the second stage, we apply a normal prior with mean 0 and standard deviation. Priors used in our models are

$$\alpha_0, \alpha_k, u_j \sim \text{Normal}(0, 1),$$

$$\beta_0, \beta_k \sim \text{Normal}(0, 1).$$

Table 5.3. Comparisons of informative criteria from different priors on coefficients in the first stage

Priors on coefficients	DIC	WAIC	LOOIC
Individual (Normal)	2183.4	2186.6	2186.6
Individual (Cauchy)	2144.4	2147.1	2147.1
Individual (Normal) + State (Normal)	2022.2	2028.1	2028.2
Individual (Cauchy) + State (Cauchy)	2022.0	2028.0	2028.1

5.4.4. Model-checking

For model-checking, 10 percent of the sample test is used. This cross-validation method is to predict and account for out-of-sample values that are left out for checking purposes, using the estimated parameters obtained from the other samples. As the three models are based on different dependent variables, the informative criteria are not compared with each other, but other goodness-of-fit tests, such as a Bayesian R-squared (Gelman et al., 2019) or a Bayesian posterior predictive p-value test (Gelman, 2005; Lunn et al., 2012) can be used. A Bayesian R-squared is slightly different from a classical R-squared as it is a ratio of variance of predicted values to the sum of variance of predicted values and expected variance of residuals (Gelman et al., 2019)[46]. It informs how much the fitted model explains the variability. A Bayesian p-value indicates whether there are extreme discrepancies between observations and predictions based on the models. In this study, we use a Bayesian mid p-value as a testing statistic[47]. A Bayesian p-value between 0.05 and 0.95 is considered acceptable (Gelman et al., 2013; Neelon et al., 2013).

Table 5.4 shows the result of the cross-validation study using a different set of 10-percent out-of-samples. In Model 1, a Bayesian R-squared reaches about 50 percent, and in

---

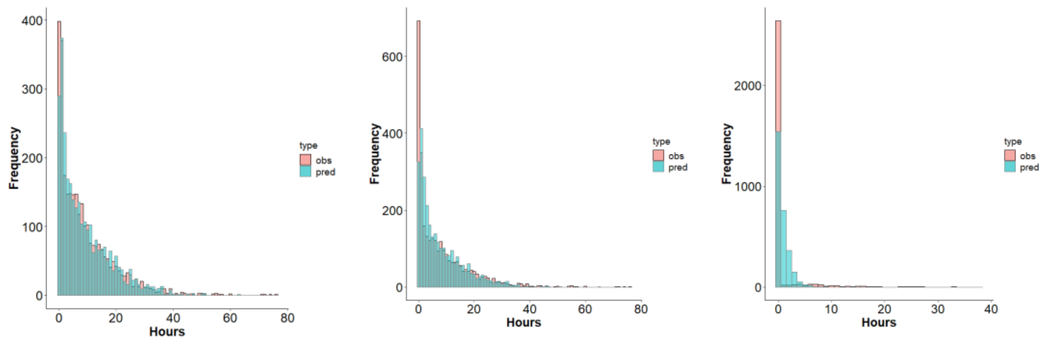
[46] Bayesian $R^2 = \frac{var^s_{fit}}{var^s_{fit}+var^s_{res}}$ in which s stands for simulations (Gelman et al., 2019).

[47] A Bayesian mid p-value refers to *pr(y.pred>y.obs) +1/2·pr(y.pred=y.obs)* (Lunn et al., 2012, p.149).

Model 2, it is around 30–40 percent, which supports the reliability of both models. However, Model 3 does not explain the variation well. The Bayesian p-values in Models 1 and 2 suggest that there is not a severe discrepancy between predictions and observations, but the Bayesian p-value in Model 3 does show extremes. The discrepancy between predictions and observations exists at near zero in Models 2 and 3 (Figure 5.4). Although adding a zero-inflation component within hurdle Poisson models can solve the overdispersion problem, the models become too complex, and the final results are not significantly changed. Therefore, we keep Models 2 and 3 without any extra parameters.

Table 5.4. Test statistics from the cross-validation study

	Bayesian R-squared			Bayesian mid p-value		
Models	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Simulation 1	0.485	0.365	0.269	0.845	0.926	0.004
Simulation 2	0.521	0.387	0.148	0.469	0.528	0.998
Simulation 3	0.482	0.398	0.212	0.07	0.004	0.001
Simulation 4	0.48	0.352	0.144	0.859	0.369	1.000



Total labour hours (Model 1) | Informal labour hours (Model 2) | Formal labour hours (Model 3)

Figure 5.4. Comparisons of observed and predicted frequencies by weekly hours

Source: IHDS 2011/12
Notes: In Model 1, predicted frequency is given as labour hours > 0, but because of a small p, some predictions are rounded as zeros; *obs* means observed frequency of each hour·p.

5.5. Results

5.5.1. The initial decision about child labour

Table 5.5 shows the regression results with child labour-force participation and a log of child labourers’ total labour hours. The result shows that girls are less likely to become child labourers than boys. The female coefficient is highly negative, indicating that fewer girl children participate in the labour force (coefficient= –0.41, 95-percent predictive intervals, PI [–0.51, –0.32]). The first part of the regression explains that social group is weakly related to labour-force participation except for children in the scheduled caste. However, there is a clear interaction between social group and girls' labour-force participation. First, the probability of participating in labour is higher among girls from other backward castes (OBC*female interaction = 0.21[0.10, 0.33]). Second, there is an even stronger interaction between tribal girls and labour-force participation (Adivasi*female = 0.37 [0.21, 0.52]). The results confirm that social group strongly affects girls' participation in the labour market.

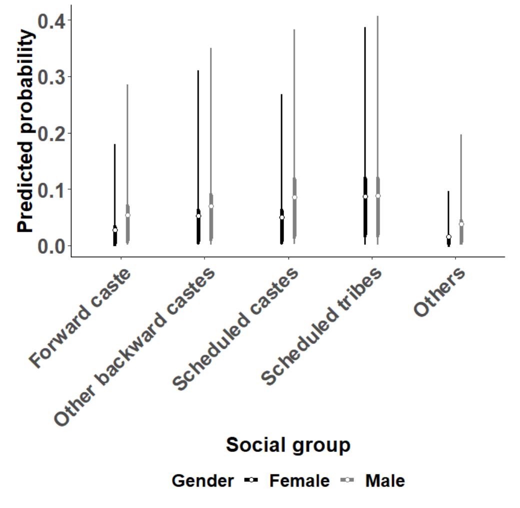
Table 5.5. Characteristics of the posterior distribution of parameters in both stages

	Stage I (for Models 1–3)				**Stage II (Model 1)**			
	(Labour participation)				**(Log of total labour hours)**			
	Coef	S.D.	2.5%	97.5%	Coef	S.D.	2.5%	97.5%
Female	**-0.41**	**0.05**	**-0.51**	**-0.32**	0.00	0.02	-0.03	0.03
Age	**0.13**	**0.00**	**0.12**	**0.14**	**0.21**	**0.00**	**0.21**	**0.22**
Age^2	**0.01**	**0.00**	**0.00**	**0.01**	**-0.02**	**0.00**	**-0.02**	**-0.02**
Urban	**-0.07**	**0.03**	**-0.12**	**-0.01**	**-0.03**	**0.01**	**-0.05**	**-0.02**
Migration	0.05	0.04	-0.04	0.13				
Female-headed household	**0.08**	**0.03**	**0.02**	**0.15**				
Social group (ref.: forward caste)								
Other backward caste (OBC)	0.05	0.04	-0.02	0.13	**0.02**	**0.01**	**0.00**	**0.03**
Scheduled caste (SC, Dalit)	**0.13**	**0.04**	**0.05**	**0.21**	**0.03**	**0.01**	**0.01**	**0.05**
Scheduled tribes (ST, Adivasi)	**0.15**	**0.05**	**0.05**	**0.26**	**0.16**	**0.01**	**0.14**	**0.18**
Others	-0.10	0.15	-0.40	0.18	**0.11**	**0.03**	**0.04**	**0.17**
Asset index category (ref.: 3rd quintile)								
1st quintile					**0.04**	**0.01**	**0.02**	**0.06**
2nd quintile					0.01	0.01	-0.01	0.03
4th quintile					**-0.05**	**0.01**	**-0.08**	**-0.02**
5th quintile					-0.02	0.02	-0.05	0.02
HH occupation (ref.: Non-agr. labourers)								
Agr. labourers	**0.11**	**0.03**	**0.05**	**0.17**	**0.03**	**0.01**	**0.01**	**0.05**
Workers	**-0.24**	**0.03**	**-0.30**	**-0.18**	0.00	0.01	-0.02	0.02
Marginal farmers (<1 hec.)	0.02	0.03	-0.04	0.09	**0.11**	**0.01**	**0.08**	**0.13**
Small farmers (1-2 hec.)	-0.03	0.05	-0.12	0.07	**0.03**	**0.02**	**0.00**	**0.06**
Middle farmers (2-5 hec.)	**-0.16**	**0.05**	**-0.26**	**-0.06**	**0.08**	**0.02**	**0.05**	**0.12**
Large farmers (5+ hec.)	0.07	0.09	-0.11	0.24	**-0.12**	**0.04**	**-0.21**	**-0.04**
Professionals/managers	**-0.55**	**0.06**	**-0.68**	**-0.43**	0.01	0.03	-0.04	0.06
Others	**-0.23**	**0.04**	**-0.32**	**-0.14**	**0.06**	**0.01**	**0.03**	**0.08**
Interactions								
OBC*Female	**0.21**	**0.06**	**0.10**	**0.33**				
Dalit*Female	0.05	0.07	-0.08	0.18				
Adivasi*Female	**0.37**	**0.08**	**0.21**	**0.52**				
Others*Female	-0.04	0.25	-0.57	0.43				
Agr. Labourers*Female					**-0.03**	**0.02**	**-0.06**	**0.00**
Regular workers/self-emp.*Female					**-0.15**	**0.02**	**-0.18**	**-0.11**
Marginal farmers*Female					**-0.04**	**0.02**	**-0.08**	**-0.01**
Small farmers*Female					0.04	0.03	-0.01	0.1
Middle Farmers*Female					0.01	0.03	-0.05	0.06
Large Farmers*Female					**0.17**	**0.05**	**0.06**	**0.28**
Professionals/managers*Female					**0.1**	**0.06**	**0.00**	**0.22**
Others*Female					**-0.06**	**0.02**	**-0.1**	**-0.01**
1st quintile*Female					-0.01	0.02	-0.04	0.02
2nd quintile*Female					0.00	0.02	-0.03	0.04
4th quintile*Female					-0.03	0.02	-0.08	0.02
5th quintile*Female					**-0.14**	**0.04**	**-0.22**	**-0.06**
Constant	**-2.19**	**0.19**	**-2.57**	**-1.84**	**3.42**	**0.01**	**3.39**	**3.45**

Source: Estimates using the Indian Human Development Survey (2012)
Notes: N of the first stage=51551; N of the second stage=2926; Posterior mean estimates and 95-percent PIs; economic activity only; ages 5 to 17; bold numbers indicate predictive intervals that do not contain zero.

Our model predicts the child labour participation rate as 0.06 (mean, 95-percent PI [0.00, 0.32]), which is slightly higher than the observed rate (0.057). Predictive intervals (PIs) are based on simulations, meaning the ranges that future observations fall within, which is usually wider than the confidence intervals for the mean. The same figure for girls is 0.049 (95-percent PI [0.00, 0.28]) and for boys is 0.07 [0.00, 0.35].

Figure 5.5 displays the predicted probabilities of individuals being child labourers by social group and gender, obtained from the first stage of Model 1 (which is the same in Models 2 and 3). There is a stark gender gap in the probability of child labour among most social groups, except tribes. 5% of Dalit girls work as labourers 9% of the same group boys are labourers. Labour participation of OBC girls is lower than that of boys in the same group (5% of girls and 7% of boys). Forward caste that includes both Brahmins and forward/general caste could have a small proportion of children in labour if the households belong to farmers.



Males	Mean	Median	LL(25%)	HL(75%)
Forward caste	5.4%	2.4%	0.8%	7.0%
OBC	7.0%	3.2%	1.1%	8.9%
SC	8.5%	4.5%	1.6%	11.7%
ST	8.9%	4.9%	1.8%	11.9%
Others	3.7%	1.4%	0.5%	4.3%
Females	**Mean**	**Median**	**LL**	**HL**
Forward caste	2.8%	0.9%	0.2%	3.2%
OBC	5.3%	2.1%	0.6%	6.2%
SC	5.0%	2.1%	0.6%	6.2%
ST	8.7%	4.7%	1.8%	11.9%
Others	1.8%	0.5%	0.1%	1.5%

Figure 5.5. Predicted probability (%) of children's labour participation by social group and gender

Source: Estimates using the Indian Human Development Survey (2012)
Notes: Model 1; mean, 50% PIs (thick lines) and 95% PIs (thin lines)

On the other hand, tribal girls share a high probability of being child labourers of about 8.7%, which is as high as that of boys (8.9%). In villages, many tribal girls participate in labour, and this is possible because they are less constrained by seclusion norms. That contrasts with the experiences of most other girls, who rarely work outside the home. Caste-related social norms are strong in Indian society, and it means both upper- and lower-caste girls are less likely to work outside the home; however, tribal girls are apparently exceptional.

5.5.2. Estimated working hours of child labourers

Contrary to the results of labour participation, girls' labour hours might be longer than those of boys. The second stage of Model 1 shows the estimated hours of labour among child labourers. According to the estimates in the second stage of the first model, the coefficient of gender does not show a meaningful result. However, there is a clear trend among girls in the agricultural sector. Children from households headed by agricultural labourers or farmers seem to experience particularly intensive working hours. Appendix Table B.5 outlines the posterior predictive estimates for children's total labour hours in terms of the hurdle Poisson and truncated Poisson. In the truncated Poisson, especially, girls from small, middle, and large farmers' households have longer mean labour hours than boys have: 47.9 (mean, 95-percent PI [11.1, 79.3]), 54.8 ([15.4, 81]) and 45.6 ([7.5, 76]) hours per week, respectively. In the hurdle Poisson, expected counts are strongly affected by the probability (p) obtained from the first stage; girls have greater variability than boys but do not show a typically higher mean in agricultural labour hours.

It is also shown that girl child labourers from farming households work for longer hours as land size increases. Remarkably, large farmers have a higher positive interaction (β=0.17, 95-percent PI [0.06, 0.28]) with girls' labouring hours compared with other groups. Previous studies find an inverted-U-shape relationship between family wealth and child labour (Basu et al., 2010; Kambhampati and Rajan, 2006); we find a similar inverted-U-shape relationship between land size and children's labour hours.

5.5.3. Sectoral division of child labour

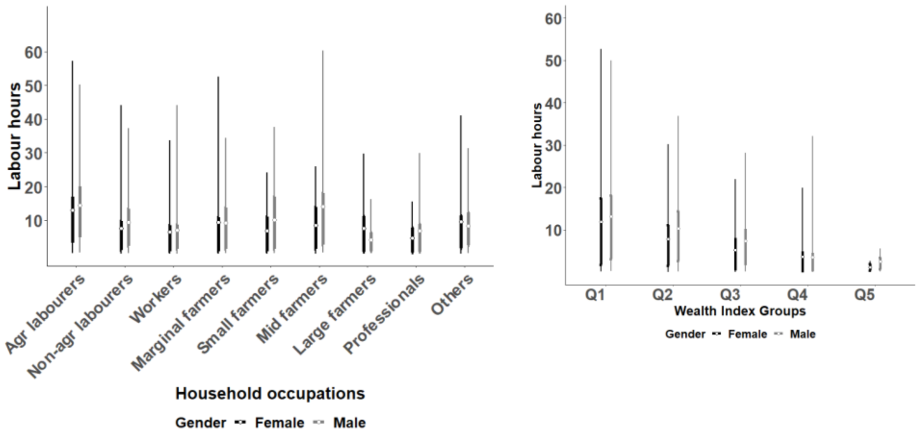
The differences in time allocation between male and female child labourers become more evident when we separate the informal and formal sectors. Models 2 and 3 provide a comparison of children's working hours in each sector, given that children are child labourers (Table 5.5). The first stage of these models is the same as Model 1, which we describe in section 5.1. The second stage of the models details in which sector girls work longer hours. It is shown that female child labourers tend to work longer hours in the informal sector (β=0.06, 95-percent PI [0.03, 0.1]) and fewer hours in the formal (β=–0.48 [–0.58, –0.37]). In the informal sector, the predicted number of labour hours is 44 hours (truncated Poisson mean, [9.8, 75.7]) for female child labourers and 43 ([10.1, 70.8]) hours for male child labourers. The difference is not large, and when considering the labour-force participation rate (p), the result is reversed: 8 (hurdle Poisson mean, [0.1, 45.5]) hours per week for females, and 11 hours ([0.2, 40.3]) per week for males. Girls show a larger variability in labour hours in the informal sector than boys do. On the other hand, male child labourers tend to work longer hours in the formal sector: they work for 7 hours a week (truncated Poisson mean, 95-percent PI [1.1, 29]), while female child labourers work 3 hours a week ([1.0, 14.3]).

Table 5.5. Posterior estimates of regression coefficients (second stage, Models 2 and 3)

	Model 2 (Log of informal labour hours)				**Model 3 (Log of formal labour hours)**			
	Coef	S.D.	2.5%	97.5%	Coef	S.D.	2.5%	97.5%
Female	**0.06**	**0.02**	**0.03**	**0.1**	**-0.48**	**0.05**	**-0.58**	**-0.37**
Age	**0.2**	**0.00**	**0.2**	**0.21**	**0.45**	**0.01**	**0.42**	**0.48**
Age^2	**-0.02**	**0.00**	**-0.02**	**-0.01**	**-0.04**	**0.00**	**-0.04**	**-0.03**
Urban	**-0.27**	**0.01**	**-0.28**	**-0.25**	**0.89**	**0.02**	**0.86**	**0.93**
Social group (ref.: forward caste)								
Other background caste (OBC)	**0.03**	**0.01**	**0.02**	**0.05**	**-0.07**	**0.02**	**-0.11**	**-0.02**
Scheduled caste (SC, Dalit)	**0.05**	**0.01**	**0.03**	**0.07**	**0.16**	**0.02**	**0.11**	**0.21**
Scheduled tribes (ST, Adivasi)	**0.22**	**0.01**	**0.2**	**0.24**	**-0.28**	**0.03**	**-0.35**	**-0.21**
Others	-0.01	0.04	-0.09	0.07	**0.62**	**0.06**	**0.49**	**0.74**
HH occupation (ref.: Non-agr. labourers)								
Agr. labourers	**0.04**	**0.01**	**0.02**	**0.06**	**-0.13**	**0.03**	**-0.18**	**-0.07**
Workers	**-0.21**	**0.01**	**-0.23**	**-0.18**	**0.49**	**0.02**	**0.44**	**0.53**
Marginal farmers (<1 hec.)	**0.12**	**0.01**	**0.1**	**0.14**	**-0.11**	**0.03**	**-0.17**	**-0.04**
Small farmers(1-2 hec.)	0.02	0.02	-0.01	0.05	**0.09**	**0.04**	**0.00**	**0.18**
Middle farmers(2-5 hec.)	**0.19**	**0.02**	**0.15**	**0.22**	**-2.3**	**0.15**	**-2.61**	**-2.02**
Large farmers(5+ hec.)	0.05	0.04	-0.03	0.13	-	-	-	-
Professionals/managers	**-0.13**	**0.03**	**-0.19**	**-0.07**	**-0.14**	**0.06**	**-0.26**	**-0.02**
Others	0.00	0.02	-0.03	0.03	**0.18**	**0.03**	**0.12**	**0.25**
Asset Index Category (ref.: 1st quintile)								
1st quintile	**0.08**	**0.01**	**0.06**	**0.1**	**-0.19**	**0.03**	**-0.24**	**-0.14**
2nd quintile	**0.04**	**0.01**	**0.02**	**0.06**	**-0.05**	**0.03**	**-0.11**	**0.00**
4th quintile	**-0.16**	**0.02**	**-0.19**	**-0.13**	**0.09**	**0.03**	**0.03**	**0.15**
5th quintile	**-0.19**	**0.02**	**-0.23**	**-0.15**	**0.32**	**0.03**	**0.27**	**0.38**
Interactions								
Agr. labourers*Female	0.02	0.02	-0.01	0.05	**-1.41**	**0.11**	**-1.62**	**-1.19**
Regular workers/self-emp.*Female	-0.03	0.02	-0.07	0.01	-0.08	0.06	-0.19	0.03
Marginal farmers*Female	-0.03	0.02	-0.06	0.00	0.06	0.08	-0.1	0.21
Small farmers*Female	**0.09**	**0.03**	**0.04**	**0.14**	-		-	-
Middle farmers*Female	**-0.06**	**0.03**	**-0.11**	**0.00**	-		-	-
Large farmers*Female	0.08	0.05	-0.02	0.18	-		-	-
Professionals/managers*Female	**0.22**	**0.06**	**0.09**	**0.35**	**1**	**0.15**	**0.71**	**1.28**
Others*Female	**-0.1**	**0.03**	**-0.15**	**-0.05**	**0.58**	**0.07**	**0.44**	**0.72**
1st quintile*Female	**-0.04**	**0.02**	**-0.07**	**-0.01**	**-0.93**	**0.06**	**-1.06**	**-0.8**
2nd quintile*Female	-0.01	0.02	-0.05	0.02	**-0.56**	**0.07**	**-0.69**	**-0.43**
4th quintile*Female	**-0.05**	**0.03**	**-0.1**	**0.00**	**0.14**	**0.06**	**0.02**	**0.27**
5th quintile*Female	-0.06	0.05	-0.15	0.03	-0.04	0.08	-0.21	0.11
Constant	**3.28**	**0.01**	**3.26**	**3.31**	**0.67**	**0.04**	**0.59**	**0.75**

Source: Estimates using the Indian Human Development Survey (2012)
Notes: Posterior mean estimates and 95-percent PIs; economic activity only; ages 5 to 17; bold numbers indicate predictive intervals that do not contain zero

Figure 5.6 (a) indicates expected informal labour hours of individuals by occupational class by hurdle Poisson models. There is great variability in the estimates of individuals' labour hours. Girls from households headed by agricultural and non-agricultural labourers, and by marginal farmers, show a notably large range of weekly labour hours. In the households of marginal and large farmers, in particular, girls work a longer mean number of hours than boys. In Figure 5.6 (b), there is a clear downward trend in child labourers' working hours against an increase in family assets. This trend is the opposite to what is observed in the formal sectors.



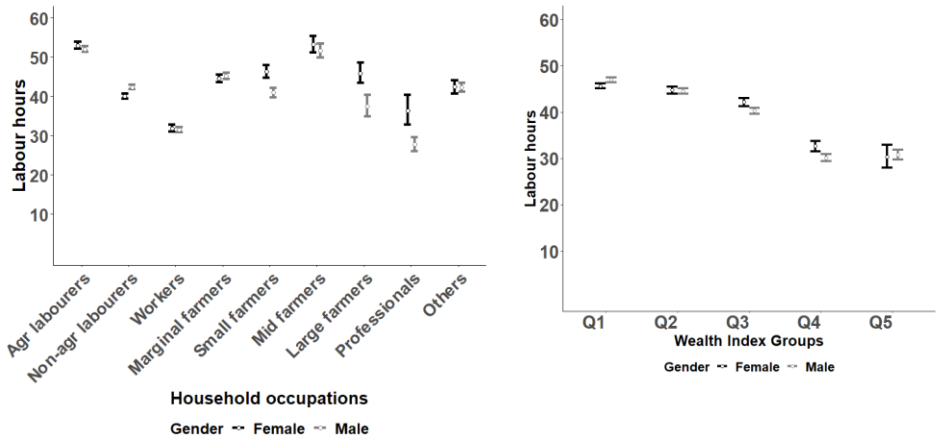
(a) Predictive expected children's working hours per week by household occupational class

(b) Predictive expected children's working hours per week by household assets

Figure 5.6. Predicted individual children's informal labour hours by class and gender

Notes: Model 2; mean, 50-percent PIs (thick lines) and 95-percent PIs (thin lines); (a) and (b) indicate posterior predictive intervals of expected count, $E(y) = p*\mu/(1-e^{-\mu})$

Meanwhile, Figure 5.7 (a) and (b) include a predicted mean and 95-percent PIs for average informal labour hours obtained by sampling from a predictive truncated Poisson distribution. They show a clear picture of the relationship between gender and informal labour hours. Girls from agricultural labourers' and farmers' households share long working hours and are employed mostly in the informal sector. In particular, farmers' girls work more hours than boys in the same group. For example, in small, middle and large farmers' households, girl child labourers work for 46.3 (95-percent PI [44.7, 48]), 53.3 ([51.2, 55.5]), and 45.9 ([43.4, 48.5]) average weekly hours, respectively, while boy child labourers work for 41 ([39.8, 42.2]), 51.6 ([49.9, 53.5]), and 37.5 ([34.8, 40.3]) hours a week. There is also a tendency for girls' workloads to increase as a household owns more land (from marginal to medium-sized land ownership). In the middle-wealth households, girl children work longer hours than children from other groups (because of the positive interaction between female and middle-wealth factors), although the difference is marginal. Among these middle-wealth groups, typically in rural areas, girls have a higher workload than boys. This different pattern of boys and girls among households with middle-level wealth supports the idea that a middle-class social norm might be related to longer hours for girls than boys.



(a) Predicted average working hours per week by household occupational class

(b) Predicted average working hours per week by household assets

Figure 5.7. Predicted average children's informal labour hours by class and gender

Notes: Model 2; (a) and (b) show predicted mean and 95-percent PIs for average expected count in a truncated Poisson, which is $\mu/(1-e^{-\mu})$.

To sum up, our analysis shows that girls may work longer than boys in the informal sector, particularly in agriculture, while female child labourers from most classes have lower working hours in the formal sector. Only a limited number of girls appear in manufacturing and other services. Conversely, male child labourers have comparatively higher labour hours in the formal sector. They are engaged in diverse types of work such as wholesale and trade retail, manufacturing and other commercial work.

5.5.4. Other social and economic effects

Children who live in urban areas work fewer hours in the informal sector than children in rural areas (β= –0.27, 95-percent PI [-0.28, -0.25], Model 2) but more hours in the formal sector (β=0.89 [0.86, 0.93], Model 3). Children in urban areas typically have jobs in construction (mostly in the informal sector) and low-skilled manufacturing, which require long working hours. They are generally boys, and many of them work within formal boundaries. At state level, Himachal Pradesh, Madhya Pradesh, Rajasthan and Chhattisgarh demonstrate a comparably high intercept point, which contributes to an increasing proportion of child labourers. This trend is consistent with the findings of a previous study based on Indian census data (Samantroy, Sekar, and Pradhan, 2017). Those states are rural-based areas with small-scale industries reliant on cheap labour. Furthermore, two instrumental variables return different results. In the regression, the migration effect on child labour participation was not clear, while children from female-headed households showed a clear association with labour-force participation (coefficient=0.08 [0.02, 0.15]). In future research, their association with gender could be considered.

5.6. Discussion and Conclusion

This study provides a very accurate prediction of the probability of Indian girls being child labourers (in the sense of damaging forms and amounts of child labour) and their working hours. It introduces a carefully designed two-stage model, making it possible to compare labour participation and labour hours at the same time. Moreover, the Bayesian hurdle Poisson model can efficiently handle data with many zero counts. As it accommodates both the proportion of participation and working hours, it can generate less biased results. The observed mean labour hours are 3 hours a week for all children, and yet are 50 hours a week for child labourers. The hurdle model estimates the average to be 11 hours a week (95-percent PI [0.14, 47.27]) among all children by incorporating both labour-force participation and hours of work. Furthermore, in this study, truncated Poisson results are shown together with the results of the hurdle Poisson model, helping at the methodological level to increase our understanding of the working-hour patterns of all working children including, specifically, 'child labourers'. According to the estimation, the mean number of girl child labourers in 2011/12 reaches 7.7 million, which is the estimated rate multiplied by census population, while boy child labourers number about 12.2 million among children aged 5–17 in India. Estimated mean hours of work are 44 hours a week for girl child labourers and 43 hours a week for boy child labourers in the informal sector. The hurdle model estimates the average to be 8 hours a week among girls and 11 hours for boys by incorporating both labour-force participation and hours of work.

The results account for why female children are less visible in the Indian labour market but work for longer, factoring in the gender and development (GAD) theoretical approach. People allocate different work types and working hours for children according to household class, institutional practices and children's gender. Girls are less likely to work outside than boys, but their workload is no smaller than boys' workloads in the informal sector. Our regressions used only the child labour hours spent in the labour market, based on the limitations of the IHDS dataset. If we include unpaid household services, girls' workload would be seen as even higher. Girls often carry dual burdens of both household and economic work. Considering these factors, girls' contributions to the family economy should be better recognised.

This study also provides firm evidence that social norms, such as social group norms and gender norms, operate on child labour. The interactions between social group (OBC and tribes) and girl factors are strong on labour hours while controlling family wealth and occupation. The class difference in girl/boy hours was very noticeable – suggesting that GAD should not be ignored by labour experts. For children in most social groups, except tribal children, there is a notable distinction between girls and boys in labour-force participation. Supposedly 'forward' social group backgrounds (i.e., certain privileged caste groups) strongly limit girls working outside the home. Furthermore, among scheduled caste children, there is a considerable gender gap in child labour participation. If social groups follow the upper caste social norm, they limit their daughters' outside activities, which is Sanskritisation. On the other hand, scheduled tribe girls have the highest probability of being child labourers, across India, and there is almost no gender gap in their labour-force participation, which represents an egalitarian group norm.

Elsewhere in India, within groups that follow patriarchal social norms, girls are expected to help with household work, and they are limited to work in the domestic sphere, such as farming or home-based production. As girls are bound to work near to their home, they become less visible. The traditional gender norm is typically stronger within farmers' households, in which the gender division of boys and girls becomes more robust. Casual labourer households have the highest levels of child labour; while boys are involved in broad areas, girls are highly concentrated in agricultural jobs. This finding supports previous literature on the feminisation of the agricultural sector (Garikipati, 2009; Olsen and Mehta, 2006b). Not only female adults but also girls work as farmers or agricultural labourers.

The concentration of girls in the informal sector is a result of gender stereotypes and norms based within the social structure and culture, which is worrying for many reasons. First, girls suffer from working long hours and receiving low wages in the informal sector. Girls working as agricultural labourers have the longest average working hours and the lowest wage rate among child labourers. Although we emphasise female child labourers' indecent work situations in this study, boy children also suffer from them. Secondly, the informal sector employs the largest number of people in India, and there is a clear trend for women to be kept within it. As mentioned, girls' areas of work within the agricultural sector might deepen the division of work by gender among adults in future. Lastly, as informal labourers, girls carry a dual burden of household chores and economic activities and also of agricultural casual labour and farming. In this way, they can suffer even longer working hours and are likely to drop out of school. Girls also work for domestic duties together with informal labour. This study is limited in capturing the time pattern of girl child labour, including time spent on household activities, because the dataset does not cover them.

Furthermore, girls' long working hours are also related to the attitude of son preference in households that allocate different types and hours of work to boys and girls. Typically, low-middle or middle-wealth families are likely to give more hours of work to girls than boys. This finding proves that there is a tendency among middle-wealth farming households to invest in their sons' educations. This is supported by a positive relationship between land size and girls' working hours among farmers' households and typically longer working hours for girls in large-land holding households compared with the boys.

Overall, the evidence of this article suggests that gendered patterns of child labour and children's working hours are compounded results of social structure, institutions and norms, and children's gender. Gender relations woven within other social relations have explanatory power regarding both child labour-force participation and intensity of labour. Female child labourers are fewer in number, but the degree to which they contribute to the informal economy is no smaller. The lack of protection might more significantly affect girls who belong to lower classes. Highest priority should be given to girl child labourers from low socioeconomic households, as they are at high risk of exploitation in the informal sector. Parental awareness of the importance of girls' education is necessary. At the same time, there should be consideration for girls from low-middle and middle-wealth households who might work long hours within family spheres.

Chapter 6. Predicting Child Labour Risks by Gender and Norms in India

Abstract

This article aims to understand how social and gender norms affect girls' and boys' participation in child labour in India, which is mainly defined by a work-hours threshold. It develops a regression model using two datasets – the Indian Human Development Survey, 2011/2012 and the World Value Survey India, 2012 – to predict child-labour risks based on such norms. The gender and development (GAD) approach provides a theoretical foundation for applying norms in association with social and gender relations. The results of the regression model have revealed that a norm supportive of women's work helps reduce the risk of child labour, especially among girls. In contrast, seclusion norms have the opposite effect on child labour. Child-labour practices are varied because agents accept or deny norms as belonging to social structure. Our findings confirm that transforming gender norms through empowering women and girls is essential for reducing child labour in India.

Keywords**:** Benevolence norm, Child labour, Gender and Development, Institutionalised norms, Seclusion norm, Time threshold, India, Youth employment

6.1. Introduction

Child labour is regarded as employment harmful to children, especially in that it lowers school attendance and increases dropout numbers. Despite the decreasing trend in child labour in India, children from the lower social class are still employed as agricultural labourers, construction workers and low-skilled production workers. Child labour occurs not only as a result of an unequal social hierarchy but also the embedded roles and expectations according to children's gender and social status. Social norms play a key role for the children, parents and families who make decisions regarding child labour. Much research has focused on economic reasons for child labour, such as 'moral economy' (Basu and Van, 1998; Lima et al., 2015) or structural inequality (Venkateshwarlu and Da Corta, 2001; Phillips et al., 2014). Less attention has been paid to the role of institutions such as norms and values in attitudes to child labour. Burra (2001) has explained that child labour among girls can be a result of long-standing 'beliefs, traditions and mind-sets'. Bhat (2010) has found that culture, such as gender roles and expectations, play a key role in lowering the educational level of girls, thereby resulting in gender bias in child labour. Girls are expected to help their parents, spending long hours in household work and missing out on education, proper nutrition and future employment (ibid.).

The aim of this study is to establish a clear link between norms and child labour in economic activities across the regions of India that have been formed in gender and social relations. The GAD approach provides a comprehensive view of child labour. Gender, as a social construct, and its embedded social relations, explains the prevalence of child labour. Based on the GAD approach, this study explores how child labourers are affected by gender and social norms according to their socioeconomic positions. Multiple norms affect agents' decision-making regarding children's participation in labour; thus, it is significant to reveal what and how these norms have increased the risks of child labour.

This study also suggests how to measure working hours in order to define child labour accurately. In general, the time threshold used to define child labour has relied on international standards; therefore, it does not reflect the differences between countries. The question of how many hours of work are harmful to children is not a simple one; it depends on the relationship between potential harm and children's labour hours. This study uses a model to measure the educational harms caused by children working long hours, and it generates a best-time threshold: a certain time point at which the harm to children's education increases to a significant level. This method suggests a novel method that does not rely on a conventional threshold.

This study answers four research questions regarding the situation in India:

(1) how do gender norms affect child labour risks?

(2) how does the benevolence norm affect these risks, in comparison with the effect of the gender norm?

(3) how do norms interact with children's genders?

(4) how does the best estimation of thresholds of working hours help obtain an accurate child-labour rate?

The next section provides a theoretical background for this study, explaining the relevance of the gender and development approach to the study of child labour. A methodological framework follows whereby the study estimates a time threshold and applies it to define child labour. Next, the incidence of child labour is regressed on norms as well as other structural and institutional variables, using a Bayesian Poisson model to predict the most precise child-labour risks. A risk map shows the predicted rate of child labour among types of household occupation, gender and state. Ultimately, based on our analysis of child labour risks in India, some policy suggestions are provided.

6.2. 'Gender and Development' and Child Labour

Before discussions of the gender and development (GAD) approach, the political economy and the social construction approach to child labour are compared. The political economy approach understands child labour as exploitation of children in an unequal power relation. Child labour is incorporated into global production networks in forms of subsistence production, such as outsourcing and subcontracting work (Phillips et al., 2014; Bhaskaran et al., 2014). Thus, in political economy, the key is 'how production networks are embedded in particular kinds of social and power relations' (Phillips et al., 2014, p.431). Nieuwenhuys (2007, p.153) notes that 'global child labour' is concentrated in the reproductive realm of the Global South, as 'new forms of labour control'. Children's care-work and subsistent production result from the global market, and girls' contribution to global production is often invisible and ignored (ibid.).

However, at the same time, child labour is a socially constructed phenomenon. Socially constructed childhood has become a crucial sociological concept (James and James, 2004; James et al., 1998; James and Prout, 1997). The institutions of childhood vary across cultures and are not separated from other aspects such as class, gender and ethnicity (James and Prout, 1997). In this sense, child-labour rates rely on the social, cultural and economic context, which should be understood within a life-course perspective (James et al., 1998). There can be different causes of child labour at local, regional and national levels. Nieuwenhuys (1994) asserts that assumptions should be avoided regarding the universalities of capitalist forms of child labour by looking at children's work routines in their everyday lives. Abebe and Bessell (2011) suggest the 'socio-cultural discourse', such as gender analysis and children's points of view, enhance understanding of the complexities of child labour.

The GAD accounts for both aspects of child labour – structural inequality and social construction. It is based on a holistic overview of social, economic and political approaches, considering both productive and reproductive areas (Young, 1993). Kabeer (1994, p.57) states that gender relations are mutually constituted with other social relations and that, therefore, gender relations can explain diverse forms of social inequalities. Based on gender and gender relations, the GAD approach focuses on revealing how development reshapes power relations, and women are considered agents of change within this structure of power (Momsen, 2010, p.29).

In the GAD approach, gender and gendered practices are social constructions, which means that there are roles played by norms and institutions alongside the structure. Kabeer

(1994, p.309) describes the role of institutions in gender relations, including rules, norms and customs. The same author (1999, p.441) emphasises that agents' choices are derived from a 'deeper reality', which is 'inscribed in the taken-for-granted rules, norms and customs within which everyday life is conducted'[48]. Therefore, in the GAD approach, how norms are reflected by agents located in the social structure is a key to understanding the complex child-labour decision process.

There are diverse definitions and attributions associated with norms from different academic disciplines. Bicchieri (2014) defines norms as social constructs that are supported by people's beliefs. López and Scott (2000, pp.25-26) describe them as 'a rule of conduct that is shared by a particular set of people'. Olsen and Morgan (2010) suggest norms make institutional changes in the informal sector; for example, 'be a dignified permanent labourer' or 'be an honourable rural casual labouring wife' are the norms associated with employed children and their parents. Elder-Vass (2007) emphasises the existence of the 'causal power' of social structure, by which a group conforms to the normative standard. However, norms are diverse and they will vary depending on individuals and relations. Norms are also subject to gradual changes. While Kandiyoti (1998) insists that 'patriarchal negotiation' occurs, Olsen and Morgan (2010) explain that variations of norms are essential because agents can reproduce and change them. Agents make decisions by accepting and applying norms tacitly (ibid.). They have choices of strategies, and they also change strategies (ibid.).

In summary, following the GAD approach, child-labour practices require an understanding of how girls and boys are differently integrated into the labour market during the process of development. Moreover, any gendered practices are driven by agents' norms and beliefs, which are generated by their diverse experiences and positions within social and gender relations. Thus, boys' and girls' participation in the labour force is evidently affected by structural and institutional features, as well as agents' social or gender norms.

6.3. Norms and Child Labour in India

This study focuses on different expectations regarding women's economic roles in society ('norms of women's work'). Females are often recognised as housewives or secondary income earners in India. Male and female division of labour in India relates to cultural norms at a deep level (Garikipati, 2009). Increased access to assets or land can help women's agency, but this is not necessarily the result (Garikipati, 2009; Jejeebhoy and

---

[48] Kabeer (1999, pp.440-441) uses Bourdieu's (1977, p.164) concept doxa to refer to a self-evident part of the social world; 'a more critical consciousness' reflects the passage from 'doxa' to discourse, which allows individuals to form a critical perspective of the social order, rather than accepting it without question.

Sathar, 2001). Indian women in some cultural groups are considered housewives despite their participation in farming and domestic work (Olsen and Mehta, 2006b). Similarly, there are different expectations for girls and boys in terms of their economic responsibilities.

This study assumes that norms supportive of women's economic participation can help reduce child-labour risks. This seemingly contradicts previous findings that higher employment among women increases the incidence of child labour, including both market and domestic work (Ray, 2000; Francavilla and Giannelli, 2007). Some considerations address how this contradiction can be resolved. First, this study is not examining women's economic participation but cultural norms regarding supporting women's work. Norms are sometimes different from practices. Indian women in some cultural groups are considered housewives despite their participation in farming and domestic work (Olsen and Mehta, 2006a), implying that norms and practices are not aligned. This study supports the view that if women's work is less valued, despite women's high economic participation rate, then child labour could arise from a lack of resources and is additionally affected by unequal gender norms. A similar hypothesis is found in the work of Kambhampati (2009), which is an attempt to prove the effect of mothers' autonomy (a mother's expenditure contribution) on the level of gender equity. For another example, in households where a mother's education level was high, the probability of child labour was reduced; on the other hand, the mother's labour force participation increased the risk of child labour outside the home (Self, 2011). In both these studies, the effects of norms are indirectly examined.

This study assumes that norms supportive of women's economic participation can help reduce child-labour risks for both boys and girls, for two reasons. First, if adult women in the home participate in the labour market, children might be less likely to take part in economic activities. Second, an egalitarian perception of gender roles and stereotypes encourages girls to continue secondary or higher education. Burra (2001) asserts that parents have 'false consciousness' regarding girls that suggests they should be socialised into economic roles; as a result, they disregard the importance of girls' formal education. In this regard, parents adopting gender-equality norms might play a key role in reducing child labour among girls.

Conversely, seclusion norms can bring about the opposite result in children's labour participation. In northern areas, such as Rajasthan, Uttar Pradesh and Bihar, women live under severe seclusion, whereas in southern areas, seclusion is less common. Seclusion norms restrict women to work locations inside the home (Kantor, 2003). Under the female seclusion norms, both boys and girls may take part in economic activities when their mothers are not allowed to work outside. In such cases, participation of girls in the labour force arises in restricted areas, thereby generating gender segregation among children: i.e., in rural areas, boys can work in both agricultural and non-agricultural sectors, while girls work only in the agricultural sector. While women's outside activities are culturally inhibited, girls can be sent to work as domestic workers in other households (Chakravarty et al., 2012).

While this study focuses on the effects of gender norms on child labour, the benevolence norm ('doing good for society') might possibly be associated with reduced incidence of child labour. The World Value Survey (2012) shows that India has a robust benevolence norm compared to other countries. The question is asked whether people value 'doing something for the good of society'. More than 42 percent of people very much agree with that benevolent idea, and 33 percent answer that they somewhat agree with it. A 'benevolent parent' is a strong assumption, which supports the view that parents decide to send their children to work only if adults' wages are too low (Basu and Van, 1998). Although this study does not employ benevolence norms among parents towards their children as an indicator, it takes a norm of benevolence at state level into account. Benevolence is a strong form of social capital as well as a parental moral obligation; therefore, it might reduce the incidence of child labour.

Figure 6.1 shows the conceptual framework of this study. The three norms mentioned above are supposed to affect agents' decision-making regarding child labour. The process of child labour occurs within a broad social structure. Class and other social relations significantly impact whether agents accept or reject norms. Gender is interwoven with social relations, and gender and norms reflect one another; therefore, the role of gender appears in all agents' decision-making processes regarding child labour.

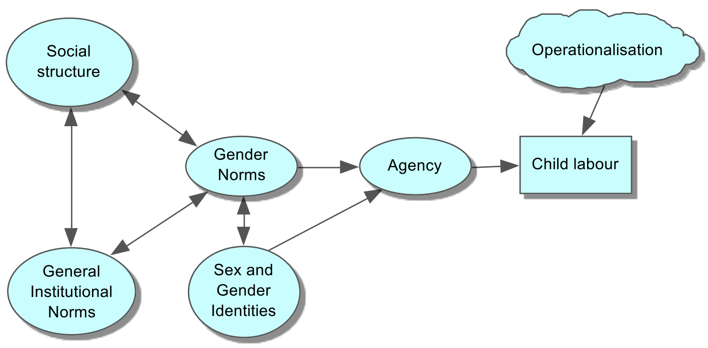


Figure 6.1. Conceptual framework: norms, structure, and gender model on child labour

Note: Caste and lineage norms in India act both as structural factors and as sets of institutionalised norms, which are not fixed and, indeed, are highly controversial in a situation where laws contradict some practices.

6.4. Operationalisation

This article defines child labour as the economic activity of children aged between 5 and 17 that deprives them of primary or secondary education. Child labour requires a considerable amount of time and thus potentially harms children's education. This study uses the ILO's (2017) definition of hazardous occupations to separate child labour from accepted child work. Hazardous occupations or industries are also categorised as child labour regardless of working hours or workers' ages. In addition, the ILO applies different working-hours thresholds for different age groups, such as one hour per week for ages 5–11, 14 hours for ages 12–14, and 43 hours for ages 15–17.

This study follows a similar categorisation to the ILO's in terms of working hours, but it uses a statistical model to estimate harmful working hours for Indian children. The model estimates a time threshold over which labour can bring significant harm to children. It removes the reliance on conventional time thresholds typically assumed to define labour as harmful to children in India. Children's working hours differ in each country, and the work-hour thresholds can significantly affect the estimation of child labour rates. Guarcello et al. (2004) provide a model to measure physical and educational harm associated with working hours as an independent variable. However, their method does not provide a way to detect a critical and statistically meaningful time-change point. Our proposed model allows incorporation and prediction of an unknown change point (a time threshold) of working hours. It relies on the assumption that educational harm, such as not being enrolled in or not attending schools, sharply increases after a certain number of working hours per week. This model enables us to learn about the relationship between working hours and the potential risks of child labour, providing accurately predicted time thresholds to define child labour in the Indian context.

6.5. Data and Models

This study used two datasets from the same year: the Indian Human Development Survey (IHDS) 2011/2012 and the World Value Survey India in 2012. The IHDS is a longitudinal survey covering multiple topics, in 33 states and 384 districts (Desai and Vanneman, 2018). It provides detailed information on child labour, such as working hours, industries or occupations, and individual and family backgrounds. The IHDS 2011/2012 was conducted in 42,152 households, including 40,018 households who were surveyed in the IHDS 2005 (ibid.).

The WVS India (2012) provided information on norms and values as part of World Value Survey Wave 6 (2010–2014; Inglehart et al., 2014), including data from 4,078 cases randomly selected from 17 states. The 17 states were Andhra Pradesh, Bihar, Chhattisgarh, Delhi, Gujarat, Haryana, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orrisa, Punjab, Rajasthan, Uttar Pradesh, Uttarakhand, and West Bengal. The indicators of norms and values were the averages of cases in each selected state.

It should be noted that in this study, child labour accounted for involvement in economic activity only, mainly due to a lack of information (Kim et al., 2020). Thus, household production – farming or working in a family business – was included in the categories of child labour but household service (domestic chores) was not. Furthermore, economic activity was taking to mean working more than 30 days a year in an attempt to exclude short-term seasonal work from the category of child labour.

The key task of this study is the measurement of social norms, and Table 1 presents three selected norms. People were asked two questions to assess to what degree they agreed with norms: 1) doing good for society (WVS 2012 question no. V74) and 2) norms supportive of women's work (WVS, 2012, a latent variable). A supportive norm for women's work shows the degree to which people think that women's work is a way to be independent. The details of how to construct this latent variable are provided in Appendix Table C.2. The other social norm was obtained from the IHDS 2011/12, which was a seclusion norm indicating the proportion of women who practice ghungat/burkha/purdah/pallu : the customs and traditions intended to keep women unseen by men outside their own family.

Table 6.1. Descriptive summary of variables

Variables	Mean	Min	Max
Child labourers (all)	0.06	0	1
Child labourers (girls)	0.05	0	1
Child labourers (boys)	0.07	0	1
Urban	0.26	0	1
Asset (proportion of households in the lowest asset index category)	0.32	0	1
Land size category (0=no land, 1=upto 1 hectare, 2= 1–2 hectares, 3=2–5 hectares, 4=above 5 hectares)	0.74	0	4
Dalit (scheduled caste, SC)	0.23	0	1
Adivasi (scheduled tribes, ST)	0.08	0	1
(Norm 1) Practicing *ghungat/burkha/purdah/pallu* (women only)	0.63	0	1
(Norm 2) Doing something for the good of society* (0=Not at all like me, 1=Not like me, 2=Somewhat like me, 3= Very much like me)	2.25	0	3
(Norm 3) Support for women's work* (A latent factor score)	0.00	-1.46	2.02

Source: (*) World Value Survey India (2012); Indian Human Development Survey (2012) for all other variables
Notes: the sample size is 44,042, which is aggregated into 422 (13 occupational groups * 17 states * gender); 17 States only; urban areas are identified by the 2011 census (Desai and Vanneman, 2018)

6.5.1. A threshold model

The model used to estimate critical points (time thresholds) related to the prevalence of out-of-school children is a simple Poisson regression with an endogenous time-index parameter. Out-of-school children indicate children who are currently not enrolled in schools (ED5[49]=0 or missing). UNESCO (n.d) defines them as "children and young people … who are not enrolled in pre-primary, primary, secondary or higher levels of education." This indicator might also include children not in schools (non-attendance) (UNESCO, 2017), but the IHDS does not provide information on current attendance. The numbers of these children in India in 2011/12 and their observed working hours are available in Appendix Table C.1.

The threshold model[50] regresses the count of children who are out-of-school at each hourly point: $\boldsymbol{t}$ denotes the number of daily working hours, and $\mathbf{Z_t}$ indicates the weighted count of out-of-school children at each hour of work. The estimates of the time index are used as thresholds to count child labour as per the norm models.

$$Zt \sim \text{Poisson}(\lambda t * nt),\ t = 0, \ldots, 16 \qquad (1)$$

$$\lambda_t = \begin{cases} \lambda_1, & \text{if } t \in [0, \Theta] \\ \lambda_2, & \text{if } t \in [\Theta, 16], \end{cases} \qquad (2)$$

$$\lambda_1 \sim \text{uniform}(0, 1), \qquad (3)$$

$$\lambda_2 \sim \text{uniform}(0, 1), \qquad (4)$$

$$\Theta \sim \text{uniform}\ (0,16). \qquad (5)$$

A Poisson distribution is applicable to counting data: the parameter $\boldsymbol{\lambda_t}$ represents the risk rate of children being out of school when they work for $\boldsymbol{t}$ hours per day. The time index parameter $\boldsymbol{\theta}$ indicates daily working hours after which educational harms increase sharply. Thus, the risk parameter $\boldsymbol{\lambda_t}$ relies on two given conditions: one is $\boldsymbol{\lambda_1}$, which is the risk parameter when children's daily working hours number less than $\boldsymbol{\theta}$, and the other is $\boldsymbol{\lambda_2}$, the risk parameter when working hours are equal or more than $\boldsymbol{\theta}$. If the two parameters, $\boldsymbol{\lambda_1}$ and $\boldsymbol{\lambda_2}$, are not identified, $\boldsymbol{\lambda_t}$ will have diverged numbers. Otherwise, it has two values, which are $\boldsymbol{\lambda_1}$, if $\mathbf{t} < \boldsymbol{\theta}$, or $\boldsymbol{\lambda_2}$, if $\mathbf{t} \geq \boldsymbol{\theta}$. Two conditional parameters, $\boldsymbol{\lambda_1}$ and $\boldsymbol{\lambda_2}$, are positive rates; therefore, a uniform distribution of (0, 1) is assumed. The parameter $\boldsymbol{\theta}$ is supposed to have an integer uniform prior between 0 and 16 hours (maximum daily working hours given by the data). It was assumed that a continuous distribution for the time index $\boldsymbol{\theta}$ as working hours is, in principle, a continuous variable, operationalised in the data by integers. A continuous prior

[49] "Is [NAME] currently enrolled in school or college?" (Indian Human Development Survey, 2012)
[50] The original model, which measures a change in the number of text messages, is available in Davidson-Pilon (2015).

permits a more precise estimation of the time index than when using, for example, an integer uniform prior.

The same model is repeated for three age groups: ages 5–11, ages 12–14 and ages 15–17. After running the model, estimates of **θ** are multiplied by 7 under the assumption that children work for seven days a week and then they are considered as weekly time thresholds of child labour in the norm models (Section 5.2).

6.5.2. Norm models

The purpose of the norm model is to predict a state-level child-labour risk by household- and state-level covariates. The regression is made with weighted aggregates of child labourers ($\mathbf{y_{ijg}}$), including both hazardous industries or occupations and any work done for long hours exceeding a threshold. They are aggregates by household occupations (***i***), states (***j***) and gender (***g***). As this model uses the count of child labour per group (non-negative integers), a Poisson distribution is a suitable choice. Household occupational groups included agricultural labourers, non-agricultural casual labourers in manufacturing, construction, and sales and services separately, workers in manufacturing and sales and services separately, marginal (land ≤ 1 hectare), small (1<land ≤ 2 hectares), and middle (2<land ≤ 5 hectares) or large (land > 5 hectares) farmers, clerical workers, managers, professionals, and others. $\mathbf{y_{ijg}}$ is the weighted count of child labourers for each group, ***ijg***, which are assumed to have a Poisson distribution with the key parameter $\boldsymbol{\mu_{ijg}}$ that is the child labour risk rate. The sample size of each group, $\mathbf{n_{ijg}}$, is used as an offset parameter. A natural logarithm of $\boldsymbol{\mu_{ijg}}$ is then explained by household- and state-level covariates $\mathbf{X_{ijg}}$ and $\mathbf{X_j}$, respectively. The constant $\boldsymbol{\beta_0}$ and the vectors of coefficients $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ are assumed to be normally distributed with a large variance. Finally, the model predicts the number of child labourers ($\mathbf{y.pred_{ijg}}$) using the key parameter $\boldsymbol{\mu_{ijg}}$. Models 1, 2, and 3 are run on different sets of variables: Model 1 using only household-level variables and Model 2 adding two state-level variables (benevolence and support for women's work). Models 1 and 2 can show how two seemingly related variables – seclusion norm and supportive norm for women's work – could work on child labour without interrupting each other. Model 3 further included the interaction variables between norms and gender, helping not only to describe the effect of norms on gender but also predict the number of child labourers based on their gender. Thus, a Poisson distribution sufficiently describes the variability of the data. The model is as follows:

$$y_{ijg} \sim \text{Poisson}(\mu_{ijg}), \quad (6)$$

$$\log(\mu_{ijg}) = \beta_0 + \beta_1 * X_{ijg} + \beta_2 * X_j + \log(n_{ijg}), \quad (7)$$

$$y.pred_{ijg} \sim \text{Poisson}(\mu_{ijg}). \quad (8)$$

Household-level aggregates ($\mathbf{X_{ijg}}$) include diverse socioeconomic variables: proportion of households in the lowest asset index category, size of household-owned land (hectare), proportion of urban residence, proportion of the scheduled caste (SC) or scheduled tribe (ST) populations, and proportion of women wearing *ghungat/burkha/purdah/pallu*. All of those variables are weighted[51] means for each group, ***ijg***, taken from the IHDS 2011/12. State-level aggregates ($\mathbf{X_j}$) are the degree of the benevolence norm and norm for women's work taken from the WVS (2012). An asset index and the norm for women's work is a latent variable constructed through confirmatory factor analysis (see Appendix Table C.2). According to the correlation test, the variables do not indicate strong collinearity among them (all $r < 0.6$, see Appendix Table C.3).

6.6. Results

6.6.1. Model checking

A Bayesian posterior predictive p-value of the likelihood ratio statistics showed that the threshold models demonstrated no significant discrepancy between predictions and observations (Lunn et al., 2012; Zhang, 2014; see Appendix Table C.4). As the threshold model did not include any explanatory variables, a Bayesian R-squared test[52] was not applied. For children aged 5 to 11 years, the data were sparse and contained some missing values as well as zero counts that are likely an effect of low exposures (see Table C.1). This did not allow the model parameters to be identified (the MCMC algorithm did not achieve convergence). Therefore, the results for this age group are not presented in the following sections, and this investigation is left for future research.

In the norm models, an out-of-sample test was implemented; for this, one occupational group (each occupational group included 17 cases, i.e. 17 states) was excluded from the total sample. In Appendix Table C.5, each simulation predicted the number of child labourers from the left-out samples. The test results showed that around 40 percent of simulations predicted the number of child labourers with a Bayesian posterior predicted p-value between 0.05–0.95, indicating that the models did not generate extreme predictions. Furthermore, a Bayesian R-squared confirmed that all three models showed high enough Bayesian R-squared values: more than 50 percent of the variation was explained by each model. Model 3 provided a high value for explained variance at around 70–80 percent. In

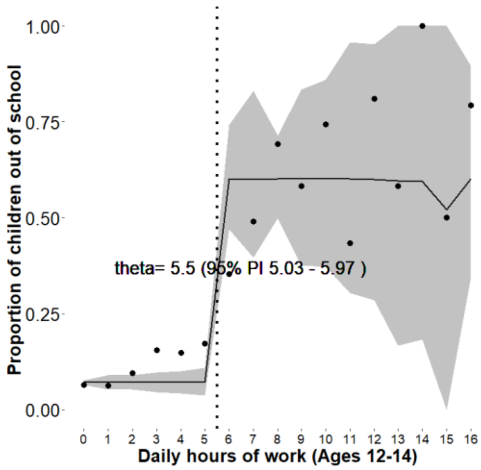
[51] Weights of the IHDS are adjusted for district populations and urban-rural differences according to the 2001 Indian Census. In this paper, relative weights, that is sampling weights divided by the mean of weights, are used.

[52] A Bayesian R-squared is 'the variance of the predicted values divided by the variance of predicted values plus the expected variance of the errors' (Gelman et al., 2019, p.307).

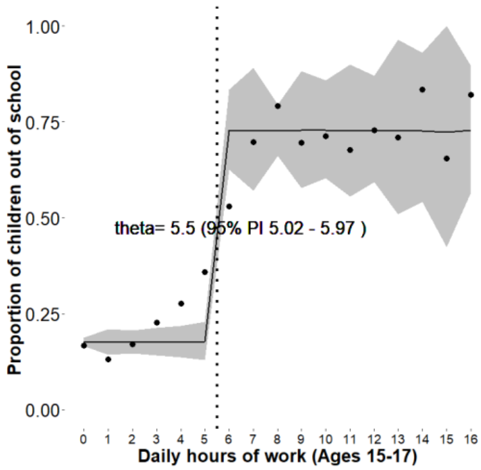
terms of DIC, Model 3, which included state-level norms and interactions, was the best performing model as it had the lowest value of DIC. There is no clear sign of overdispersion, which was checked by residual diagnostics.[53] Overdispersion occurs when a Poisson model assumption is not met. The assumption is that a mean of the parameter equals its variance. The dispersion test showed that observed residuals of the simulations are not significantly different from expected residuals.

6.6.2. Revealing working hours that potentially cause harm to children

The threshold models indicated that out-of-school risk values sharply increase when children work more than 5.5 hours a day (95-percent predictive intervals, PI [5, 6] hours), which was equivalent to 38.5 ([35, 42]) hours per week for children aged 12 to 17. In particular, the estimated proportion of out-of-school children (λ_2) among those aged 12 to 14 increased to 0.6 (95-percent PI [0.58, 0.66]) if they work longer than 5.5 hours (95-percent PI [5.02, 5.98]) a day. Owing to the small number of cases of children working over 12 hours a day, there was a large variance in results. The same figure reached 0.73 ([0.72, 0.76]) for children aged 15 to 17. Both age groups (ages 12–14 and 15–17) showed a similar relationship between working hours and educational harms. Not enrolling in school increases sharply among children aged 12 to 17 if they work longer than the estimated time parameter, 38.5 hours a week, which supported the idea that it can be used as a time threshold without any problems. However, it was difficult to prove a clear change-point in working hours for children under 12 years old.



(a) Predicted proportion of out-of school children (ages 12–14)



(b) Predicted proportion of out-of school children (ages 15–17)

Figure 6.2. Potential educational harm by daily working hours

Source: Estimates using the Indian Human Development Survey (2012)

[53] The 'DHARMa' package in R was used (https://cran.r-project.org/web/packages/DHARMa/index.html, accessed 12 July 2021).

While the ILO (2017) suggest a time threshold of at least 43 hours a week when defining child labour[54], our estimates indicated that it should be lower than that in India. The impact on children of working long hours was not so different for ages 12–14 and 15–17. Among children aged 5 to 11 years, the ILO's suggested threshold of at least one hour of economic activity was suitable (ILO, 2017). Thus, our suggested time thresholds for Indian child labour were as follows: for ages 5–11, at least one hour of economic activity a week, and for ages 12–17, 38.5 hours a week. Further study is required to reveal the diverse harmful aspects relating to the length of children's working hours.

6.6.3. Results of the child labour incidence model with norms

The results of Models 1–3 with child-labour risks regressed on social norms are shown in Table 6.2. Regression results showed that norms strongly affected decisions regarding child labour, and norm variables explained a large part of the variance. A norm supporting women's work reduced the child-labour risk, while female seclusion norms increased it. The coefficient of norms on women's work was negative (β= –0.42, Model 2), implying that child-labour risks recede as more people support women's employment. In particular, a norm supporting women's work showed a strong interaction with girls' participation in the labour force, indicated by the negative interaction effect between them in Model 3 (β= –0.42 in Model 3). This implied that norms supporting women's work are more relevant to the incidence of child labour by girls. On the other hand, female seclusion increased the risks of child labour (β=0.51 in Model 1 and β=0.37 in Model 2). When more women are restricted from outside activities because of tradition or religion, children take more roles in economic activities in their stead. There was no evidence that this was the case with girls more than boys; if women were secluded, there was a high risk that not only girls but also boys would participate in child labour.

Table 6.2. The characteristics of the posterior distribution estimated from the norm models

	Model 1 Log of child labour rate				Model 2 Log of child labour rate				Model 3 Log of child labour rate			
	Coef.	S.D.	2.5%	97.5%	Coef.	S.D.	2.5%	97.5%	Coef.	S.D.	2.5%	97.5%
Constant	**-3.04**	**0.11**	**-3.25**	**-2.83**	**-2.07**	**0.29**	**-2.63**	**-1.51**	**-1.86**	**0.33**	**-2.49**	**-1.22**
Female	**-0.36**	**0.04**	**-0.42**	**-0.29**	**-0.36**	**0.04**	**-0.42**	**-0.29**	**-0.90**	**0.41**	**-1.70**	**-0.09**
Urban	**-1.19**	**0.12**	**-1.43**	**-0.95**	**-1.40**	**0.14**	**-1.67**	**-1.14**	**-1.41**	**0.14**	**-1.67**	**-1.14**
Asset (the lowest asset category)	0.09	0.10	-0.12	0.29	-0.19	0.13	-0.43	0.06	-0.18	0.12	-0.42	0.07
Land size category	**-0.09**	**0.04**	**-0.16**	**-0.01**	**-0.12**	**0.04**	**-0.19**	**-0.05**	**-0.12**	**0.04**	**-0.19**	**-0.05**
Dalit (or SC)	**0.77**	**0.16**	**0.46**	**1.09**	**0.86**	**0.16**	**0.54**	**1.18**	**0.84**	**0.16**	**0.53**	**1.16**
Adivasi (or ST)	**1.34**	**0.17**	**1.00**	**1.67**	**1.37**	**0.17**	**1.03**	**1.70**	**1.37**	**0.17**	**1.03**	**1.70**
(Norm 1) Practicing seclusion	**0.51**	**0.08**	**0.35**	**0.67**	**0.37**	**0.09**	**0.19**	**0.55**	**0.29**	**0.11**	**0.07**	**0.52**
(State-level effects)												
(Norm 2) Doing good for society (*)					**-0.33**	**0.10**	**-0.52**	**-0.14**	**-0.39**	**0.12**	**-0.62**	**-0.17**

[54] Thresholds of 43 hours have been used throughout the ILO estimates, with 43 hours regarded as the mid-point between working hours regulated by national legislations, ranging from 40 to 44 hours (ILO, 2017).

(Norm 3) Support for women's work (*)					**-0.42**	**0.12**	**-0.65**	**-0.20**	-0.26	0.14	-0.54	0.02
(Interactions)												
Seclusion*Female									0.19	0.17	-0.14	0.53
Doing good for society * Female									0.17	0.16	-0.14	0.48
Support for women's work * Female									**-0.42**	**0.21**	**-0.83**	**-0.01**
Bayesian R-squared	**0.81**	**0.02**	**0.77**	**0.85**	**0.83**	**0.02**	**0.79**	**0.86**	**0.83**	**0.02**	**0.79**	**0.86**
DIC	**2,324**				**2,310**				**2,296**			

Sources: (*) World Value Survey India (2012); Indian Human Development Survey (2012) for all other variables
Notes: The number of observations is 442 (13 occupational groups * 17 states * gender); bold letters indicate that 95-percent PIs do not include 0; Urban - the proportion of people living in urban areas; Dalit - the proportion of people having a scheduled caste (SC); Adivasi – the proportion of people with scheduled tribe (ST) status

Meanwhile, the coefficient of the benevolence norm was negatively related to the incidence of child labour (β= –0.33 for Model 2 and β= –0.39 for Model 3). The norm of 'doing good for society' decreased the incidence of child labour among both girls and boys, and there was no clear interaction between females and the benevolence norm. The benevolence norm was relevant, then, as it seemingly controlled child-labour risks. Concerning this result, however, a benevolence norm requires careful appreciation as it might be too diverse to be linked to specific groups or classes.

Alongside those norms, a higher urban population in a group was associated with decreased child-labour risks for both girls and boys. This finding confirmed that child labour is more likely to occur in rural areas. Yet, the effect of the asset index was confounding. The lowest asset-index category was correlated with other variables (see Appendix Table A.4) and did not indicate a strong relationship with the incidence of child labour. Conversely, land size demonstrated an impact on child labour; as households owned more land, the risk of child labour decreased. This finding confirmed that the prevalence of child labour was related to the socioeconomic status of households.

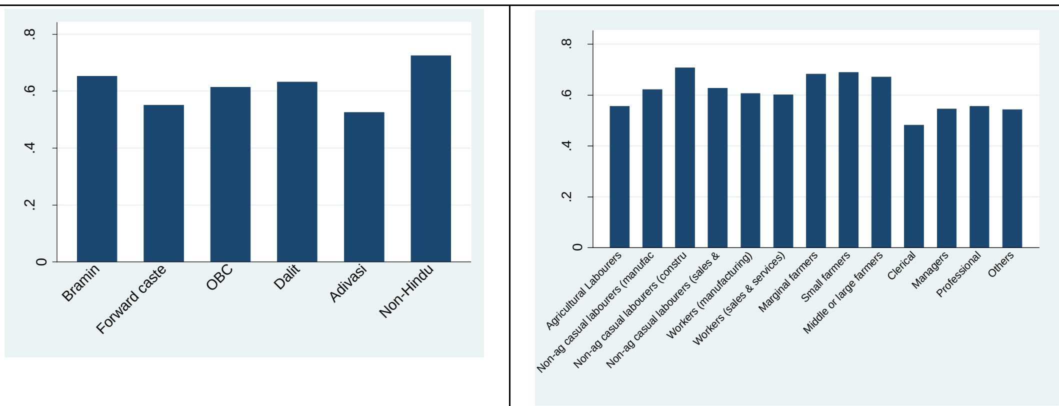
Furthermore, social group membership demonstrated a strong impact on the incidence of child labour. A high proportion of Dalits in a group was related to a higher incidence of child labour. Likewise, larger Adivasi populations within groups increased the incidence of child labour. Particularly, among Adivasi households, girls' labour force participation rate was as high as boys'. These results supported the conclusion that social-group norms also relate to the persistence of child labour.

The predicted proportion of child labour in the 17 states was 0.06 (mean, 95-percent PI [0.057, 0.063]), using the time threshold obtained through the model. This was much higher than the estimate of child labour using the international threshold of working hours, which was 0.04 of the child population in India in the same year (Kim et al., 2020). In addition, boys appeared more likely to participate in the labour force (0.07, 95-percent PI

[0.066, 0.074]) than girls (0.05 [0.046, 0.053]). Children from agricultural and casual labour households were found to be at high risk of being involved in child labour. According to the modelled outcome, in households that were headed by agricultural labourers, the risks of their children becoming child labourers increased sharply. In those households, the child-labour rate among boys reached up to 0.1 ([0.09, 0.11]), the rate among girls, 0.07 ([0.06, 0.08]). Casual construction-labourer households demonstrated child-labour risks as high as 0.09 ([0.08, 0.1]) for boys and 0.07 ([0.06, 0.08]) for girls. Households headed by marginal or small farmers and casual manufacturing labourers also involved a high risk of child labour.

6.6.4. Social groups, class and norms

The previous sub-section established an association between norms and child labour in India. This sub-section elaborates how norms are related to social groups or the class membership of children's households. Figure 6.3 shows the relationship between social groups, class and female seclusion. The practice of seclusion was stronger among Brahmin or Non-Hindu (mostly Muslim) households. Dalit households also demonstrated a comparatively high degree of female seclusion, indicating that Dalit women are subjected to seclusion or segregation in their economic activities. Moreover, seclusion norms were stronger among households headed by farmers or casual labourers in the non-agricultural sector, e.g. construction (Figure 6.3(b)). Thus, the assumption that more upper-class women than lower-class women experience seclusion might be inappropriate. In rural and traditional households, women's roles in household production are significant, but their roles are limited and restrained as they adhere strongly to the norm of seclusion. As a result, children could be led to work to help family economy.

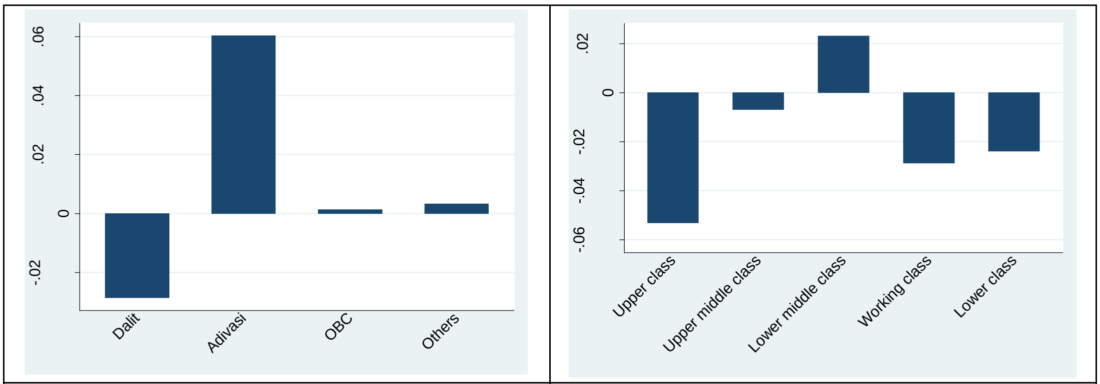


(a) Female seclusion by social group | (b) Female seclusion by hh occupational class

Figure 6.3. Descriptive information of seclusion by social group and class

Source: IHDS 2011/12

Moreover, a norm supporting women's work appeared relatively frequently within the lower-middle class (Figure 6.4(b)). As mentioned, people's attitudes to women's work underpin decision-making regarding child labour. In this regard, less supportive attitudes regarding the work of women from lower-class groups can cause a higher incidence of child labour. Working- or low-class women must participate in work, but, in actuality, their employment is a way to meet needs rather than to be independent. In such cases, women's workforce participation did not necessarily reduce child labour among the lower classes. An egalitarian gender norm was stronger among Adivasi households, which usually had both boys and girls participating in labour (Figure 6.4(a)). Adivasi households demonstrated the highest proportion of child labourers, but compared to other socio-economic motivations, gender norms had a limited impact on reducing child labour among them.



(a) Norms supportive of women's work by social group	(b) Norms supportive of women's work by subjective class

Figure 6.4. Descriptive information of gender norm by social group and class

Source: WVS 2012
Note: WVS 2012 does not have detailed information on household class, but it has a variable of subjective social class.

6.6.5. State-level analysis of child labour

This sub-section explains the structural and institutional characteristics of the key Indian states where the level of child-labour risk was highest. According to the predicted rates of child labour (Model 3), the states of Madhya Pradesh, Odisha, Jharkhand, West Bengal, Uttar Pradesh, and Rajasthan had the highest levels of child-labour risk (Figure 6.5). They are located in the Northern, Central and Eastern regions of India. The norm patterns of each state showed a relevance regarding the incidence of child labour (Figure 6.5 and Figure 6.6). For example, Rajasthan had the most robust seclusion norms and Odisha had the lowest levels of support for women's work. These reasons contributed to the high incidence of child labour in these states. Odisha showed the highest level of child labour by girls, in particular. In contrast, states in the Southern region showed lower child-labour risks. Andhra Pradesh, Karnataka and Kerala showed the lowest levels of child labour. In these states, women were less subject to seclusion, and the norm supporting women's labour participation was stronger than that in other states. Moreover, the benevolence norm was found to be stronger in Uttarakhand and Haryana, which helped reduce child labour in those states.

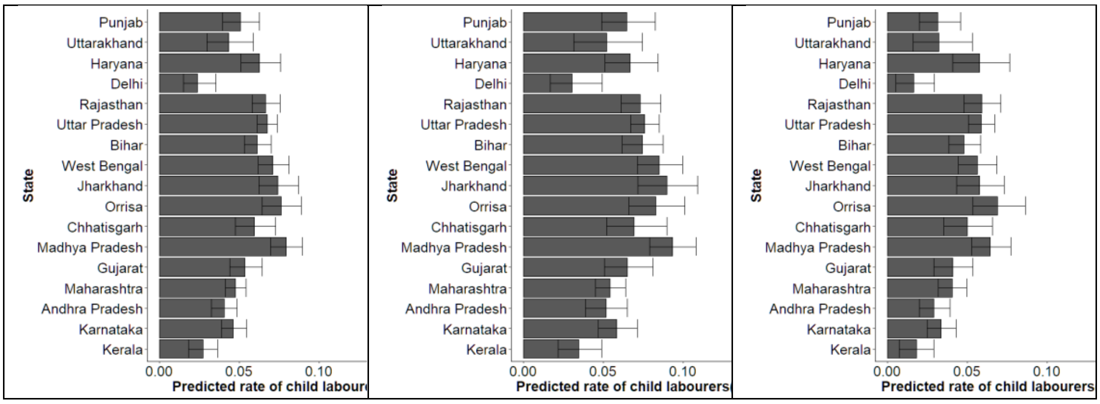


Figure 6.5. Predicted rates of child labour

Sources: IHDS 2011/12, World Value Survey India 2012

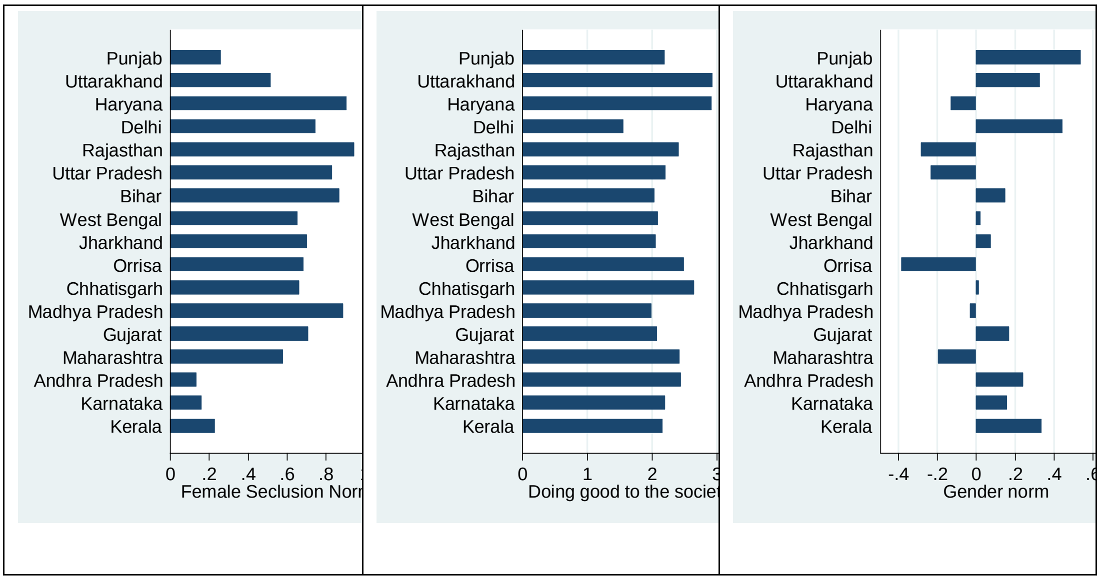


Figure 6.6. Comparisons of norms in the key states

Sources: World Value Survey India 2012; female seclusion norm – IHDS 2011/12

Figure 6.7 indicates the states with high child-labour risk in households in major industrial sectors: agriculture, manufacturing, construction, and sales and service. Madhya Pradesh had the highest level of male and female child labour in the construction and agricultural sectors. More than half of the indigenous population is concentrated in Central India, and Madhya Pradesh has the highest tribal population in India (14.7 percent in 2011, Ministry of Tribal Affairs, 2013), and poverty, unemployment and illiteracy are related to the incidence of child labour among indigenous people. In addition, widespread Purdah practices, as well as little support for women's work, increased the risks of child labour for both boys and girls (Figure 6.6).

Jharkhand and Chhattisgarh showed a large prevalence of child labour among households headed by casual workers in the low-skilled manufacturing industry (see Figure 6.7). Haryana showed a high rate of child labour in the construction industry as it absorbs many migrant workers. Migrant workers bring their families, and so the risk of children becoming labourers increases. Meanwhile, the highest proportion of child labourers occurs in the agricultural households in West Bengal. As Das et al. (2013) explained, many children in West Bengal are involved in agricultural work, such as potato cultivation, owing to their families' low socio-economic status.

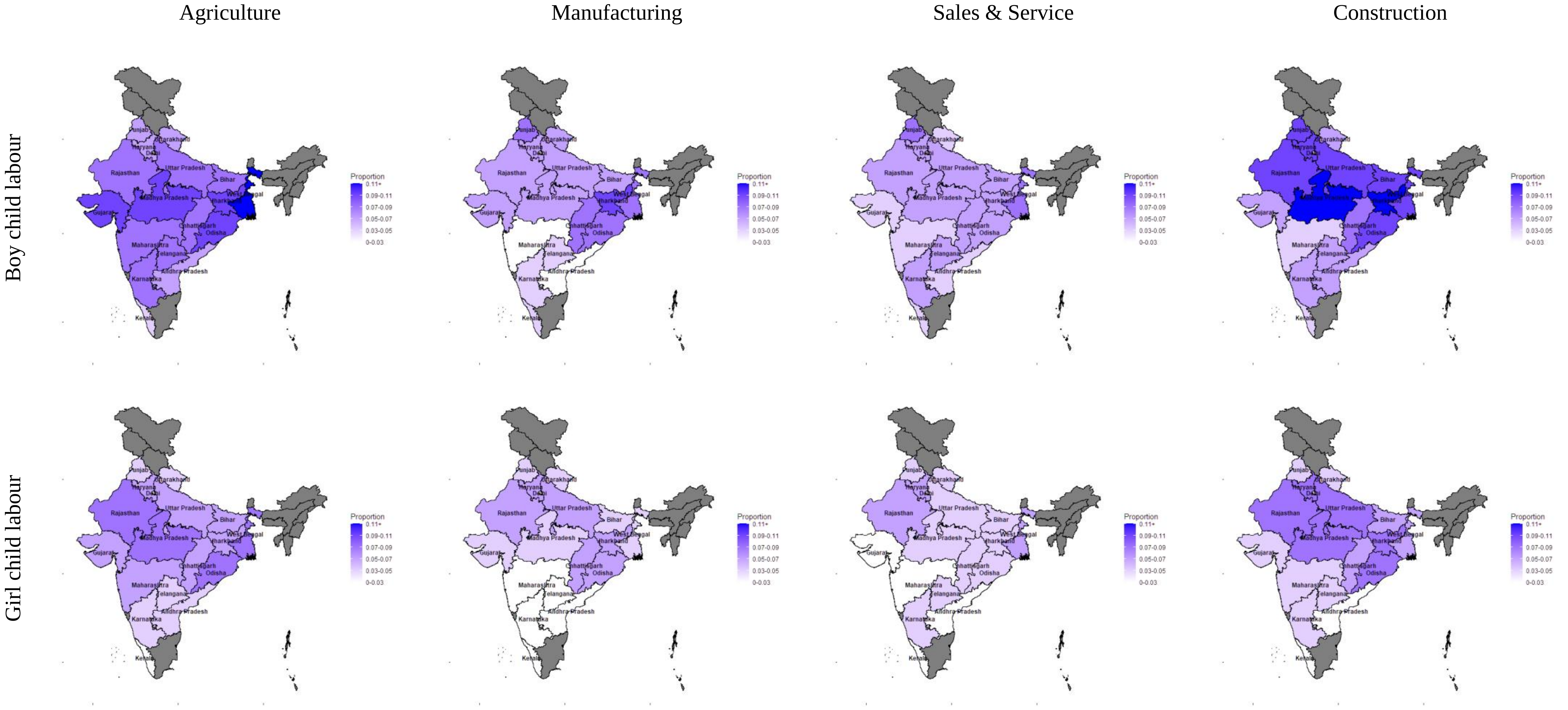


Figure 6.7. Child-labour risk mapped by household industries and gender

Source: IHDS 2011/12 (Norm variables are from WVS 2012.), Map source: http://projects.datameet.org/maps/states/
Notes: Model 3 is used; risks are categorised into six groups by predicted risk rate (0–0.03, 0.03–0.05, 0.05–0.07, 0.07–0.09, 0.09–0.11, 0.11+); grey colour indicates states excluded in the estimations.

6.7. Discussions and Policy Implications

This article has revealed how norms might affect the incidence of child labour. These norms are invisible; however, they form the beliefs and values of agents (children and parents), influence behaviours and construct a '*habitus*' (Bourdieu, 1977). Norms represent key parts of child labour decision-making. As child labour happens in an informal work environment, norms can be social institutions in the informal sphere (Olsen and Morgan, 2010). Agents accept or change norms within the social structure; therefore, consequences are changed by geographical locations and also by class and gender.

The findings of this research suggest that gender equality helps prevent child labour for both boys and girls. Support for women's right to work evokes a sense of consideration for boys' and girls' rights too. This support is also related to girls' expectations of becoming formally employed in the future. Kabeer (2018) points out that gender norms, such as seclusion norms or perceptions about sexual risks, restrict adolescent girls in developing their careers. More equal gender norms contribute to reducing girls' labour so that they might develop their capabilities for future careers. In this regard, women's empowerment, on which most GAD scholars place great emphasis, can also significantly reduce child labour.

This study insists that supportive norms of female workforce participation can make a positive impact on the prevention of child labour, with a stronger impact on outcomes for girls. This norm is different from female work participation itself but could be related more to an egalitarian perception of gender roles and stereotypes. This result is slightly different from previous studies that stated mothers' work participation increases children's labour market participation. On the other hand, the result supports the hypothesis that children's wellbeing is improved if women's work is valued as much as men's. Lower-class women are more likely to participate in the workforce but less likely to consider employment as autonomy. In this situation, child labour arises from a lack of resources and is additionally affected by unequal gender norms.

Conversely, with greater awareness, an agent's decisions could change, i.e. parents try to seek help with labour from others rather than their children. In such cases, the trade-off between women's work and children's wellbeing might not be necessary. For example, Drèze and Sen (1995, p.164) reveal that a high level of female workforce participation improves the chances of a female child's survival because it enhances women's agency. If women have proper jobs, their children might not need to work. Furthermore, regarding gender equality, girls are also educated to have better jobs and therefore child labour among girls can be reduced. In this sense, the employment status of Indian women should be improved in order to achieve the goal of reducing child labour. Women's high levels of participation in informal activities, their reproductive labour and the inaccessibility to them of paid work (Mezzadri, 2016; Olsen and Mehta, 2006b) should be addressed in order to improve children's status.

This study has proved that child labour is a compounded result of both social and gender relations by investigating the embedded norms. Subsequently, there are two suggested methods of reducing child labour, which are girls' education and women's empowerment. Despite the importance of girls' training and education, the social norm is clearly against it in India. The introduction of compulsory education has enabled a sharp rise in the enrolment rate for both boys and girls in India[55]. However, girls' secondary education attendance is significantly lower than that of boys, and there are many other challenges. Das and Khawas (2009) explain that the gender gap remains large in technical and professional education, being affected by cultural factors. The inclusion of girls in secondary or technical education should be a priority. Education policy should also help girls to reach higher education and provide opportunities for better jobs. It can reduce the incidence of child labour among girls and equip them with better skills so that they can increase their participation in the formal sector as adults. Furthermore, cultural barriers to their inclusion should be removed.

Restriction of women's roles and lower participation in formal employment increases child-labour risk for boys and girls. Thus, this paper emphasises that women's empowerment and access to active economic roles are essential to resolve the child-labour problem. Women's empowerment is a process of obtaining control over the self, and the ideology and resources that determine power (Batliwala, 1993). Expanding women's agency, through raising consciousness and improving their social positions, can help reduce child labour. To this end, social acceptance of the types of jobs available to women should be extended and childcare supported.

[55] The Right of Children to Free and Compulsory Education Act, enacted in 2009

Chapter 7. Discussion

7.1. Summary of Findings

Overall, this study has found the causes of child labour in India lie in the interrelation between structures, including class and social groups, institutions (as well as norms), and gender. In this section, I summarise the findings of the study regarding those critical components of child labour before going further to explain their implications.

First of all, this study presents plenty of evidence that child labour is a form of exploitation grounded in the social hierarchy. Child labour reflects an existing social hierarchy, such as class and caste systems, and children's disadvantages within that hierarchy. A lack of resources (land and wealth) and, accordingly, a lack of power among families and communities also affect the risk of child labour. A majority of the child labourers in India are from the poor peasant and landless labour class. Most child labour cases occur in rural areas, especially from households headed by marginal farmers and landless agricultural and non-agricultural labourers. The problem is not limited to rural areas; the urban informal sector is another key path to child labour. Children may work as unskilled service labourers, construction workers, or in unpaid household production. This sector has a lower percentage of child labour than the agricultural sector, but as the informal economy is growing, it could have more impact on child labour in future. Thus, child labour is not solely caused by economic needs or parental morality. In a broad sense, socioeconomic factors, such as a large proportion being in the informal economy and the high social inequality found in India, can cause severe child labour problems. These findings corroborate the localised studies of Phillips (2013), Phillips et al. (2014) and Nieuwenhuys (1994; 2007).

In Chapter 5, I analysed the relationship between household class and children's labour hours. For example, Indian children from landless labourers' households spend the longest hours in work. Households, due to a lack of resources, require their children's help and, consequently, children might spend long hours at work in or near their households, and lose schooling time. I found also that there is a positive effect of land size on children's working hours. Farmers' children have considerable workloads, as their households own a certain amount of land and do not have enough wealth to hire people to work it. The impact of land attenuates as the land size increases. More evidence is found in terms of asset status, which shows not only financial status but also living standards, demonstrating that the index has a negative relationship with the working hours of child labourers. These findings are consistent with the class analyses of Athreya et al. (1987; 1990) and Bardhan (1982), which give a rounded measurement of the size of effects. Without regression modelling, effects of social class, land and wealth could be exaggerated in bivariate tables. In this thesis, I was controlling for other simultaneous factors.

Chapter 6 explores predicting child labour risks by household occupational group, again at the all-India level. The results showed that agricultural labourers, non-agricultural casual labourers (construction, sales and services), and marginal farmers face higher child-labour risk levels. Among those groups, children – in particular those who are in households headed by farmers or family business owners – are involved in the same sectors as their parents, but some (male) children work on their own. To generate results, I used sophisticated methods of data wrangling to get within – and between – household contrasts. As we can then predict the risk level by household occupation/industry, it might be possible to identify geographical locations that should be prioritised in terms of child-labour policy. This kind of result could be useful at micro levels: e.g., choosing blocks within districts or choosing a region within a state for targeted attention. For example, where a large number of agricultural labourers or migrant workers are located, we can expect to observe a high risk of child labour. Identifying more specific industrial sectors, such as construction and home-based manufacturing, should also be included. In recent years, the large-scale annual Periodic Labour Force Surveys, with their quarterly revisit sampling methods, could offer the possibility of identifying new hotspots for child-labour risk (NSSO, 2020a). The release of time-use data (NSSO, 2020b) and the NFHS Wave 5 (International Institute for Population Sciences, 2020) in the years 2019–2020 also offers new opportunities, because these are representative random sample surveys with details of personal migration status, siblings' jobs, unpaid work activities, parental health, and other relevant factors.

Together, these structural characteristics of child labour show that child labour can best be addressed by challenging structural inequality. Unequal distribution of power in a socioeconomic structure leads to different outcomes in children's lives as lower-class households are deprived of economic, social and political opportunities. In terms of social welfare and policy-making in the development process, their voices hardly reach the mainstream. Consequently, their children benefit less from economic growth as well as being deprived of chances for upward mobility.

Furthermore, in India, social hierarchy is a combination of class and social groups. Although the association of caste and occupation remains in a weaker form in wider Indian society (Jodhka, 2012), child labour remains strongly related to caste hierarchy. Chapter 5 revealed the class and caste connections influencing child labour. Caste and class are not entirely distinct, but they are never the same; caste status is inherited, while class status can change. Adivasi boys and girls are the most vulnerable groups in terms of the incidence of child labour while most other social groups do not allow children's, and especially girls', involvement in labour. The best knowledge of caste relations shows that caste norms strongly limit girls' participation in the labour force.

Owing to the gender and development approach (GAD), this study can confirm that gender inequality is involved in the structure and therefore affects the incidence of child labour. Even within a similar class or social group, children might experience diverse discriminations because of their gender. Among boys, discrimination could cause them to lose their formal primary-school education. Among girls, besides this, discriminatory stereotypes also restrict their access to skilled tasks in their later lives. Gender relations include all the relationships between boys and girls or female and male adults, and also between children and parents. Those gender relations reflect different relations with power, access to resources, and decision-making. Consequently, children have different difficulties based on their age and gender. Gender inequality among child labourers is not shown clearly, especially when we use survey data. This study has the advantage of large scale but the disadvantage of not having much interview data. The disadvantage was offset by my scoping study, early on, and wide reading of ethnographies about Indian children. Gendered child labour is also often hidden within institutions (such as forms of household labour or reproduction), but I have constructed the theoretical framework in such as a way that key issues are revealed rather than being hidden. This framework will be particularly useful for analysing India's new nationwide time-use data (NSSO, 2020b).

Chapter 4 suggested the large numbers of girl children in domestic work; boys are more often employed in paid work, while girls tend to work in the domestic sphere. We have revealed that the overall number in child labour reached almost 13.2 million (aged 5–17) in India in 2011/12 when household work was included in the definition of child labour.

Chapter 5 investigated gender roles in commercial labour. In the labour market, the number of male child labourers is greater than that of female child labourers. This is well understood within the caste system, in which girls should stay inside the home. This rule results in a much smaller number of girls becoming involved in labour outside the home. However, among Adivasi children, the incidence of girl child labour is as high as that of boy child labourers, which supports the view that gender norms influence decisions regarding child labour. In middle-asset households, girls could work slightly longer hours than boys as households are in need of additional hands while more investment in sons is preferred. Therefore, child labour is a phenomenon entangled with parental son-preference and economic needs among certain classes.

In Chapter 6, the gender norms behind child labour decisions were revealed. Employment of female adults affects the incidence of child labour, and this study supports the idea that if women are not allowed to be active economic agents, more boys and girls will be expected to become labourers. Women's limited status in the family and in society can result in their children becoming involved in labour. In contrast, norms supporting women's employment in households, in accordance with valuing gender equality, can reduce the risks of child labour. These trends are opposite to previous studies suggesting that as more women are employed, child labour increases. Women's empowered position can reduce the risk of child labour for both boys and girls.

Lastly, structural child labour is practised through institutionalised and cultural properties of the notion of 'children' held in those communities. These notions are affected by time and location; they are not fixed. As many scholars have pointed out, cultural reasons for child labour are considered less often (see Lieten, 2001; Weiner, 1991). This study shows that caste is a robust social norm as well as a social structure in India. Social group norms and values affect whether or not a family and community accept child labour. Moreover, norms and values that shared by socioeconomic groups of households shape childhood experiences. Although revealing the norms behind household class is tricky, it is found that strong seclusion norms among the peasant classes and middle-class son preferences are risk factors for child labour. Such examples show that gender norms should be integrated into the discussion of class and social groups in India. Children's behaviour in each risk group should be more carefully studied in terms of attitudes and values. This is because social and gender inequalities are exercised through biased norms, rules and procedures (Kabeer, 1994, p.225). At the same time, we should also avoid generalisations about children's lives, even if they belong to a similar group.

In Chapter 6, values and norms regarding child labour in the key states of India are explored in depth. Purdah and seclusion customs have remained in India and the results of the models reveal that seclusion norms play a significant role in child labour risks. The seclusion norm is stronger among farming households, and therefore it is found to have an increased effect on child labour risks. In addition, the positive effect of benevolence norms is explored to explain the relationship between theories of child labour: that is, gender and development versus moral economy. For example, Uttarakhand places a rich value on 'doing good to others', which has a marginal effect on reducing child labour.

Gender and social norms regarding child labour are affected by the socioeconomic conditions of the location. State-wise, trends confirm that a variety of combinations of norms exist between the states. In some, social norms might be a stronger factor than economic status in the incidence of child labour. For instance, in Chapter 6, although Gujarat is one of the wealthiest states, the high risk of child labour there is associated with less support for women’s work and high female seclusion. Conversely, the state of Chhattisgarh has high risks of child labour mainly due to low economic resources. Many women participate in work in Chhattisgarh due to economic need, and women’s participation in the labour force does not lower the risks of child labour amidst prevailing poverty.

7.2. Implications of the Gender and Development Approach

The findings of this study lead to the conclusion that the gender and development approach (GAD) is a suitable one for analysing the causes of child labour. While Marxist feminism stresses class relations, the GAD approach emphasises broader gender and social relations, and the practices of relations through institutions. Furthermore, GAD considers both males and females equally, effectively explaining gendered patterns of child labour and the close relationship between gender norms and the incidence of child labour.

However, in the GAD approach, children’s problems are not mainstream. While children are treated as objects of care, their subjectivity as individuals tends to be disregarded. There has been a clear gap between GAD and the study of child labour. Child labour is considered harmful for any children; therefore, gender inequality among child labourers is a less significant concern. The GAD approach focuses less on children’s matters, even while it emphasises the inequality in adult male and female relations. Furthermore, in general, household chores, which form an integral part of child labour and many girls are involved in, are considered not harmful to children’s development. It should be remembered that there is no consensus yet about the definition or allowable number of working hours for domestic chores.

Nevertheless, the GAD approach is advantageous for the study of child labour. The GAD approach pursues equality of human beings and works to overcome the dichotomy between men and women; thus, it has contributed to integrating the gender issue with human development (Benería, 2003, pp.161–168). Children are independent human beings and entitled to equal human rights and development. Thus, GAD is by no means limited to adults’ relations but also includes children’s. Moreover, children’s and adults’ relations are mutually influential. In particular, parental and child labour statuses are closely linked. Finally, the GAD approach avoids generalising the problem and takes into account diverse possibilities of relations. This child issue is based on huge variations that depend on location and time; therefore, it is better understood by taking in the complexity of the relations.

In the following subsections, I elaborate on how the GAD approach can develop studies of child labour within socioeconomic and cultural dimensions, mostly drawing on observations through this study. First, we should look at trends in male and female child labour by including all types of work and status into the category of child labour. Second, the impact of gender inequality in development and the consequent results for children should be considered together. Avoiding generalisation of child labour but understanding its institutionalised forms in society is crucial in addressing the issue of child labour. Lastly, I look at how empowerment can be implied in the context of child labour.

7.2.1. Measuring the prevalence of boys' and girls' labour-force participation

Measuring child labour by children's sex is an essential step to approach the gendered problem of child labour. In this subsection, I add the results of the further estimations of the numbers of male and female child labourers using the all-India data, to have clear implications. The projections are made based on the same methods measuring child labour that were applied in Chapter 4 and in Chapters 5 and 6, but this time for boys and girls, respectively. The first measurement uses the ILO time threshold for economic activities and the UNICEF time threshold for household chores in which a data-combining method using two datasets – the IHDS 2011/12 and NSS 2011/12 – is applied. Children's principal activity status is screened out, so that it is a more conservative way to calculate child labour. The second method uses its own estimated time thresholds, which are 1 hour for ages 5–11 and 38.5 hours for ages 12–17 for any work done for more than 30 days a year, based on the IHDS 2011/12. There is no information on household chores to be applied in this second method and it concerns market labour only. A summary of the results is given in Table 7.1.

Table 7.1. Estimated number of male and female child labourers in all India in 2011/12

Category	Ages	Boys		Girls	
		Mean (2.5%, 97.5%)	% of the same group pop.	Mean (2.5%, 97.5%)	% of the same group pop.
(1) Market and domestic labour, ILO & UNICEF time threshold, NSS+IHDS, principal activity status=workers	5–17	**7,096,388** (6,211,544–8,077,236)	**4.1** (3.6–4.6)	**5,696,321** (4,777,273–6,739,305)	**3.6** (3.0– 4.3)
	5–14	**1,359,667** (1,148,336–1,595,365)	**1.0** (0.8–1.2)	**1,816,879** (1,442,747–2,300,761)	**1.5** (1.2– 1.9)
(2) Market labour only, own time threshold, IHDS, total days of work > 30	5–17	**11,867,385** (11,362,377–12,384,028)	**6.8** (6.5–7.1)	**7,171,674** (6,794,097–7,557,581)	**4.6** (4.3– 4.8)
	5–14	**4,790,606** (4,479,157–5,108,510)	**3.5** (3.3–3.8)	**3,604,309** (3,336,609–3,879,773)	**2.9** (2.7– 3.1)

Source: NSS 2011/12, IHDS 2011/12; the method is available at Kim et al., 2020.
Notes: Category (1) indicates the number of labourers who work in economic activities or domestic chores (the definition of child labour used in Chapter 4); Category (2) uses its own time threshold (1 hour for ages 5–11, 38.5 hours for ages 12–17) as well as the ILO definition of hazardous occupations

and industries but without applying for principal activity status (the definition of child labour used in Chapters 5 and 6).

When using the international time threshold (Category 1), male child labourers in India in 2011/12 are estimated to number 6.2–8.1 million for ages 5 to 17, and female child labourers 4.7–6.7 million for ages 5 to 17. Noticeably, among children aged 5 to 14 years, the number of girls in the labour force is greater than the number of boys (1.8 million for girls and 1.4 million for boys) because the UNICEF 2019 definition includes domestic chores for ages 5–14. The results are consistent with other countries' examples showing that including household works in the definition of child labour significantly increases the rate of work among female children and young people (Knaul, 2001). Girls take on more domestic duties, which are composed of taking care of siblings, fetching water, firewood collection, making fuel, sewing, maintaining the kitchen garden, and looking after poultry so that other family members can participate in paid productive work.

The latter method (Category 2) does not limit children's principal status when counting child labourers; therefore, the estimated number is much higher than that of the previous definition. The increase is much higher for boys, implying that many boys do schooling and work at the same time. In this definition, the number of boy child labourers aged 5–17 years reaches from 11.3 to 12.4 million, while girl child labourers of the same age group number 6.8 to 7.6 million. Inclusion of all children's statuses obviously extends the boundary of child labour, implying that children hold two or more positions simultaneously. However, even after the inclusion of all statuses, the number of girl child labourers sees a smaller increase. Under-recognition of girls' participation in the labour force in India is not only because their usual reported status is as non-workers, i.e. students, but also because work hours in the household are excluded.

Both measurements confirm that girl child labour is not well-recognised in production activities; hence, more sophisticated definitions and measurement processes are required to measure girl child labour properly. The GAD approach is interested in extending the boundaries of work that men and women are involved in, which apparently helps more child labourers to be recognised. It is necessary to understand that girls' work is not light, even though it happens to be in care work or household production. Niewenhuys (1994) notes that the division between the productive and reproductive sectors is vague in India. In fact, some unpaid household labour occurs in the area of production. However, there is a tendency to perceive unpaid household services as less severe or less harmful. As a result, girls' chances of being recognised as child labourers become much lower because of the work types involved, but the inclusion of unpaid household work into the category of child labour increases recognition of girl child labourers.

7.2.2. Development, gender and child labour

During the development process, economic growth does not benefit children equally. Development has deepened social inequality, and child labour is evidently a by-product of it. Contrary to common belief, economic development, such as GDP growth, has been less influential in reducing child labour; the high-income states have a higher number of child labourers (Kambhampati and Rajan, 2006; Olsen and Watson, 2011). Economic growth might instead increase children's burden (Kambhampati and Rajan, 2006). Bangladesh has seen massive economic growth, but child labour is still persistent among low-income families (Kabeer, 2001). The Indian garment sector is a typical example of social inequality that places children at the bottom of global production networks (Phillips et al., 2014).

In the meantime, the oppression of child labourers is concealed in the labour market. Child labourers are somewhat vaguely identified. The current industrial structure of India explains the potential exploitation of young labourers; the most vivid example is India's growing informal economy. This study provides evidence that children are mainly limited to the informal sector and suffer from long working hours. Child labourers, as informal labourers, have a weak bargaining position, which worsens their working conditions. I emphasise that accepting and acknowledging the realities faced by child labourers, including the existence of the worst forms of labour, is necessary to bring the child-labour problem to the mainstream.

Furthermore, gender inequality exists in the social structure and also among child labourers. This problem is concealed further by the insufficient interest it attracts. In India, boys help parents work when they are young, but as they grow up, they leave home to earn money and to find independent jobs. Wage-based child labour mainly occurs among older boys who work in diverse economic activities, some of them in hazardous industries such as construction and mining. Conversely, girl child labourers are concentrated in the agricultural sector, where the percentage of girl child labourers becomes almost the same as the percentage of boy child labourers. The feminisation of the agricultural sector evidently happens among children. As a result, girls might be gradually withdrawn from formal, skilled employment as they become grown up. Many girls have dual jobs (agricultural labour and farming) and are simultaneously involved in household chores. These heavy workloads may increase the likelihood of girls not enrolling in school. In particular, gender norms and division of labour contribute to girl child labourers' isolated positions as agricultural labours, farmers, or domestic workers. This finding is slightly different from the view that exploitation of women is a parallel to child exploitation (which Marxist feminists insist). Rather, adult gender relations affect child labour, but it does not replace child relations as children themselves form different relations within and outside families.

Furthermore, gender inequality based on adult male-female relations offers diverse implications for the incidence of child labour in India. Women’s limited economic position can be problematic for children, as children usually work with mothers, especially when they are young. In addition, in India, women have less access to productive work, which increases the risks of child labour among both boys and girls (Chapter 6). A few observations are particularly relevant in India: first, the labour-force participation rate of Indian women is extremely low and even shows a declining trend; second, women are highly concentrated in the informal and unorganised sectors; and third, women who are withdrawn from the labour market are moved into the realm of reproduction (Rao, 2018; Bhandari and Dubey, 2019). The mismatch between education and labour market and the lack of support for caring activities have deepened the problem. These situations are factors that continue to threaten children’s wellbeing in India.

In summary, child labour is a form of discrimination that is rooted in social structures based on class and gender hierarchy. Children’s contributions to development, as well as the unequal distribution of its benefits, need better recognition. Structural gender inequality deepens the gender gap and as well as increasing child labour, which is a strong risk factor for children. The key question is now how the voices of children who are subject to labour exploitation can be heard better during decision-making, and how society can share the benefits of development more equally with its children.

7.2.3. Institutional and cultural influences on child labour

Child labour happens through diverse institutions (households, firms, and markets). Among them, households are the most direct and straightforward institutions in terms of child labour decision-making. Parents’ conceptions of children’s work significantly impact on their children’s status. Parents decide the type and intensity of duties and responsibilities for boys and girls, according to internalised norms and expectations in their mindsets. Furthermore, parents and families are affected by kinship pressure and other stakeholders in communities. Accordingly, this study is in line with Nieuwenhuys’ (1994) study that showed that child labour is a form of exploitation occurring within the household and kinship relations, especially for boys in early ages and girls at any ages. While regulations are imposed, child labour is hidden inside households and cannot be reached properly. Meanwhile, changing norms can be more effective in this household setting, as norms are given to individuals by religions and family traditions. Girls’ education, for example, is not only decided by households but also by the availability of institutional support and awareness at community level. As such, how strictly child labour is controlled is largely dependent on institutions such as rules and norms and the given conditions at household level.

Norms and culture play a pivotal role in explaining gendered patterns of labour. Children encounter a diverse range of working lives as different economic roles are given to children according to their gender and social group. Empirical studies provide evidence that there is an interrelation between norms and gender regarding child labour. Kambhampati and Rajan (2008) find that patriarchal social norms are deeply connected to the probability of girls participating in work. The Indian garment industry, for example, shows the interlinkage of caste and gender, through which girls are only allowed to participate in labour at home (Kara, 2019). On top of that, this study provides evidence that social group membership and their gender norms affect girls' participation in the labour force, i.e., high levels of labour-market participation among tribal girls.

To be specific, this study has found a few norms that are relevant to child labour incidences or work intensity, which should be further investigated in order to address the problem of child labour in India. The norms that directly involve children are as follows: i) girls should stay at home rather than participate in public activities (girls' seclusion); ii) girls and boys can equally take part in economic activities (egalitarian gender norm); or iii) daughters, not sons, should help with household chores and production on the farm (son preference). Some norms relate to what is expected by the adults who decide which work their children will do: iv) when mothers are secluded, both boys and girls are eligible to help parents (seclusion of adult women); v) women's economic participation is important for women's autonomy (norm supportive of women's work); and vi) keeping children from harm is a way to be benevolent (benevolence norm).

The various combinations of norms and socioeconomic conditions produce a different child labour outcome in each state in India. Some states, in particular, show a strong association between some cultural features and children's economic participation: Madhya Pradesh has the largest tribal population and among them both boys' and girls' participation in the labour force is largely accepted; Rajasthan and Uttar Pradesh each show high risks of child labour, for both boys and girls, due to strong female seclusion. For another example, Gujarat is one of the states showing strong son preference (i.e., imbalanced sex ratio) together with high risks of child labour. Women in the state tend to remain in the agricultural sector, as agricultural labourers or cultivators, which is associated with low wages (Pattnaik et al., 2018). Despite the state's high economic growth, a considerable number of girls still appear in the farming sector. This affirms that girl children's intensive workloads are more likely a cultural phenomenon based on gender.

Recalling Figure 2.1, which originated in James and James (2004), to display different concepts of childhood, a complete separation of socially structural and socially constructional children might not be suitable. Relations and hierarchy operate through norms, values, traditions and gender. The reality of child labour that is rooted in the social structure is combined with the roles played by institutions and culture. In this sense, it is difficult to say that structure is universal or deterministic. The more dynamic and fluid relationships that Indian children experience are shown in Figure 7.1 below. Norms, institutions, and structure are closely connected to each other regarding the incidence of child labour, and child labour is directly influenced by the norms and values of agents in a given social structure. The social structure is built on social and gender relations within which the children and their households are located. Thus, any isolated solution that focuses on one aspect is not suitable for achieving the goal of ending child labour. It is essential to look at every relationship that surrounds children and the normative patterns governing individuals' decisions in any society.

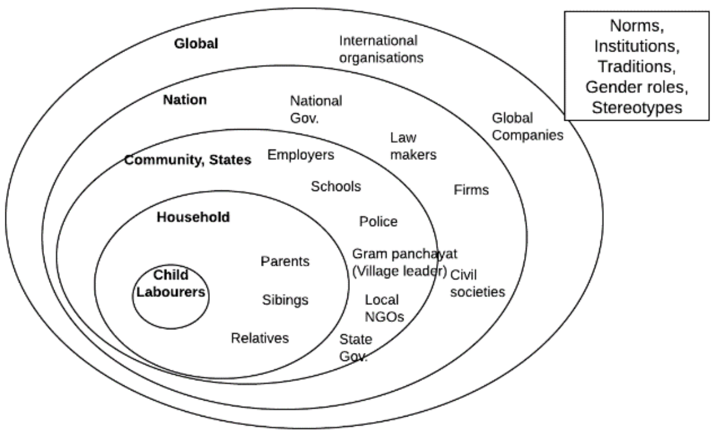


Figure 7.1. Relations, institutions, and norms regarding child labour in India

7.2.4. Empowering children and households

Empowerment is important in addressing child labour as much as in solving women's problems. In the GAD approach, empowerment is the way to recover 'the power within' (improved ability to control resources and decision-making) and reveal social rules, norms, and practices based on unequal gender ideology (Kabeer, 1994, pp.227–229). Moser (1993, pp.74–75) explains that the empowerment approach originated in the feminist movement of Third World women, emphasising women's roles in challenging oppressive social structures, which is based on ‘race, class, gender, colonial history and current position in the international economic order’. Moser (1993) considers the concept of empowerment as a key criterion in the whole process of planning and management. Kabeer (1994, pp.223–224) elaborates that empowerment is about power, and programmes can be successful when they place women as ‘competent but socially constrained actors who are capable of making choices, articulating priorities and taking responsibilities.’ In her later study (Kabeer, 1999), empowerment is defined as a process of obtaining resources, agency (decision-making), and achievement (well-being). Afshar (1998) conceptualises empowerment as a ‘process, and something which cannot be done to/for women, but which has to be their own, [raising] serious questions for their development agencies’.

Empowerment is an important step in compensating for the gap in power of agents who make decisions on child labour. It is not only relevant to children, but to parents, families and communities as well. As it was pointed out in the thesis, lack of resources – in association with norms and values – is the main cause of child labour in India. Although empowerment does not always mean the accumulation of material resources, it can help recover agency and take back control in decision-making. Empowerment relates to norms that control people’s behaviour; therefore, it makes people conscious of their choices and actions regarding child labour. Furthermore, as empowered individuals make more voices heard, they can affect policies and the distribution of resources toward addressing child labour problems. The grassroot movements against child labour represent a powerful solution that addresses the child labour problem through empowering and organising people. In future, any child-labour programmes and interventions should consider empowering children and their households and reiterating the notion of gender equality. The strong relationship between gender equality norms and the incidence of child labour is confirmed in Chapter 6. In light of these findings, empowering women is not only useful for women, but also in addressing child labour.

7.3. A Bayesian Approach to Reduce the Uncertainty of Child Labour

A Bayesian hierarchical model with socioeconomic and cultural information helps to explain the causes of child labour as well as to predict its risk level. This study suggested an alternative statistical method to incorporate the uncertainty encountered in measuring child labour: a Bayesian hierarchical model with a data-combining approach. In this study, the advantages of a Bayesian approach include the following: developing a data-combining model using multiple datasets; applying multi-level social and economic variables to allow appropriate prediction of child labour; and predicting child labour risks and harmful hours of work based on rare data. In this sub-section, I point out a few methodological issues that have been found regarding the measurement of child labour and suggest how a Bayesian approach could deal with them effectively.

7.3.1. Observed problems of data and data analysis in the child labour study

In the estimation of child labour, there are a few clear obstacles related to data collection and data analysis. First of all, this research has found that underreporting of child labour is significant as, since child labour is banned, parents might underreport child labour. The possibility of underreporting is a critical concern when estimating an accurate number of child labourers. In this thesis, 'nowhere children' are not yet considered child labourers because of the lack of evidence to prove that they are labourers. However, it seems highly possible that they are actually involved in labour, even though it is unclear how far underreporting is involved in the estimation. Janzen (2018) found that the self-reported number of those involved in child labour is about two times more than the data reported by the heads of the household in Tanzania. Underreporting might be more significant in urban areas in India (Chamarbagwala, 2008).

Moreover, the use of 'principal activity status' as the usual status for children can cause of underestimation of child labour. It is a common method in defining child labour, but it potentially excludes some significant types of child labour if children record their employment as a secondary status. Moreover, it may omit some students who attend school but also work and whose work qualifies as child labour under another heading. In data combination, applying the same criteria to two different datasets generates many complications. Thus, in Chapter 4, for a rigorous match of datasets, this study used 'principal activity status' of children as one of the conditions to classify child labour among samples of two datasets. However, in such cases, children with 'other' or 'too young' statuses might be missed out from the category of child labour (which I have already mentioned in Chapter 4). To deal with this issue, in Chapters 5 and 6, I have no longer applied principal activity status as a category of child labour but considered 30 days or more work as the usual status of labour-force participation.

Lastly, there is significant inconsistency regarding a 'reference period' of time spent at work. The NSS 2011/12 used 'current weekly status' as a reference period, and the IHDS 2011/12 asked for the time use of a 'usual day'. In both cases, the definitions of 'current' or 'usual' are uncertain. In terms of the length of recall period, the NSS used a one-week recall method while the IHDS used a one-day recall method. One-day or one-week recall methods have different advantages and disadvantages. However, if using the international definition of child labour, the one-week method is more suitable. For example, the UNICEF MICS survey uses one-week recall during the last week of the survey as a reference period.

7.3.2. Implications of using a Bayesian inference style

In this section, I explain how a Bayesian approach helps to overcome the limitations of past work. Although addressing all uncertainties is well beyond what statistics can achieve, a Bayesian analysis method can benefit in closing the gaps between data and information in several ways, such as through i) data-combination, ii) use of priors and additional parameters, iii) incorporating uncertainties into the part of predictions, and iv) dealing with small probabilities of occurrence.

First, by applying a data combination method, we can obtain a more reliable figure than the ones estimated using a single dataset. If multiple datasets have partial information about the same population, we can set a common unknown parameter based on all datasets and link them into one model. The details of how exchangeability and a hierarchical structure support the use of the data-combining method were fully explained in Chapter 3 (Section 3.2.2). As mentioned, in the study of child labour, the lack of relevant information is a critical issue – a Bayesian analysis provides one way to overcome the problem by combining different datasets. On the one hand, it is important to match datasets carefully before combining them. Some differences between datasets, such as different definitions of a reference period and distinct categories of activity status, can be addressed in the process of data cleaning and transformation. On the other hand, there is a discrepancy between datasets that arise from over- or underreporting or possible sampling errors, which are difficult to remove or control by dealing with datasets only. The use of reasonable over- or undercounting parameters in a combined data model greatly contributes to solving problems related to the uncertainty of one dataset. In Chapter 4, a possible underreporting of attending to domestic chores was estimated using an undercount parameter. In the model, it is treated as a prior that is based on our knowledge about underreporting in some significant areas of work. This adjustment parameter (of data A) works if there is a piece of reference information (from data B), and also base knowledge (prior). Thus, the use of under- or overcounting parameters relies on a data combination and priors. A Bayesian data-combining approach utilising a suitable parameter can extend the use of multiple datasets and give potentially less biased estimates than those obtained from a single dataset.

Second, a Bayesian approach provides a consistent way to incorporate priors into the posterior distribution. The use of prior knowledge is the most important feature that differentiates a Bayesian approach from a frequentist approach. For example, if a noninformative prior (a flat prior) is used in a model, the posterior distribution of the model parameters would not be much different from those of a frequentist approach. However, even a lack of information should be included in statistical analysis (Kaplan, 2014). In this study, weakly informative priors are used, and so the impact of any priors is limited. In this way, insufficient knowledge of child labour is reflected throughout the models. It should be noted that with limited information from observed data, priors make a more substantial difference. Thus, a careful model selection and a test of priors are necessary for Bayesian inference. As more knowledge is obtained, more accurately informative priors can be used to improve accuracy in predicting the risks of child labour. Priors can be modified by knowledge accumulated through data and estimations. Moreover, the use of priors based on expert knowledge and opinion is highly feasible for future research. In many other studies, expert priors are applied. For example, Wiśniowski et al. (2019) use a Bayesian hierarchical model to forecast the outcome of referendum voting, in which priors elicited from experts are used for the prediction. In his study, expert-based priors indicate the probability of the voting outcome, provided by the expert views obtained from the Delphi questionnaire (ibid.).

It is important to note that a multilevel model is the most common structure that combines likelihood and priors in a Bayesian approach. In a Bayesian model, a group-level distribution is treated as another prior information (Gelman and Hill, 2007, p.346). Group-level models are exchangeable with the common-prior (a hyper-prior; ibid.). Thus, again, the Bayesian exchangeability provides a grounding for a hierarchical structure for the model. When it comes to a frequentist approach, a multilevel model is a special model that includes a group-level variance, but a Bayesian hierarchical model is a more direct and natural way to contain knowledge about a group within a model. For example, in Chapter 5, a Bayesian hurdle model was used with individual-level predictors and state-level intercept. In Chapter 6, a model includes two state-level predictors (norm variables) as well. Using a group-level prior significantly contributes to reducing uncertainty in data and models.

Third, a Bayesian approach provides a probabilistic way of quantifying any uncertainties. In a frequentist approach, probability requires an assumption of a long-run frequency (Kaplan, 2014); in a Bayesian approach, such an assumption is unnecessary. Probability means uncertainty that is related to a lack of prior knowledge and has arisen from data and incompleteness of the models. There is a strong reliance on the p-value to indicate the significance of estimates in a frequentist approach. In contrast, a Bayesian approach does not use a significant test but provides probabilities for hypotheses, i.e. a posterior distribution. A Bayesian approach employs a number of techniques and methodologies in order to avoid relying on a classical hypothesis test using confidence intervals. For instance, predictive intervals are helpful for capturing uncertainty in measuring the number of child labourers or the intensity of child labour. Predictive intervals are obtained through simulation (or derived using predictors), and they are usually wider than confidence intervals capturing uncertainty in a more realistic way. In Chapter 4, the mean prediction interval has been used, which is similar to a confidence interval in the sense that it predicts the range of population averages. Chapter 5 uses the predictive intervals of individual observations, such as the 95-percent PIs of individual children's labour hours. The result shows the predictive intervals are much wider than the confidence intervals, in which all individual samples are considered random samples. Moreover, predictive intervals allow a direct interpretation of the ranges that future individual observations lie within, providing quantified information on uncertainties. There are many other ways for posterior predictive checks, such as information criteria and Bayes Factors, but a Bayesian p-value is one method to assess whether there is an obvious discrepancy between model predictions and observations, which does not determine the significance of estimates.

Lastly, a Bayesian approach helps to deal with an estimation based on rare events. Child labour includes a small number of occurrences, and so a large amount of uncertainty is involved. There is an ongoing discussion over whether a Bayesian approach helps solve such a rare event problem. A few studies found that the Bayesian method performs more effectively in using data with an excessive number of zeros (Ghosh et al., 2006; Baldwin and Fellingham, 2013). A frequentist approach and a Bayesian approach each work similar with a not-extreme level of a probability of y=0; yet, as the probability becomes close to one, a Bayesian approach works better, showing a tighter interval (Ghosh et al., 2006). In addition to that, the size of sample (n) is large enough, and the difference can be reduced (ibid.). As the sample size of this study is relatively large, we could expect the MLE method and Bayesian method to perform similarly. However, in social sciences, samples can be categorised into several groups, e.g. by sex, age, location, etc., in order that the number of samples might be further reduced (Gelman et al., 2013). Especially when dealing with a small sample size, the Bayesian method is better than the MLE method, as it does not depend on asymptotic approximations but uses true posterior distribution (Lee and Kim, 2008).

Overall, the benefits of a Bayesian approach in the study of child labour are enormous. Bayesian and frequentist approaches complement each other rather than compete. Frequentist methods are usually fast and tested using multiple softwares. Bayesian methods often rely on simulations and may take longer time to develop but are much more flexible and more specific to the context. Comparing and confirming results obtained from both sides are useful procedures for guaranteeing conclusions from data analysis. Nevertheless, if the subject has a large uncertainty and so accumulation and application of prior knowledge are required, a Bayesian approach is able to provide more suitable solutions.

7.4. Institutional Interventions towards Reducing Child Labour

In the following sub-sections, I take a closer look at how policy interventions and grassroots movements can address the child-labour problem. I will review current international and national policies, educational policies, and gender-based strategies, focusing on how they can help children and their families to have more power in social and gender relations, in dealing with child labour in India. Furthermore, emphasis will be given on the importance of the working hours policy for children to be built based on county-specific evidence and consideration of domestic chores.

7.4.1. International and national child-labour policies

International agencies have been running global campaigns and advocating activities to abolish child labour. The ILO has been involved in advocacy to eradicate child labour, in addition to their role of developing the standards of labour through international conventions. Furthermore, some international NGO networks try to link civil societies and influence governmental decisions regarding policy and interventions. The Global March Against Child Labour, which is one of the largest movements aiming to stop the phenomenon, is an example of how worldwide child labour issues are tackled together by international and local stakeholders. In addition, the private sector, as the key stakeholder in child labour, can engage in addressing child labour and improve the ethical performance of production. For example, the ILO initiated the Child Labour Platform, which shared good practices between employer organisations and trade unions[56]. It also launched a global checkpoint app for businesses to conduct a self-check-up in order to avoid any involvement in child labour[57]. There is a significant role for international organisations and civil societies to push companies to comply with legislation and prioritise children’s rights.

Despite the efforts of the global society to end child labour, many international programmes are limited for two reasons: first, the distance between global and local child labour; and second, difficulty in controlling or auditing through non-compulsory means. The clear concern regarding those activities is that there is less involvement from the grassroots and children’s voices are missing from current international discourses (Liebel and Invernizzi, 2019; Holzscheiter, 2016). Child labour is deeply associated with the family and

---

[56] Available at https://www.ilo.org/ipec/Action/CSR/clp/lang--en/index.htm (accessed 28 July 20)
[57] Available at https://www.ilo.org/global/publications/WCMS_460489/lang--en/index.htm (accessed 28 July 20)

community levels, and the voices of direct stakeholders are critical in addressing child labour. Child labour is difficult to resolve without proper awareness among parents, families and communities. Thus, for successful awareness and advocacy, the most urgent task is working out how to cooperate with local- and household-level stakeholders.

Furthermore, national child-labour policy interventions need to target the right population and draw more specific and strategic plans. In India, the National Child Labour Project (NCLP) is a key government scheme for child labour initiated in 1988. It aims to find children who are involved in hazardous occupations and processes and provide them with special schooling and guide them into formal schooling (Ministry of Labour and Employment, n.d.c). Special schools or rehabilitation centres provide former child labourers with non-formal education, vocational training, meals, stipends, and health care. The numbers of children who have benefitted from the scheme include 132,840, 94,657 and 125,716 in 2009/10, 2010/11 and 2011/12, respectively, across 17 states (Ministry of Labour and Employment, n.d.b). Uttar Pradesh, Bihar and Madhya Pradesh have the largest numbers of children who have been rehabilitated by the scheme (ibid.). It is noticeable that gender difference in enrolled children is not as great as expected and even in some states, such as Andhra Pradesh and Rajasthan, the number of girls outweighs the number of boys (see Appendix Table VI.1, ibid.).

The NCLP scheme has rescued child labourers and sent them back to schools, but there are some considerations regarding the scheme. The project focuses greatly on providing training through Special Training Centres and returning children to school, which targets the age group of 9–14 years. Therefore, adolescents are not well covered by the project. There is also a mismatch between states with many child-labour rescues and states with high risk. For example, in 2018, Bihar and Rajasthan do not report any children rescued through the scheme. More efforts are required to help children in high-risk areas[58]. Satpathy et al. (2010) point out that the process of selecting NCLP districts should be more rigorous, and more investment should be made in child-labour surveying, awareness, and teacher training.

Policy-level interventions are often limited in reaching the most vulnerable groups and transforming people's attitudes toward child labour. There is a more significant role for civil society and non-governmental organisations in reaching families in need and increasing people's awareness. For example, in implementing the NCLP scheme, the cooperative roles of NGOs have been recognised (Ministry of Labour and Employment, n.d.c). Since the

[58] Available at https://indiacsr.in/heights-no-of-cases-on-child-labour-reported-in-west-bengal/ (accessed 28 July 20)

Child Labour Act was enacted in 1987, many NGO programmes have moved towards implementing non-formal and part-time education provisions (Murphy, 2010). Among them, the MV Foundation has set a model aiming to achieve normative change, emphasising the norm that ‘every child must go to school’ (ibid.). This awareness programme is outstanding in that it has great involvement from parents and communities, proving the efficacy of changing norms in sending children back to schools.

7.4.2. Education and child labour

Without doubt, education policy is key to reducing child labour through providing children with increased accessibility to educational opportunities. Weiner (1991) suggests that compulsory education is the primary solution for ending child labour. It is known that other countries, such as China and South Korea that achieved compulsory education at an early stage of development, have experienced low incidences of child labour (ibid.). Conversely, compulsory education in India started late in 2009 (the Right of Children to Free and Compulsory Education Act 2009) and covers children from 6 to 14 years of age or up to class 8 (at age 12).

Regarding the relationship between education and child labour, some scholars insist that a combination of work and education is possible and provides better opportunities for children (White, 1996; Bourdillon, 2006). I argue that it is hard for children to do both schooling and working as ‘child labourers’. Children who do light economic activities and attend schools at the same time are child workers not child labourers. On the other hand, children who are neither in school nor at work, known as ‘nowhere children’, are presumably also labourers (Jayaraj and Subramanian, 2007; Chaudhri and Wilson, 2000; Giri and Singh, 2016).

National educational programmes include the Sarva Shiksha Abhiyan (SSA, meaning ‘Education for All’) programme and the National Programme for Education of Girls at Elementary Level (NPEGEL). The SSA programme began in 2000/01 and has been very successful in reducing the number of out-of-school children: from 42 million in 2000/01 to 13 million in 2005 (Premchander et al., 2012). The NPEGEL launched in 2003 as part of the SSA programme, focusing on enhancing the education level of girls in Educationally Backward Blocks (EBBs) where rural female literacy is high (Nuna, 2016). Girl-friendly schools are financially supported through this programme in 637 blocks in 161 districts (ibid.).

Many educational programmes have focused on primary education. After the Compulsory Education Act 2009 was enacted, elementary schooling became free for children. However, this study provides evidence that a large proportion of child labour occurs at ages 13–14, which indicates that child labour happens within the compulsory education age groups. To put it differently, child labour or dropout rates from schools grow around the time of entering secondary education. Chaudhri et al. (2003) and Chaudhri and Wilson (2000) support the need for secondary education to stop child labour by showing that most child labourers belong to age groups from 10 to 14. Unfortunately, secondary education is not yet free in India, and due to its high cost, many children do not go to secondary school but choose to work. Therefore, it is important to encourage families to continue their children's schooling and send the children to secondary school in order to reduce the incidence of child labour.

7.4.3. Gender-based strategies

The presence of gender inequality in child labour is never obvious as it is mixed with culture and tradition. In this regard, child-labour policymakers require an understanding of local and cultural roots and beliefs about child labour and gender roles. The lack of gender balance and recognition of institutionalised gender roles in development policies or projects have been criticised by many GAD scholars (Kabeer, 1994, pp.267–274). Similarly, agencies should consider more children's economic participation and the division of labour of gender roles among children into their policies and projects.

Child-labour policy should carefully consider both male and female child labourers. In general, girls' lack of educational opportunities is much more significant than for boys. Combining work and school attendance is more difficult for a girl in developing countries (Hill and King, 1995). Chaudhri and Wilson (2000) pointed out that rural girls represent the majority of the nowhere children who are not in school nor at work in India. Kambhampati and Rajan (2008) acknowledge that Indian girls are not only in schools less often but also less often at work. Furthermore, this study found that girls spend long hours in farming and home-based production. Any solution for child labour issues should address how to protect girls from dropping out of school, thereby preventing them from becoming labourers.

Providing more opportunities for education for girls is essential as girls' activities outside the home have been restricted in many states. The gender gap in school enrolment has been narrowing, but there is still a significant gap in secondary education, especially in rural areas. This study understands that this educational gap could increase the risk of girl child labour. The National Programme for Education of Girls at Elementary Level (NPEGEL) has been increasing its investment in girls' education, especially for lower social groups. However, the primary focus is given to elementary education; girls' secondary education and training should also be improved. According to the NSS 1999/2000, gender disparity in secondary school is the highest in Bihar and Rajasthan, where girls' secondary school enrolment is only half that of the boys (Kingdon, 2007). Madhya Pradesh, Chhattisgarh, and Uttar Pradesh also show an enormous gender inequality (ibid.). Those states mentioned present a high level of child labour risks, too. The increase of secondary school enrolment among girls in those states will eventually help reduce the risk of girl child labour.

It is important to transform attitudes towards gender roles among parents, employers, and communities. In particular, changing norms and values about girls' education and training need more interests. Mensch et al. (2004) tell their experiences of vocational training for girls in India, where the most significant barrier was notions held by parents and girls about activities outside the home. Many girls never go outside their local areas, and they even need permission to go outside their homes; meanwhile, the programmes were too short to change such embedded social norms about gender roles or expectations (Mensch et al., 2004). As Nuna (2009) points out, cultural factors play a vital role in the gender gap in education, throughout higher education and on technical or professional courses in India. Thus, there should be institutional support and efforts to reduce the reluctance to enabling girls in further education.

Changing social norms to improve gender equality might require a long time to be achieved. Although parents have slowly changed their values and attitudes to educating girls, there are still concerns regarding women's insecurity and unemployment (Dacosta, 2008). Kingdon (2007) points out that the low number of girls enrolled in secondary schools is partly because of parents' perceptions of the futility of girls in education as well as concerns for their safety. This study, in Chapter 6, emphasises how girls' low expectations for future work demotivates them from entering further education and increases their labour-force participation (before marriage). On the other hand, support for female participation in the labour market helps to decrease risks of child labour for both boys and girls. Thus, to reduce child-labour risks, there should be better opportunities for women's employment accompanied by efforts to change people's notions and attitudes about women in work.

7.4.4. Limiting maximum working hours

This study tries to explain the issues of children's working hours in terms of social class and gender relations. It reveals that household wealth and occupational class have a relationship with the working hours of children, which are longer for children of lower-class families in general, although there is some gender variance. For example, girls work longer in the informal sector than boys but for less time in the formal sector (Chapter 5). Nevertheless, revealing the long working hours of children is not an easy task. Datasets that include children's working hours are often insufficient, and revealing their relationship with potential harm is complicated. Whether the current international threshold of working hours is suitable for categorising child labourers as distinct from child workers should be further discussed.

In India, there is no clear legislation for maximum weekly working hours for adolescents (children above the ages allowed for labour). According to the Factories Act 1948, young workers (15–18 years old) are not allowed to work more than 4.5 hours a day, but this ruling only applies to children working in factories[59]. In the Child Labour Act 1986, children's maximum working day is 6 hours; however, there is no rule controlling their maximum weekly hours. According to the IHDS 2011/12, the number of working hours of child labourers reaches over 50 hours a week. Some children indicate that they work 16 hours a day. There is an urgent need for the government to set out a more explicit working hours policy in order to control children's working hours.

This study reflects recent changes in time thresholds in child labour. It uses the latest working hours suggested by the ILO and UNICEF for children, which are 1, 14, and 43 hours for non-domestic work (for ages 5–11, 12–14, and 15–17, respectively), and 21 hours for domestic work (for ages 5–14). Usually, girls work in unpaid household services and boys in non-domestic work; therefore, a large gap in standards between non-domestic and domestic work (1 hour for the former, 28 hours for the later) might bring gender bias into the calculations regarding child labour (Chaubey et al., 2007, p.3). It was only recently that time thresholds for domestic work were reduced to 21 hours for ages 5–14 (UNICEF, 2019).

As of this change by UNICEF, a slightly greater number of girls are now included as child labourers. Another change in the UNICEF 2019 definition of child labour is the exclusion of adolescents aged 15–17 who work in domestic chores from the categories of child labour. However, long hours spent in domestic work can cause harm to adolescents, too. It is clear that under the ILO policy, domestic work for long hours is regarded as a hazardous activity. The study points out that, in India, many products are manufactured at household-level. There is not a clear boundary in reality between production and non-production. Thus, disregarding domestic chores of children aged 15–17 might lead to an underestimation of the number of child labourers, particularly among female adolescents.

[59] Adults cannot work for more than 48 hours in a week and not more than 9 hours in a day.

Working-hour policy should be built upon the relationship between the length of working hours and hazards to children. Studies suggest that 20 hours of weekly employment have a limited impact on school attendance, but there is a marked increase in risk beyond this point (Brown, 2011). Jayaraj and Subramanian (2007, p.194) use working for at least 6 hours a day (42 hours a week) as the standard to qualify a child as a 'worker'. This study reveals that in India working longer than 38.5 hours a week increases the rate of school absence among children aged 12 to 17 by up to 70 percent. Child labour under the age of 14 is already completely banned in India. However, exceptions are allowed including children performing household work (including farming). How to define child labour among those children who are under the age of 14 and help with family work requires further discussion.

Obviously, the ILO standards on working time (+43 hours) are far more extensive than the time thresholds that this study finds risky to children's education. Looking at some examples in other countries, adolescents' maximum working hours are set lower than that: for example, at 40 hours in the UK (ages 16–17). Under these international regulations, people might perceive that fewer than 43 weekly working hours are not harmful to children. Thus, the Indian government needs to set a reasonable maximum number of weekly hours for children. To this end, there should be more studies of outcomes from long working hours in terms of diverse aspects of children's development and wellbeing.

7.5. Limitations and Further Study Suggestions

While this study tries to provide accurate estimations of child labour using country-specific analysis, there are some limitations regarding insufficient data and in dealing with the hidden features of child labour. This sub-section discusses the shortcomings of this study and the issues that should be addressed in future research. In addition, I will explain how the use of evidence-based priors in the model can improve estimations in future work.

Firstly, this study might exclude some children who do labour from a model or estimation due to limited data availability. Children who are age under five could not be included in this study. The NSS 2011/12 does not provide working hours for children who are under age five, and their activity status is recorded as "too young" (activity code 99). The IHDS 2011/12 provides limited information on working hours of children under age 5. The age category ranging from 5 to 17 is most commonly used in defining child labour, but further study is necessary to explore the effect of child labour at earlier ages. Furthermore, in Chapters 6 and 7, domestic working hours were excluded in the count of child labour. The IHDS 2011 lacks the information documenting the working hours of children in unpaid household services.

Second, more investigation into the accuracy of children's labour 'hours' and 'days' is needed. It is assumed that all child labourers work seven days a week, and this should be carefully investigated. One hour of economic activity for children aged 5–11 as a threshold of child labour might be too rigid. Furthermore, domestic working hours need more statistical analysis; the time threshold of 21 hours a week of household chores (of UNICEF 2019 definition) needs to be backed by further research. Domestic work for ages 15–17 is not recognised as part of child labour by the UNICEF 2019 definition so that children who do a mixture of economic activity and domestic chores might not be recognised as labourers.

Third, although the impact of long working hours was included in this study, other categories, such as investigations into hazardous industrial or occupational sectors or working conditions, have not thoroughly considered. There is even less known here than about the effect of working hours, although the ILO has used them as categories of child labour. The lack of relevant data is the main cause for there being fewer studies of those components of child labour. Particularly, definitions of hazardousness can differ in each country. Therefore, exactly how we can clarify definitions of harmful industries, occupations, and working conditions for children should be addressed in further studies.

Fourth, this study did not include 'nowhere children' (who are not recognised as either as labourers or students) as child labourers, although much of the literature includes them as potential child labourers. For instance, the incidence of 'nowhere children' is the most critical factor reinforcing child labour supply (Chaudhri et al., 2003, p.13) with rural girls mainly categorised as 'nowhere children' (Chaudhri and Wilson, 2000, p.19). Jayaraj and Subramanian (2007) used the concept of NWNAS (Non-Workers and Not Attending School), a more restrictive form of 'nowhere children' which excludes children who work in household service from the categorisation. They can be considered a separate category of child labour in further research, as a way to include child labourers who are not recognised in the dataset, which might extend the number of girl child labourers significantly.

Fifth, this study used only the educational impact of proxies for the harmful results of long working hours. The physical or mental impacts and other related problems of child labour have been studied more recently. For instance, Das et al. (2013) found in West Bengal that children spending more time in agricultural activities results in acute pain and discomfort (using their own data collection). Ahmed and Ray (2014) reveal that child labour is associated with injury and illness in the construction and manufacturing sectors in Bangladesh. Trinh (2020) proves that being involved in labour can affect children's mental health in Vietnam, such as emotional symptoms and behavioural problems, although the results are mixed depending on gender and the measures used. The IHDS provides information regarding children's ill-health: in this case mostly chronic diseases and based on rare events. Consequently, this study relies on education information (children being out of schools) to estimate the potential harm of child labour. In future, more comprehensive data is needed to reveal the diverse aspects of harm from child labour.

Sixth, family structure, such as the marital status of children and the sibling effect, were excluded in the models. The impact of early marriage on child labour is difficult to be proven due to its small proportion. Still, early marriage could be the reason for the increased burden of reproductive work among girls. The sibling effect, such as the presence of younger children in a household, could affect the decision of child labour. Due to its complicated relations between siblings, this study does not include it in the models. A future study might be necessary to explore the sibling effect on child labour in association with gender preferences.

Another limitation of this study is its reliance on the 2011 population and the datasets in 2011/12. With Bayesian modelling, it is possible to apply a newly projected population size in order to predict the number of child labourers in the near future. The projection of the population requires further demographic studies on mortality, fertility and migration. The projection of child-labour numbers is possible through using the projected population size of children in the child-labour risk model that is introduced in this study. In addition, the newly available datasets can work with the suggested model in this thesis. The PLFS 2018/19 (already released on the website, NSSO, 2020a) and the India Human Development Survey, Wave III (under construction, the University of Maryland and the National Council of Applied Economic Research, 2020) can be considered. Both datasets are based on a random sampling procedure and include the information on children's employment, working hours and industries or occupations.

Lastly, this study of child labour can benefit from the development of diverse priors. So far, this study has used non-informative and weakly informative priors in the models. This means that this study limits the effects of the priors but weighs the influences of observations, given the lack of knowledge about the number of child labourers among child populations. However, as more knowledge is accumulated on the size or shape of child-labour distribution, we might consider the application of more informative priors in future models. This study implies that the uncertainty of child labour is a mixture of its variations in definition, limitations in the datasets and the sensitive nature of the issue. More knowledge-based priors, in particular, might help reduce any uncertainty arising from characteristics of child labour. It is possible through using diverse measures, such as expert surveys, to collect knowledge on less known characteristics of child labour: for example, in unpaid household services or home-based productions.

Chapter 8. Conclusion

This study is grounded in the gender and social relations of child labour and explains the norms behind agents' decisions. Agents are positioned within the social structure and are affected by norms embodied in it. They can accept or reject norms on the basis of given social conditions and also consciously. Thus, changing norms is a natural and feasible way to help reduce child labour, thereby developing a new 'habitus' in society of avoiding child labour. Through this study, specific investigations about gender norms regarding child labour in combination with social and cultural aspects in India were accomplished. In order to eliminate child labour in India, more strategic and gender-focused interventions will be required.

This study proposes the importance of applying theories regarding child labour and linking them with empirical evidence while also introducing methodological innovation to strengthen the point of view. The contributions of this research are several. This study has shown that the gender and development (GAD) approach is valid and advantageous in analysing child labour, indicating it is decided by social and gender relations and by individuals' internalised norms. Furthermore, it provides a way to predict the numbers of child labourers based on a Bayesian data combination. Its predictions helped to gauge the significance of child labour in India. Moreover, it suggests a child labour risk model that accounts for the socioeconomic and cultural factors affecting child labour, which could help to improve the accuracy of predictions. The suggested method can be applied to other countries or to micro-localities in future research. Lastly, it measured the unknown time threshold beyond which child labourers' risk of harm from being out of school increases, supporting the idea that working-time regulation is necessary to prevent children from becoming full-time workers in India.

Indian child labour is based on various structural and relational problems. This point of view has been confirmed through a class analysis and sectoral breakdown of informal and formal labour. In addition, gender and class appear to have explanatory power in terms of the intensity of child labour. A strong relationship was found between household occupations and the incidence of child labour, while children's gender critically affects their working hours and the sectors in which they work in association with their household class positions. All of these findings suggest that child labour is formed within a class or social structure, having group-based patterns in type, sector or hours of labour. Within the social structure, gender and gender roles or stereotypes are the key aspects to explain the variations in children's economic roles.

Furthermore, this study recognises the institutional and cultural characteristics of child labour. Norms are institutionalised but also changed in accordance with agents' interactions with relations and consciousness. India is a patriarchal society, in which gender norms limit women's outside activities, but to some degree, child labour is accepted. Institutionalised norms directly impact on the decisions regarding child labour. A high concentration of girls in farming and domestic chores is a mixed outcome of people's gender norms and sociocultural factors. Agents have diverse social and gender norms, according to the structural conditions and internalisation of norms. Thus, child-labour practices can only be reduced when norms embedded in a society are challenged and shifted.

Child labour is not a gender-neutral nor a uniform phenomenon. Girls' work is considered less important because of the types of work. This study rejects the traditional prejudice that girls' economic roles are trivial by proving that they make considerable economic contributions to the family. Girls' economic activities should not be valued any less because they occur at home. In revealing gender inequality among child labourers, the GAD approach provides a comprehensive view. Two pillars of child labour – social relations and gender relations – contribute enormously to understanding child labour as a socially structured problem based on gender, norms and institutions.

Both aspects should be considered jointly in addressing child labour. On the one hand, child labour is a structural and global phenomenon, defined by age, gender, and class relations. In this sense, child labour is not only a problem for developing countries but also of developed countries, despite their different trends and scales. Addressing structural inequality requires arduous work. While this study focuses on the case of India, the global efforts to reduce child labour are significant. On the other hand, it is a problem compounded by cultural and institutional features. Specific locations and times, in which norms and values are grounded, are important factors that shape children's economic roles and responsibilities. Therefore, there should be an inclusive and long term intervention to change norms and beliefs regarding child labour in society.

Bibliography

Abebe, T. and Bessell, S. (2011) ‘Dominant discourses, debates and silences on child labour in Africa and Asia’, *Third World Quarterly*, 32(4), pp. 765–786.

Abernethie, L. (1998) ‘Child labour in contemporary society: why do we care?’, *The International Journal of Children s Rights*, 6(1), pp. 81–114.

Abraham, V. (2013) *Missing labour force or ‘de-feminization’ of labour force in India?* Centre for Development Studies Thiruvananthapuram.

Ades, A., Welton, N. J., Caldwell, D., Price, M., Goubar, A. and Lu, G. (2008) ‘Multiparameter evidence synthesis in epidemiology and medical decision-making’, *Journal of Health Services Research & Policy*, 13(3_suppl), pp. 12–22.

Afshar, H. (ed.) (1998) *Women and empowerment: illustrations from the Third World*. Basingstoke: Macmillan Press (Women’s studies at York series).

Agarwal, B. (2016) *Gender challenges*. New Delhi, India: Oxford University Press.

Ahmed, S. and Ray, R. (2014) ‘Health consequences of child labour in Bangladesh’, *Demographic research*, 30, pp. 111–150.

Alcaraz, C., Chiquiar, D. and Salcedo, A. (2012) ‘Remittances, schooling, and child labor in Mexico’, *Journal of Development Economics*, 97(1), pp. 156–165.

Andres, L. A., Dasgupta, B., Joseph, G., Abraham, V. and Correia, M. (2017) *Precarious drop: Reassessing patterns of female labor force participation in India*. The World Bank.

Ashagrie, K. (1993) ‘Statistics on child labor’, *Bulletin of Labour Statistics*, 3, pp. 11–24.

Athreya, V., Böklin, G., Djurfeldt, G. and Lindberg, S. (1987) ‘Identification of agrarian classes: a methodological essay with empirical material from south India’, *The Journal of Peasant Studies*, 14(2), pp. 147–190.

Athreya, V. B. (1990) *Barriers broken: production relations and agrarian change in Tamil Nadu*. New Delhi: Sage Publications.

Baden, S. and Goetz, A. M. (1998) ‘Who needs [sex] when you can have [gender]? conflicting discourses on gender at Beijing’, in Jackson, C. and Pearson, R. (eds) *Feminist Visions of Development: Gender, analysis and policy*. London: Routledge.

Balagopalan, S. (2014) *Inhabiting ‘childhood’: children, labour and schooling in postcolonial India*. Basingstoke: Palgrave Macmillan.

Baldwin, S. A. and Fellingham, G. W. (2013) ‘Bayesian methods for the analysis of small sample multilevel data with a complex variance structure.’, *Psychological methods*, 18(2), pp. 151–164. doi: 10.1037/a0030642.

Bandara, A., Dehejia, R. and Lavie-Rouse, S. (2015) ‘The Impact of income and non-income shocks on child labor: evidence from a panel survey of Tanzania’, *World Development*, 67, pp. 218–237. doi: https://doi.org/10.1016/j.worlddev.2014.10.019.

Bardhan, P. (1982) 'Agrarian class formation in India', *The Journal of Peasant Studies*, 10(1), pp. 73–94.

Barrientos, S. and Smith, S. (2007) 'Do workers benefit from ethical trade? Assessing codes of labour practice in global production systems', *Third World Quarterly*, 28(4), pp. 713–729.

Basu, K. (1999) 'Child labor: cause, consequence, and cure, with remarks on international labor standards', *Journal of Economic Literature*, xxxvii(3), pp. 1083–1119.

Basu, K., Das, S. and Dutta, B. (2010) 'Child labor and household wealth: theory and empirical evidence of an inverted-U', *Journal of Development Economics*, 91(1), pp. 8–14.

Basu, K. and Van, P. H. (1998a) 'The Economics of Child Labor', *The American Economic Review*, 88(3), pp. 412–427.

Basu, K. and Van, P. H. (1998b) 'The economics of child labor', *The American Economic Review*, 88(3), pp. 412–427.

Batliwala, S. (1993) *Women's empowerment in South Asia: concepts and practices*. Asian South Pacific Bureau of Adult Education.

Bayarri, M. and Berger, J. O. (2000) 'P values for composite null models', *Journal of the American Statistical Association*, 95(452), pp. 1127–1142.

Benería, L. (1979) 'Reproduction, production and the sexual division of labour', *Cambridge Journal of Economics*, 3(3), pp. 203–225.

Benería, L. (2003) *Gender, development, and globalization: economics as if all people mattered*. New York: Routledge.

Benería, L. and Sen, G. (1981) 'Accumulation, reproduction, and "women's role in economic development": Boserup revisited', *Signs: Journal of Women in Culture and Society*, 7(2), pp. 279–298.

Berger, J. (2006) 'The case for objective Bayesian analysis', *Bayesian Analysis*, 1(3), pp. 385–402. doi: 10.1214/06-BA115.

Bernardo, J. M. (1996) 'The concept of exchangeability and its applications', *Far East Journal of Mathematical Sciences*, 4, pp. 111–122.

Bernardo, J. M. and Smith, A. F. M. (2000) *Bayesian theory*. Chichester, Eng. ; Wiley. doi: 10.1002/9780470316870.

Bessell, S. (2011) 'Influencing international child labour policy: the potential and limits of children-centred research', *Children and Youth Services Review*, 33(4), p. 564,568.

Bhalotra, S. and Heady, C. (2003) 'Child farm labor: the wealth paradox', *The World Bank Economic Review*, 17(2), pp. 197–227.

Bhandari, L. and Dubey, A. (2019) *Emerging employment patterns of 21st century India*. Available at: https://www.indicus.org/pdf/Emerging-Employment-Patterns-in-21st-Century-Indiav-v181019.pdf (Accessed: 24 July 2020).

Bhaskaran, R., Nathan, D., Phillips, N. and Choragudi, U. (2014) ‘Vulnerable workers and labour standards (non-) compliance in global production networks: home-based child labour in Delhi’s garment sector’, in Rossi, A., Luinstra, A., and Pickles, J. (eds) *Towards Better Work*. Springer, pp. 172–190.

Bhat, B. A. (2010) ‘Gender, education and child Labour: a sociological perspective’, *Educational Research and Reviews*, 5(6), pp. 323–328.

Bhat, B. and Rather, T. (2009) ‘Child labour in the handicrafts home industry in Kashmir: a sociological study’, *International NGO Journal*, 4(9), pp. 391–400.

Bhowmik, S. K. (2002) ‘Productivity and labour standards in tea plantation sector in India’, in Sivananthiran, A. and Venkata Ratnam, C. S. (eds) *Labour and social issues in plantations in South Asia: role of social dialogue*. Geneva: International Labour Organisation, pp. 133–166.

Bicchieri, C. (2014) ‘Norms, conventions, and the power of expectations’, in Cartwright, N. and Montuschi, E. (eds) *Philosophy of Social Science: a new introduction*. New York: Oxford University Press.

Bijak, J. and Bryant, J. (2016) ‘Bayesian demography 250 years after Bayes’, *Population Studies*, 70(1), pp. 1–19. doi: 10.1080/00324728.2015.1122826.

Blasco, M. (2005) ‘Mobilising family solidarity? Rights, responsibilities and secondary schooling in urban Mexico’, in Goddard, J. et al. (eds) *The Politics of Childhood*. London: Palgrave Macmillan, pp. 127–144.

Bose, S. (2012) ‘A contextual analysis of gender disparity in education in India: the relative effects of son preference, women’s status, and community’, *Sociological Perspectives*, 55(1), pp. 67–91. doi: 10.1525/sop.2012.55.1.67.

Boserup, E. (1970) *Woman’s role in economic development*. London: Allen & Unwin.

Bourdieu, P. (1977) *Outline of a theory of practice*. Cambridge: Cambridge University Press (Cambridge studies in social and cultural anthropology, 16).

Bourdieu, P. (1990) *The logic of practice*. Cambridge: Polity.

Bourdillon, M. (2006) ‘Children and work: a review of current literature and debates’, *Development and Change*, 37(6), pp. 1201–1226.

Bourdillon, M. F. ., White, B. and Myers, W. E. (2009) ‘Re-assessing minimum-age standards for children’s work’, *International Journal of Sociology and Social Policy*, 29(3/4), pp. 106–117.

Breen, R. (2005) ‘Foundations of a neo-Weberian class analysis’, in Wright, E. O. (ed.) *Approaches to class analysis*. Cambridge: Cambridge University Press, pp. 31–50.

Breman, J. (1976) ‘A dualistic labour system? A critique of the “informal sector” concept II: a fragmented labour market’, *Economic and Political Weekly*, 11(49), pp. 1905–1908.

Breman, J. (1996) *Footloose labour: working in India’s informal economy*. Cambridge: Cambridge University Press (Contemporary South Asia, 2).

Breman, J. (1999) 'Industrial labour in post-colonial India II: employment in the informal-sector economy', *International Review of Social History,* 44(3), pp. 451–483.

Brezis, E. S. and Young, W. (2016) 'Population and economic growth: ancient and modern', *The European Journal of the History of Economic Thought*, 23(2), pp. 246–271.

Brezis, E. and Young, W. (2003) 'The new views on demographic transition: a reassessment of Malthus's and Marx's approach to population', *European Journal of the History of Economic Thought*, 10(1), pp. 25–45.

Brown, G. (2011) *Child labor & educational disadvantage – breaking the link, building opportunity*. London: UN Special Envoy for Global Education.

Burr, V. (1995) *An introduction to social constructionism*. Available at: http://search.proquest.com/docview/38825123/.

Burra, N. (2001) 'Cultural stereotypes and household behaviour: girl child labour in India', *Economic and Political Weekly,* 36(5/6), pp. 481–488.

Burra, N. (2005) 'Crusading for children in India's informal economy', *Economic and Political Weekly*, 40(49), pp. 5199–5208.

Burrone, S. and Giannelli, G. C. (2020) 'Child labour, gender and vulnerable employment in adulthood. Evidence for Tanzania', *The Journal of Development Studies*, pp. 1–16.

Cameron, A. and Trivedi, P. (2005) *Microeconometrics: methods and applications*. New York: Cambridge University Press. doi: 10.1017/CBO9780511811241.

Carpenter, B. *et al.* (2017) 'Stan: A probabilistic programming language', *Journal of Statistical Software*, 76(1).

Central Statistics Office (2018) *Children in India 2018 - a statistical approach*. Ministry of Statistics and Programme Implementation, Government of India. Available at: http://www.mospi.gov.in/sites/default/files/publication_reports/Children%20in%20India%202018%20%E2%80%93%20A%20Statistical%20Appraisal_26oct18.pdf (Accessed: 21 July 2020).

Chakravarty, D., Chakravarty, I. and others (2012) *When daughters migrate and mothers do not: girl children's paid outside work in West Bengal, India*. SOAS Department of Economics Working Paper Series 175. SOAS, University of London.

Chamarbagwala, R. (2008) 'Regional returns to education, child labour and schooling in India', *The Journal of Development Studies*, 44(2), pp. 233–257.

Chaubey, J., Perisic, M., Perrault, N., Laryea-Adjei, G. and Khan, N. (2007) *Child labour, education and policy options*. New York: UNICEF.

Chaudhri, D. P., Nagar, A. L., Tauhidur, R. and Wilson, E. J. (2003) 'Determinants of child labour in Indian states: some empirical explorations (1961–1991)', *Journal of Quantitative Economics,* 1(1), pp. 1–19.

Chaudhri, D. P. and Wilson, E. J. (2000) *The challenge of child labour in rural India: a multi-dimensional problem in need of an orchestrated policy response*. School of Economics, University of Wollongong. Available at: https://ro.uow.edu.au/commwkpapers/23/ (Accessed: 13 January 2021).

Chaudhuri, S. K. and Chaudhuri, S. S. (2005) *Primitive tribes in contemporary India: concept, ethnography and demography*. New Delhi: Mittal Publications.

Chen, M. (1995) ‘A matter of survival: women’s right to employment in India and Bangladesh’, in Nussbaum, M. C. and Glover, J. (eds) *Women, Culture, and Development: A study of human capabilities*. Oxford: Oxford University Press.

Chib, S., Greenberg, E. and Jeliazkov, I. (2009) ‘Estimation of semiparametric models in the presence of endogeneity and sample selection’, *Journal of Computational and Graphical Statistics*, 18(2), pp. 321–348. doi: 10.1198/jcgs.2009.07070.

Christensen, R. (2011) *Bayesian ideas and data analysis: an introduction for scientists and statisticians*. Boca Raton, Fla.: CRC (Chapman & Hall/CRC texts in statistical science series).

Chua, P., Bhavnani, K. K. and Foran, J. (2000) ‘Women, culture, development: a new paradigm for development studies?’, *Ethnic and Racial Studies*, 23(5), pp. 820–841. doi: 10.1080/01419870050110913.

Cowles, M. K. (2013) *Applied Bayesian statistics: with R and OpenBUGS examples*. New York: Springer.

Da Corta, L. and Venkateshwarlu, D. (1999) ‘Unfree relations and the feminisation of agricultural labour in Andhra Pradesh, 1970–95’, *Journal of Peasant Studies*, 26(2–3), pp. 71–139.

Dacosta, D. (2008) ‘Spoiled sons and sincere daughters: schooling, security, and empowerment in rural West Bengal, India’, *Signs: Journal of Women in Culture and Society*, 33(2), pp. 283–308.

Dammert, A. C. (2010) ‘Siblings, child labor, and schooling in Nicaragua and Guatemala’, *Journal of Population Economics*, 23(1), pp. 199–224. doi: 10.1007/s00148-008-0237-0.

Das, B. and Das, D. N. (2009) ‘Changing socio-demographic composition of child labour: evidence from the Indian Census 1991-2001’, *Social Change*, 39(3), pp. 351–364.

Das, B. and Khawas, V. (2009) *Gender issues in development: concerns for the 21st century*. Jaipur and New Delhi: Rawat Publications.

Das, D. K., Erumban, A. A., Aggarwal, S. and Das, P. C. (2015) *Measuring productivity at the industry level. The India KLEMS database*. Available at: https://rbidocs.rbi.org.in/rdocs/PublicationReport/Pdfs/KLEMSD4B66D90B4904EB981E18B129108DBA0.PDF.

Das, S. (2012) ‘Incidence of child labour and child schooling in India: pattern and determinants’, *ISRN Economics*, 2012, pp. 1–7.

Das, S. and Mukherjee, D. (2007) ‘Role of women in schooling and child labour decision: the case of urban boys in India’, *Social Indicators Research*, 82(3), pp. 463–486.

Das, S. and Mukherjee, D. (2011) ‘Measuring deprivation due to child work and child labour: a study for Indian children’, *Child Indicator Research*, 4(3), pp. 453–466. doi: 10.1007/s12187-010-9097-8.

Davidson-Pilon, C. (2015) *Bayesian methods for hackers: probabilistic programming and Bayesian inference*. Addison-Wesley Professional.

De Finetti, B. (1974) 'Bayesianism: its unifying role for both the foundations and applications of statistics', *International Statistical Review / Revue Internationale de Statistique*, 42(2), pp. 117–130. doi: 10.2307/1403075.

De Lange, A. (2009) 'Gender dimensions of rural child labour in Africa', in. *FAO-IFAD-ILO Workshop*. Available at: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.565.5141&rep=rep1&type=pdf (Accessed: 16 November 2020).

Dempster, A. (1997) 'The direct use of likelihood for significance testing', *Statistics and Computing*, 7(4), pp. 247–252. doi: 10.1023/A:1018598421607.

Desai, N. and Krishnaraj, M. (2002) 'An overview of the status of women in India', in Mohanty, M. (ed.) *Class, caste, gender*. New Delhi: Sage.

Desai, S. and Vanneman, R. (2018) *India Human Development Survey-II (IHDS-II), 2011–12*. Inter-university Consortium for Political and Social Research. Available at: http://doi.org/10.3886/ICPSR36151.v6 (Accessed: 10 December 2019).

Deshpande, A. and Ramachandran, R. (2014) *How backward are the other backward classes? Changing contours of caste disadvantage in India*. Available at: http://www.cdedse.org/pdf/work233.pdf (Accessed: 16 November 2020).

Dessy, S. E. (2000) 'A defense of compulsive measures against child labor', *Journal of Development Economics*, 62(1), pp. 261–275.

Dessy, S. E. and Pallage, S. (2005) 'A theory of the worst forms of child labour', *The Economic Journal*, 115(500), pp. 68–87.

Dorman, P. (2008) *Child labour, education and health: a review of the literature*. Geneva: International Labour Organization.

Drèze, J. and Sen, A. (1995) *India: Economic development and social opportunity*. Oxford: Oxford University Press.

Dubey, A., Olsen, W. and 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), pp. 589–612.

Dumas, C. (2012) 'Does work impede child learning? The case of Senegal', *Economic Development and Cultural Change*, 60(4), pp. 773–793.

Edmonds, E. V. (2006) 'Understanding sibling differences in child labor', *Journal of Population Economics*, 19(4), pp. 795–821.

Edmonds, E. V. (2009) *A review of the definitions of child labour in policy research*. International Labour Organization.

Edmonds, E. V. and Pavcnik, N. (2005) 'Child labor in the global economy', *Journal of Economic Perspectives*, 19(1), pp. 199–220.

Ekbrand, H. and Nandy, S. (2017) ‘Inter-state differences in caste-specific risks for child deprivation in India’, *Journal of Poverty and Social Justice*, 25(1), pp. 3–20.

El-Bushra, J. (2000) ‘Rethinking gender and development practice for the twenty-first century’, *Gender and Development,* 8(1), pp. 55–62. doi: 10.1080/741923413.

Elder-Vass, D. (2007) ‘Social structure and social relations’, *Journal for the Theory of Social Behaviour,* 37(4), pp. 463–477.

Elder-Vass, D. (2008) ‘Integrating institutional, relational and embodied structure: an emergentist perspective 1’, *The British Journal of Sociology*, 59(2), pp. 281–299.

Ellis, F. (2000) *Rural livelihoods and diversity in developing countries*. Oxford: Oxford University Press.

Elson, D. (1982) ‘The differentiation of children’s labour in the capitalist labour market’, *Development and Change,* 13(4), pp. 479–497.

Francavilla, F. and Giannelli, G. C. (2007) *The relation between child labour and mother’s work: The case of India*. IZA Discussion Papers.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M. and Gelman, A. (2019) ‘Visualization in Bayesian workflow’, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2), pp. 389–402.

Gamlin, J., Camacho, A. Z., Ong, M. and Hesketh, T. (2015) ‘Is domestic work a worst form of child labour? The findings of a six-country study of the psychosocial effects of child domestic work’, *Children’s Geographies*, 13(2), pp. 212–225.

Garikipati, S. (2009) ‘Landless but not assetless: female agricultural labour on the road to better status, evidence from India’, *The Journal of Peasant Studies*, 36(3), pp. 517–545.

Garikipati, S. and Olsen, W. (2008) ‘The role of agency in development planning and the development process: introduction to the Special Issue on Agency and Development’, *International Development Planning Review*, 30(4), pp. 327–338.

Gelman, A. (2005) ‘Comment: Fuzzy and Bayesian p-values and u-values’, *Statistical Science*, 20(4), pp. 380–381.

Gelman, A. (2006) ‘Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper)’, *Bayesian Analysis*, 1(3), pp. 515–534.

Gelman, A., Jakulin, A., Pittau, M. G. and Su, Y. S. (2008) ‘A weakly informative default prior distribution for logistic and other regression models’, *Annals of Applied Statistics*, 2(4), pp. 1360–1383.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2013) *Bayesian data analysis*. Chapman and Hall/CRC.

Gelman, A., Goodrich, B., Gabry, J. and Vehtari, A. (2019) ‘R-squared for Bayesian regression models’, *The American Statistician*, 73(3), pp. 307–309.

Gelman, A. (2020) *Prior choice recommendations*. Available at: https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations (Accessed: 3 January 2021).

Gelman, A. and Hill, J. (2007) *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press (Analytical methods for social research).

Gelman, A., Meng, X. L. and Stern, H. (1996) 'Posterior predictive assessment of model fitness via realized discrepancies', *Statistica Sinica*, pp. 733–760.

Geyer, J. C. (2006) *The K-Truncated Poisson Distribution*. University of Minnesota. Available at: http://www.stat.umn.edu/geyer/aster/library/aster/doc/ktp.pdf (Accessed: 16 November 2020).

Giri, A. K. and Singh, S. . (2016) 'Child labour and "nowhere" children in post-reforms India', *Indian Journal of Human Development*, 10(1), pp. 97–110. doi: 10.1177/0973703016654562.

Goldstein, M. and Wooff, D. (2007) *Bayes linear statistics: theory and methods*. Chichester, England: John Wiley. doi: 10.1002/9780470065662.

Gómez-Paredes, J., Alsamawi, A., Yamasue, E., Okumura, H., Ishihara, K. N., Geschke, A. and Lenzen, M. (2016) 'Consuming childhoods: an assessment of child labor's role in Indian production and global consumption', *Journal of Industrial Ecology*, 20(3), pp. 611–622.

Goose, N. and Honeyman, K. (2013) *Childhood and child labour in industrial England diversity and agency, 1750-1914*. Farnham, Surrey: Ashgate Pub. Ltd.

Government of India (1986) *Child Labour (Prohibition and Regulation) Act, 1986*. Available at: https://clc.gov.in/clc/acts-rules/child-labour-prohibition-and-regulation-act-1986 (Accessed: 21 July 2020).

Government of India (no date) *Note on Sample Design and Estimation Procedure of NSS 68th Round*. Available at: www.icssrdataservice.in › index.php › catalog › download (Accessed: 24 July 2020).

Grootaert, C. and Kanbur, R. (1995) 'Child labour: an economic perspective', *International labour review*, 134(2), pp. 187–204.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. and Tourangeau, R. (2011) *Survey methodology*. John Wiley & Sons.

Guarcello, L., Lyon, S., Rosati, F. and others (2004) *Impact of working time on children's health*. Understanding Children's Work (UCW Programme).

Guarcello, L., Lyon, S., Rosati, F., Valdivia, C. and others (2007) *Children's non-market activities and child labour measurement: a discussion based on household survey data*. Understanding Children's Work (UCW Programme).

Harriss-White, B. (2003) *India working: Essays on society and economy*. Cambridge University Press.

Harriss-White, B. and Gooptu, N. (2001) 'Mapping India's world of unorganized labour', *Socialist Register*, 37.

van Hasselt, M. (2014) 'Bayesian analysis of sample selection models', in Jeliazkov, I. and Yang, X. S. (eds) *Bayesian Inference in the Social Sciences*. Hoboken, New Jersey: John Wiley & Sons, Inc., pp. 91–113.

Hilbe, J. M. (2011) *Negative binomial regression*. Cambridge University Press.

Hill, M. A. and King, E. (1995) 'Women's education and economic well-being', *Feminist Economics*, 1(2), pp. 21–46.

Hirway, I. (2002) 'Understanding children's work in India: an analysis based on time use data', in Ramachandran, N. and Massun, L. (eds) *Coming to grips with rural child work: A food security approach*. New Delhi: Institute for Human Development.

Hirway, I. and Jose, S. (2011) 'Understanding women's work using time-use statistics: the case of India', *Feminist Economics*, 17(4), pp. 67–92.

Holzscheiter, A. (2016) 'Representation as power and performative practice: global civil society advocacy for working children', *Review of International Studies*, 42(2), pp. 205–226.

Humphreys, B. R. (2013) 'Dealing with zeros in economic data', *University of Alberta, Department of Economics*, pp. 1–27.

Humphreys, B. R., Lee, Y. S. and Soebbing, B. P. (2010) 'Consumer behaviour in lottery: the double hurdle approach and zeros in gambling survey data', *International Gambling Studies*, 10(2), pp. 165–176.

Humphries, J. (2010) *Childhood and child labour and the British Industrial Revolution*. Cambridge: Cambridge University Press (Cambridge studies in economic history. Second series).

Ibrahim, J. G., Chen, M. H., Gwon, Y. and Chen, F. (2015) 'The power prior: theory and applications', *Statistics in Medicine*, 34(28), pp. 3724–3749. doi: 10.1002/sim.6728.

Ibrahim, J. G., Chen, M. H. and Ryan, L. M. (2000) 'Bayesian variable selection for time series count data', *Statistica Sinica*, pp. 971–987.

Ilahi, N. (2001) *Children's work and schooling: does gender matter? Evidence from the Peru LSMS*. The World Bank.

ILO (2013) 'Unpaid household services and child labour', in. *19th International Conference of Labour Statisticians*, Geneva: International Labour Organization. Available at: https://www.ilo.org/ipec/Informationresources/WCMS_IPEC_PUB_25524/lang--en/index.htm (Accessed: 13 January 2021).

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 Organization.

ILO (2018a) *Hours of work*. International Labour Organization. Available at: https://www.ilo.org/ilostat-files/Documents/description_HRS_EN.pdf (Accessed: 11 January 2020).

ILO (2018b) *Resolution to amend the 18th ICLS Resolution concerning statistics of child labour*. ILO. Available at: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/meetingdocument/wcms_648624.pdf (Accessed: 11 January 2020).

ILO (2018c) *Survey methods to improve measurement of paid and unpaid work: country practices in time-use measurement*. Geneva: International Labour Organization. Available at: https://www.ilo.org/stat/Publications/WCMS_636055/lang--ja/index.htm (Accessed: 22 July 2020).

ILO (2020) *The ILO STAT database*. Available at: https://www.ilo.org/shinyapps/bulkexplorer58/?lang=en&segment=ref_area&id=IND_A (Accessed: 13 January 2021).

International Institute for Population Sciences (2020) *The National Family Health Survey 2019-20 (NFHS-5)*. Available at: http://rchiips.org/nfhs/NFHS-5sub_presentation.shtml (Accessed: 14 December 2020).

Iversen, V. (2002) 'Autonomy in child labor migrants', *World Development*, 30(5), pp. 817–834. doi: https://doi.org/10.1016/S0305-750X(02)00007-4.

Iversen, V., Krishna, A. and Sen, K. (2017) 'Rags to riches? Intergenerational occupational mobility in India', *Economic & Political Weekly*, 52(44), pp. 107–114.

Jackman, S. (2009) *Bayesian analysis for the social sciences*. Chichester, UK: Wiley. doi: 10.1002/9780470686621.

Jackson, C. and Pearson, R. (eds) (1998) *Feminist Visions of Development: Gender, analysis and policy*. London: Routledge.

James, A. and James, A. L. (2004) *Constructing childhood: theory, policy, and social practice*. Houndmills: Palgrave Macmillan.

James, A., Jenks, C. and Prout, A. (1998) *Theorizing childhood*. Cambridge: Polity Press.

James, A. and Prout, A. (1997) *Constructing and reconstructing childhood: contemporary issues in the sociological study of childhood*. 2nd edn. London: Falmer.

Janzen, S. A. (2018) 'Child labour measurement: whom should we ask?', *International Labour Review*, 157(2), pp. 169–191.

Jayaraj, D. and Subramanian, S. (2007) 'Out of school and (probably) in work: child labour and capability deprivation in India', *Journal of South Asian Development*, 2(2), pp. 177–226.

Jejeebhoy, S. J. and Sathar, Z. A. (2001) 'Women's autonomy in India and Pakistan: the influence of religion and region', *Population and Development Review*, 27(4), pp. 687–712.

Jodhka, S. S. (2012) *Caste*. New Delhi: Oxford University Press (Oxford India Short Introductions).

Joshi, K. (2010) 'Indigenous children of India: enrolment, gender parity and drop-out in school education', *International Journal of Sociology and Social Policy*, 30(9/10).

Kabeer, N. (1994) *Reversed realities: gender hierarchies in development thought*. London: Verso.

Kabeer, N. (2001) *Deprivation, discrimination and delivery: competing explanations for child labour and educational failure in South Asia*. IDS working paper 135. Brighton: Institute of Development Studies.

Kabeer, N. (2018) *Gender, livelihood capabilities and women's economic empowerment: reviewing evidence over the life course*. Working paper. London: Gender and Adolescence: Global Evidence (GAGE).

Kaizer, A. M., Koopmeiners, J. S. and Hobbs, B. P. (2018) 'Bayesian hierarchical modeling based on multisource exchangeability', *Biostatistics*, 19(2), pp. 169–184. doi: 10.1093/biostatistics/kxx031.

Kak, S. (2004) 'Magnitude and Profile of Child Labour in the 1990s: evidence from the NSS Data', *Social Scientist*, 32(1/2), pp. 43–73. doi: 10.2307/3518327.

Kambhampati, U. S. (2009) 'Child schooling and work decisions in India: The role of household and regional gender equity', *Feminist Economics*, 15(4), pp. 77–112.

Kambhampati, U. S. and Rajan, R. (2006) 'Economic growth: a panacea for child labor?', *World Development*, 34(3), pp. 426–445.

Kambhampati, U. S. and Rajan, R. (2008) 'The "nowhere" children: patriarchy and the role of girls in India's rural economy', *The Journal of Development Studies*, 44(9), pp. 1309–1341.

Kandiyoti, D. (1988) 'Bargaining with patriarchy', *Gender & Society*, 2(3), pp. 274–290. doi: 10.1177/089124388002003004.

Kandiyoti, D. (1998) 'Gender, power and contestation: "bargaining with patriarchy" revisited', in Jackson, C. and Pearson, R. (eds) *Feminist Visions of Development: Gender, analysis and policy*. London and New York: Routledge.

Kantor, P. (2003) 'Women's empowerment through home-based work: evidence from India', *Development and Change*, 34(3), pp. 425–445.

Kaplan, D. (2014) *Bayesian statistics for the social sciences*. New York: The Guilford Press (Methodology in the Social Sciences).

Kelkar, G. and Nathan, D. (1991) *Gender and tribe: women, land and forests in Jharkhand*. Zed Books.

Kennedy, M. C. and O'Hagan, A. (2001) 'Bayesian calibration of computer models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3), pp. 425–464.

Kingdon, G. G. (2007) 'The progress of school education in India', *Oxford Review of Economic Policy*, 23(2), pp. 168–195.

Kirby, P. (2003) *Child labour in Britain, 1750-1870*. London: Macmillan (Social history in perspective).

Knaul, F. M. (2001) 'The impact of child labor and school dropout on human capital: gender differences in Mexico', in Katz, E. G. and Correia, M. C. (eds) *The Economics of gender in Mexico: Work, family, state, and market*. Washington, DC: The World Bank (Directions in development), pp. 46–84.

Koissy-Kpein, S. (2012) 'Gender and competition between economic or non-economic labor and schooling: evidence from EPAM Mali', *African Development Review*, 24(1), pp. 107–123.

Kumar, A. (2013) 'Preference based vs. market based discrimination: implications for gender differentials in child labor and schooling', *Journal of Development Economics*, 105(1), pp. 64–68.

Lavalette, M. (1994) *Child employment in the capitalist labour market*. Aldershot: Avebury.

LeBaron, G., Lister, J. and Dauvergne, P. (2017) 'Governing global supply chain sustainability through the ethical audit regime', *Globalizations*, 14(6), pp. 958–975.

Lee, K. S. and Kim, S. U. (2008) 'Identification of uncertainty in low flow frequency analysis using Bayesian MCMC method', *Hydrological Processes*, 22(12), pp. 1949–1964.

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

Lewis, M. A. and Lockheed, M. E. (2007) 'Social exclusion: the emerging challenge in girls' education', in Lewis, M. A. and Lockheed, M. (eds) *Exclusion, gender and education: case studies from the developing world*. Washington, D.C.: Center for Global Development.

Li, K. (1998) 'Bayesian inference in a simultaneous equation model with limited dependent variables', *Journal of Econometrics*, 85(2), pp. 387–400. doi: 10.1016/S0304-4076(97)00106-1.

Li, M. and Tobias, J. L. (2014) 'Bayesian analysis of treatment effect models', in Jeliazkov, I. and Yang, X. S. (eds) *Bayesian inference in the social sciences*. Wiley Online Library, pp. 63–90.

Liddle, J. and Joshi, R. (1989) *Daughters of Independence: gender, caste, and class in India*. Rutgers University Press.

Liebel, M. (2004) *A will of their own: cross-cultural perspectives on working children*. Zed Books.

Liebel, M. and Invernizzi, A. (2019) 'The movements of working children and the International Labour Organization. A lesson on enforced silence', *Children & Society*, 33(2), pp. 142–153.

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

Lieten, K. (2001) 'Child labour: questions and magnitude', in White, B. and Lieten, K. (eds) *Child labour: policy options*. Amsterdam: Aksant Academic Publishers.

Lima, L., Mesquita, S. and Wanamaker, M. (2015) ‘Child labor and the wealth paradox: the role of altruistic parents’, *Economics Letters*, 130, pp. 80–82.

Lin, T. and Adserà, A. (2013) ‘Son preference and children’s housework: the case of India’, *Population Research and Policy Review*, 32(4), pp. 553–584.

Little, R. J. A. and Rubin, D. B. (1987) *Statistical analysis with missing data*. New York: Wiley.

López, J. and Scott, J. (2000) *Social structure*. Buckingham: Open University Press (Concepts in the Social Sciences).

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

Lynch, S. M. (2007) *Introduction to applied Bayesian statistics and estimation for social scientists*. New York: Springer (Statistics for Social and Behavioral Sciences). doi: 10.1007/978-0-387-71265-9.

Majumdar, M. (2001) ‘Child labour as a human security problem: evidence from India’, *Oxford Development Studies*, 29(3), pp. 279–304.

Mandelbaum, D. G. (1993) *Women’s seclusion and men’s honor: sex roles in North India, Bangladesh, and Pakistan*. Tucson: University of Arizona Press.

McNeish, D. (2016) ‘On using Bayesian methods to address small sample problems’, *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), pp. 750–773. doi: 10.1080/10705511.2016.1186549.

Mensch, B. S., Grant, M. J., Sebastian, M. P., Hewett, P. and Huntington, D. (2004) *The effect of a livelihoods intervention in an urban slum in India: do vocational counseling and training alter the attitudes and behavior of adolescent girls?* Policy Research Division Working Paper 194. New York: Population Council.

Mezzadri, A. (2016) ‘Class, gender and the sweatshop: on the nexus between labour commodification and exploitation’, *Third World Quarterly*, 37(10), pp. 1877–1900.

Mies, M. (1980a) ‘Capitalist development and subsistence reproduction: rural women in India’, *Bulletin of Concerned Asian Scholars*, 12(1), pp. 2–14. doi: 10.1080/14672715.1980.10405557.

Mies, M. (1980b) *Indian women and patriarchy: conflicts and dilemmas of students and working women.* New Delhi: Concept Publishing.

Ministry of Home Affairs (2011) *Census data 2011*. Available at: https://censusindia.gov.in/2011census/population_enumeration.html (Accessed: 11 January 2020).

Ministry of Labour (2001) *Second Report of the National Labour Commission*. New Delhi: Ministry of Labour.

Ministry of Labour and Employment (no date a) *Census data on child labour*. Available at: https://labour.gov.in/childlabour/census-data-child-labour (Accessed: 11 January 2020).

Ministry of Labour and Employment (no date b) *Children's mainstreamed under NCLP scheme*. Available at: https://labour.gov.in/childlabour/childrens-mainstreamed (Accessed: 26 July 2020).

Ministry of Labour and Employment (no date c) *The policy of the government on the issue of child labour*. Available at: https://labour.gov.in/sites/default/files/PolicyofGovernmenontheissueofChildLabour.pdf (Accessed: 26 July 2020).

Mohanty, M. (2002) *Class, caste and gender*. New Delhi: Sage.

Momsen, J. (2010) *Gender and development*. 2nd ed. Abingdon: Routledge (Routledge perspectives on development).

Moser, C. O. N. (1993) *Gender planning and development: theory, practice, and training*. London: Routledge.

Mosse, D. (2010) 'A relational approach to durable poverty, inequality and power', *The Journal of Development Studies*, 46(7), pp. 1156–1178.

Mosse, D. (2018) 'Caste and development: contemporary perspectives on a structure of discrimination and advantage', *World Development*, 110, pp. 422–436.

Mukherjee, D. (2012) 'Schooling, child labor, and reserve army evidences from India', *Journal of Developing Societies*, 28(1), pp. 1–29.

Mullahy, J. (1986) 'Specification and testing of some modified count data models', *Journal of Econometrics*, 33(3), pp. 341–365.

Murphy, D. (2010) *Eliminating child labour through education: the potential for replicating the work of the MV Foundation in India*. Secunderabad: MV Foundation.

Murray, U. (2004) *Girl child labour in agriculture, domestic work and sexual exploitation: a comparative analysis, cases of Ghana, Ecuador, Philippines (vol. 2)*. Geneva: ILO.

Narayan, A. (1988) 'Child labour policies and programmes: the Indian experience', in Bequele, A. and Boyden, J. (eds) *Combating child labour*. International Labour Organization.

Nardinelli, C. (1980) 'Child labor and the Factory Acts', *The Journal of Economic History*, 40(4), pp. 739–755.

National Sample Survey Office (2013) *India - Employment and unemployment July 2011-June 2012*. Available at: http://microdata.gov.in/nada43/index.php/catalog/127/study-description (Accessed: 11 January 2020).

National Sample Survey Office (2020a) *Time use survey*. Available at: http://mospi.nic.in/time-use-survey-0 (Accessed: 14 December 2020).

National Sample Survey Office (2020b) *Unit level data of Periodic Labour Force Survey (PLFS), July 2018-June 2019*. Available at: http://mospi.nic.in/unit-level-data-periodic-labour-force-survey-plfs-july-2018-june-2019 (Accessed: 14 December 2020).

Neelon, B., Ghosh, P. and Loebs, P. F. (2013) ‘A spatial Poisson hurdle model for exploring geographic variation in emergency department visits’, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(2), pp. 389–413.

Newman, C., Henchion, M. and Matthews, A. (2003) ‘A double-hurdle model of Irish household expenditure on prepared meals’, *Applied Economics*, 35(9), pp. 1053–1061.

Nieuwenhuys, O. (1994) *Children’s lifeworlds: gender, welfare and labour in the developing world*. London: Routledge.

Nieuwenhuys, O. (2007) ‘Embedding the global womb: Global child labour and the new policy agenda’, *Children’s Geographies*, 5(1–2), pp. 149–163.

Nuna, A. (2009) ‘Bridging the gender gaps: an analysis of the educational situation with special emphasis on higher, technical and professional education’, in Das, B. and Khawas, V. (eds) *Gender issues in development: concerns for the 21st century*. Jaipur and New Delhi: Rawat Publications.

Nuna, A. (2016) *Education of girls - NCERT*. Available at: http://www.ncert.nic.in/departments/nie/der/publication/pdf/Anita_nuna_2016.pdf (Accessed: 26 July 2020).

Oakley, A. (2002) ‘Women and children first and last: parallels and differences between children’s and women’s studies’, in Mayall, B. (ed.) *Children’s childhoods: observed and experienced*. London, Washington D.C.: Routledge, pp. 19–38.

Office of the Registrar General & Census Commissioner (2011) ‘Census of India 2011 - meta data’. Available at: https://censusindia.gov.in/2011census/HLO/Metadata_Census_2011.pdf (Accessed: 9 July 2021).

Okin, S. M. (1994) ‘Gender inequality and cultural differences’, *Political Theory*, 22(1), pp. 5–24. doi: 10.1177/0090591794022001002.

Olsen, W. (2009) ‘Exploring practical horizons of beyond sociology: structure, agency, and strategy among tenants in India’, *Asian Journal of Social Science*, 37(3), pp. 366–390.

Olsen, W. (2011) *Data collection: Key debates and methods in social research*. London: Sage.

Olsen, W. K. (1996) *Rural Indian social relations: a study of southern Andhra Pradesh*. Delhi: Oxford University Press.

Olsen, W. and Mehta, S. (2006a) ‘Female labour participation in rural and urban India: does housewives’ work count?’, *Radical Statistics*, 93, p. 57.

Olsen, W. and Mehta, S. (2006b) ‘The right to work and differentiation in Indian employment’, *Indian Journal of Labour Economics*, 49(3), pp. 389–406.

Olsen, W. and Morgan, J. (2015) ‘The entrapment of unfree Labor: theory and examples from India’, *Journal of Developing Societies*, 31(2), pp. 184–203.

Olsen, W. and Neff, D. F. (2007) *Informal agricultural work, habitus and practices in an Indian context*. ESRC Global Poverty Research Group.

Olsen, W. and Zhang, M. (2015) *How to statistically test attitudes over space: answers from the gender norms and labour supply project*. University of Manchester. Available at: http://hummedia.manchester.ac.uk/institutes/cmist/InnovationsProject/BP3OlsenZhang-1.pdf (Accessed: 16 November 2020).

Omoike, E. (2010) 'Child domestic labour: fostering in transition?', in Craig, G. (ed.) *Child slavery now: a contemporary reader*. Bristol, U.K.: Policy Press, pp. 203–214.

Ostergaard, L. (1992) *Gender and development: a practical guide*. London: Routledge.

Parish, W. L. and Willis, R. J. (1993) 'Daughters, education, and family budgets Taiwan experiences', *Journal of Human Resources*, pp. 863–898.

Parry, J., Breman, J. and Kapadia, K. (eds) (1999) *The worlds of Indian industrial labour*. New Delhi: Sage (Contributions to Indian sociology series, 9).

Parsons, T. (1952) *The Social System*. Glencoe IL: Free Press.

Patrinos, H. and Psacharopoulos, G. (1997) 'Family size, schooling and child labor in Peru - An empirical analysis', *JOURNAL OF POPULATION ECONOMICS*, 10(4), pp. 387–405. doi: 10.1007/s001480050050.

Pattnaik, I., Lahiri-Dutt, K., Lockie, S. and Pritchard, B. (2018) 'The feminization of agriculture or the feminization of agrarian distress? Tracking the trajectory of women in agriculture in India', *Journal of the Asia Pacific Economy*, 23(1), pp. 138–155.

Phillips, N. (2013) 'Unfree labour and adverse incorporation in the global economy: comparative perspectives on Brazil and India', *Economy and Society*, 42(2), pp. 171–196.

Phillips, N., Bhaskaran, R., Nathan, D. and Upendranadh, C. (2014) 'The social foundations of global production networks: towards a global political economy of child labour', *Third World Quarterly*, 35(3), pp. 428–446.

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*. Vienna, Austria: Distributed Statistical Computing, pp. 1–10. Available at: https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf (Accessed: 17 November 2020).

Poddar, H. A., Verma, R. P., Roy, S., Sancheti, S. and Saha, S. (no date) *Analysis of statistics relating to child labour in India*. Child Rights and You. Available at: https://www.cry.org/resources/pdf/Kol_PaperOnChildLabour.pdf (Accessed: 29 June 2020).

Pollard, D. (1997) 'Chapter 8: Poisson approximations'. Available at: http://www.stat.yale.edu/~pollard/Courses/241.fall97/Poisson.pdf (Accessed: 9 July 2021).

Premchander, S., Kanneganti, S. and Pandey, B. (2012) *The report of interim independent evaluation - convergence against child labour: support for India's model*. Interim report. Available at: https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---sro-new_delhi/documents/publication/wcms_440096.pdf.

Puhani, P. (2000) 'The Heckman correction for sample selection and its critique', *Journal of Economic Surveys*, 14(1), pp. 53–68. doi: 10.1111/1467-6419.00104.

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

Raju, S. and Bagchi, D. (1993) *Women and work in South Asia: regional patterns and perspectives*. London: Routledge.

Rao, S. (2018) 'Gender and class relations in rural India', *The Journal of Peasant Studies*, 45(5–6), pp. 950–968.

Rathgeber, E. M. (1990) 'WID, WAD, GAD: trends in research and practice', *The Journal of Developing Areas*, 24(4), pp. 489–502.

Ravallion, M. and Wodon, Q. (2000) 'Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy', *Economic Journal*, 110(462), pp. 158–175. doi: 10.1111/1468-0297.00527.

Ray, Ranjan (2000) 'Child labor, child schooling, and their interaction with adult labor: empirical evidence for Peru and Pakistan', *The World Bank Economic Review*, 14(2), p. 347.

Ray, R (2000) 'Child labor, child schooling, and their interaction with adult labor: Empirical evidence for Peru and Pakistan', *WORLD BANK ECONOMIC REVIEW*, 14(2), pp. 347–367. doi: 10.1093/wber/14.2.347.

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

Reeves, H. and Baden, S. (2000) *Gender and development: concepts and definitions*. BRIDGE 55. Institute of Development Studies.

Robson, E. (2004) 'Children at work in rural northern Nigeria: patterns of age, space and gender', *Journal of Rural Studies*, 20(2), pp. 193–210.

Roemer, J. E. (1982) 'Methodological individualism and deductive Marxism', *Theory and Society*, 11(4), pp. 513–520.

Rubin, D. B. (1996) 'Multiple imputation after 18+ years', *Journal of the American Statistical Association*, 91(434), pp. 473–489.

Sakshaug, J. W., Wiśniowski, A., Perez Ruiz, D. A. and Blom, A. G. (2019) 'Supplementing small probability samples with nonprobability samples: a Bayesian approach', *Journal of Official Statistics*, 35(3), pp. 653–681. doi: 10.2478/jos-2019-0027.

Samantroy, E. and Sekar, H. R. (2016) 'Work participation of children: analysis of gender dimensions and determinants', *Labour & Development*, 23(2).

Samantroy, E., Sekar, H. R. and Pradhan, S. (2017) *State of child workers in India*. Delhi: V.V.Giri National Labour Institute.

Satpathy, A., Sekar, H. R. and Karan, A. K. (2010) *Rehabilitation of child labour in India: lessons learnt from the evaluation of NCLPs*. VV Giri National Labour Institute.

Save the Children (2015) *The hidden workforce - A study on child labour in the garment industry in Delhi*. Delhi: Save the Children. Available at: https://resourcecentre.savethechildren.net/node/9346/pdf/the_hidden_workforce_15-06-2015.pdf (Accessed: 16 November 2020).

Scott, J. (2011) *Conceptualising the social world: principles of sociological analysis*. Cambridge: Cambridge University Press.

Self, S. (2011) 'Market and non-market child labour in rural India: The role of the mother's participation in the labour force', *Oxford Development Studies*, 39(3), pp. 315–338.

Sen, A. (2005) 'Human rights and capabilities', *Journal of Human Development*, 6(2), pp. 151–166.

Sharma, M. (1985) 'Caste, class, and gender: production and reproduction in North India', *The Journal of Peasant Studies*, 12(4), pp. 57–88.

Soares, R. R., Kruger, D. and Berthelon, M. (2012) 'Household choices of child labor and schooling: a simple model with application to Brazil', *Journal of Human Resources*, 47(1), pp. 1–31.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (1998) *Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models*. Research Report, 98-009.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van Der Linde, A. (2002) 'Bayesian measures of model complexity and fit', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), pp. 583–639. doi: 10.1111/1467-9868.00353.

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

Srinivas, M. N. (1956) 'A note on Sanskritization and Westernization', *The Far Eastern Quarterly*, 15(4), pp. 481–496.

Srinivas, M. N. (1984) 'Some reflections on the nature of caste hierarchy', *Contributions to Indian Sociology*, 18(2), pp. 151–167.

Srivastava, R. (2012) 'Changing employment conditions of the Indian workforce and implications for decent work', *Global labour journal*, 3(1), pp. 63–90. doi: 10.15173/glj.v3i1.1113.

Stewart, J. (2013) 'Tobit or not Tobit?', *Journal of Economic and Social Measurement*, 38(3), pp. 263–290.

Tilly, C. (1998) *Durable inequality*. Berkeley: University of California Press.

Trinh, T. A. (2020) ‘Mental health impacts of child labour: evidence from Vietnam and India’, *The Journal of Development Studies*, pp. 1–15.

UNESCO (2017) *Estimation of the numbers and rates of out-of-school children and adolescents using administrative and household survey data*. UNESCO Institute for Statistics. Available at: http://uis.unesco.org/sites/default/files/documents/ip35-estimation-numbers-rates-out-of-school-children-adolescents-household-2017-en.pdf (Accessed: 17 January 2021).

UNESCO (no date) *The UNESCO Institute for Statistics Glossary*. Available at: http://uis.unesco.org/en/glossary-term/out-school-rate-1-year-primary-primary-education-lower-secondary-education-upper (Accessed: 17 January 2021).

UNICEF (2017) *Access the data - child Labour*. UNICEF. Available at: 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*. Available at: https://data.unicef.org/topic/child-protection/child-labour/ (Accessed: 11 January 2020).

UNICEF (no date) *India statistics, UNICEF*. Available at: https://www.unicef.org/infobycountry/india_statistics.html (Accessed: 16 May 2019).

University of Maryland and the National Council of Applied Economic Research (2020) *The India Human Development Survey website*. Available at: https://ihds.umd.edu/about (Accessed: 15 December 2020).

Vaid, D. (2012) ‘The caste-class association in India: an empirical analysis’, *Asian Survey*, 52(2), pp. 395–422.

Vaid, D. and Heath, A. (2010) ‘Unequal opportunities: Class, caste and social mobility’, in *Proceedings of the British Academy*.

Van Daalen, E. and Hanson, K. (2019) ‘The ILO’s shifts in child labour policy’, *International Development Policy*, pp. 133–150.

Venkateshwarlu, D. and Da Corta, L. (2001) ‘Transformations in the age and gender of unfree workers on hybrid cotton seed farms in Andhra Pradesh’, *The Journal of Peasant Studies*, 28(3), pp. 1–36. doi: 10.1080/03066150108438782.

Venkateswarlu, D. (2007) *Child bondage continues in Indian cotton supply chain*. Hyderabad: Global Research and Consultancy Services.

Walby, S. (1990) *Theorizing patriarchy*. Oxford: Basil Blackwell.

Webbink, E., Smits, J. and de Jong, E. (2012) ‘Hidden child labor: determinants of housework and family business work of children in 16 developing countries’, *World Development*, 40(3), pp. 631–642. doi: https://doi.org/10.1016/j.worlddev.2011.07.005.

Webbink, E., Smits, J. and de Jong, E. (2013) ‘Household and context determinants of child labor in 221 districts of 18 developing countries’, *Social Indicators Research*, 110(2), pp. 819–836. doi: 10.1007/s11205-011-9960-0.

Webbink, E., Smits, J. and de 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), pp. 84–98.

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

Weiner, M., Burra, N. and Bajpai, A. (2006) *Born unfree: child labour, education, and the state in India*. New Delhi: Oxford University Press.

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

White, B. (1982) ‘Child Labour and Population Growth in Rural Asia’, *Development and Change*, 13(4), pp. 587–610. doi: 10.1111/j.1467-7660.1982.tb00139.x.

White, B. (1996) ‘Globalization and the child labor problem’, *Journal of International Development: The Journal of the Development Studies Association*, 8(6), pp. 829–839.

White, B. (1999) ‘Defining the intolerable: child work, global standards and cultural relativism’, *Childhood*, 6(1), pp. 133–144.

White, S. C. (1997) ‘Men, masculinities, and the politics of development’, *Gender & Development*, 5(2), pp. 14–22.

Widaman, K. F. (1993) ‘Common factor analysis versus principal component analysis: differential bias in representing model parameters?’, *Multivariate Behavioral Research*, 28(3), pp. 263–311.

Winkelmann, R. (2003) *Econometric analysis of count data*. 4th ed. Berlin: Springer.

Winkelmann, R. and Boes, S. (2006) *Analysis of microdata*. Springer Science & Business Media.

Wiśniowski, A., Bijak, J., Christiansen, S., Forster, J. J., Keilman, N., Raymer, J. and Smith, P. W. (2013) ‘Utilising expert opinion to improve the measurement of international migration in Europe’, *Journal of Official Statistics*, 29(4), pp. 583–607.

Wiśniowski, A., Forster, J. J., Smith, P. W. F., Bijak, J. and Raymer, J. (2016) ‘Integrated modelling of age and sex patterns of European migration’, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(4), pp. 1007–1024.

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), pp. 185–202. doi: 10.1111/rssa.12189.

Wiśniowski, A., Bijak, J., Forster, J. J. and Smith, P. W. . (2019) ‘Hierarchical model for forecasting the outcomes of binary referenda’, *Computational Statistics and Data Analysis*, 133, pp. 90–103. doi: 10.1016/j.csda.2018.09.007.

Woodhead, M. (1999) ‘Combatting child labour: listen to what the children say’, *Childhood*, 6(1), pp. 27–49.

Wyness, M. G. (2006) *Childhood and society*. 2nd edn. Basingstoke: Palgrave Macmillan.

Young, K. (1993) *Planning development with women: making a world of difference*. London: Macmillan.

Young, K. (1997) ‘Gender and development’, in Visvanathan, N. et al. (eds) *The Women, Gender and Development Reader*. London: Zed.

Zapata, D., Contreras, D. and Kruger, D. (2011) ‘Child labor and schooling in Bolivia: who’s falling behind? The roles of domestic work, gender, and ethnicity’, *World Development*, 39(4), pp. 588–599. doi: https://doi.org/10.1016/j.worlddev.2010.08.022.

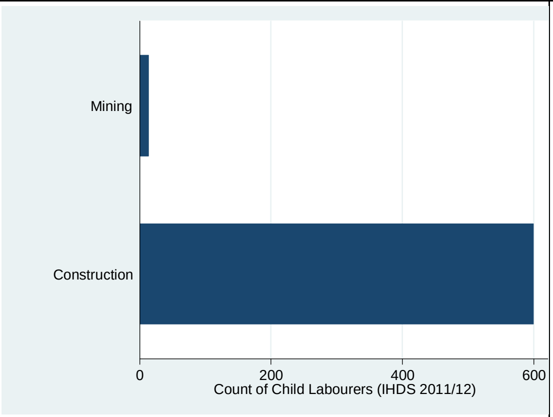
Zhang, J. L. (2014) ‘Comparative investigation of three Bayesian p values’, *Computational Statistics & Data Analysis*, 79, pp. 277–291.

Zhao, Y., Lee, A. H. and Barnes, T. (2010) ‘On application of the empirical Bayes shrinkage in epidemiological settings’, *International Journal of Environmental Research and Public Health*, 7(2), pp. 380–394.

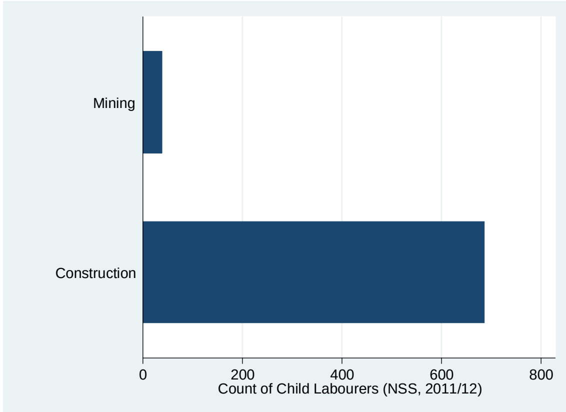
Appendices

A. Appendix for Chapter 4

Hazardous Industries

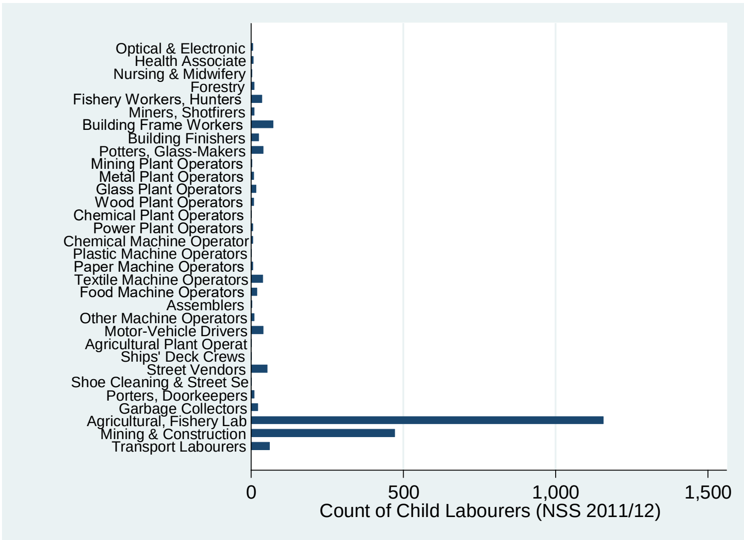


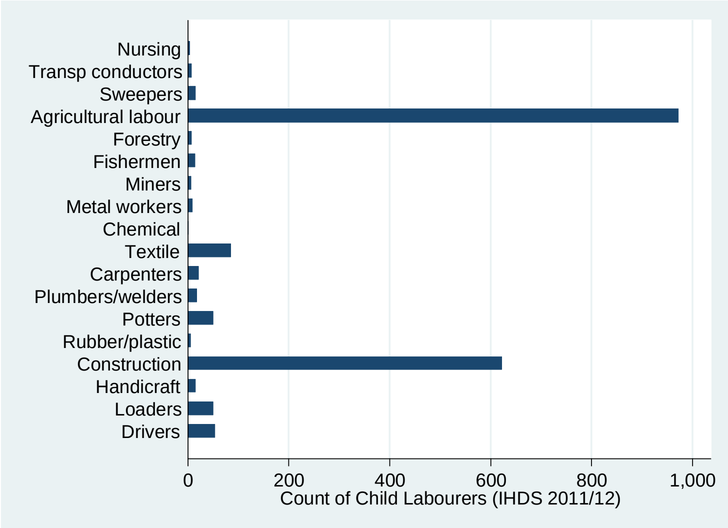
IHDS 2011/12



NSS 2011/12

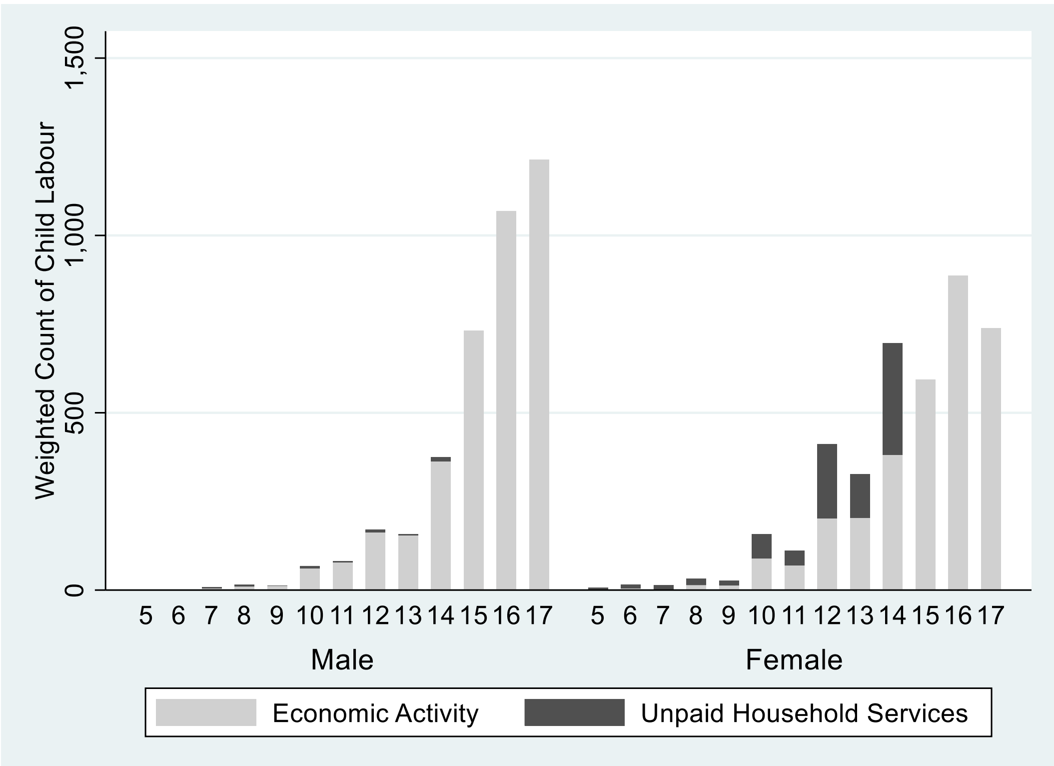
Hazardous occupations





Appendix Figure A.1. 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



Appendix Figure A.2. 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”). Source: NSS, 2011/12

Appendix Table A.1. Over-dispersion and undercount

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

Over-dispersion
The use of an over-dispersion 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 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 any records of children's domestic work. Because the NSS does include 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 the lack of domestic-work measures in one survey vs the other survey.

Appendix Table A.2. 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	$y.a_{ij} \sim \text{Poisson}(\mu.a_{ij} * n.a_{ij})$	-	$y.a_{ij} \sim \text{Poisson}(\mu_{ij} * n.a_{ij})$	$y.a_{ij} \sim \text{Poisson}(\upsilon * \mu.a_{ij} * n.a_{ij})$
	-	$y.b_{ij} \sim \text{Poisson}(\mu.b_{ij} * n.b_{ij})$	$y.b_{ij} \sim \text{Poisson}(\mu_{ij} * n.b_{ij})$	$y.b_{ij} \sim \text{Poisson}(\mu.b_{ij} * n.b_{ij})$
Overdispersion	-	-	-	$\log(\mu.a_{ij}) = \log(\psi_{ij}) + \lambda.a_{ij}$ $\lambda.a_{ij} \sim \text{Normal}(0, \tau.a)$
	-	-	-	$\log(\mu.b_{ij}) = \log(\psi_{ij}) + \lambda.b_{ij}$ $\lambda.b_{ij} \sim \text{Normal}(0, \tau.b)$
Prediction	$\hat{y}_{ij} \sim \text{Poisson}(\mu.a_{ij} * N_{ij})$ $\hat{y}_{i+} = \sum_i \hat{\mathbf{y}}_{\mathbf{ij}}$	$\hat{y}_{ij} \sim \text{Poisson}(\mu.b_{ij} * N_{ij})$ $\hat{y}_{i+} = \sum_i \hat{\mathbf{y}}_{\mathbf{ij}}$	$\hat{y}_{ij} \sim \text{Poisson}(\mu_{ij} * N_{ij})$ $\hat{y}_{i+} = \sum_i \hat{\mathbf{y}}_{\mathbf{ij}}$	$\hat{y}_{ij} \sim \text{Poisson}(\psi_{ij} * N_{ij})$ $\hat{y}_{i+j} = \sum_i \hat{\mathbf{y}}_{\mathbf{ij}}$

Model for true child labour rate	$\log(\mu.a_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij})$	$\log(\mu.b_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij})$	$\log(\mu_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij})$	$\log(\psi_{ij}) = \beta_0 + \beta_1 * x_i + \beta_2 * \log(z_{ij}) + \lambda_{ij}$ $\lambda_{ij} \sim \text{Normal}(0, \tau)$

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

Appendix Table A.3. 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	Population(N)	No. of Main Workers[3)]	No. of Main and Marginal Workers[4)]
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			
5	1	4,876	5,342.12	4,908	8	13,589	15,335	8,303	26,048,171	223,354	430,785
6	-	4,831	-	-	18	14,152	32,622	27,000	25,647,854	211,068	442,565
7	4	5,457	18,193	22,419	22	13,027	41,917	32,000	24,820,355	214,041	491,150
8	5	4,684	28,780	22,525	47	15,932	79,537	78,000	26,961,440	234,439	583,419
9	9	4,196	50,230	48,085	40	10,833	86,471	69,000	23,418,444	225,906	585,719
10	17	5,640	92,066	87,647	225	17,988	382,059	440,000	30,544,351	344,651	925,032
11	29	4,196	170,944	129,918	193	11,144	428,360	390,000	24,733,883	409,365	1,003,678
12	70	6,504	299,949	365,350	583	18,292	888,254	1,200,000	27,869,538	663,856	1,579,741
13	140	5,240	648,541	660,536	485	12,502	941,679	940,000	24,273,967	699,458	1,636,946
14	324	5,669	1,443,139	1,613,758	1,071	15,342	1,762,695	2,000,000	25,250,481	1,127,109	2,449,628
15	436	4,691	2,406,495	2,063,072	1,326	14,548	2,359,957	2,500,000	25,891,864	1,975,126	3,865,154
16	612	4,815	3,124,739	2,956,737	1,956	14,617	3,289,798	3,600,000	24,584,341	2,552,054	4,738,080
17	737	4,715	3,315,434	3,450,573	1,951	11,595	3,568,955	3,500,000	21,210,681	2,911,827	5,047,586
Total	2,384	65,514	11,598,511	11,425,529	7,925	183,561	13,877,638	14,784,303	331,255,370	11,792,254	23,779,483

Source: Desai and Vanneman, 2018; Ministry of Home Affairs, 2011; National Sample Survey Office, 2013

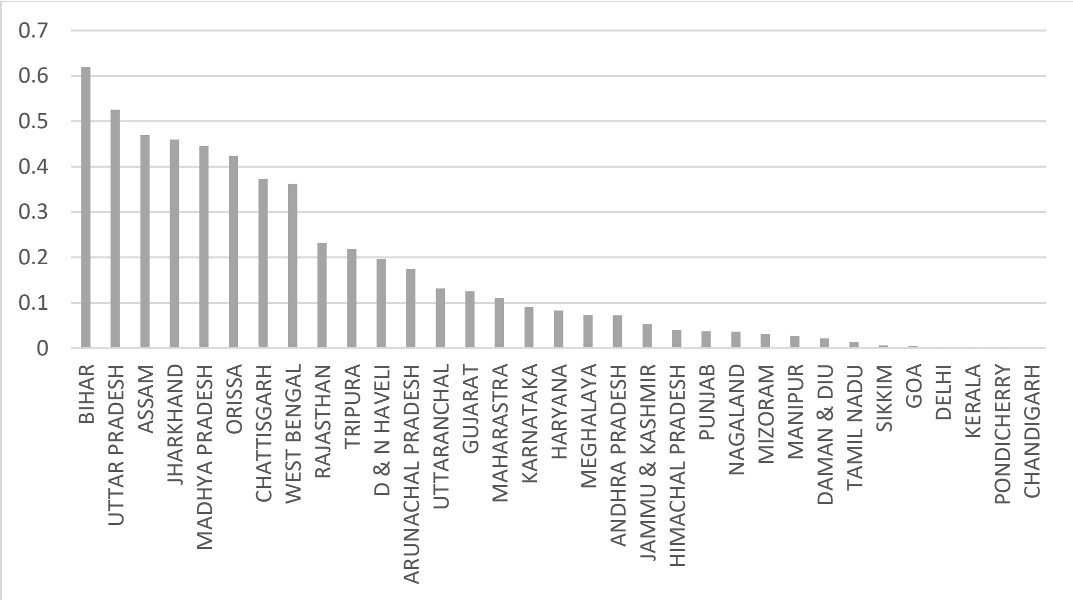
Notes: 1), 2) Our definition of child labour is applied; 3) main workers (working more than six months); 4) marginal workers (working less than six months); the definitions of child workers of the Indian Census do not correspond to the definitions used in IHDS or NSS.; Andaman & Nicobar and Lakshadweep are not included.

Appendix Table A.4. Estimates of child labour by Bayesian models

	Model 1 (IHDS)			Model 2 (NSS)			Model 3 (Combination, Poisson)			Model 4.2(Combination, Poisson Log-Normal)		
Age	No. of Child Labourers1)	95% PI		No. of Child Labourers1)	95% PI		No. of Child Labourers1)	95% PI		No. of Child Labourers1)	95% PI	
		LL2)	HL3)		LL2)	HL3)		LL2)	HL3)		LL2)	HL3)
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	939,958	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%)

B. Appendix for Chapter 5



Appendix Figure B.1. The proportion of the population in the lowest wealth quintile by state, India, 2011/12

Source: IHDS, 2011/12

Appendix Table B.1. The dependent variables of Models 2 and 3

a. Model 2 (informal labour)	b. Model 3 (formal labour)
Working hours of child labourers in any of these activities: family work in small family enterprises (≤ 4 employees) *or* casual agricultural and non-agricultural waged labour (paid daily) *or* farming	Working hours of child labourers in any of these activities: family work in large family firms (> 5 employees) *or* salaried work (paid monthly or annually)

Appendix Table B.2. Estimating a time threshold using educational harm of long working hours

Long working hours increase the proportion of out-of-school children. Using a threshold model, we can estimate the numbers of working hours at which the number of children who are out of schools is significantly extended. The dependent variables are the numbers of out-of-school children for each age group (ages 5–11, 12–14 and 15–17), grouped by the number of hours spent on work. We use a Poisson regression for each age group, using a risk rate (λ) and population. Most importantly, we fit the time index parameter, Θ, as a prior of λ, which indicates the working hours show which educational harms increase to a significant level. As a result, we found that among ages 12–14 and 15–17, the rates of being out-of-school grow to 0.6 (95-percent predictive intervals [0.58, 0.66]) and 0.75 ([0.72, 0.76]) each if children work longer than 5.5 hours a day. This estimated time index parameter ($\Theta = 5.5$, 95% PI [5.02–5.98] for ages 12–14, and [5.02, 5.97] for ages 15–17) is multiplied by seven; therefore, 38.5 hours a week can be used as a time threshold for those age groups. Among the youngest age group (ages 5–11), a time change point is not identified; therefore, we decide to keep the ILO time-threshold, which is at least one hour a week.

Appendix Table B.3. Summary of household occupational categories

Categories	Codes
Agricultural labourers	Paid daily (WS9=1) & agricultural labourers (WS4=63-67)
Non-agricultural labourers	Paid daily (WS9=1) & other labourers
Workers	1) Workers who are paid monthly or annually (WS9=2 or 3, excluding farmers and professionals) or 2) Self-employment workers excluding farmers and professionals
Marginal farmers (<1 hec.)	Farmers 1) FM38 2) WS4 or NF1B=61–62
Small farmers (1-2 hec.)	
Middle farmers (2-5 hec.)	
Large farmers (5+ hec.)	
Professionals	Professional workers (WS4 or NF1B = 0–29)

Notes: Categories indicate areas where household heads spend most of working hours; if the household heads do not have working hours, then we use the primary status code (RO7) instead.

Appendix Table B.4. Summary of items for wealth index

This study uses a confirmatory factor analysis with a logit function to calculate the wealth index. The selected items are drinking water source, electricity, toilet facilities, floor materials, telephone and TV. They are the items used in other related studies such as Demographics and Health Surveys (DHS) Wealth Index or International Wealth Index (IWI). Land ownership is excluded from this study, following the IWI method. We use factor analysis (FA) rather than a PCA. FA is more accurate in capturing a latent factor because a PCA can cause an underrepresentation of inter-correlations due to the overrepresentation of loadings (Widaman, 1993). Missing items are imputed using multiple imputation methods. Cronbach alpha test shows that the variables are strongly correlated (0.75). The ownership of a phone was constrained as 1. Electricity (coef=2.00), TV (coef=1.95) and floor (coef=1.73) have a strong positive coefficient value. Toilet

facilities (coef =1.13), water (coef=0.85), and roof (coef=0.62) also explained the latent variable. The variance of the latent variable is significant, indicating the appropriateness of the use of the variable as a proxy of the asset index.

(a). Summary of items

Variable name	Description	Obs	Mean	S.D.
CG17	Owning mobile phone=1, otherwise=0	204,553	0.85	0.36
CGTV	TV (black & white or colour)=1, otherwise=0	204,553	0.67	0.47
Toilet	Household toilet (semi-flush+)=1, otherwise=0	204,553	0.40	0.49
FU1	Electricity=1, otherwise=0	204,553	0.88	0.33
HQFLOOR	Floor - wood=1, mud=0	204,553	0.64	0.48
HQROOF	Roof - asbestos, metal, brick, stone, concrete=1, otherwise=0	204,553	0.66	0.47
WATER	Indoor piped drinking water=1, otherwise=0	204,553	0.31	0.46

Appendix Table B.5. Posterior predictive estimates of children's labour hours

	Hurdle Poisson				Truncated Poisson			
	Mean	Median	2.50%	97.50%	Mean	Median	2.50%	97.50%
Total labour hours, all	11.0	7.0	0.1	47.3	50.3	57.3	10.4	75.7
Total labour hours, girls	13.3	9.5	0.2	55.7	47.2	53.1	10.1	74.9
Total labour hours, boys	16.2	12.7	0.4	54.1	51.8	59.0	10.5	75.9
Informal labour hours, all	9.7	5.8	0.1	43.3	43.0	45.0	9.6	73.3
Informal labour hours, girls	8.2	3.9	0.1	45.5	43.5	46.4	9.8	75.7
Informal labour hours, boys	10.9	7.9	0.2	40.3	42.9	45.2	10.1	70.8
Formal labour hours, all	1.0	0.5	0.0	4.6	5.9	3.7	1.0	27.6
Formal labour hours, girls	9.9	4.9	0.1	55.6	2.6	1.6	1.0	14.3
Formal labour hours, boys	10.2	7.6	0.1	37.5	7.3	5.5	1.1	29.6
Total labour hours by household class								
Agr. Lab., Female	7.0	3.9	0.1	24.2	54.1	58.9	10.6	74.2
Agr. Lab., Male	11.4	6.0	0.2	40.5	58.3	62.8	15.4	77.5
Non-agr. Lab., Female	8.7	8.0	0.1	25.6	43.9	51.5	9.56	74.1
Non-agr. Lab., Male	14.5	8.8	0.4	61.5	50.1	57.4	10.4	72.8
Marginal farmers, Female	7.5	2.7	0.1	29.7	46.7	54.8	9.3	77.7
Marginal farmers, Male	4.1	2.1	0.2	15.5	50.7	59.8	10.7	81.4
Small farmers, Female	13.3	9.5	0.2	55.7	47.9	53.5	11.1	79.3
Small farmers, Male	16.2	12.7	0.4	54.1	46.5	52.2	7.2	77.9
Middle Farmers, Female	8.2	3.9	0.1	45.5	54.8	63.0	15.4	81.0
Middle Farmers, Male	10.9	7.9	0.2	40.3	53.8	64.0	18.8	71.5
Large Farmers, Female	9.9	4.9	0.1	55.6	45.6	54.9	7.5	76.0
Large Farmers, Male	10.2	7.6	0.1	37.5	37.4	42.8	8.8	63.0

Notes: Expected count in a hurdle Poisson = $p*\mu/(1-e^{-\mu})$; expected count in a Truncated Poisson = $\mu/(1-e^{-\mu})$

C. Appendix for Chapter 6

Appendix Table C.1. Observed numbers of out-of-school children by working hours (Zt)

Daily working hours[1)]	Count of out-of-school children			Sample no. of children		
	Ages 5–11	Ages 12–14	Ages 15–17	Ages 5–11	Ages 12–14	Ages 15–17
0	3,094	908	1,564	32,057	14,127	9,344
1	18	51	88	705	816	674
2	25	82	146	571	865	855
3	4	64	119	234	411	524
4	18	49	119	163	330	430
5	6	33	97	38	192	271
6	4	52	153	51	147	289
7	3	26	83	7	53	119
8	12	196	699	29	283	883
9	0	28	89	8	48	128
10	2	32	131	6	43	184
11	1	10	67	5	23	99
12	0	17	118	3	21	162
13	0	7	39	1	12	55
14	0	11	60	1	11	72
15	NA	1	17	NA	2	26
16	2	23	87	3	29	106
Total	3,189	1,590	3,676	33,882	17,413	14,221

Appendix Table C.2. Latent variables

The norm for women's work
Confirmatory factor analysis with an ordinal logit function was implemented to identify gender norms explaining people's attitudes about women's work. In implementing this, four items were chosen from the World Value Survey 2012. V45 took value 1 if people agree, 2 if people neither agree nor disagree and 3 if people disagree. V50, V52, and V54 were ordinal variables (1-4) indicating whether people agree/disagree or strongly agree/disagree. Missing values were excluded. The list of variables are below: (V45) When jobs are scarce, men should have more right to a job than women. (1=agree, 2=neither agree nor disagree, 3=disagree) (V50) When a mother works for pay, the children suffer. (1=strongly agree, 2=agree, 3=disagree 4=strongly disagree) (V52) A university education is more important for a boy than for a girl. (1=strongly agree, 2=agree, 3=disagree 4=strongly disagree) (V54) Being a housewife is just as fulfilling as working for pay (1=strongly agree, 2=agree, 3=disagree 4=strongly disagree) All variables had significant positive coefficients for the latent variable. The latent variable is associated with higher support for the higher education of girls (V52, coef=1.33); mother's work for pay (V50, coef=0.7); women's working for pay (V54, coef=0.52) and women's right to a job (V45 was constrained as 1). Cronbach alpha test showed the items are weakly correlated (0.44), but the reason for the index is not to have a great fit but to reduce the complexity of regression. The covariance of the gender norm index is significant, indicating that it represents well whether people have a supportive norm on women's employment.
Asset index category
Similarly, confirmatory factor analysis was used to create the asset index category based on the IHDS 2011/12. The selected variables were toilet facilities, floor materials, the ownership of a vehicle, phone, and TV. Cronbach alpha test showed that the variables are moderately correlated (0.66). The ownership of a phone was constrained as 1. TV (coef=1.56) and floor (coef=1.55) have a strong positive coefficient value. Toilet facilities (coef =1.05), roof (coef=0.63) and vehicle (coef=0.32) also explained the latent variable. The variance of the latent variable is significant, indicating the appropriateness of the use of the variable as a proxy of the asset index.

Appendix Table C.3. Correlation matrix

	Child labour	Female	Urban	Lowest assets	Land category	Dalit	Adivasi	Norm 1	Norm 2	Norm 3
Child labour	1.00									
Female	-0.11	1.00								
Urban	-0.37	0.00	1.00							
Lowest assets	0.54	0.00	-0.49	1.00						
Land category	0.00	0.00	-0.49	0.02	1.00					
Dalit	0.17	0.00	-0.10	0.12	-0.38	1.00				
Adivasi	0.20	0.00	-0.31	0.48	0.14	-0.14	1.00			
Norm 1 (seclusion)	0.22	0.00	-0.10	0.43	0.13	-0.10	0.09	1.00		
Norm 2 (benevolence)	-0.04	0.00	-0.30	-0.11	0.07	0.09	0.01	-0.08	1.00	
Norm 3 (women's work)	-0.25	0.00	0.19	-0.37	-0.07	0.14	-0.26	-0.52	-0.32	1.00

Appendix Table C.4. Test statistics of threshold models

Models	Bayesian p-value	MSE	DIC	Deviances	pD
Ages 5–11	0.45	375.62	173.6	138.95	34.7
Ages 12–14	0.48	190.49	212.7	210.70	2.0
Ages 15–17	0.65	429.00	221.1	219.10	2.0

Notes: a Bayesian p-value of the likelihood ratio statistics; MSE=Mean squared errors; DIC=deviance+pD; pD = var(deviance)/2

Appendix Table C.5. Out-of-sample test using the norm models

	Bayesian R-squared test			Bayesian P-value test			DIC		
Models	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Simulation 1	0.69	0.75	0.76	1.00	0.99	0.97	2176.70	2117.60	2099.30
Simulation 2	0.59	0.62	0.66	0.81	0.80	0.75	2273.60	2208.30	2188.20
Simulation 3	0.81	0.85	0.83	0.72	0.62	0.78	2160.10	2111.70	2093.40
Simulation 4	0.51	0.56	0.56	0.98	0.98	0.96	2153.10	2092.50	2056.20
Simulation 5	0.30	0.33	0.29	1.00	1.00	1.00	2252.00	2188.10	2162.30
Simulation 6	0.74	0.80	0.83	1.00	0.99	0.98	2204.20	2150.00	2131.40
Simulation 7	0.91	0.88	0.89	0.05	0.01	0.02	2231.20	2151.10	2129.30
Simulation 8	0.51	0.55	0.60	0.67	0.65	0.58	2250.30	2188.80	2171.60
Simulation 9	0.73	0.69	0.68	0.01	0.01	0.01	2249.00	2177.40	2157.30
Simulation 10	0.76	0.81	0.83	0.94	0.83	0.69	2220.30	2145.00	2121.40

Notes: a Bayesian p-value of the likelihood ratio statistics (Lunn et al, 2013)

Annexes

A. Summary of Interviews with the Stakeholders

1. The VV Giri National Labour Institute
2. Centre for Women's Development Studies
3. Centre for the Study of Regional Development, Jawaharlal Nehru University
4. Save the Children India
5. Carpet industry
6. The Child Rights and You (CRY)

1. The VV Giri National Labour Institute (Visiting date: 9 January 2018)
There is a large number of child labourers in India, estimated at 10.1 million in 2011 in the recent report from the VV Giri National Labour Institute (Samantroy et al., 2017). The report provides an investigation into child labour mainly using the Indian census and shows the regional differences in child labour over all states in India. UP and Bihar are the states with the highest numbers of child labourers. However, there is a great possibility of undercounting child labourers. As awareness increases, parents are less likely to report their children's participation in work. (Interviewees referred to their Nagaland project.) The institute has been actively engaged in amending the Child Labour Act in 2016, which was authorised in June 2016 and notified on September 2016. In January 2017, a technical advisory meeting was held, and the operational procedure is currently going on. A key concern is that any child work that is done voluntarily to help the family is not regarded as child labour. There are still many child workers or labourers who are engaged in home-based work or some jobs that do not require higher skill levels or education (e.g. moulding glass). Education is closely related to child work, and occasional or frequent absence in schooling must be closely related to its incidence. There are a few categories used to classify children's status: children who work, children who study, children who do both, or children who do not belong to any of these categories. Lack of schools and less accessibility due to social exclusion are also significant causes of child labour. At around age 12, children make a decision as to whether they drop out of school and start a job or continue their studies. Note that children's dropout rates vary by whether a survey is done before or after the enrolment. Religion and the caste system strongly affect the incidence of child labour, and therefore social dimensions should be considered in order to understand child labour. Key causes relating to child labour include female-headed households (especially regarding domestic

labour), child marriage (which is also regarded as child labour), and migration (as workers bringing their children who also engage in labour).
2. Centre for Women's Development Studies (Visiting date: 10 January 2018)
Child labour is a matter closely related to geographic characteristics such as poverty, education, and cultural background. The BIMARU (Bihar, Madhya Pradesh, Rajasthan, and Uttar Pradesh) States show a higher prevalence of child labour compared to others. Taking geographical information at the state level, as well as the district level, explains the causes of child labour. Child-labour marriage or migrants could provide a partial explanation of child labour. Figuring out hot spots of child labour and carefully assessing those areas could provide a deeper understanding of child labour. At ages 12 or 13 (around grade 8 in school), children decide whether to move onto higher education or not. Because of financial concerns, children have to decide to continue to study or drop out and work as full-time workers. Thus, the prevalence of child labour by age correlates with the drop-out rate.
3. Centre for the Study of Regional Development, Jawaharlal Nehru University (Visiting date: 16 January 2018)
Child labour should be understood in relation to education. Universal primary schooling in India was accepted in 2001, and the bill for the right to education was passed in 2009. It should be investigated that after grade 5, many students, especially girls, tend to drop out of school. Many children do not go beyond elementary education (dropping out at age 14). The low-quality education of the public schools, especially in rural areas, results in higher drop-out rates. Many of those who drop out of school become labourers; they are involved in diverse industrial sectors such as cottonseed, carpet, football, or brass work industries (Moradabad). Child-labour survey results are diverse according to the timing of the survey. The interviewee recommends using the recent information from the Ministry of Education, which provides more recent information on children’s drop-out rates from schools.

4. Save the Children India (Visiting dates: 12 and 17 January 2018)

Save the Children India has been supporting child labourers in the garment sector in Delhi. I visited the project site, Madanpur Khadar, in south-east Delhi, near the NOIDA industrial site where many readymade garment industries are located. This particular project currently includes 57 children across five streets who are involved in labour. The organisation aims to send children between ages 6 to 14 back to school, and children over age 15 to start vocational training.

Children normally work 3-4 hours daily, including Saturday and Sunday, typically helping their parents. If they go to school, they work less than those hours, for about 2 hours per day. Despite the support, many children either drop out or stay out of school. Most children are from households headed by migrant workers that are originally from UP, Bihar, West Bengal, and Assam. Enrolment in secondary education requires them to submit a certificate from their original schools, but because of distance, this is hardly possible. The most common characteristics shared by those child workers or labourers in these urban slums are as follows: having many siblings and many family members (5 to 6); dropping out from school or never having been to school; high incidence of parental unemployment or parents working in a lower-skilled sector with a total earning of about 7,000-8,000 rupees a month (rent for a small house costs 3,000 rupees per month); and having a Muslim family background.

5. Carpet industry (Visiting date: 19 January 2018)
Visit Site: Mirzapur Carpet industry belt – 1) Project Mala Higher Secondary School, and 2) Obeetee Private Limited factory, Varanasi, UP
In 1989, the Mala project began in Guria and Mirzapur, districts in Uttar Pradesh, where child labourers were spread over the whole carpet industry. The Mala project runs six schools covering primary (grade 1–5), middle (grade 6–8), secondary (grade 9–10) and higher secondary education (grade 11–12). Targets are child labourers and any children who belong to the families working in the carpet industry. The founder of this project is the CEO of the carpet company (Obeetee Private Limited), and so many of the students are from the families working in that company. According to the project implementers, no more children are working in their carpet industry. There is greater awareness regarding the illegality of child labour, and the company oversees the involvement of children in household handlooms using regular auditing and video investigation. Apparently, the carpet industry has been experiencing recession due to people favouring machine-made carpets instead of handmade ones. That could also be the reason for the decreasing number of child labourers in the carpet industry. However, this area is not entirely free from child labour, as children work in household production or in the service sector (e.g. cycle repairing, teashop, hotels, etc.). Children in the project schools tend to work after school within limited hours (1–3 hours), and so they are not child labourers. Children are sometimes absent, to help family farming during a harvest season. Some students, after the age of 14, decide to go for vocational training or start light work as preparation for full-time employment. Nonetheless, some children drop out from schools: for example, the girls' drop-out rate (at around age 12–13 years) is higher than the boys'.

6. Child Rights and You (CRY) (Visiting date: 24 January 2018)
CRY (Child Rights and You) has been working on supporting children's rights focusing on stopping child labour in India since 1979. The organisation applies a rights-based approach, looking at a close relationship between education and work. High drop-out rates among Indian children are caused by their participation in the labour force. Dropping out of schools occurs in grades 5–6 and then again in grades 8–9. There is intense family pressure on children to work and financial demand. According to the CRY report that uses the Indian censuses, child labour had decreased only by 2.2 percent between 2001 and 2010, which is too slow to meet the targets of ending child labour in India (5 to 14: 10 million and 5 to 18: 22 million in 2011/12). Over the same period, the total number of child labourers aged 5–9 years increased owing to increased migration from rural to urban areas during those years. Child labourers are spread over many industries such as agriculture, the service sector, handicrafts, and textiles. Industrial sectors that employ child labourers are more diverse in the cities. Economic growth in urban areas, such as Delhi, Mumbai, Surat, Lucknow, and Chennai, has resulted in higher demand for child labour. Small-scale units and production sites (e.g. production of bags or leather in a small factory) are more related to child labour. Few children can be paid workers in a semi-skilled unit. In addition, children could work at the household level. To address child labour in India, transforming the perceptions of parents has the most critical role in preventing child labour, along with efforts from the government and community.

B. Wealth Thresholds of the NSS 2011/12

Annex Table B.1. Wealth thresholds of the NSS 2011/12 in rural and urban areas

SSS[1)]	Composition of SSS	Number of households to be surveyed	
		FSU[2)] without hg/sb formation	FSU with hg/sb formation (for each hg/sb)
Rural			
SSS 1	Relatively affluent households	2	1
SSS 2	Of the remaining, households having principal earning from non-agricultural activity	4	2
SSS 3	Other households	2	1
Urban			
SSS 1	Households having MPCE[3)] of top 10% of urban population (MPCE> B[4)])	2	1
SSS 2	Households having MPCE of middle 60% of urban population (A[5)] ≤ MPCE ≤ B)	4	2
SSS 3	Households having MPCE of bottom 30% of the urban population (MPCE < A)	2	1

Source: GOI (n.d.)

Note: [1)] SSS: Second Stage Stratum, [2)] FSU: First Stage Units, [3)] MPCE: Monthly Per Capita Expenditure, [4)] B: each NSS state-region's top 10 percent MPCE from the NSS 2009/10 data, [5)] A: each NSS state-region's bottom 30 percent MPCE from the NSS 2009/10 data

C. Selection of Statistical Package

Fitting a complex model is another advantage from using a Bayesian analysis method. This study utilises the software packages R, JAGS and Stan to run the Bayesian model. JAGS and Stan are the most commonly used programming languages for Bayesian data analysis, and open sources, written in C++, can be run easily in R. In this thesis, I used both packages, RJags and Rstan, which can run JAGS and Stan in the R environment. Chapters 4 and 6 are based on R2Jags, and Chapter 5 uses Rstan. In the following paragraphs, I sum up the different processes and strengths of both programs.

JAGS is based on Gibbs sampling (Gelfand and Smith, 1990), which is a useful tool in dealing with large sample sizes, flexible in writing and less limited when adding models or parameters. In JAGS, adding a few more parameters or calculations does not significantly affect the running time. It is a clear difference from Stan, in which speed is strongly reliant on the number of parameters or functions. Because of its flexibility in setting parameters, using a simple model that requires more diverse sets of parameters is efficient in JAGS. The number of available models is less than that of Stan. Thus, in JAGS, if a specific model, e.g. Hurdle Poisson model, is not available, diversions should be made. JAGS is efficient in handling a large number of samples, but it requires a large number of iterations. Usually, more than over 100,000 iterations are required. I used R2Jags to run models in Chapters 4 and 6, which are based on aggregate samples and comparably simple models, and it can thus efficiently generate samples without any convergence issues.

On the other hand, Stan, based on Hamiltonian Monte Carlo sampling (Girolami and Calderhead, 2011), uses a small number of iterations to converge, such as 1000 or more. Thus, it uses a smaller amount of memory space and takes a shorter time if the model is not too complex, and the sample size is reasonably large. The time needed to run a model depends on how efficiently the code is written. In this study, a Bayesian hurdle model with a large sample size was required for the selection of the suitable programming language in running the model. I have compared both JAGS and Stan in each stage: the first stage (with N=51,551) takes longer in R2jags than in Rstan, but the second stage (N=2,926) is similar in both models. Again, the speed of Stan is strongly affected by the number of parameters, and to reduce the time consumed to run a model, removing unnecessary parameters is the most important thing.

Annex Table C.1. Comparisons of R2jags and Rstan

	Programming language	Number of data inputs	(Iterations – burn-in) * No. of chains	Approx. time
Chapter 4	R2jags	910	(300,000 – 70,000) * 2	15-20 min.
Chapter 5	Rstan	1st stage: 51,551	1st stage: (1,500 – 500) * 2	1st stage: 5 hrs
		2nd stage: 2,926	2nd stage: (1,200 – 500) * 2	2nd stage: 2 hrs
Chapter 6	R2jags	442	(300,000 – 100,000) * 4	15-20 min.

D. Examples of Model Codes

1. Chapter 4, Model 4-2, R2jags
2. Chapter 5, Model 1 – first stage, Rstan
3. Chapter 5, Model 1 – second stage, Rstan
4. Chapter 6, Threshold model, R2jags
5. Chapter 6, Norm Model 1, R2jags

1. Chapter 4, Model 4-2, R2jags

```
# Run the Poisson
# using R2jags
# Requires:  R2jags, dplyr, lattice

library("R2jags")
library("dplyr")
library("lattice")

###########################################################
################## Download data   ######################
###########################################################

data <- dget("E:/…/data.R")

###########################################################
##################### Set parameters  ###################
###########################################################

params <- c("undercount", "y.pred", "total.rate", "rate.pred", "beta", "tau", "sigma",
"tau.a", "sigma.a", "tau.b", "sigma.b", "sum.y.pred", "upsilon", "total.y.pred.under15",
"total.y.pred")

###########################################################
################### Model 4.2 ########################
###########################################################

model4.2<- function() {
   for( i in 1 : 13) { # i:age
   for( j in 1 : 35) {# j:state

    y.a[i, j] ~ dpois(upsilon[i]*n.a[i, j]*mu.a[i,j]) #y.a: the number of child labour in IHDS
    mu.a[i,j] <- exp(log.mu.a[i,j])
    log.mu.a[i,j] <- log(psi[i,j]) + lambda.a[i,j]

    y.b[i, j] ~ dpois(n.b[i, j]*mu.b[i,j]) #y.b: the number of child labour in NSS
    mu.b[i,j] <- exp(log.mu.b[i,j])
    log.mu.b[i,j] <- log(psi[i,j]) + lambda.b[i,j]

    log(psi[i, j]) <- beta[1]  + beta[2] * x[i]  + beta[3] * log(K[i, j]/N[i, j]) + lambda[i,j]
```

```
      }
  }
  for( i in 1 : 13) {   for( j in 1 : 35) {# j:state
     y.pred[i,j] ~ dpois (N[i,j] * psi[i,j])
        }
  }

  # Calculate the number of child labourers per age per state

   for (i in 1 : 13) {
    sum.y.pred[i]  <- sum(y.pred[i,])
    rate.pred[i] <- sum.y.pred[i]/sum(N[i,])
    }

    total.y.pred.under15 <-
sum.y.pred[1]+sum.y.pred[2]+sum.y.pred[3]+sum.y.pred[4]+sum.y.pred[5]+sum.y.pred[
6]+sum.y.pred[7]+sum.y.pred[8]+sum.y.pred[9]+sum.y.pred[10]
    total.y.pred <- sum(sum.y.pred[])

########################################################
#################### Priors ############################
########################################################

   for( i in 1 : 3) { beta[i]  ~ dnorm(0.0,1.0E-6)}

  # Inverse Gamma priors
  for( i in 1 : 13) {   for( j in 1 : 35) {
  lambda[i,j] ~  dnorm(0.0, tau)
  lambda.a[i,j] ~ dnorm(0.0, tau.a)
  lambda.b[i,j] ~  dnorm(0.0, tau.b)}}
  tau ~  dgamma(0.5, 0.5)
  tau.a ~  dgamma(0.5, 0.5)
  tau.b ~  dgamma(0.5, 0.5)
  sigma <- 1/sqrt(tau)
  sigma.a <- 1/sqrt(tau.a)
  sigma.b <- 1/sqrt(tau.b)

  for (i in 1 : 13) {
  upsilon[i] ~ dunif(0,1)      } #undercount
    }

########################################################
################### Run model #######################
########################################################

jags.fit.model4.2 <- jags.parallel(data=data, inits=NULL, params, model.file=model4.2,
n.chains=2, n.thin=8, n.iter=350000, n.burnin=30000, DIC=TRUE, jags.module = c("glm",
"dic"))
```

2. Chapter 5, Model 1 – first stage, Rstan

```
# Run Probit
# using STAN
# Requires:rstan, tidyverse, tictoc, readstata13, dplyr, lattice
rm(list=ls(all=TRUE))
library("rstan")
library("tidyverse")
library("tictoc")
pkgbuild::has_build_tools(debug = TRUE)
writeLines('PATH="${RTOOLS40_HOME}\\usr\\bin;${PATH}"', con = "~/.Renviron")
Sys.which("make")
library("readstata13")
library("tictoc")
library("dplyr")
library("lattice")
library("tidyverse")

########################################################
################## Download data   ######################
########################################################

setwd("D:/…/")
datanew1 <- read.dta13("data.dta")

########################################################
############## Data cleaning for models ###################
########################################################

line.data <-list(

female = datanew1$female,
agec = datanew1$agec,
agecsq = datanew1$agecsq,
urban = datanew1$urban,
migration = datanew1$migration,
femheadedhh = datanew1$femheadedhh,
sgroup2 = datanew1$sgroup2,
sgroup3 = datanew1$sgroup3,
sgroup4 = datanew1$sgroup4,
sgroup5 = datanew1$sgroup5,
sgroupinter3 = datanew1$sgroupinter3,
sgroupinter4 = datanew1$sgroupinter4,
sgroupinter5 = datanew1$sgroupinter5,
sgroupinter6 = datanew1$sgroupinter6,
classcategory2 = datanew1$classcategory2,
classcategory3 = datanew1$classcategory3,
classcategory4 = datanew1$classcategory4,
classcategory5 = datanew1$classcategory5,
classcategory6 = datanew1$classcategory6,
```

```
classcategory7 = datanew1$classcategory7,
classcategory8 = datanew1$classcategory8,
classcategory9 = datanew1$classcategory9,
asset2 = datanew1$asset2,
asset3 = datanew1$asset3,
asset4 = datanew1$asset4,
asset5 = datanew1$asset5,
state = datanew1$state,
fwt = datanew1$fwt,
z = datanew1$z,
N=51551)

#######################################################
##################Model 1: 1st stage ###################
#######################################################


model <- "
data {
  // Covariates
  int  N;
  int <lower=0, upper=1> female[N];
  real agec[N];
  real agecsq[N];
  int <lower=0, upper=1> urban[N];
  int <lower=0, upper=1> migration[N];
  int <lower=0, upper=1> femheadedhh[N];
  int <lower=0, upper=1> sgroup2[N];
  int <lower=0, upper=1> sgroup3[N];
  int <lower=0, upper=1> sgroup4[N];
  int <lower=0, upper=1> sgroup5[N];
  int <lower=0, upper=1> sgroupinter3[N];
  int <lower=0, upper=1> sgroupinter4[N];
  int <lower=0, upper=1> sgroupinter5[N];
  int <lower=0, upper=1> sgroupinter6[N];
  int <lower=0, upper=1> classcategory2[N];
  int <lower=0, upper=1> classcategory3[N];
  int <lower=0, upper=1> classcategory4[N];
  int <lower=0, upper=1> classcategory5[N];
  int <lower=0, upper=1> classcategory6[N];
  int <lower=0, upper=1> classcategory7[N];
  int <lower=0, upper=1> classcategory8[N];
  int <lower=0, upper=1> classcategory9[N];
  int <lower=0, upper=1> asset2[N];
  int <lower=0, upper=1> asset3[N];
  int <lower=0, upper=1> asset4[N];
  int <lower=0, upper=1> asset5[N];
  int <lower=1> fwt[N];
  int <lower=1, upper=33> state[N];
  int <lower=0, upper=1> z[N];
}
```

```
parameters {
real alpha[23];
real a[33];
}

model {

  alpha ~ normal(0, 1);
  a ~ normal(0, 1);

  for (i in 1:N) {
  z[i] ~ bernoulli(Phi_approx(alpha[1] + alpha[2]*female[i] + alpha[3]*agec[i]
      + alpha[4]*agecsq[i]  + alpha[5]*urban[i] + alpha[6]*migration[i]
      + alpha[7]*femheadedhh[i]
      + alpha[8]*sgroup2[i]
      + alpha[9]*sgroup3[i]
      + alpha[10]*sgroup4[i]
      + alpha[11]*sgroup5[i]
      + alpha[12]*sgroupinter3[i]
      + alpha[13]*sgroupinter4[i]
      + alpha[14]*sgroupinter5[i]
      + alpha[15]*sgroupinter6[i]
      + alpha[16]*classcategory2[i]
      + alpha[17]*classcategory3[i]
      + alpha[18]*classcategory4[i]
      + alpha[19]*classcategory5[i]
      + alpha[20]*classcategory6[i]
      + alpha[21]*classcategory7[i]
      + alpha[22]*classcategory8[i]
      + alpha[23]*classcategory9[i]
      + a[state[i]]
      + log(fwt[i])));
 }
 }
"

init <- function() list(alpha=rep(0.05, 23), a=rep(0.05, 33))

tic()
fit <- stan(model_code = model, data = line.data, init=init, iter = 1500, warmup=500,
chains = 2)
toc()
```

3. Chapter 5, Model 1 – second stage, Rstan

```
# Run the Poisson model
# using STAN
# Requires:  rstan, tidyverse, tictoc, readstata13, dplyr, lattice

rm(list=ls(all=TRUE))
```

```
library(rstan)
library("tidyverse")
library("tictoc")
pkgbuild::has_build_tools(debug = TRUE)
writeLines('PATH="${RTOOLS40_HOME}\\usr\\bin;${PATH}"', con = "~/.Renviron")
Sys.which("make")
library("readstata13")
library("dplyr")
library("lattice")

####################################################
################## Download data   #####################
####################################################
setwd("D:/…/")
datanew1 <- read.dta13("Data/newdata0808.dta")

####################################################
############## Data cleaning for models ###################
####################################################
line.data <-list( "CLhours" = datanew1$CLhours,
p = datanew1$p,
female = datanew1$female,
agec = datanew1$agec,
agecsq = datanew1$agecsq,
urban = datanew1$urban,
sgroup2 = datanew1$sgroup2,
sgroup3 = datanew1$sgroup3,
sgroup4 = datanew1$sgroup4,
sgroup5 = datanew1$sgroup5,
classcategory2 = datanew1$classcategory2,
classcategory3 = datanew1$classcategory3,
classcategory4 = datanew1$classcategory4,
classcategory5 = datanew1$classcategory5,
classcategory6 = datanew1$classcategory6,
classcategory7 = datanew1$classcategory7,
classcategory8 = datanew1$classcategory8,
classcategory9 = datanew1$classcategory9,
asset2 = datanew1$asset2,
asset3 = datanew1$asset3,
asset4 = datanew1$asset4,
asset5 = datanew1$asset5,
interclass2 = datanew1$interclass2,
interclass3 = datanew1$interclass3,
interclass4 = datanew1$interclass4,
interclass5 = datanew1$interclass5,
interclass6 = datanew1$interclass6,
interclass7 = datanew1$interclass7,
interclass8 = datanew1$interclass8,
interclass9 = datanew1$interclass9,
asset_f2 = datanew1$asset_f2,
asset_f3 = datanew1$asset_f3,
```

```
asset_f4 = datanew1$asset_f4,
asset_f5 = datanew1$asset_f5,
fwt = datanew1$fwt,
N=2926)

######################################################
###################Model 1: 2nd stage #################
######################################################
model <- "
 data {

   // Covariates
   int  N;
   int <lower=0, upper=1> female[N];
   real agec[N];
   real agecsq[N];
   int <lower=0, upper=1> urban[N];
   int <lower=0, upper=1> sgroup2[N];
   int <lower=0, upper=1> sgroup3[N];
   int <lower=0, upper=1> sgroup4[N];
   int <lower=0, upper=1> sgroup5[N];
   int <lower=0, upper=1> classcategory2[N];
   int <lower=0, upper=1> classcategory3[N];
   int <lower=0, upper=1> classcategory4[N];
   int <lower=0, upper=1> classcategory5[N];
   int <lower=0, upper=1> classcategory6[N];
   int <lower=0, upper=1> classcategory7[N];
   int <lower=0, upper=1> classcategory8[N];
   int <lower=0, upper=1> classcategory9[N];
   int <lower=0, upper=1> asset2[N];
   int <lower=0, upper=1> asset3[N];
   int <lower=0, upper=1> asset4[N];
   int <lower=0, upper=1> asset5[N];
   int <lower=0, upper=1> interclass2[N];
   int <lower=0, upper=1> interclass3[N];
   int <lower=0, upper=1> interclass4[N];
   int <lower=0, upper=1> interclass5[N];
   int <lower=0, upper=1> interclass6[N];
   int <lower=0, upper=1> interclass7[N];
   int <lower=0, upper=1> interclass8[N];
   int <lower=0, upper=1> interclass9[N];
   int <lower=0, upper=1> asset_f2[N];
   int <lower=0, upper=1> asset_f3[N];
   int <lower=0, upper=1> asset_f4[N];
   int <lower=0, upper=1> asset_f5[N];
   int <lower=1, upper=11> fwt[N];
   int <lower=1, upper=112> CLhours[N];
   real <lower=0, upper=1>  p[N];

}
```

```
parameters {
real beta[33];
}

transformed parameters  {

  real  <lower=0> mu[N];
  real  <lower=0> CLhours_pred[N];

  for (i in 1:N) {
    // Linear predictor
    mu[i] = exp(beta[1]+beta[2]*female[i] + beta[3]*agec[i]
     + beta[4]*agecsq[i]  + beta[5]*urban[i]
     + beta[6]*sgroup2[i]
     + beta[7]*sgroup3[i]
     + beta[8]*sgroup4[i]
     + beta[9]*sgroup5[i]
     + beta[10]*classcategory2[i]
     + beta[11]*classcategory3[i]
     + beta[12]*classcategory4[i]
     + beta[13]*classcategory5[i]
     + beta[14]*classcategory6[i]
     + beta[15]*classcategory7[i]
     + beta[16]*classcategory8[i]
     + beta[17]*classcategory9[i]
     + beta[18]*asset2[i]
     + beta[19]*asset3[i]
     + beta[20]*asset4[i]
     + beta[21]*asset5[i]
     + beta[22]*interclass2[i]
     + beta[23]*interclass3[i]
     + beta[24]*interclass4[i]
     + beta[25]*interclass5[i]
     + beta[26]*interclass6[i]
     + beta[27]*interclass7[i]
     + beta[28]*interclass8[i]
     + beta[29]*interclass9[i]
     + beta[30]*asset_f2[i]
     + beta[31]*asset_f3[i]
     + beta[32]*asset_f4[i]
     + beta[33]*asset_f5[i]);

    CLhours_pred[i]=p[i]*mu[i]/(1-exp(-mu[i]));

}
}

model {
  for (i in 1:N) {
  CLhours[i] ~ poisson(mu[i])T[0, 112];
 }
```

```
}

"
init <- function() list(beta=rep(0.1, 33))

########################################################
#################### Run model ########################
########################################################
tic()
fit <- stan(model_code = model, data = line.data, init=init, iter = 1200, warmup=500,
chains = 2)
toc()
```

4. Chapter 6, Threshold model, R2jags

```
# Run the Poisson model
# using R2jags
# Requires:  R2jags, dplyr, lattice

library("R2jags")
library("dplyr")
library("lattice")

########################################################
################## Download data   #####################
########################################################
data2 <- dget("E:/…/data_edu_02032020.R")

########################################################
######## Model3 out-of-school children among ages 15-17#####
########################################################
model3<- function() {

  for( h in 1 : 17) {
     z3[h] ~ dpois(lambda[h]*n3[h])
     lambda[h] <- lambda1*step(tau - hour[h]- 0.0000000001) +  lambda2*step(hour[h]-
tau)
     z.pred[h] ~ dpois(n3[h] * lambda[h])
     rate0.pred[h] <- z.pred[h]/n3[h]
     rate.pred[h] <- min(rate0.pred[h], 1)
     ll[h] <- z3[h]*log(lambda[h]*n3[h]) - lambda[h]*n3[h] - loggam(z3[h]+1)-log(1-exp(-
lambda[h]*n3[h]))
  }

  lambda1 ~ dunif(0,1)
  lambda2 ~ dunif(0,1)
  tau ~ dunif(0, 16)

 #mse
  mse <- sum((z.pred[]-z3[]))^2/16
```

```
#Bayesian P-value of likelihood ratio
 for (k in 1:17) {
  logE[k] <- ll[k]
  m[k] <- z3[k]
  m.pred[k] <- z.pred[k]
  LR[k] <- 2*m[k]*(log(m[k]+0.0000001)-logE[k])
  LR.pred[k] <-2*m.pred[k]*(log(m.pred[k]+0.0000001)-logE[k])
 }

 obs <- sum(LR[])
 sim <- sum(LR.pred[])
 p.value <- step(obs - sim)
 }

########################################################
##################### Set parameters  ###################
########################################################
params <- c("m", "m.pred", "mse", "p.value", "lambda", "lambda1", "lambda2", "tau",
"z.pred", "rate.pred")

########################################################
#################### Run model ########################
########################################################
jags.model3 <- jags.parallel(data=data2, inits=NULL, params, model.file=model3,
n.chains=2, n.thin=4, n.iter=150000,n.burnin=30000, DIC=TRUE, jags.module = c("glm",
"dic"))
```

5. Chapter 6, Model 1, R2jags

```
# Run the Poisson model
# using R2jags
# Requires: R2jags, dplyr, lattice
library("R2jags")
library("dplyr")
library("lattice")

########################################################
################## Download data   #####################
########################################################
data1 <- dget("E:/…/Data_mainmodel_1208.R")

########################################################
############## Data cleaning for models ###################
########################################################
data <-list( "cl" = rbind(data1$cl.m, data1$cl.f),
             "state" = rbind(data1$state, data1$state),
             "urban" = rbind(data1$urban, data1$urban),
             "purdah" = rbind(data1$purdah, data1$purdah),
             "land" = rbind(data1$land, data1$land),
             "n" = rbind(data1$n.m, data1$n.f),
             "sgroup1" = rbind(data1$sgroup1, data1$sgroup1),
```

```
              "sgroup2" = rbind(data1$sgroup2, data1$sgroup2),
              "asset1" = rbind(data1$asset1, data1$asset1),
              "noland" = rbind(data1$noland, data1$noland),
              "v74" = data1$v74,
              "genderwork" = data1$genderwork)

female0 <- matrix( 0 , 13 , 17 , byrow = TRUE )
female1 <- matrix( 1 , 13 , 17 , byrow = TRUE )
data$female <- rbind(female0, female1)

#######################################################
######################### Model 1#####################
#######################################################
# i=household class, j=district

model1 <- function() {

  for(i in 1 : 26) {
    for (j in 1 : 17) {
      cl[i,j] ~ dpois(mu[i,j])
      log(mu[i,j]) <- beta[1] + beta[2]*female[i,j] + beta[3]*urban[i,j]+ beta[4]*asset1[i,j]+
beta[5]*land[i,j]+ beta[6]*sgroup1[i,j]+ beta[7]*sgroup2[i,j] + beta[8]*purdah[i,j]
       + log(n[i,j])
  }}

# multivariate normal prior
beta[1:8]~dmnorm(m.a[1:8],t.a[1:8,1:8])

  t.a[1,1]<-0.1
  t.a[2,2]<-0.1
  t.a[3,3]<-0.1
  t.a[4,4]<-0.1
  t.a[5,5]<-0.1
  t.a[6,6]<-0.1
  t.a[7,7]<-0.1
  t.a[8,8]<-0.1

for (i in 1:7)  {

  for (j in (i+1):8)
  {t.a[j,i]   <-0 }

  for (j in 1:i)
  { t.a[j,i+1]  <-0}}

 for(i in 1 : 8) {
   m.a[i] <- 0 }

# predicted count
  for( i in 1 : 26)
  {for( j in 1 : 17) {
```

```
    y.pred[i,j] ~ dpois (mu[i,j])
  }}

#Bayesian R-squared
for( i in 1 : 26)
  { for( j in 1 : 17) {
err[i,j] <- cl[i,j] - y.pred[i,j]
}
}

r_sq <- pow(sd(y.pred), 2) / (pow(sd(y.pred), 2) + (pow(sd(err), 2)))
}

inits1 <- function(){list("beta"=rep(0.1, 8))}

#########################################################
##################### Set parameters  ###################
#########################################################
params <- c("y.pred", "r_sq", "beta")

#########################################################
#################### Run model ########################
#########################################################
jags.model1 <- jags.parallel(data=data, inits=inits1, params, model.file=model1,
n.chains=4, n.thin=4, n.iter=300000,n.burnin=100000, DIC=TRUE, jags.module = c("glm",
"dic"))
```