HARMFUL CHILD LABOR: A THEORETICAL AND EMPIRICAL ANALYSIS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Dhushyanth Raju

January 2011

HARMFUL CHILD LABOR: A THEORETICAL AND EMPIRICAL ANALYSIS

Dhushyanth Raju, Ph. D.

Cornell University 2011

This dissertation comprises of three chapters on the economics of harmful child labor. The first chapter is theoretical. Using household survey data from the Philippines, the second and third chapters examine whether a model assumption and an explanation for a model prediction in the first chapter are supported empirically.

In Chapter 1, I model the labor market and welfare effects of banning harmful child labor. The effects are examined in two informational cases: (1) the parent has perfect information on harmful child labor and (2) the parent does not. The effects under both cases are contrasted between when the parent is the welfare evaluator for the household and when the child is, given imperfect parental altruism. Under both informational cases, the ban generates re-equilibrating labor market adjustments that expand employment and reduce wages in the non-harmful child labor market, as well as reduce child labor force participation. Under the first case, the ban is welfare-reducing. Under the second case, it is generally welfare-reducing; under special conditions, it can be welfare-improving. Under both informational cases and when the child is the welfare evaluator, a ban is generally welfare-ambiguous; under special conditions, it can be welfare-improving.

In Chapter 2, I examine the existence and magnitude of positive compensating wages for harmful child labor. Among the various harmful child labor measures examined, I find consistent evidence of a large and significant earnings premium for physically-strenuous labor at both the conditional mean and median. The result at the

conditional mean is largely driven by the large and increasing premia as one moves down the lower half of the conditional earnings distribution.

In Chapter 3, proxying the asymmetry in preferences and power statuses between the parent and child by the contradictory response of the parent when the child reports a work injury or illness, I examine whether parent-child injury report mismatches have an impact on the probability of harmful child labor. I find consistent evidence that mismatches have a large and significant positive effect on the probability of harmful child labor.

BIOGRAPHICAL SKETCH

Dhushyanth Raju is from India. He received a B.A. degree in economics from Grinnell College, USA in 1997 and a M.A. degree in economics from Cornell University, USA in 2005.

ACKNOWLEDGMENTS

I would like to thank my Special Committee members, Gary S. Fields, Kaushik Basu, Nancy Chau and George Jakubson, for their guidance and encouragement. I would also like to express my gratitude to my wife, Ruchira Bhattamishra, for her constant support, encouragement, and assistance. Responsibility for all findings and any errors are my own.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

BANNING HARMFUL CHILD LABOR:
A LABOR MARKET AND WELFARE ANALYSIS

1.1. Introduction

Child labor remains a mass phenomenon in much of the developing world, particularly in its poorer parts such as sub-Saharan Africa and South Asia. The International Labor Organization (ILO), the major source of statistics on the extent of child labor worldwide, estimates that 191 million children ages 5–14 years were economically active in 2004, of which, 166 million were categorized as child laborers (Hagemann et al. 2006).[1] More disconcertingly, 74 million of these 166 million child workers (or over 40%) were considered to be employed in harmful or exploitative situations or conditions. Of these, roughly 8 million children were deemed to be in what are termed as unconditional worst forms of child labor (ILO 2002).[2]

Though informative, child labor statistics were not always produced along the lines reported above, and, in fact, are a rather recent development made possible by changes on two mutually related fronts: the evolution in the notion of harmful child labor in policy discussions (demand), and the collection of more detailed data on child labor (supply). In the early stages of the modern policy debate on harm and child labor, it was commonly perceived that *all* child labor was harmful. Any distinctions drawn between different types of child labor were, for the most part, seen as artificial

---

[1] Economic activity covers all market production (paid work) and certain types of non-market production (unpaid work), including production of goods for own use. Child labor consists of all children under age 15 years who are economically active excluding (1) those who are under age 5 years and (2) those between ages 12–14 years who spend less than 14 hours a week on their jobs, unless their activities or occupations are hazardous by nature or circumstance. Added to this are children ages 15–17 years in the worst forms of child labor (ILO 2002).

[2] The worst forms of child labor refers to child labor in the context of slavery or slave-like conditions, in prostitution or pornography, in illicit activities such as drug trafficking, or in conditions that are likely to adversely affect the health or safety of the children involved. Unconditional worst forms of child labor exclude the last type.

and, thus, irrelevant. This perspective was primarily based on the fact that child labor often occurred at the expense of schooling (i.e., human capital accumulation), which, in turn, likely adversely affected the employment and earnings prospects of the child in adulthood. This view has since been superseded by the growing recognition that not all child labor necessarily interferes with schooling and that some child labor might actually *enable* school attendance by augmenting scarce household income. Moreover, there are cases of child labor that function basically as apprenticeships or on-the-job training, bestowing the child with potentially valuable, durable skills in a particular trade. Notwithstanding, policymakers have increasingly become aware that child labor is, in fact, a highly complex and diverse phenomenon that belies a ready solution and that attempting to distinguish between different forms of child labor would indeed be a meaningful exercise, especially given its bearing on the formulation and prioritization of appropriate policy.

In light of this, the course of the debate has gradually but perceptibly shifted from viewing child labor per se as inherently malignant to focusing more directly on certain distinct forms of child labor, specifically, those which (potentially) impair the physical and psychological development and health of the child. The adoption of ILO Convention No. 182 in 1999, which calls for immediate and effective steps to bring about the elimination of the worse forms of child labor, as well as its rapid and broad ratification by countries, strongly reflects this shift in the policy discussion and the growing international agreement that such practices are even less tolerable.[3] The intensification of efforts by individual governments, either independently or in partnership with international agencies, to address harmful and exploitative forms of

[3] As of the end of 2002, 132 countries had ratified Convention No. 182. Several countries are currently in the process of taking the necessary internal steps that will lead to eventual ratification. The pace of ratification of Convention No. 182 has been the fastest of any ILO Convention.

child labor also demonstrates the special status afforded to combating this insidious strain in national child labor eradication campaigns.

This chapter focuses on harmful child labor for two reasons. First, harmful child labor is qualitatively different from child labor generally because employment no longer functions purely as a means of acquiring valuable income, where the possible disruption in schooling is the only issue of concern, but, rather, also potentially constitutes a direct source of both short- and long-term adverse health effects. Consequently, this aspect of the costs tied to child labor merits its own independent examination. Second, understanding the causes of harmful child labor, its character, and the settings in which it arises is necessary in order to inform the ongoing policy discussion in line with its current focus on harmful child labor.

The aim of this chapter is to examine the labor market and welfare effects of banning harmful child labor (with perfect enforcement) because it arises in open labor market settings as opposed to situations in which the child labor decision by households is a fait accompli (e.g., child servitude, child trafficking). Harmful child labor is defined in a dynamic sense, namely work-related harm experienced as a child results in detrimental health effects which manifest themselves in the future with certainty, impairing the earnings potential of the child as an adult. The labor market analysis consists of examining the labor market participation, wage and employment patterns of children before and after the imposition of the ban. The welfare analysis comprises of examining the utilities of households before and after the ban using the method of welfare dominance. Household welfare is evaluated by taking the parent's utility function as the household welfare function; it is also evaluated by taking the child's utility function as the household welfare function, where the preferences of the parent and child within the household are considered to differ due to imperfect parental altruism. The economic effects of a ban are studied in two distinct settings.

The first setting, which the benchmark model is based on, is one in which parents are fully aware of the nature of harm associated with different child employment environments and accurately forecast the effect of this harm on the future wages of children. The second setting is one in which parents have problems in obtaining and properly processing relevant information regarding work-related harm, consequently, hampering their ability to correctly forecast future wages.

The remaining sections of the chapter are organized as follows. Section 1.2 presents a brief review of the related literature. Section 1.3 provides our operational definition of harmful child labor. Section 1.4 develops the benchmark model in which informational and perceptional problems related to work-related harm are absent. Section 1.5 discusses the labor market and welfare effects of banning harmful child labor in the context of our benchmark model. Section 1.6 extends the benchmark model to examine the labor market and welfare implications of banning harmful child labor in a setting in which informational and perceptional problems related to work-related harm are prevalent. Section 1.7 examines the welfare implications of banning harmful child labor under both informational settings when assessed by the child given divergent preferences between the parent (the decision-maker) and the child (the decision-implementer). Finally, Section 1.8 summarizes the main results and discusses some policy implications.

1.2. Related literature

There is a sizable and well-developed corpus of economic literature on child labor generally. Motivated principally by the desire to inform the ongoing policy discussion on child labor, the literature delves into the causes and consequences of child labor. A segment of this literature also examines the welfare implications of various proposed measures to address child labor, chief among which is its outright prohibition. It is now well-established empirically that households, given the lack of viable options, resort to child labor as a risk-coping strategy in the face of chronic economic deprivation or sudden, acute economic distress (see, for example, Basu 1999; Beegle et al. 2006; Edmonds 2006; Grootaert and Kanbur 1995; Ray 2000). Although child labor has been a longstanding subject of economic thought, formal economic models of child labor have appeared only more recently. In their seminal contribution, Basu and Van (1998) were among the first to formalize the relationship between child labor and poverty, and analyze the effects of imposing a ban on child labor in a labor market environment where adult and child workers are (imperfectly) substitutable in the production process. Several other researchers have since followed suit, offering alternative analytical explanations for the existence of child labor ranging from, for example, the interaction between imperfect credit markets (borrowing constraints) and poverty (Baland and Robinson 2000; Ranjan 1999; Ranjan 2001) to coordination failure between the decision of households to invest in the schooling of their children and the decision of firms to invest in skill-intensive technologies (Dessy and Pallage 2001) to a social norms argument based on the mutual interdependence between the degree of social stigma associated with child work and the incidence of child labor in a society (López-Calva 2000).

In contrast to the theoretical literature on child labor in general, analytical research on harmful child labor is relative scant albeit nascent. To the best of our knowledge, there is less than a handful of studies which strive to describe harmful child labor leave alone investigate the welfare implications of banning such forms of labor. To be sure, many of the same reasons why child labor arises are equally applicable to why harmful child labor arises. However, the main contribution of this branch of the literature lies in describing the specific processes and conditions under which harmful child labor emerges.

The study that is most pertinent to ours is Dessy and Pallage (2005). Other studies such as Rogers and Swinnerton (2008) and Dessy and Pallage (2003) are only tangentially relevant because they focus on child exploitation in settings where the household is fully aware of the availability and likelihood of different child employment outcomes but has no control over which employment outcome is realized for the child and, moreover, is unable to remove the child from an unfavorable outcome (exploitative child labor) once realized. We are more interested in understanding harm as it relates to work and working conditions in firms operating in open labor markets rather than in employment situations where the right of exit is effectively surrendered.

More in line with this perspective of the child labor market, Dessy and Pallage (2005) analyze the case where harmful child labor—although beneficial in that it is accompanied by higher (compensating) child wages, thereby improving the financial resources of the household and making schooling more affordable for the child—is detrimental in that it reduces the innate endowment of human capital by impairing the ability of child workers to learn in school. They show that a ban on harmful child labor with perfect effectiveness is under most circumstances welfare-reducing for households, even though it raises the average human capital level of children in the economy and spurs technological advancement of firms which translates into a higher future remuneration for human capital accumulated during childhood.

Though our model has several features in common with that of Dessy and Pallage (2005), there are also certain substantive differences which distinguish this chapter from their study. We draw attention to three such differences, in ascending order of importance. First, although the benefits of harmful child labor are described in similar fashion—namely, higher compensating wages for higher levels of harm—the mechanism through which the costs of harmful child labor materialize differ. Here, the costs of harm arise in a more direct manner, that is, higher levels of work-related harm as a child are associated with reduced wages in adulthood (via lower labor productivity) due to the long-term detrimental effects on the health and development of the child. That is, the transmission mechanism is one in which the costs of harm are intrinsically tied to the nature of work. Second, unlike Dessy and Pallage (2005) who consider households to be homogenous in both preferences and resources, we consider households which differ in terms of parental income. The introduction of household heterogeneity greatly enriches the description of household behavior, particularly with respect to job sorting. In addition, it enables us to examine the differential impact of the ban on welfare across households. Third and last, like Dessy and Pallage (2005), we begin by analyzing the economic behavior of households and firms and the welfare effects of banning harmful child labor in an environment of perfect information and perception regarding child occupational harm. However, unlike them, we proceed to analyze the same issues in an environment of information asymmetry and errors in information processing by households.

1.3. Defining harm

We now turn briefly to a discussion of our operational definition of harmful child labor in this chapter. As previously alluded to, the ILO definition of the "worst forms of child labor" encompasses several disparate forms of child labor. However, there is an important common element that ties all of these forms of child labor together, which is that they all fundamentally relate to employment conditions that impair or can potentially impair the physiological or psychological well-being of the child worker. It is this common element that essentially defines harm in this context. Admittedly, certain ambiguities remain. Consequently, the definition of harmful child labor is open to normative interpretation, which we do not discuss here because it lies outside the scope of this chapter.

Clearly, harm is a multi-dimensional concept. However, instead of constructing and using a rich description of harm (which would make the concept too nebulous to be practicable), we formulate and apply a specific characterization of harm. The notion of harm we use in this chapter essentially has four main stylized features. First, harm is a direct consequence of the nature of work which the child is engaged in. That is, we abstract from indirect "harmful" effects such as on human capital development where the channels are as simple as child work taking time away from schooling or as relatively complex as the reduction in ability of children who work part-time to learn in school. Second, child work in certain workplaces results in unavoidable harm, that is, harm in a nonstochastic occurrence. Third, harm from child work does not manifest itself concurrently with the period of child work or in the short-term; rather, child work in certain workplaces results in detrimental physical or physiological effects later on as an adult, that is, child work has a delayed or long-term harmful effect on the health of the individual. Third, these adverse health effects, once they present themselves in adulthood, are considered to be irreversible and permanent.

1.4. The benchmark model

In this section, we develop a simple benchmark model of harmful child labor. The model consists of two periods, so as to allow a dynamic representation of harm, and is entirely deterministic. We begin by describing the informational setting related to harm. We then characterize the environment, the decision problem, and decisionmaking behavior of households and, in analogous fashion, for firms. Finally, we combine the optimal decision rules from the two sides of the market in order to characterize the market equilibrium.

Informational structure related to harm: Parents are considered to be *fully aware* of the extent of detrimental health effects that manifest themselves in the future from sending their children to work. Furthermore, parents are able to perfectly judge what effects child occupational harm has on the earnings potential of the child as an adult. In other words, parents are considered to have perfect foresight (present beliefs about outcomes in the future perfectly match the actual outcomes realized in the future), and fully take into account these effects in their decisionmaking process.

Households

Environment: We consider an economy populated by a continuum of households of measure I. With a slight abuse of notation, we shall also refer to the set of all households as I. For the sake of simplicity, each household $i \in I$ consists of two individuals, an adult (the parent) and a child. The model spans two periods. In the first period, both the parent and child are present, while, in the second period, only the child survives, having become an adult. In the first period, both the parent and child are endowed with one unit of time, which the parent allocates entirely to wage employment, while the child's endowment of time is allocated entirely to schooling or wage employment.[4] In the second period, the child as an adult allocates her endowment of one unit of time entirely to wage employment. We assume that households also obtain some nonlabor income. Borrowing is however not possible; this constraint is often broadly consistent with the reality of poor households in developing countries.

We assume that adult and child workers are employed in separate labor markets. This is a key departure from some previous theoretical studies on child labor

[4] The decision to abstract from intermediate cases of both schooling and work by children was made to simplify the analysis. The results in this chapter do not change in any qualitative manner as a result of this decision.

which treat adult and child workers as substitutable (albeit imperfectly) in the production process (see, e.g., Basu and Van 1998). We argue that a situation of exclusive child labor markets can arise if child workers are always more "cost-effective" than adult workers in the production of certain goods and services, that is, child workers have a higher marginal product *per* unit of compensation than adult workers. Note, however, that this assumption does not necessarily imply that child workers are more productive than adult workers.

The parent is considered to be the sole decisionmaker in the household. This decisionmaking power is exercised in the following paternalistic fashion: the parent decides on behalf of the child whether the child should work or not; if the former, the parent also decides what type of environment the child works in with respect to the presence of harmful child labor. The child is assumed to fully abide by the parent's decision. Given that the parent is the decisionmaker for the household, here, the parent's utility function is used in characterizing the optimizing behavior of the household. Furthermore, parents are considered to be altruistic towards their children, albeit imperfectly. This is demonstrated by the parent caring *positively* for the health of the child in the manner to be delineated shortly.

Preferences: All parents are assumed to have identical preferences. Denote total household consumption in period 1, the simple addition of adult and child consumption, by c, the earnings of the child as an adult in period 2 as ω', and the parental altruism factor by β, where $\beta \in (0,1)$ indicates partial altruism. The utility function u[5] for each parent $i \in I$ is given by

$$u = \ln c + \beta \ln \omega' \ . \tag{1}$$

[5] The utility function is a monotonic transformation of a standard Cobb-Douglas utility function. The use of linear(ized) utility functions is common in the child labor literature (see, for example, Basu and Van 1998, Basu 1998, and Baland and Robinson 2000).

There are two important aspects related to the parental utility function that deserve some explanation. First, the parental utility function possesses as one of its arguments the child's earnings as an adult (ω') and *not* the child's consumption as an adult. The reason we employ this approach is two-fold: (1) it allows us to sever the parent's intergenerational link with all noncontiguous future generations in the decisionmaking process, thereby providing significantly greater analytical tractability and (2) from a pragmatic standpoint, it is highly plausible that parents, in making decisions that have intergenerational consequences, consider (if at all) only their immediate children (or, at most, their grandchildren) and not *all* future generations. That is, their horizons are much shorter than the infinite horizon decision problems which households are often presumed to perform. Second, parental altruism enters instrumentally through the parent's valuation of the child's future earnings ω', where working in a harmful workplace as a child causes poorer health as an adult, consequently lowering adult labor productivity and, in turn, lowering adult labor earnings. Hence, although the parent cares explicitly about the earnings of the child as an adult, given the intertemporal link between child work in harmful workplaces and future adult earnings, the parent cares implicitly about the nature of child employment. It is in this manner that parental altruism towards children is incorporated into the model. Notwithstanding, the parent is assumed to only care imperfectly for the earnings of the child as an adult.

Resource constraint: Denote the wages earned by the parent and the child in household $i \in I$ by ω_i^P and ω_i, respectively. Note that ω_i^P and ω_i are determined endogenously as equilibrium wages, as will be shown later. Parental nonlabor income earnings , denoted by N_i^P, are considered to be exogenously determined, where $N^P \in \left[0, \bar{N}^P\right]$ and is distributed according to the continuously differentiable cumulative distribution function *K*, where $K\left(N^P\right)$ gives us the proportion of households with parental income equal to or below N^P. Denote the addition of parental wages $\left(\omega^P\right)$ and parental nonlabor income N^P by total parental income y^P, where $y^P \in \left[\underline{y}^P, \overline{y}^P\right]$ and is distributed accordingly to the continuously differentiable cumulative distribution function *F* (with the corresponding probability density function *f*), where $F\left(y^P\right)$ gives us the proportion of households with parental income equal to or below y^P, and The budget constraint for household $i \in I$ is given by

$$c \leq y_{\mathrm{i}}^P + \omega_{\mathrm{i}}. \quad (2)$$

Since the marginal utility of household consumption is strictly increasing in its argument, the budget constraint holds with equality, which permits us to rewrite the utility function Since the marginal utility of household consumption is strictly increasing in its argument, the budget constraint holds with equality, which permits us to rewrite the utility function for household *i* as

$$u = \ln\left(y_i^P + \omega_i\right) + \beta \ln \omega_i' . \quad (3)$$

For simplicity, suppose there are three options for the child: (1) to not work and go to school (denoted by *N*), (2) to work in an unharmful environment (denoted by *U*), or (3) to work in a harmful environment (denoted by *H*). Denote ω_U and ω_H as the child wages associated with the type-*U* and type-*H* environments, respectively. In the next period, depending on whether the child does not work, works in a type-*U* environment, or works in a type-*H* environment, the child (as an adult) can expect to earn ω_N', ω_U', or ω_H', respectively, where $\omega_N' > \omega_U' > \omega_H'$. Working as a child in the type-*H* environment as opposed to the type-*U* environment produces adverse health effects in the future which reduce labor productivity as an adult, and hence, labor earnings as an adult. Further, as discussed previously, these adverse health effects and their effects on adult productivity and earnings are perfectly known to the parent. On the other hand, not working and going to school leads to the human capital development of the child, resulting in higher labor market earnings as an adult than is possible if the child did not go to schools and worked.

Optimization

The utility maximization problem for the parent of household $i \in I$ is simply

$$\max_{(\omega,\omega')\in\{(0,\omega'_N),(\omega_U,\omega'_U),(\omega_H,\omega'_H)\}} \ln\left(y_i^P + \omega_i\right) + \beta \ln \omega'_i. \qquad (4)$$

Assumption 1: *All three choices, namely no child work, child work in the type-U environment, and child work in the type-H environment, are simultaneously optimal for distinctly different subsets of households.*

In order for assumption 1 to be satisfied, the following conditions have to hold:

Condition 1: $\omega_U > \dfrac{\left(\omega'_N / \omega'_U\right)^{\beta} - 1}{\left(\omega'_N / \omega'_H\right)^{\beta} - 1}\omega_H$, and

Condition 2: $\omega_H < \dfrac{1 - \left(\omega'_U / \omega'_N\right)^{\beta}}{\left(\omega'_H / \omega'_U\right)^{\beta} - \left(\omega'_H / \omega'_N\right)^{\beta}}\omega_U$.

These conditions set the lower and upper bounds for ω_U and ω_H, respectively, as a function of ω'_N, ω'_U, ω'_H and $\beta \in (0,1)$.

Given assumption 1, solving the above utility maximization problem, those households $i \in I$ with parental income y_i^P such that $y_i^P \geq \hat{y}^P$ will choose not to send their children to work, while those households $i \in I$ such that $y_i^P < \hat{y}^P$ will choose to send their children to work, where $\hat{y}^P = \dfrac{\omega_U}{\left(\dfrac{\omega'_N}{\omega'_U}\right)^{\beta} - 1}$. We shall refer to parental income $\hat{y}^P$ as the "labor force participation cutoff". Hence, the decision to send the child to work is driven by insufficient household resources consistent with the luxury axiom established by Basu and Van (1998). To ensure that choosing to send or not to send children to work are feasible options for households, we make the following assumption:

Assumption 2: *The variation in parental earnings spans the labor force participation cutoff*

$$\hat{y}^P = \frac{\omega_U}{\left(\frac{\omega'_N}{\omega'_U}\right)^{\beta} - 1}.$$

Turning our attention to those households $i \in I$ with parental income $y_i^P < \hat{y}^P$, assume that there is at least one household $i \in I$ with parental income $\tilde{y}^P$ such that

$$\ln\left(\tilde{y}_i^P + \omega_U\right) + \beta \ln \omega'_U = \ln\left(\tilde{y}_i^P + \omega_H\right) + \beta \ln \omega'_H.$$

This implies that the household with parental income $\tilde{y}^P$ is indifferent between sending the child to work in the type-*U* or type-*H* environment. We shall refer to the parental income level $\tilde{y}^P$ as the "*U-H* cutoff". Since we assume that some households choose to send their children to work in the type-*U* environment, it is easy to establish that the *U-H* cutoff $\tilde{y}^P$ is lower than the labor force participation cutoff $\hat{y}^P$. To ensure that households find it feasible to choose different work environments for their children, we make the following assumption:

Assumption 3: *The variation in parental earnings spans the U-H cutoff*

$$\tilde{y}^P = \frac{\left[\omega_H - \left(\frac{\omega'_U}{\omega'_H}\right)^{\beta} \omega_U\right]}{\left[\left(\frac{\omega'_U}{\omega'_H}\right)^{\beta} - 1\right]}.$$

Given Assumptions 1-3, we obtain the following two results:

Proposition 1: *Given that* $\omega'_U > \omega'_H$ *, in equilibrium, it must be the case that* $\omega_H > \omega_U$ *. This implies that a positive compensating wage differential arises for child workers between the type-H and type-U environments.*

Proof: Take the household with parental income at the labor force participation cutoff $\tilde{y}_P$. We know that this household is indifferent between sending the child to work in the type-*U* or type-*H* environments, that is,

$$\ln\left(\tilde{y}^P + \omega_H\right) + \beta \ln \omega'_H = \ln\left(\tilde{y}^P + \omega_U\right) + \beta \ln \omega'_U .$$

Rearranging and simplifying, we obtain

$$\frac{\tilde{y}^P + \omega_H}{\tilde{y}^P + \omega_U} = \left(\frac{\omega'_U}{\omega'_H}\right)^\beta .$$

Given $0 < \beta < 1$ and $\omega'_U > \omega'_H$, we have that

$$\frac{\tilde{y}^P + \omega_H}{\tilde{y}^P + \omega_U} = \left(\frac{\omega'_U}{\omega'_H}\right)^\beta > 1 \Rightarrow \omega_H > \omega_U . \blacksquare$$

The exact value of the compensating wage differential between working in the type-*H* environment and the type-*U* environment, denoted by ω_{H-U} ($\equiv \omega_H - \omega_U$), is determined by the marginal household with parental income $\tilde{y}^P$, and is given by

$$\omega_{H-U} = (\lambda - 1)\left(\tilde{y}^P + \omega_U\right),$$

where $\lambda = \left(\frac{\omega'_U}{\omega'_H}\right)^\beta > 1$, and the partial derivatives $\frac{\partial \omega_{H-U}}{\partial \tilde{y}^P} = \frac{\partial \omega_{H-U}}{\partial \omega_U} = \lambda - 1 > 0$.

As expected, the partial derivatives indicate that the compensating wage differential for working in the type-*H* environment increases as the parental income of the marginal household increases and as the wage associated with working in the type-*U* environment increases. The intuition is straightforward: if a parent observes an increase in her own personal income or a higher child wage associated with work in the type-*U* environment, a higher wage in the type-*H* environment is required to induce the parent to switch the child's work environment from type-*U* to type-*H*, ceteris paribus.

Proposition 2: *Households sort themselves systematically on the basis of parental income between the three available choices of (1) no child work, (2) child work in the type-U environment, and (3) child work in the type-H environment. More specifically, the population of households can be split into three distinct parental income segments: low-income, middle-income, and high-income. Low-income households (i.e., the poorest households) choose to send their children to work in the type-H environment, middle-income households choose to send their children to work in the type-U environment, and, finally, high-income households (i.e., the richest households) choose not to send their children to work.*

Proof: We have already established that all households $i \in I$ such that parental income y_i^P is equal to or greater than the labor force participation cutoff $\hat{y}^P$ will choose not to send their children to work. Now, among those households $i \in I$ such that parental income

y_i^P is less than $\hat{y}^P$, consider a household $i \in I$ with parental income at the *U*-*H* cutoff $\tilde{y}^P$. We know that this household is indifferent between sending the child to work in the type-*U* and type-*H* environments, that is,

$$\ln\left(\tilde{y}^P + \omega_H\right) + \beta \ln \omega'_H = \ln\left(\tilde{y}^P + \omega_U\right) + \beta \ln \omega'_U .$$

Rearranging and simplifying, we obtain

$$\frac{\tilde{y}^P + \omega_H}{\tilde{y}^P + \omega_U} = \left(\frac{\omega'_U}{\omega'_H}\right)^{\beta} .$$

If the household's parental income y^P is greater than the *U*-*H* cutoff $\tilde{y}^P$, then

$$\frac{y^P + \omega_H}{y^P + \omega_U} < \frac{\tilde{y}^P + \omega_H}{\tilde{y}^P + \omega_U} = \left(\frac{\omega'_U}{\omega'_H}\right)^{\beta} \Rightarrow$$
$$\ln\left(y^P + \omega_U\right) + \beta \ln \omega'_U > \ln\left(\tilde{y}^P + \omega_H\right) + \beta \ln \omega'_H .$$

This household obtains higher utility from choosing to send the child to work in the type-*U* environment. If the household's parental income y^P is less than the *U*-*H* cutoff $\tilde{y}^P$, we obtain the reverse relationship: the household obtains higher utility from choosing to send the child to work in the type-*H* environment. ■

To further clarify, all households $i \in I$ such that $y_i^P \in \left[\tilde{y}^P, \hat{y}^P\right)$, will choose to send their children to work in the type-*U* environment, while all those households $i \in I$ such that $y_i^P \in \left[\overline{y}^P, \tilde{y}^P\right)$ will choose to send their children to work in the type-*H* environment.

Thus, just as how insufficient household resources or material poverty is the reason why parents choose to send their children to work, insufficient household resources also serves as the reason why parents choose to send their children to work in the type-*H* environment. Proposition 2 carries particular weight because it implies that the poorest households choose to send their children to work in the type-*H* environment when they are *equally* averse to the harmful child labor experienced by their children and *equally* altruistic towards their children as households which choose not to send their children to work in the type-*H* environment. What drives the result is the variation in parental income across households in the economy.

Corollary 1: *Among households that choose to send their children to work, for households with parental income* y^P *such that* $y^P > \tilde{y}^P$, *the higher* y^P *is, the higher the utility "rent" obtained from sending their children to work in the type-U environment instead of in the type-H environment. Conversely, for households with parental income* y^P *such that* $y^P < \tilde{y}^P$, *the lower* y^P *is, the higher the utility "rent"* *obtained from sending their children to work in the type-H environment instead of in the type-U environment.*

Proof: First, consider households with parental income y^P greater than the *U*-*H* cutoff $\tilde{y}^P$. As we saw in the proof for Proposition 2, this implies that
$\ln\left(y^P + \omega_U\right) + \beta \ln \omega'_U > \ln\left(y^P + \omega_H\right) + \beta \ln \omega'_H$.
Define $\vartheta_U = \ln\left(y^P + \omega_U\right) + \beta \ln \omega'_U - \ln\left(y^P + \omega_H\right) - \beta \ln \omega'_H > 0$.
Differentiating ϑ_U with respect to y^P, we obtain

$$\frac{\partial \vartheta_U}{\partial y^P} = \frac{1}{y^P + \omega_U} - \frac{1}{y^P + \omega_H}.$$

Since $\omega_H > \omega_U$ (Proposition 1), $\frac{\partial \vartheta_U}{\partial y^P}$ is positive. This implies that as y^P increases, the positive utility difference between choosing the type-*U* environment over the type-*H* environment, given by the value of ϑ_U, increases.

For the case where $y^P < \tilde{y}^P$, in which ϑ_H is defined as $\vartheta_H = \ln\left(y^P + \omega_H\right) + \beta \ln \omega'_H - \ln\left(y^P + \omega_U\right) - \beta \ln \omega'_U > 0$, we obtain that $\frac{\partial \vartheta_H}{\partial y^P}$ is negative.

In this case, this implies that as y^P *decreases*, the positive utility difference between choosing the type-*H* environment over the type-*U* environment, given by the value of ϑ_H, increases. ■

Child labor supplies: We now can characterize the child labor supplies to the type-*U* and type-*H* environments. Denote the set of all children that seek work in the type-*U* environment (or more accurately, the households that these children belong to) as $\Theta_U = \left\{i : y_i^P \in \left[\tilde{y}^P, \hat{y}^P\right)\right\}$. Likewise, denote the set of all children that seek work in the type-*H* environment as $\Theta_H = \left\{i : y_i^P \in \left[\underline{y}^P, \tilde{y}^P\right)\right\}$. Accordingly, the child labor supplies to the type-*U* and type-*H* environments are given by

$$S_U = I \times \left(F\left(\hat{y}^P\left(\omega_U\right)\right) - F\left(\tilde{y}^P\left(\omega_U, \omega_H\right)\right)\right) \text{ and}$$
$$S_H = I \times \left(F\left(\tilde{y}^P\left(\omega_U, \omega_H\right)\right)\right), \text{ respectively,} \tag{5}$$

where $\hat{y}^P\left(\omega_U\right) = \dfrac{\omega_U}{\left(\dfrac{\omega'_N}{\omega'_U}\right)^{\beta} - 1}$ and $\tilde{y}^P\left(\omega_U, \omega_H\right) = \left[\dfrac{\omega_H - \left(\dfrac{\omega'_U}{\omega'_H}\right)^{\beta} \omega_U}{\left(\dfrac{\omega'_U}{\omega'_H}\right)^{\beta} - 1}\right]$.

The signs of the own- and cross-wage first derivatives of the supply functions are (1) $\dfrac{\partial S_U}{\partial \omega_U} > 0$, (2) $\dfrac{\partial S_U}{\partial \omega_H} < 0$, (3) $\dfrac{\partial S_H}{\partial \omega_H} > 0$, and (4) $\dfrac{\partial S_H}{\partial \omega_U} < 0$. The signs of the derivatives of total child labor supply $\left(S_U + S_H\right)$ with respect to the wages in the type-*U* and type-*H* labor markets are $\dfrac{\partial\left(S_U + S_H\right)}{\partial \omega_U} > 0$ and , $\dfrac{\partial\left(S_U + S_H\right)}{\partial \omega_H} > 0$, respectively (see the Appendix for the workings out).

Firms

Environment: All firms are considered to produce a single identical consumption good with a fixed price normalized to one. Suppose that there are a measure *J* and *K* firms with type-*U* and type-*H* environments, respectively. The type-of-environment choice is considered to be a historical decision – in the current period, we begin with the ex post equilibrium number of type-*U* and type-*H* firms. Both sets of firms are owned by employers in the economy who are distinct from the households in the economy. All profits generated from producing the good are fully consumed by the employers (i.e., profits are not shared with the set of households *I*). Both sets of firms are also considered to operate in a competitive market taking output and input prices as given.

Production technology: Type-*U* firms produce the good according to the continuously differentiable production technology f_U defined over child labor denoted by c_U. Similarly, type-*H* firms produce the same good according to the continuously differentiable production technology f_H defined over child labor denoted by c_H. These technologies are represented by

$$f_U\left(c_U\right) = c_U^{\alpha} \text{ and}$$
$$f_H\left(c_H\right) = c_{H,}^{\gamma} \tag{6}$$

where the exponents α and γ lie strictly between 0 and 1.

Optimization

Each type-*U* firm solves the following profit maximization problem:

$$\max_{c_U \geq 0} c_U^{\alpha} - \omega_U c_U .$$

The first order condition is as follows:

$$c_U : \alpha c_U^{\alpha-1} - \omega_U \leq 0 \quad \left(= \text{ if } c_U > 0\right).$$

By solving the analogous profit maximization problem for the type-*H* firms, we obtain the following first order condition:

$$c_H : \gamma c_H^{\gamma-1} - \omega_H \leq 0 \quad \left(= \text{ if } c_H > 0\right).$$

Child labor demands: Consequently, the individual firm demand function by type-*U* and type-*H* firms are $d_U = \left(\frac{\omega_U}{\alpha}\right)^{\frac{1}{\alpha-1}}$ and $d_H = \left(\frac{\omega_H}{\gamma}\right)^{\frac{1}{\gamma-1}}$, respectively.

Accordingly, the market child labor demands by the type-*U* and type-*H* firms are defined as

$$D_U = Jd_U = J\left(\frac{\omega_U}{\alpha}\right)^{\frac{1}{\alpha-1}} \text{ and}$$
$$D_H = Kd_H = K\left(\frac{\omega_H}{\gamma}\right)^{\frac{1}{\gamma-1}}, \text{ respectively.} \tag{7}$$

As expected, the signs of the first derivatives of the labor demand functions with respect to their wages are (1) $\frac{\partial D_U}{\partial \omega_U} < 0$ and (2) $\frac{\partial D_H}{\partial \omega_H} < 0$.

Future adult labor demand: Denote *e* as an efficiency unit of the future adult labor of children, where the amount of efficiency units of labor that a child possesses as an adult is a function of whether or not the child worked (*N*) and, if the child worked, whether it was in a harmful setting (*H*) or not (*U*), namely $e_N > e_U > e_H$. Children as adults observe the following constant returns to scale (CRS) production function for their labor services

$$f(e) = Ae, \tag{8}$$

where *A*, the efficiency parameter, is positive.

Denote ω_e as the wage rate for one efficient unit of labor. In equilibrium, *A* equals ω_e. Consequently, ω'_N, ω'_U, and ω'_H equal Ae_N, Ae_U, and Ae_H, respectively. Given the form of the production function, the future adult labor market clears trivially. Further, changes in future adult labor supply only affect equilibrium employment and not the wage rate for efficiency labor. This feature of the future adult labor market should be noted because key aspects of the general equilibrium results of banning of harmful child labor hinges on this particular feature.

Present adult labor demand: Demand for parental labor in the present period is determined analogously to the demand for future adult labor of children discussed above. Denote e^P as an efficiency unit of parental labor, where, like for children, the amount of efficiency units possessed by parents is a function of whether or not the parent worked as a child (*N*) and, if the parent worked, whether it was in a harmful setting (*H*) or not (*U*), namely $e_N^P > e_U^P > e_H^P$. Parents observe an identical CRS production function to (8), and observe the same equilibrium wages, namely $\omega_N^P = Ae_N^P$, $\omega_U^P = Ae_U^P$, and $\omega_H^P = Ae_H^P$.

Market equilibrium

Child labor market

Definition 1: An equilibrium is a pair of child labor wages (ω_U^*, ω_H^*)such that:

1) $D_U(\omega_U^*) = S_U(\omega_U^*, \omega_H^*)$ (Market clearing condition in the type-*U* child labor market);
2) $D_H(\omega_H^*) = S_H(\omega_U^*, \omega_H^*)$ (Market clearing condition in the type-*H* child labor market); and
3) $\omega_H^* > \omega_U^*$ (Compensating wage differential for harmful child labor).

See Proof 1 in the Appendix for existence, uniqueness, and stability proofs.

1.5. Ban on harmful child labor with perfect enforcement

In this section, we examine the labor market and welfare effects of a ban on harmful child labor with perfect enforcement. A ban on harmful child labor with perfect enforcement implies that child workers can no longer be employed in the type-*H* environment; further, the ban is introduced early enough for child workers in the type-*H* environment to escape the future adverse health effects. After the ban comes into effect, each type-*H* firm faces three distinct choices: (1) shut down operations; (2) maintain the same work environment (i.e., the same occupational health and safety standards) but employ adult workers instead; or (3) raise the occupational health and safety standards to at least the minimum level considered to be acceptable by the authorities and continue to hire child workers (this transformation will certainly involve a cost). We consider the scenario where the ban on harmful child labor results in *all* type-*H* firms shutting down. Admittedly extreme, we can conceive of such a case arising when the cost of raising the occupational health and safety standards up to the minimum acceptable level is prohibitively high or, alternatively, maintaining the same standards and hiring adult workers to replace child workers are neither possible nor economically viable.

Labor market effects: If all the type-*H* firms cease to operate after the ban on harmful child labor, the only remaining firms that hire child workers are the type-*U* firms. Let ω_U^{*B} denote the market-clearing wage for work in the type-*U* environment after the ban on harmful child labor is introduced and effectively enforced, that is, $D_U\left(\omega_U^{*B}\right) = S_U\left(\omega_U^{*B}\right)$.

Proposition 3: *The post-ban equilibrium wage in the type-U child labor market,* ω_U^{*B}*, is less than the pre-ban equilibrium wage in the same market,* ω_U^**. In other words, after the ban, the equilibrium wage in the type-U child labor market falls.*

Proof: Suppose instead that ω_U^{*B} is greater than or equal to ω_U^*. Using (5), it is straightforward to determine that $S_U\left(\omega_U^{*B}\right)$ is strictly larger than $S_U\left(\omega_U^*, \omega_H^*\right)$. Since, in equilibrium, the amount of labor demanded equals the amount of labor supplied, $D_U\left(\omega_U^{*B}\right)$ must also be larger than $D_U\left(\omega_U^*\right)$. Given that D_U is decreasing in ω_U, this implies that ω_U^{*B} is less than ω_U^*, which contradicts our initial claim. On the other hand, we do not necessarily arrive at a contradiction when we consider ω_U^{*B} is less than ω_U^*. If ω_U^{*B} is less than ω_U^*, then $D_U\left(\omega_U^{*B}\right)$ is greater than $D_U\left(\omega_U^*\right)$. Thus, in equilibrium, $S_U\left(\omega_U^{*B}\right)$ must be greater than $S_U\left(\omega_U^*, \omega_H^*\right)$. Using (5) again, it is easy to verify that this holds so long as $S_H\left(\omega_U^*, \omega_H^*\right)$ is greater than the decline in total child employment resulting from the ban. ■

Corollary 2: *Child employment in the type-U child labor market is larger after the ban.*

Proof: See latter part of proof for Proposition 3.

The intuition for Proposition 3 and Corollary 2 is as follows. Suppose prior to the ban there were E_U^* children working in the type-*U* environment and receiving ω_U^*, while there were E_H^* children working in the type-*H* environment and receiving ω_H^*. Immediately after the ban, the total supply of child workers to the type-*U* environment will increase to $E_U^* + E_H^*$, as the parents who chose to send their children to work in the type-*H* environment before the ban will now choose to send their children to work in the type-*U* environment (a result related to Proposition 2). This means that at ω_U^* the supply of child labor to type-*U* firms will exceed the demand for child labor by these firms. As a result, the equilibrium wage will adjust downwards until the type-*U* child labor market clears once again. This happens when the equilibrium wage falls to ω_U^{*B}. Clearly, at ω_U^{*B}, there will be more than E_U^* child workers employed in the type-*U* environment.

Corollary 3: *Total post-ban child employment is less than total pre-ban child employment.*

Proof: From Section 1.4 we know that the lower bound on parental income for choosing no child work is given by $\hat{y}^P(\omega_U)$. Since ω_U^{*B} is less than ω_U^* (from Proposition 3), and $\hat{y}^P$ is increasing in ω_U, $\hat{y}^P(\omega_U^{*B})$ is less than $\hat{y}^P(\omega_U^*)$. The lower child labor force participation cutoff implies that less parents choose to send their children to work after the ban.

Which children work after the ban? Consider a parent with own income $\tilde{\tilde{y}}_i^P$ who chooses to send her child to work. Denote $\tilde{\tilde{\omega}}_U$ as the wage associated with employment in the type-*U* environment such that the parent with own income $\tilde{\tilde{y}}_i^P$ sets the utility obtained from sending the child to work in the type-*U* environment equal to the utility obtained from not sending the child to work. This condition is satisfied for parent income $\tilde{\tilde{y}}_i^P$ when

$$\tilde{\tilde{\omega}}_U^P = \tilde{\tilde{y}}_i^P \left(\left(\frac{\omega_N'}{\omega_U'} \right)^{\beta} - 1 \right).$$

Differentiating $\tilde{\tilde{y}}_i^P$ with respect to $\tilde{\tilde{\omega}}_U$, we see that parental income $\tilde{\tilde{y}}_i^P$ is increasing in $\tilde{\tilde{\omega}}_U$. This indicates that as ω_U falls (in our case, as a result of the ban), households with the *highest* parental income which initially chose to send their children to work will be the first to choose to remove their children from work followed in order by those with degressively lower parental income. This process will continue until ω_U settles at ω_U^{*B}. At ω_U^{*B}, those households $i \in I$ such that parental income y_i^P is equal to or greater than $\hat{y}^P\left(\omega_U^{*B}\right)$ will choose not to send their children to work at all, while those households $i \in I$ such that parental income y_i^P is less than $\hat{y}^P\left(\omega_U^{*B}\right)$ will choose to send their children to work in the type-*U* environment.

To sum up, the labor market effects of the ban on harmful child labor which results in the shutdown of type-*H* firms are as follows:

1. *The equilibrium wage associated with the type-U environment falls* $\left(\omega_U^{*B} < \omega_U^{*}\right)$.
2. *The number of children not working increases. In other words, total child employment falls in the economy.*
3. *The number of child workers in the type-U environment increases.*
4. *While the number of households in each relevant labor market choice have changed, the household sorting pattern across labor market choices remains:*

richer households choose not to send their children to work, while the poorer households choose to send their children to work in the type-U environment.

Welfare effects: In order to examine the welfare consequences of the ban on harmful child labor, we compare the distribution of utility levels of households before and after the ban on harmful child labor; note, here, we use the parent's utility function as the household's welfare function. We use the method of welfare dominance to rank the distributions of household utilities under different situations, when possible. The criterion we consider is first-order (or first-degree) dominance.[6]

Definition of first-order dominance: Consider situations *A* and *B*. Denote the cumulative distribution of household utilities under situation *A* as G_A and the cumulative distribution of household utilities under situation *B* as G_B. Situation *A* is considered to "first-order dominate" situation *B* if

$$G_A(u_i) \le G_B(u_i) \text{ for all } u_i \in U \text{ and } G_A(u_i) < G_B(u_i) \text{ for some } u_i \in U.$$

Furthermore, if situation *A* first-order dominates situation *B*, then any welfare measure *w* that belongs to the class of social welfare functions *W* which are anonymous and increasing in its arguments will assign a higher rank to situation *A* than to situation *B* (Fields 2001).[7]

[6] This criterion is fairly stringent, though not as stringent as Pareto dominance.

[7] The property of *anonymity* indicates that social welfare depends on the level of utilities of households and not on which households has which level. That is, all households are treated identically and no other information other than their utility levels matter. The property of *increasing* indicates that social welfare is increasing in the utility level of a household, holding the utility levels of all other households under consideration as constant.

Proposition 4: *In the case of perfect information,* a *ban on harmful child labor results in an unambiguous decrease in the welfare status of households. That is, the pre-ban situation first-order dominates the post-ban situation.*

Instead of presenting a formal proof involving the comparison of utilities of households before and after the ban on harmful child labor, we discuss the basic intuition behind Proposition 4 because it serves the purpose just as well, and is, in our view, more illuminating. Consider a set of alternatives $X^0 = \{x_N, x_U, x_H\}$ available to a collection of households. Households choose among the alternatives in X^0 and we have that each alternative is preferred over the other two by a subset of households. Suppose set X^0 is permanently substituted with a new set of alternatives $X^B = \{x_N, \tilde{x}_U\}$. In this new set X^B, alternative x_H is no longer available and alternative x_U has been replaced by $\tilde{x}_U$ where alternative $\tilde{x}_U$ is not as attractive as alternative x_U. Alternative x_N is the same in both sets. Households now choose among the alternatives in set X^B. Clearly, households that chose either alternatives x_U or x_H from set X^0 will experience a decline in their welfare when set X^0 is replaced by set X^B. The only households that experience no change in their welfare are those households that chose alternative x_N from set X^0 and continue to choose the same alternative from set X^B.

In our study, sets X^0 and X^B represent the choice sets before and after the ban on harmful child labor, where alternative x_N corresponds to the no-work choice, x_U to the pre-ban type-*U* child work choice, $\tilde{x}_U$ to the post-ban type-*U* work choice (where $\tilde{x}_U$ is not as attractive as x_U due to market adjustments which reduce the equilibrium wage in the type-*U* child labor market), and x_H to the type-*H* child work choice. The subset of households unaffected by the ban are those households $i \in I$ such that

$\hat{y}_i^P \geq \hat{y}^P\left(\omega_U^*\right)$, while the subset of household adversely affected by the ban are those households $i \in I$ such that $\hat{y}_i^P < \hat{y}^P\left(\omega_U^*\right)$.

Hence, clearly, $u_i^B \leq u_i$ for all households $i \in I$ and $u_i^B < u_i$ for some households $i \in I$. It is easy to see that this implies that

$G_0(u_i) \leq G_B(u_i)$ for all $u_i \in U$ and $G_0(u_i) < G_B(u_i)$ for some $u_i \in U$,

where G_0 and G_B are the cumulative distribution functions of utilities before and after the ban on harmful child labor, respectively.

Corollary 4: *Among households that initially chose to send their children to work in the type-H environment, the poorer the household is in terms of parental earnings, the larger the welfare loss experienced by the household as a result of the ban.*

Among households that initially chose to send their children to work in the type-*H* environment, given that the utility "rent" from choosing the type-*H* environment rather than the type-*U* environment increases monotonically as parental income falls (see Corollary 1), following the ban on harmful child labor, the situation is reversed: utility loss increases monotonically as parental income falls. The fact that the equilibrium wage ω_U^* falls in the type-*U* environment as a result of adjustments in labor market supply only serves to enlarge the utility loss experienced by these households.

To summarize, the individual policy of banning harmful child labor results in a distribution of household utilities which is first-order dominated by the distribution of household utilities prior to the ban. In fact, the result is much stronger: the pre-ban distribution of household utilities *Pareto-dominates* the post-ban distribution of household utilities. This indicates the policy has been unambiguously welfare-reducing, that is, although not all households experience a decline in utility as a result of the ban, *no* household experiences an increase in utility. Further, among households that initially chose to send their children to work in the type-*H* environment, the poorer the household is in terms of parental income, the larger the welfare loss experienced by the household as a result of the ban.

As a final point, recall that future adult labor demand was derived from a production process that exhibits constant returns to scale. Allowing the future wages of children to instead adjust to shifts in aggregate adult labor supply does not alter the welfare result in any qualitative way as the ban, resulting in the shutdown of firms which impair the future productivity of child workers, results in a higher aggregate supply of adult efficiency units in the future, thus reducing equilibrium adult wages. This downward adjustment in equilibrium adult wages only acts to reinforce our welfare result.

1.6. Systematic informational and perceptional biases in relation to harmful child labor

In this section, we retain the same stylized description of harm described in Section 1.3, that is, child work in the type-*H* environment results in adverse health effects that are certain, manifest themselves in adulthood, and are permanent in nature. We also maintain the core features and mechanisms of the benchmark model. Here, however, we relax the assumption that households possess correct and complete information about the nature of occupational harm associated with type-*H* work and, in addition, that households are able to process the available information properly.

Motivation: There are several reasons why we expect informational problems or perceptional errors to play an especially important role in this context of harmful child labor. We present three possible reasons.

Employee-worker information asymmetry: Parents might be unaware about the exact nature of harm suffered by their working children due to employers intentionally withholding vital information regarding the extent of occupational harm in their workplaces. Because most child labor occurs in workplaces that are entirely unregulated or at very best loosely regulated by the authorities, employers often have wide latitude with respect to what work-related information to divulge to their workers, both incumbent and prospective. Unchecked by any real regulation, there are clear incentives for employers to limit or even misrepresent information on the nature of occupational harm. Such informational malpractices on the part of employers might be particularly egregious in cases where the adverse health effects from occupational harm suffered during childhood manifest themselves only in the future, such as in adulthood, as employers stand to benefit greatly from exploiting this temporal separation in cause and effect. When the informational asymmetry related to occupational harm between employers and households is finally resolved in the future, the adverse health effects are already irreversibly present. Furthermore, if the adverse health effects are sufficiently far ahead in the future, it might be difficult for households to draw a clear link between past child employment and the poor health status of the adult.

Parent-child information asymmetry: Even if we take for granted that employers do in fact voluntarily and fully reveal information on occupational harm to their workers (i.e., the children), the fact that the parent decides what type of employment setting (in terms of the associated level of occupational harm) to send her child to, but does not personally experience the health-related consequences of her decision, creates another source of potential information asymmetry, in this case engendered by the physical separation between the decision-maker (the parent) and the decision-implementer (the child). Child workers in harmful employment settings may not be able to perfectly convey information about the nature of their work and working conditions to their parents, even if they are able to accurately assess for themselves the extent of harm they encounter. Harm in such cases may be more appropriately viewed as an experience good where unless the individual *personally* experiences the good, an accurate assessment of the "qualities" of the good is difficult, if not impossible, to make. There is some evidence from the social psychology literature that individuals do a better job of assessing the risk of some negative event when they have personally experienced that event (e.g., food poisoning, heart attacks, crime). Given the highly dynamic nature of labor markets with the emergence of new, unfamiliar employment opportunities as well as changing work and working conditions in old, familiar employment opportunities, parents might not have the relevant personal work experience to draw from in ascertaining the occupational risks that their children might be facing in their particular jobs.

Imperfect information processing (optimism bias): In our context, optimism bias refers to the tendency by individuals to view certain negative events as less likely to occur to themselves than to others. There is growing survey-based evidence in the social psychology literature which shows that the majority of individuals believe that they are less likely than others to experience certain negative events (for reviews, see Helweg-Larsen and Sheppard 2001; Klein and Weinstein 1997). The various risks examined include pregnancy, sexually-transmitted diseases, cancer, smoking, substance abuse, environmental, and general health. This systematic misperception of the risks leads to behaviors that can actually contribute to increasing the probability of a negative event. In the case of harmful child labor, unrealistic optimism can conceivably take the form of the parent underestimating the level of occupational harm suffered by her own child even when the parent correctly assesses the level of harm suffered by other child workers in the same employment activity. Some of the reasons why researchers believe individuals exhibit optimism bias include the lack of previous personal experience and defensive denial. Defensive denial refers to the phenomenon where individuals deliberately deny that they are at risk of experiencing a negative event to psychologically cope with their risk-prone behavior.

Pre-ban labor market equilibrium: Suppose, for the reasons discussed above, households uniformly underestimate the true level of occupational harm associated with child employment in the type-*H* environment. That is, parents believe that children employed in the type-*H* environment experience adverse health effects that are *less* injurious to their physical health than actually turns out to be the case. Consequently, parents overestimate the future adult earnings of child workers employed in the type-*H* environment.

Hereinafter, let the subscript 1 on a variable indicate its value under the present case of informational and perceptional problems regarding occupational harm while the subscript 0 indicates its value under the benchmark case of perfect information and information processing analyzed in Sections 1.4 and 1.5. Further, notation sans the 0 or 1 subscript which previously denoted specific values of the variable will hereinafter denote the variable itself. Applying these notational rules, we have that ω'_{H1} is greater than ω'_{H0}. Note that since households only underestimate the level of occupational harm associated with work in the type-*H* environment, the future labor market earnings of child workers in the type-*U* environment remains unchanged, that is, ω'_{U1} equals ω'_{U0}. We, however, assume that ω'_{U0} is greater than ω'_{H1}, that is, even though parents overestimate the future labor earnings of child workers in the type-*H* environment, the estimate does not exceed the future labor earnings of child workers in the type-*U* environment.

The decision problem for household $i \in I$ remains the same as in Section 1.4 except that the parent takes into consideration ω'_{H1} instead of ω'_{H0}. The decision problem for type-*U* and type-*H* firms remains unchanged by this misjudgment by households. The misjudgment however affects the equilibrium child wages and employment in both the type-*U* and type-*H* labor markets.

In order to determine the labor market outcomes under the imperfect information case relative to the perfect information case, we need to investigate how a change in ω'_{H1} affects the equilibrium wages ω^*_U and ω^*_H. We determine that ω^*_U is increasing in ω'_H while ω^*_H is decreasing in ω'_H (see the Appendix for the workings out). Thus, given $\omega'_{H0} < \omega'_{H1} < \omega'_{U1} = \omega'_{U0}$, relative to the perfect information case, we obtain the following equilibrium wage results:

1. *Equilibrium wage in the type-U environment is higher* $\left(\omega^*_{U1} > \omega^*_{U0}\right)$;
2. *Equilibrium wage in the type-H environment is lower* $\left(\omega^*_{H1} < \omega^*_{H0}\right)$; *and*
3. *The compensating wage differential for harmful child labor is smaller* $\left(\omega^*_{H1} - \omega^*_{U1} < \omega^*_{H0} - \omega^*_{U0}\right)$.

In addition, the equilibrium wage pair $\left(\omega^*_{U1}, \omega^*_{H1}\right)$ implies the following employment results relative to the perfect information case:

1. *Equilibrium child employment in the type-U environment is lower;*
2. *Equilibrium child employment in the type-H environment is higher; and*
3. *Total child employment is higher* $\left(\hat{y}^P\left(\omega^*_{U1}\right) > \hat{y}^P\left(\omega^*_{U0}\right)\right)$.

The intuition for these equilibrium wage and employment results relative to the perfect information case is probably most straightforward to explain if we consider the economy shifting from the labor market equilibrium under the perfect information case to that under the imperfect information case. That is, suppose ω'_H increases from

ω'_{H0} to ω'_{H1}. This increase in ω'_H will induce a flow of child labor from the type-*U* to the type-*H* labor market, decreasing child labor supply in the former and increasing child labor supply in the latter labor market. This flow raises the equilibrium wage in the type-*U* labor market $\left(\omega_U^*\right)$ while lowering the equilibrium wage in the type-*H* labor market $\left(\omega_H^*\right)$. The increase in the equilibrium wage in the type-*U* labor market serves to induce a flow of child labor from out-of-labor-force and into the type-*U* labor market. Given the nature of the equilibrium wage changes, the flow of child labor into the type-*U* labor market is exceeded by the flow out of the type-*U* labor market and into the type-*H* labor market.

Ban on harmful child labor

Labor market effects: A ban on harmful child labor with perfect effectiveness (which entails the shutdown of type-*H* firms) generates the same labor market outcomes discussed in Section 1.5, because the ban eliminates the child labor market beset by informational problems regarding occupational harm. Thus, the ban on harmful child labor returns the economy to the post-ban perfect information case, where the post-ban equilibrium wage in the type-*U* child labor market is ω_{U0}^{*B}. Although, the labor market effects of banning harmful child labor under the imperfect information case is qualitatively the same as under the perfect information case, the extent of the changes in the equilibrium wage in the type-*U* child labor market, child employment, and child labor force participation are larger.

Welfare effects: In order to determine the welfare effects of banning harmful child labor, we examine the utilities of households before and after the ban by separately considering three distinct groups of households: (1) those households that initially (pre-ban) chose not to send their children to work; (2) those households that initially chose to send their children to work in the type-U environment; and (3) those households that initially chose to send their children to work in the type-H environment.

1. *Households which initially chose not to send their children to work will experience no change in their utility from the ban on harmful child labor.*

Before the ban, all households $i \in I$ such that parental income y_i^P was equal to or exceeded $\hat{y}^P\left(\omega_{U1}^*\right)$ chose not to send their children to work. Define $I_{N1} = \left\{i : y_i^P \in \left[\hat{y}^P\left(\omega_{U1}^*\right), \overline{y}^P\right]\right\} \subset I.$ After the ban, the equilibrium wage in the type-U environment falls from ω_{U1}^* to $\omega_{U0}^{*B},$ which implies that the level of parental earnings that constitutes the child labor force participation cutoff also falls from $\hat{y}^P\left(\omega_{U1}^*\right)$ to $\hat{y}^P\left(\omega_{U0}^{*B}\right),$ indicating that the set of households which choose not to send their children to work expands as a result of the ban. The fall in the equilibrium wage in the type-U environment only acts to reinforce the pre-ban choice of households in I_{N1} as it widens the utility difference between their pre-ban choice and the remaining alternative (child work in the type-U environment). Thus, households in I_{N1} continue to choose not to send their children to work, and their utilities before and after the ban remain unchanged, which implies that these households are unaffected by the ban on harmful child labor.

2. *Households which initially chose to send their children to work in the type-U environment will experience a fall in their utility from the ban on harmful child labor.*

Before the ban, all households $i \in I$ such that parental income y_i^P was lower than the child labor force participation cutoff $\hat{y}^P\left(\omega_{U1}^*\right)$ and equal to or higher than the *U-H* cutoff $\tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)$ chose to send their children to work in the type-*U* environment. Define $I_{U1} = \left\{i : y_i^P \in \left[\tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right), \hat{y}^P\left(\omega_{U1}^*\right)\right)\right\} \subset I$. After the ban, households $i \in I_{U1}$ are forced to reevaluate their choice as a result of the fall in the equilibrium wage in the type-*U* environment caused by the flow of child labor from the now-nonexistent type-*H* environment. Those households $i \in I_{U1}$ such that the post-ban utility from sending their children to work in the type-*U* environment, $\ln\left(y_i^P + \omega_{U0}^{*B}\right) + \beta \ln \omega'_{U0}$, is less than or equal to the utility from their children not working, $\ln y_i^P + \beta \ln \omega'_{N0}$, will choose to remove their children from the type-*U* environment. Whereas, those households $i \in I_{U1}$ such that $\ln\left(y_i^P + \omega_{U0}^{*B}\right) + \beta \ln \omega'_{U0}$ exceeds $\ln y_i^P + \beta \ln \omega'_{N0}$ will choose to continue to send their children to work in the type-*U* environment. Regardless of their respective choices, all households in I_{U1} will experience a fall in utility after the ban as

$$\ln\left(y_i^P + \omega_{U1}^*\right) + \beta \ln \omega'_{U0} > \max\left\{\ln y_i^P + \beta \ln \omega'_{U0}, \ln\left(y_i^P + \omega_{U0}^{*B}\right) + \beta \ln \omega'_{U0}\right\}.$$

3. *Households which initially chose to send their children to work in the type-H environment can potentially experience an increase in their utility from the ban on harmful child labor.*

Before the ban, all households $i \in I$ such that parental earnings y_i^P were lower than the *U-H* cutoff $\tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)$ chose to send their children to work in the type-*H* environment. Define $I_{H1} = \left\{i : y_i^P \in \left[\underline{y}^P, \tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)\right)\right\} \subset I$. The ban on child work in the type-*H* environment makes households realize that the true level of occupational harm was such that the future earnings of child workers in the type-*H* environment before the ban was actually ω'_{H0} and not ω'_{H1}, where $\omega'_{H0} < \omega'_{H1}$. Households in I_{H1}

discover that the utility level from their choice was $\ln\left(y_i^P+\omega_{H1}^*\right)+\beta\ln\omega'_{H0}$ and not $\ln\left(y_i^P+\omega_{H1}^*\right)+\beta\ln\omega'_{H1}$ as previously thought, where $\ln\left(y_i^P+\omega_{H1}^*\right)+\beta\ln\omega'_{H0}<\ln\left(y_i^P+\omega_{H1}^*\right)+\beta\ln\omega'_{H1}$. After the ban, even though the equilibrium wage in the type-*U* environment falls to ω_{U0}^{*B}, those households $i\in I_{H1}$ such that $\ln\left(y_i^P+\omega_{H1}^*\right)+\beta\ln\omega'_{H0}<\max\left\{\ln\left(y_i^P+\omega_{U0}^{*B}\right)+\beta\ln\omega'_{U0},\ln y_i^P+\beta\ln\omega'_{N0}\right\}$ will experience an increase in utility as a result of the ban. On the other hand, those households $i\in I_{H1}$ such that

$\ln\left(y_i^P+\omega_{H1}^*\right)+\beta\ln\omega'_{H0}>\max\left\{\ln\left(y_i^P+\omega_{U0}^{*B}\right)+\beta\ln\omega'_{U0},\ln y_i^P+\beta\ln\omega'_{N0}\right\}$

will experience a decrease in their utility.

The extent of the equilibrium wage decline in the type-*U* environment as a result of the ban determines the incidence of households in I_{H1} which experience an increase in utility after the ban. At one extreme, if $\omega_{U0}^{*B}>\lambda\omega_{H1}^*$, where $\lambda=\left(\dfrac{\omega'_{H0}}{\omega'_{U0}}\right)^{\beta}<1$, then all households in I_{H1} experience an increase in utility from the ban on harmful child labor. At the other extreme, if $\omega_{U0}^{*B}<(\lambda-1)\,\tilde{y}^P\left(\omega_{U1}^*,\omega_{H1}^*\right)+\lambda\omega_{H1}^*$, then all households in I_{H1} experince a decrease in utility as a result of the ban on harmful child labor. Finally, if ω_{U0}^{*B} assumes some intermediate value, that is, $(\lambda-1)\,\tilde{y}^P\left(\omega_{U1}^*,\omega_{H1}^*\right)+\lambda\omega_{H1}^*<\omega_{U0}^{*B}<\omega_{H1}^*$, some households in I_{H1} experience an increase in utility while other households in I_{H1} experience a decrease in utility.

4. *If the post-ban equilibrium wage in the type-U environment,* ω_{U0}^{*B}, *lies in the range* $\left((\lambda-1)\,\tilde{y}^P\left(\omega_{U1}^*,\omega_{H1}^*\right)+\omega_{H1}^*,\lambda\omega_{H1}^*\right)$, *then those households in* I_{H1} *who have relatively higher parental income experience an increase in welfare from the ban, while those households in* I_{H1} *who have relatively lower parental income experience a decrease in welfare. In other words, if the post-ban equilibrium wage in the type-U environment,* ω_{U0}^{*B}, *is such that some households in* I_{H1} *observe a*

decrease in welfare, these households are invariably the poorest ones in terms of parental income. Furthermore, among these households, the poorer the household is in terms of parental income, the larger the welfare loss experienced by the household.

5. *Using any welfare measure w belonging to the class of social welfare functions W which are anonymous and increasing, the ban on harmful child labor will either be unambiguously welfare-reducing or, at best, welfare-ambiguous for households in the economy. Which welfare outcome result is realized hinges on the size of the wage decline in the type-U environment.*

In the case of informational and perceptional problems regarding occupational harm analyzed in this section, a ban on harmful child labor may still be first-order *dominance* worsening, but unlike the case analyzed in Section 1.5, this result is no longer guaranteed. Under the case examined in this section, the ban on harmful child labor may improve the welfare status of *some* households while hurting the welfare status of others, rendering the ban welfare-ambiguous. In no case is a ban on harmful child labor first-order *welfare* dominant.

1.7. Welfare effects of banning harmful child labor: the child as the welfare evaluator

Thus far, the welfare effects of banning harmful child labor with perfect enforcement under both informational settings have been evaluated by using the parent's utility function as the household's welfare function. In the benchmark model, the parent is considered the sole decisionmaker of the household, and the child is assumed to fully abide by the choices made by the parent. In addition, the parent is assumed to only care imperfectly for the future welfare of children as captured in the parental altruism factor β, where $0 < \beta < 1$. Consequently, the parent's and the child's utility functions may well diverge, in this case, in how they individually evaluate the future welfare of the child. That is, while the parent's utility function for household $i \in I$ is given by $\ln c_i + \beta \ln \omega_i'$ (see (1)), the child's utility function for household $i \in I$ is given by

$$\ln c_i + \ln \omega_i'. \tag{9}$$

In this section, we evaluate the welfare effects of banning harmful child labor from the *perspective of the child* given the divergence in preferences between parent and child specified above. As in the benchmark model, the parent remains the decisionmaker of the household and the optimizing behavior of the household follows the parent's preferences. This implies that the pre- and post-ban labor market choices of households under both informational settings are identical to those discussed in Sections 1.5 and 1.6. That is, the labor market undergoes, both qualitatively and quantitatively, the same wage and employment changes as a result of the ban. Table 1.1 summarizes the cases examined in this chapter, the welfare analysis completed thus far, and the analysis to follow.

Before discussing the labor market and welfare effects for cases 3 and 4, we first have to establish the counterfactual pre-ban situation if the child was the decisionmaker instead of the parent. The next section does this.

Counterfactual pre-ban employment choices: To facilitate the analysis of whether the welfare assessment of the ban differs between the parent and the child, it is useful to answer the following counterfactual question: *Given the child's utility function in (9) and supposing the child was the decisionmaker in the household, to what extent would the pre-ban labor market choices of the household be similar or different between the parent and the child?*

To begin answering this question, we know from Section 1.4. that if we used the parent's utility function and the parent was treated as the decisionmaker, the parental income y^P such that the household is indifferent between the child not working and the child working is given by $\hat{y}^P(\omega_U) = \dfrac{\omega_U}{\left(\dfrac{\omega'_N}{\omega'_U}\right)^{\beta} - 1}$.

Table 1.1. Welfare effects of a ban on harmful child labor under alternative cases—partial

Information settings regarding child occupational harm		Welfare evaluator in the household: Parent	Welfare evaluator in the household: Child
	Perfect information processing by parent	CASE 1 *Welfare reducing.* Discussed in Section 1.5.	CASE 3 *?* To be discussed in Section 1.7.B.
	Imperfect information processing by parent	CASE 2 *In general, welfare reducing.* *In special case, welfare ambiguous* Discussed in Section 1.6.	CASE 4 *?* To be discussed in Section 1.7.C.

Alternatively, if we use the child's utility function *and* the child is the decisionmaker, the "counterfactual labor participation cutoff" in terms of parental income y^P would be $\hat{y}_c^P(\omega_U) = \dfrac{\omega_U}{\left(\dfrac{\omega'_N}{\omega'_U}\right) - 1}$. Given that $\omega'_N > \omega'_U > 0$ and $0 < \beta < 1$, $\hat{y}_c^P(\omega_U) < \hat{y}^P(\omega_U)$. That is, a smaller number of households would choose child work if the child was the decisionmaker. Specifically, for households i such that $y_i^P \in \left(\hat{y}_c^P(\omega_U), \hat{y}^P(\omega_U)\right)$, the parent chose child work the type-*U* environment but the child would choose no work. The share of such households in the economy is given by $F_{\hat{y}_c^P - \hat{y}^P} = F\left(\hat{y}^P(\omega_U;\beta)\right) - F\left(\hat{y}_c^P(\omega_U)\right)$.[8]

[8] It is easy to see that $F_{\hat{y}_c^P - \hat{y}^P}$ is decreasing in β, indicating that as the level of parental altruism falls, the number of households suboptimally in the type-*U* environment (based on the child's assessment)

In analogous fashion, if we used the parent's utility function and the parent was treated as the decisionmaker, the level of parental income y^P such that the household is indifferent between the child working in the type-*U* environment and in the type-*H* environment is given by $\tilde{y}^P(\omega_U, \omega_H) = \frac{\left[\omega_H - \left(\frac{\omega'_U}{\omega'_H}\right)^\beta \omega_U\right]}{\left[\left(\frac{\omega'_U}{\omega'_H}\right)^\beta - 1\right]}$. Alternatively, if we use the child's utility function and the child is the decisionmaker, the "counterfactual *U*-*H* cutoff" in terms of parental income y^P would be $\tilde{y}^P_c(\omega_U, \omega_H) = \frac{\left[\omega_H - \left(\frac{\omega'_U}{\omega'_H}\right)\omega_U\right]}{\left[\left(\frac{\omega'_U}{\omega'_H}\right) - 1\right]}$. Given that $\omega'_U > \omega'_H$, $\omega_U < \omega_H$, and $0 < \beta < 1$, $\tilde{y}^P_c(\omega_U, \omega_H) < \tilde{y}^P(\omega_U, \omega_H)$. That is, a smaller number of households would choose to send their children to work in the type-*H* environment if the child was the decisionmaker. Specifically, for households i such that $y^P_i \in \left(\tilde{y}^P_c(\omega_U, \omega_H), \tilde{y}^P(\omega_U, \omega_H)\right)$, the parent chose child work in the type-*H* environment but the child would choose child work in the type-*U* environment. The share of such households in the economy is given by $F_{\tilde{y}^P_c - \tilde{y}^P} = F\left(\tilde{y}^P(\omega_U, \omega_H; \beta)\right) - F\left(\tilde{y}^P_c(\omega_U, \omega_H)\right)$.[9]

While possible, not all households necessarily experience a mismatch between the parent and the child in child employment choices. Specifically, for households i such that $y^P_i \in \left[\hat{y}^P(\omega_U), \overline{y}^P\right]$, the parent chose no child work and the child would do the same; for households i such that $y^P_i \in \left[\tilde{y}^P(\omega_U, \omega_H), \hat{y}^P_c(\omega_U)\right]$, the parent chose

---

increases.

[9] The share $F_{\tilde{y}^P_c - \tilde{y}^P}$ is decreasing in β, indicating that as the level of parental altruism falls, the number of households suboptimally in the type-*H* environment (based on the child's assessment) increases.

child work in the type-U environment and the child would do the same; and, finally, for households i such that $y_i^P \in \left[\underline{y}^P, \tilde{y}_c^P(\omega_U, \omega_H)\right]$, the parent chose child work in the type-H environment and the child would do the same.

Thus, in total, there are five types of households in this economy: three types of households where the child employment choices of the household match between the child and the parent and two types of households where they do not.

The above results are summarized in Table 1.2, with all households in the economy organized by parental income ranges from highest to lowest as we go from the top of the table to the bottom.

Table 1.2. Distribution of households by match and mismatch in parent and child decisions on child work— Perfect information processing case

Parental income range	Parent choice	Child choice	Result
$y_i^P \in \left[\hat{y}_c^P(\omega_U), \bar{y}^P\right]$	No child work	No child work	Match
$y_i^P \in \left(\hat{y}_c^P(\omega_U), \hat{y}^P(\omega_U)\right)$	Child work in type-U environment	No child work	Mismatch
$y_i^P \in \left[\tilde{y}^P(\omega_U, \omega_H), \hat{y}_c^P(\omega_U)\right]$	Child work in type-U environment	Child work in type-U environment	Match
$y_i^P \in \left(\tilde{y}_c^P(\omega_U, \omega_H), \tilde{y}^P(\omega_U, \omega_H)\right)$	Child work in type-H environment	Child work in type-U environment	Mismatch
$y_i^P \in \left[\underline{y}^P, \tilde{y}_c^P(\omega_U, \omega_H)\right]$	Child work in type-H environment	Child work in type-H environment	Match

Welfare effects of banning harmful child labor under perfect information processing

As a reminder, a ban on harmful child labor with perfect enforcement implies that child workers can no longer be employed in the type-H environment. This triggers a process of labor market adjustment, in which equilibrium wages in the type-U labor market falls, equilibrium employment in the type-*U* labor market rises, and total child employment falls (see Section 1.5 for a detailed discussion).

1. *Generally, given imperfect altruism by the parent, for households for which the child employment choice matches between the parent and the child, household welfare (assessed by either the parent or the child) either remains the same or falls as a result of the ban on harmful child labor.*

There are three types of households for which the decision made by the parent matches with that which would be made by the child if the child was the decisionmaker (see Table 1.2). The welfare results for these households follow the same logic as the welfare effects discussion in Section 1.5. Specifically, for households i such that $y_i^P \in \left[\hat{y}^P(\omega_U), \bar{y}^P\right]$, for which the parent chose no child work, there is no change in welfare as a result of the ban on harmful child labor. For households i such that $y_i^P \in \left[\tilde{y}^P(\omega_U, \omega_H), \hat{y}_c^P(\omega_U)\right]$, for which the parent chose child work in the type-*U* environment, the fall in equilibrium wages in the type-*U* labor market reduces the welfare of these households, even for those households for which the fall in wages precipitates the exit of the children from the labor market altogether. Finally, for households i such that $y_i^P \in \left[\bar{y}^P, \tilde{y}_c^P(\omega_U, \omega_H)\right]$, for which the parent chose child work in the type-*H* environment, the elimination of this choice from their choice set reduces their welfare, irrespective of whether the children end up shifting to the type-*U* environment or exiting from the labor market altogether.

2. *Generally, given imperfect altruism by the parent, for households for which the child employment choice made by the parent does not match with that which would*

be made by the child, household welfare (as assessed by the child) increases as a result of the ban if the following two conditions simultaneously hold:

(a) $\hat{y}^P\left(\omega_U^{*B}\right) \leq \hat{y}_c^P\left(\omega_U^*\right)$, *and*

(b) *for all i such that* $y_i^P \in \left[\tilde{y}_c^P\left(\omega_U^*, \omega_H^*\right), \tilde{y}^P\left(\omega_U^*, \omega_H^*\right)\right]$, $\ln\left(y_i^P + \omega_U^{*B}\right) + \ln \omega_U' > \ln\left(y_i^P + \omega_H^*\right) + \ln \omega_H'$.

Condition (a) is relevant for households i such that $y_i^P \in \left(\hat{y}_c^P\left(\omega_U^*\right), \hat{y}^P\left(\omega_U^*\right)\right)$. It states that so long as the post-ban labor force participation cutoff $\hat{y}^P\left(\omega_U^{*B}\right)$ is equal to or less than the counterfactual labor force participation cutoff $\hat{y}_c^P\left(\omega_U^*\right)$, that is, the labor market adjusts so that the children that exit the labor market at least include all those for whom the type-*U* environment was suboptimal to begin with, these households experience a gain in welfare as a result of the ban. This condition bounds the post-ban equilibrium wage in the type-*U* labor market $\left(\omega_U^{*B}\right)$ from above.

Likewise, Condition (b) pertains to households i such that $y_i^P \in \left(\tilde{y}_c^P\left(\omega_U^*, \omega_H^*\right), \tilde{y}^P\left(\omega_U^*, \omega_H^*\right)\right)$. So long as the post-ban equilibrium wage in the type-*U* labor market $\left(\omega_U^{*B}\right)$ yields a higher level of welfare than that observed in the type-*H* labor market, these households also experience a gain in welfare as a result of the ban. This condition bounds the post-ban equilibrium wage in the type-*U* labor market $\left(\omega_U^{*B}\right)$ from below. The bounding from above and below sets a range of values for ω_U^{*B} for which Conditions (a) and (b) simultaneously hold.

In general, the net welfare effect of banning harmful child labor under the perfect information setting and where the child is the welfare evaluator for the household is *ambiguous*. As shown above, some households experience a gain in welfare, others a loss in welfare, and yet others experience no change in their welfare. However, under a special case, a ban on harmful child labor in this setting can yield a welfare improvement.

3. *Specifically, given imperfect altruism by the parent, household welfare (as assessed by the child) either increases or remains unchanged for all households in the economy if the following four conditions simultaneously hold:*
   (a) $\tilde{y}_c^P\left(\omega_U^*, \omega_H^*\right) = \underline{y}^P$,
   (b) $\hat{y}_c^P\left(\omega_U^*\right) = \tilde{y}^P\left(\omega_U^*, \omega_H^*\right)$,
   (c) $\hat{y}^P\left(\omega_U^{*B}\right) = \tilde{y}^P\left(\omega_U^*, \omega_H^*\right)$, *and*
   (d) *for all i such that* $y_i^P \in \left[\underline{y}^P, \tilde{y}^P\left(\omega_U^*, \omega_H^*\right)\right]$, $\ln\left(y_i^P + \omega_U^{*B}\right) + \ln \omega_U' > \ln\left(y_i^P + \omega_H^*\right) + \ln \omega_H'$.

   Thus, under the above four conditions, using any welfare measure w belonging to the class of social welfare functions W which are anonymous and increasing, the ban on harmful child will be unambiguously welfare-improving.

Conditions (a) and (b) state that the parental incomes associated with the counterfactual *U-H* and labor force participation cutoffs ($\tilde{y}_c^P\left(\omega_U^*, \omega_H^*\right)$ and $\hat{y}_c^P\left(\omega_U^*\right)$) perfectly coincide with the lower bound on parental income in the economy $\underline{y}^P$ and the actual pre-ban *U-H* cutoff $\tilde{y}^P\left(\omega_U^*, \omega_H^*\right)$, respectively. Condition (a) states that, if the child was the decisionmaker, all households which chose child work in the type-*H* environment would have instead chosen child work in the type-*U* environment. Likewise, Condition (b) states that, if the child was the decisionmaker, all households which chose child work in the type-*U* environment would have instead chosen no child work. That is, all households which sent their children into the labor market face a situation of incompatible choices between the child and the parent. Condition (c) states that when child employment adjusts as a result of the ban, it adjusts to the point where the post-ban labor force participation cutoff $\hat{y}^P\left(\omega_U^{*B}\right)$perfectly coincides with the pre-ban *U-H* cutoff $\tilde{y}^P\left(\omega_U^*, \omega_H^*\right)$. That is, all child workers initially in the type-*H* environment flow into the type-*U* environment, and all child workers initially in the type-*U* environment exit the labor market. Finally, Condition (d) pertains to child workers initially in the type-*H* environment. It states that, despite the downward adjustment, the post-ban equilibrium wage in the type-*U* labor market $\left(\omega_U^{*B}\right)$ yields a higher level of welfare than that observed in the type-*H* labor market.

Welfare effects of banning harmful child labor under imperfect information processing: If parents in the economy uniformly overestimate the future earnings level of children who works in the type-*H* environment—that is, under imperfect information processing, they assume future earnings of ω'_{H1}, but under perfect information processing, they would assume future earnings of ω'_{H0}, where $\omega'_{H1} > \omega'_{H0}$—and the parents are decisionmakers for the households, this causes a distortion in the labor market (see Section 1.6 for the discussion of the distortionary effects of imperfect information processing on the labor market).

Combining imperfect information processing by the parent with imperfect parental altruism exacerbates the problem of a mismatch in choices between the parent and the child; it also spreads the problem to a larger number of households. Specifically, for households i such that $y_i^P \in \left(\hat{y}_c^P\left(\omega_{U0}\right), \hat{y}^P\left(\omega_{U1}\right)\right)$, where $\hat{y}^P\left(\omega_{U1}\right) > \hat{y}^P\left(\omega_{U0}\right)$, the parent chose child work in the type-*U* environment but the child would choose no work. Similarly, for households i such that $y_i^P \in \left(\tilde{y}_c^P\left(\omega_{U0}, \omega_{H0}\right), \tilde{y}^P\left(\omega_{U1}, \omega_{H1}\right)\right)$, where $\tilde{y}^P\left(\omega_{U1}, \omega_{H1}\right) > \tilde{y}^P\left(\omega_{U0}, \omega_{H0}\right)$, the parent chose child work in the type-*H* environment but the child would choose child work in the type-U environment.

However, as under the perfect information processing setting, these are three subsets of households in the economy where, despite the imperfect information processing and imperfect altruism by the parent, the choice made by the parent matches with the one the child would have made if the child was the decisionmaker. Specifically, for households i such that $y_i^P \in \left[\hat{y}^P\left(\omega_{U1}\right), \overline{y}^P\right]$, the parent chose no child work and the child would choose the same; for households i such that $y_i^P \in \left[\tilde{y}^P\left(\omega_{U1}, \omega_{H1}\right), \hat{y}_c^P\left(\omega_{U0}\right)\right]$, the parent chose child work in the type-U environment and the child would choose the same; and, finally, for households i such that $y_i^P \in \left[\underline{y}^P, \tilde{y}_c^P\left(\omega_{U0}, \omega_{H0}\right)\right]$, the parent chose child work in the type-H environment and the child would choose the same.

The above results are summarized in Table 1.3, with all households in the economy organized by parental income ranges from highest to lowest as we go from the top of the table to the bottom.

Table 1.3. Distribution of households by match and mismatch in parent and child decisions on child work—imperfect information processing case

Parental income range	Parent choice	Child choice	Result
$y_i^P \in \left[\hat{y}^P\left(\omega_{U1}\right), \overline{y}^P\right]$	No child work	No child work	Match
$y_i^P \in \left(\hat{y}_c^P\left(\omega_{U0}\right), \hat{y}^P\left(\omega_{U1}\right)\right)$	Child work in type-U environment	No child work	Mismatch
$y_i^P \in \left[\tilde{y}^P\left(\omega_{U1}, \omega_{H1}\right), \hat{y}_c^P\left(\omega_{U0}\right)\right]$	Child work in type-U environment	Child work in type-U environment	Match
$y_i^P \in \left(\tilde{y}_c^P\left(\omega_{U0}, \omega_{H0}\right), \tilde{y}^P\left(\omega_{U1}, \omega_{H1}\right)\right)$	Child work in type-H environment	Child work in type-U environment	Mismatch
$y_i^P \in \left[\underline{y}^P, \tilde{y}_c^P\left(\omega_{U0}, \omega_{H0}\right)\right]$	Child work in type-H environment	Child work in type-H environment	Match

Given this, the welfare results for the imperfect information processing case are *qualitatively* the same as for the perfect information processing case. They are as follows.

1. *Generally, in a setting of imperfect information processing and given imperfect altruism by the parent, for households for which the child employment choice matches between the parent and the child, household welfare (as assessed by either the parent or the child) either remains the same or falls as a result of the ban on harmful child labor.*

2. *Generally, in a setting of imperfect information processing and given imperfect altruism by the parent, for households for which the child employment choice made by the parent does not match with that which would be made by the child, household welfare (as assessed by the child) increases as a result of the ban if the following two conditions simultaneously hold:*
   (a) $\hat{y}^P\left(\omega_{U0}^{*B}\right) \leq \hat{y}_c^P\left(\omega_{U0}^*\right)$*; and*
   (b) *for all i such that* $y_i^P \in \left[\tilde{y}_c^P\left(\omega_{U0}^*, \omega_{H0}^*\right), \tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)\right]$, $\ln\left(y_i^P + \omega_{U0}^{*B}\right) + \ln \omega_{U0}' > \ln\left(y_i^P + \omega_{H1}^*\right) + \ln \omega_{H0}'$.

3. *Specifically, in a setting of imperfect information processing and given imperfect altruism by the parent, household welfare (as assessed by the child) either increases or remains unchanged for households in the economy if the following four conditions simultaneously hold:*
   (a) $\tilde{y}_c^P\left(\omega_{U0}^*, \omega_{H0}^*\right) = \underline{y}^P$*;*
   (b) $\hat{y}_c^P\left(\omega_{U0}^*\right) = \tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)$*;*
   (c) $\hat{y}^P\left(\omega_{U0}^{*B}\right) = \tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)$*; and*
   (d) *for all i such that* $y_i^P \in \left[\underline{y}^P, \tilde{y}^P\left(\omega_{U1}^*, \omega_{H1}^*\right)\right]$, $\ln\left(y_i^P + \omega_{U0}^{*B}\right) + \ln \omega_{U0}' > \ln\left(y_i^P + \omega_{H1}^*\right) + \ln \omega_{H0}'$.

   Thus, using any welfare measure w belonging to the class of social welfare functions W which are anonymous and increasing, the ban on harmful child will be unambiguously welfare-improving.

Thus, to summarize, in general, in a setting where the parent is imperfectly altruistic and is the decisionmaker but the child is the household welfare evaluator, a ban on harmful child labor with perfect enforcement can result in some households experiencing a gain in welfare but *not all*. Specifically, those households for which the child employment choice made by the parent does not match with what the child would have made can potentially experience a welfare gain; the numbers of such households increase when, under the imperfect information processing case, parents also underestimate the future consequences of the harm suffered by the child in the type-*H* environment. However, under certain special conditions, all households in the economy can experience a welfare gain. Specifically, if the parental decision to send the child to work in either the type-*U* or type-*H* environment is suboptimal from the point of the view of the child, a ban on harmful child labor which results in these children ending up in their optimal employment choices (no work for type-*U* workers and work in type-*U* environment if type-*H* workers) yields an unambiguous welfare improvement. These results hold under the case of perfect as well as imperfect information processing by the parent regarding the nature of occupational harm in the type-*H* environment. The full set of welfare results for the different cases is summarized in Table 1.4.

1.8. Conclusion

This chapter examines the labor market and welfare effects of banning harmful child labor with complete effectiveness. Harm is defined as certain, irreversible adverse health effects that emerge later in life as a result of child work undertaken in particular (type-*H*) workplaces, and where a ban on harmful child labor leads to the shutdown of such workplaces. These effects of the ban are analyzed in two different informational settings.

Table 1.4. Welfare effects of a ban on harmful child labor under alternative cases—complete

		Welfare evaluator in the household	
		Parent	Child
Information settings regarding child occupational harm	Perfect information processing by parent	CASE 1 *Welfare reducing.* Discussed in Section 1.5.	CASE 3 *In general, welfare ambiguous.* *In special case, welfare improving.* Discussed in Section 1.7.B.
	Imperfect information processing by parent	CASE 2 *In general, welfare reducing.* *In special case, welfare ambiguous* Discussed in Section 1.6.	CASE 4 *In general, welfare ambiguous.* *In special case, welfare improving.* Discussed in Section 1.7.C.

In the first setting (our benchmark case), households possess perfect information on the nature of occupational harm associated with different types of child employment environments and correctly assess the effect of such harm on the future wages of child workers. Under the perfect information case, in terms of labor market effects, the ban reduces the equilibrium wage and increases employment in the type-*U* child labor market, although it reduces overall child employment. In terms of welfare effects, given the labor market adjustments that occur, the ban unambiguously reduces the welfare of all households which initially chose to send their children to work; those households which initially chose not to send their children to work experience no change in the welfare. In addition, among households which initially chose to send their children to work in type-*H* firms, the poorer the household is in terms of parental income, the larger the welfare loss experienced by the household due to the ban on harmful child labor.

The extension to the benchmark case relaxes the assumption of perfect information availability and information processing regarding occupational harm experienced by child workers and alternatively considers a setting in which informational problems are rife, causing all households to uniformly underestimate the adverse health effects of child work in type-*H* firms. This systematic misjudgment on the part of households creates a distortion in the labor market. Relative to the perfect information case, the labor market equilibrium that emerges is characterized by higher wages and lower employment in the type-*U* child labor market, lower wages and higher employment in the type-*H* child labor market, lower compensating wage differentials, and higher overall child employment.

Banning harmful child labor, by precipitating the shutdown of type-*H* firms, eliminates the segment of the child labor market beset by the informational and perceptional problems. It follows then that, in terms of labor market effects, the ban under the imperfect information case has qualitatively the same set of effects as under the perfect information case, although the magnitude of the labor market adjustments that occur are larger than under the former.

In terms of welfare effects, depending on the size of the resulting wage decline in the type-*U* child labor market, the ban can potentially raise the welfare of some if not all households that initially chose to send their children to work in the type-*H* child labor market. The more moderate is the wage decline, the larger is the share of households that experience welfare gains. However, as the wage falls in the type-*U* child labor market, those households that initially chose to send their child to work in the type-*U* child labor market will experience welfare losses. Furthermore, among those households that initially chose to send their child to work in the type-*H* child labor market, if the wage decline is such that some fraction of these households experience welfare gains while other households experience welfare losses, the households that experience welfare losses are invariably the poorest. Thus, under the imperfect information case, the ban on harmful child labor is, at best, welfare-ambiguous.

The benchmark model also assumes that the parent, who is the decisionmaker for the household, cares imperfectly for the welfare of the child. Specifically, the parent is considered not to care for the future welfare of the child to the same extent that the child does. Given these divergent preferences, treating the child as the welfare evaluator for the household, a ban on harmful child labor with perfect enforcement can result in some households experiencing a gain in welfare but not all. Specifically, those households for which the employment choice made by the parent does not match with what the child would have made can potentially experience a welfare gain; the numbers of such households increase when, under the imperfect information processing case, parents also underestimate the future consequences of the harm suffered by child workers in the type-*H* environment. Furthermore, under certain special conditions, all households in the economy can experience a welfare gain. Specifically, if the parental decision to send the child to work in either the type-*U* or type-*H* environment is suboptimal from the point of the view of the child, a ban on harmful child labor which results in these children ending up in their optimal employment choices (no work for type-*U* workers and work in type-*U* environment if type-*H* workers) yields an unambiguous welfare improvement. These results hold under both informational cases.

In general, the welfare results that emerge critically hinge on the post-ban adjustments that take place in the labor market as it recalibrates. The welfare improvement results in the chapter rely on the wage in the type-*U* environment not falling below a certain threshold (where the threshold varies depending on the particular case under examination). The wage decline that occurs as a result of the ban is more likely to be moderate if the size of the type-*H* child labor market is small relative to the type-*U* child labor market. It also helps if labor demand in the type-*U* child labor market is relatively elastic or if labor demand is increasing. Some of this increase in labor demand can conceivably come about if instead of type-*H* firms shutting down as a result of the ban, some of them raise their occupational health and safety standards to become type-*U* firms and continue to employ children.

APPENDIX

Proof 1: Market equilibrium existence, uniqueness, and stability

Dropping the star superscripts from the wage variables, define

$F_U(\omega_U, \omega_H) = D_U(\omega_U) - S_U(\omega_U, \omega_H) = 0,$ and

$F_H(\omega_U, \omega_H) = D_H(\omega_H) - S_H(\omega_U, \omega_H) = 0.$

Solving for partial derivatives, we obtain the Jacobian matrix

$$J = \begin{pmatrix} \frac{\partial F_U}{\partial \omega_U} & \frac{\partial F_U}{\partial \omega_H} \\ \frac{\partial F_H}{\partial \omega_U} & \frac{\partial F_H}{\partial \omega_H} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial D_U}{\partial \omega_U} - \frac{\partial S_U}{\partial \omega_U} & -\frac{\partial S_U}{\partial \omega_H} \\ -\frac{\partial S_H}{\partial \omega_U} & \frac{\partial D_H}{\partial \omega_H} - \frac{\partial S_H}{\partial \omega_H} \end{pmatrix}.$$

The determinant of J is

$$|J| = \frac{\partial F_U}{\partial \omega_U}\frac{\partial F_H}{\partial \omega_H} - \frac{\partial F_H}{\partial \omega_U}\frac{\partial F_U}{\partial \omega_H}$$

$$= \left(\frac{\partial D_U}{\partial \omega_U} - \frac{\partial S_U}{\partial \omega_U}\right)\left(\frac{\partial D_H}{\partial \omega_H} - \frac{\partial S_H}{\partial \omega_H}\right) - \left(-\frac{\partial S_H}{\partial \omega_U}\right)\left(-\frac{\partial S_U}{\partial \omega_H}\right)$$

$$= \frac{\partial D_U}{\partial \omega_U}\frac{\partial D_H}{\partial \omega_H} - \frac{\partial D_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H} - \frac{\partial D_H}{\partial \omega_H}\frac{\partial S_U}{\partial \omega_U} + \frac{\partial S_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H} - \frac{\partial S_H}{\partial \omega_U}\frac{\partial S_U}{\partial \omega_H}.$$

It is straightforward to see that

$\frac{\partial D_U}{\partial \omega_U}\frac{\partial D_H}{\partial \omega_H} = (-)(-) = (+),$

$\frac{\partial D_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H} = (-)(+) = (-),$ and

$\frac{\partial D_H}{\partial \omega_H}\frac{\partial S_U}{\partial \omega_U} = (-)(+) = (-).$

For the others, we obtain the following detailed derivatives:

$$\frac{\partial S_U}{\partial \omega_U} = I\left(f\left(\hat{y}^P\right)\frac{\partial \hat{y}^P}{\partial \omega_U} - f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_U}\right),$$

$$\frac{\partial S_H}{\partial \omega_H} = I\left(f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H} \right),$$

$$\frac{\partial S_U}{\partial \omega_H} = -I\left(f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H} \right), \text{ and}$$

$$\frac{\partial S_H}{\partial \omega_U} = I\left(f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_U} \right).$$

Thus,

$$\frac{\partial S_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H} - \frac{\partial S_H}{\partial \omega_U}\frac{\partial S_U}{\partial \omega_H}$$

$$= I^2\left(f\left(\hat{y}^P\right) f\left(\tilde{y}^P\right)\frac{\partial \hat{y}^P}{\partial \omega_U}\frac{\partial \tilde{y}^P}{\partial \omega_H} - f^2\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_U}\frac{\partial \tilde{y}^P}{\partial \omega_H} + f^2\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_U}\frac{\partial \tilde{y}^P}{\partial \omega_H} \right)$$

$$= I^2 f\left(\hat{y}^P\right) f\left(\tilde{y}^P\right)\frac{\partial \hat{y}^P}{\partial \omega_U}\frac{\partial \tilde{y}^P}{\partial \omega_H}.$$

Given $\frac{\partial \hat{y}^P}{\partial \omega_U} > 0$ and $\frac{\partial \tilde{y}^P}{\partial \omega_H} > 0,$

$$\frac{\partial S_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H} - \frac{\partial S_H}{\partial \omega_U}\frac{\partial S_U}{\partial \omega_H} = I^2 f\left(\hat{y}^P\right) f\left(\tilde{y}^P\right)\frac{\partial \hat{y}^P}{\partial \omega_U}\frac{\partial \tilde{y}^P}{\partial \omega_H} > 0.$$

Thus,

$$|J| = \left(\frac{\partial D_U}{\partial \omega_U}\frac{\partial D_H}{\partial \omega_H}\right) - \left(\frac{\partial D_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H}\right) - \left(\frac{\partial D_H}{\partial \omega_H}\frac{\partial S_U}{\partial \omega_U}\right) + \left(\frac{\partial S_U}{\partial \omega_U}\frac{\partial S_H}{\partial \omega_H} - \frac{\partial S_H}{\partial \omega_U}\frac{\partial S_U}{\partial \omega_H}\right)$$

$$= (+) - (-) - (-) + (+) > 0.$$

The determinant of J is nonzero for all positive pairs of child wages. Since the determinant of J is not identically zero for all values of child wages, the functions F_U and F_H are nonlinearly independent, which means that a nontrivial solution exists. Further, since the determinant of J is always positive, this ensures that the system of nonlinear equations has a unique solution which is stable.

Claim: ω_U^* *is increasing in* ω_H' *;* ω_H^* *is decreasing in* ω_H' *.*

Proof 2: Dropping the star superscripts from the variables, define

$$F_U\left(\omega_U, \omega_H, \omega_H'\right) = D_U\left(\omega_U\right) - S_U\left(\omega_U, \omega_H, \omega_H'\right) = 0, \text{ and} \qquad (.10)$$

$$F_H\left(\omega_U,\omega_H,\omega'_H\right)=D_H\left(\omega_H\right)-S_H\left(\omega_U,\omega_H,\omega'_H\right)=0. \quad (.11)$$

By totally differentiating equations (A.1) and (A.2) with respect to ω_U, ω_H, and ω'_H, we obtain

$$\frac{\partial F_U}{\partial \omega_U}d\omega_U+\frac{\partial F_U}{\partial \omega_H}d\omega_H+\frac{\partial F_U}{\partial \omega'_H}d\omega'_H=0,\ \text{and} \quad (.12)$$

$$\frac{\partial F_H}{\partial \omega_U}d\omega_U+\frac{\partial F_H}{\partial \omega_H}d\omega_H+\frac{\partial F_H}{\partial \omega'_H}d\omega'_H=0. \quad (.13)$$

By simultaneously solving (A.3) and (A.4) for $d\omega_U$ and $d\omega_H$, we obtain

$$d\omega_U=\left[\frac{\dfrac{\partial F_H}{\partial \omega'_H}\dfrac{\partial F_U}{\partial \omega_H}-\dfrac{\partial F_U}{\partial \omega'_H}\dfrac{\partial F_H}{\partial \omega_H}}{|J|}\right]d\omega'_H$$

$$d\omega_H=\left[\frac{\dfrac{\partial F_U}{\partial \omega'_H}\dfrac{\partial F_H}{\partial \omega_U}-\dfrac{\partial F_H}{\partial \omega'_H}\dfrac{\partial F_U}{\partial \omega_U}}{|J|}\right]d\omega'_H,$$

where $|J|=\dfrac{\partial F_U}{\partial \omega_U}\dfrac{\partial F_H}{\partial \omega_H}-\dfrac{\partial F_H}{\partial \omega_U}\dfrac{\partial F_U}{\partial \omega_H}>0$ (see Proof 1 in the Appendix).

It follows that

$$\frac{\partial \omega_U}{\partial \omega'_H}=\frac{\dfrac{\partial F_H}{\partial \omega'_H}\dfrac{\partial F_U}{\partial \omega_H}-\dfrac{\partial F_U}{\partial \omega'_H}\dfrac{\partial F_H}{\partial \omega_H}}{|J|},\ \text{and} \quad (.14)$$

$$\frac{\partial \omega_H}{\partial \omega'_H}=\frac{\dfrac{\partial F_U}{\partial \omega'_H}\dfrac{\partial F_H}{\partial \omega_U}-\dfrac{\partial F_H}{\partial \omega'_H}\dfrac{\partial F_U}{\partial \omega_U}}{|J|}. \quad (.15)$$

The numerator in equation (A.5) can be expanded to

$$\frac{\partial F_H}{\partial \omega'_H}\frac{\partial F_U}{\partial \omega_H}-\frac{\partial F_U}{\partial \omega'_H}\frac{\partial F_H}{\partial \omega_H}$$

$$=\frac{\partial S_H}{\partial \omega'_H}\frac{\partial S_U}{\partial \omega_H}+\frac{\partial S_U}{\partial \omega'_H}\left(\frac{\partial D_H}{\partial \omega_H}-\frac{\partial S_H}{\partial \omega_H}\right)$$

$$=\frac{\partial S_H}{\partial \omega'_H}\frac{\partial S_U}{\partial \omega_H}+\frac{\partial S_U}{\partial \omega'_H}\frac{\partial D_H}{\partial \omega_H}-\frac{\partial S_U}{\partial \omega'_H}\frac{\partial S_H}{\partial \omega_H}.$$

We have that

$$\frac{\partial S_U}{\partial \omega_H'} = -I\left(f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\right) < 0,$$

$$\frac{\partial S_H}{\partial \omega_H'} = I\left(f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\right) > 0,$$

$$\frac{\partial S_U}{\partial \omega_H'}\frac{\partial D_H}{\partial \omega_H} = (-)(-) > 0, \text{where } \frac{\partial \tilde{y}^P}{\partial y_H'} > 0, \text{ and}$$

$$\frac{\partial S_H}{\partial \omega_H'}\frac{\partial S_U}{\partial \omega_H} - \frac{\partial S_U}{\partial \omega_H'}\frac{\partial S_H}{\partial \omega_H} = I^2\left(f^2\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\frac{\partial \tilde{y}^P}{\partial \omega_H} - f^2\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\frac{\partial \tilde{y}^P}{\partial \omega_H}\right) = 0.$$

Thus,

$$\frac{\partial F_H}{\partial \omega_H'}\frac{\partial F_U}{\partial \omega_H} - \frac{\partial F_U}{\partial \omega_H'}\frac{\partial F_H}{\partial \omega_H} = \frac{\partial S_U}{\partial \omega_H'}\frac{\partial D_H}{\partial \omega_H} > 0, \text{which implies that } \frac{\partial \omega_U}{\partial \omega_H'} > 0.$$

Similarly, the numerator in equation (A.6) can be expanded to

$$\frac{\partial F_U}{\partial \omega_H'}\frac{\partial F_H}{\partial \omega_U} - \frac{\partial F_H}{\partial \omega_H'}\frac{\partial F_U}{\partial \omega_U}$$
$$= \frac{\partial S_U}{\partial \omega_H'}\frac{\partial S_H}{\partial \omega_U} + \frac{\partial S_H}{\partial \omega_H'}\left(\frac{\partial D_U}{\partial \omega_U} - \frac{\partial S_U}{\partial \omega_U}\right)$$
$$= \frac{\partial S_U}{\partial \omega_H'}\frac{\partial S_H}{\partial \omega_U} + \frac{\partial S_H}{\partial \omega_H'}\frac{\partial D_U}{\partial \omega_U} - \frac{\partial S_H}{\partial \omega_H'}\frac{\partial S_U}{\partial \omega_U}.$$

We have that

$$\frac{\partial S_H}{\partial \omega_H'}\frac{\partial D_U}{\partial \omega_U} = (+)(-) < 0, \text{ and}$$

$$\frac{\partial S_U}{\partial \omega_H'}\frac{\partial S_H}{\partial \omega_U} - \frac{\partial S_H}{\partial \omega_H'}\frac{\partial S_U}{\partial \omega_U}$$
$$= I^2\left(-f^2\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\frac{\partial \tilde{y}^P}{\partial \omega_U} - f\left(\hat{y}^P\right)f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\frac{\partial \hat{y}^P}{\partial \omega_U} + f^2\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\frac{\partial \tilde{y}^P}{\partial \omega_U}\right)$$
$$= -I^2\left(f\left(\hat{y}^P\right)f\left(\tilde{y}^P\right)\frac{\partial \tilde{y}^P}{\partial \omega_H'}\frac{\partial \hat{y}^P}{\partial \omega_U}\right) < 0.$$

Thus, $\dfrac{\partial F_U}{\partial \omega_H'}\dfrac{\partial F_H}{\partial \omega_U} - \dfrac{\partial F_H}{\partial \omega_H'}\dfrac{\partial F_U}{\partial \omega_U} < 0,$ which implies that $\dfrac{\partial \omega_H}{\partial \omega_H'} < 0.$

REFERENCES

Baland, Jean-Marie, and James A. Robinson. 2000. Is Child Labor Inefficient? *Journal of Political Economy*. v108 (n4):663-79.

Basu, Kaushik, and Pham Hoang Van. 1998. The Economics of Child Labor. *American Economic Review*. v88 (n3):412-27.

Basu, Kaushik. 1999. Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards. *Journal of Economic Literature*. v37 (n3):1083-1119.

Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti. 2006. Child Labor and Agricultural Shocks. *Journal of Development Economics*. v81 (n1):80-96.

Dessy, Sylvain E., and Stephane Pallage. 2001. Child Labor and Coordination Failures. *Journal of Development Economics*. v65 (n2):469-76.

Dessy, Sylvain, and Stephane Pallage. 2003. *The economics of child trafficking*. CIRPEE Working Paper No. 03-23.

Dessy, Sylvain E., and Stephane Pallage. 2005. A Theory of the Worst Forms of Child Labour. *Economic Journal*. v115 (n500):68-87.

Edmonds, Eric V. 2006. Child Labor and Schooling Responses to Anticipated Income in South Africa. *Journal of Development Economics*. v81 (n2):386-414.

Fields, Gary S. 2001. *Distribution and development: A new look at the developing world*. New York: Russell Sage Foundation; Cambridge and London: MIT Press.

Grootaert, Christiaan, and Ravi Kanbur. 1995. Child Labour: An Economic Perspective. *International Labour Review*. v134 (n2):187-203.

Hagemann, Frank, Yacouba Diallo, Alex Etienne, and Farhad Mehran. 2006. *Global Child Labour Trends 2000 to 2004.* Geneva: International Labour Organization.

Helweg-Larsen, Marie and Shepperd, James A. 2001. Do Moderators of the Optimistic Bias affect Personal or Target Risk Estimates? A Review of the Literature. *Personality and Social Psychology Review.* v5 (n1):74-95.

ILO. 2002. *Every Child Counts: New Global Estimates of Child Labour.* Geneva: International Labour Office.

Klein, William, M., and Neil D. Weinstein. 1997. Social comparison and unrealistic optimism about personal risk. In Bram P. Buunk and Frederick X. Gibbons (eds.) *Health, coping and well-being: Perspectives from social comparison theory*. Hillsdale, NJ: Lawrence Erlbaum.

López-Calva, Luis. 2000. *Child Labor, Compulsory Schooling, and Social Norms: Statics and Dynamics.* Ph.D. dissertation, Cornell Univeristy.

Ranjan, Priya. 1999. An Economic Analysis of Child Labor. *Economics Letters*. v64 (n1):99-105.

Ranjan, Priya. 2001. Credit Constraints and the Phenomenon of Child Labor. *Journal of Development Economics*. v64 (n1):81-102.

Ray, Ranjan. 2000. Child Labor, Child Schooling, and Their Interaction with Adult Labor: Empirical Evidence for Peru and Pakistan. *World Bank Economic Review*. v14 (n2):347-67.

Rogers, Carol Ann, and Kenneth A. Swinnerton. 2008. A Theory of Exploitative Child Labor. *Oxford Economic Papers*. v60 (n1):20-41.

CHAPTER 2

DO WAGES COMPENSATE FOR HARMFUL CHILD LABOR?

QUANTILE REGRESSION EVIDENCE FROM THE PHILIPPINES

2.1. Introduction

Child labor remains a mass phenomenon in much of the developing world, particularly in its poorer parts such as sub-Saharan Africa and South Asia. The International Labor Organization (ILO), the major source of statistics on the extent of child labor worldwide, estimates that 191 million children ages 5-14 years were economically active in 2004, of which, 166 million child workers were categorized as child laborers (Hagemann et al. 2006).[10] More disconcertingly, 74 million of these 166 million child workers (or over 40%) were considered to be employed in harmful or exploitative situations or conditions. Of these, roughly 8 million children were deemed to be in what are termed as unconditional worst forms of child labor (ILO 2002).[11]

In order to gain some insight into the reasons behind harmful child labor and the socioeconomic conditions underlying its existence, in this chapter, I examine whether child workers in harmful employment settings are compensated monetarily in the form of higher labor market earnings. This is an important question, especially from a policy perspective, as it indicates the extent to which market mechanisms operate to compensate child workers for the disutility of experiencing harmful child labor. In particular, it also provides indirect evidence of the extent to which child workers (or their parents as decisionmakers on behalf of their children) are informed

[10] Economic activity covers all market production (paid work) and certain types of non-market production (unpaid work), including production of goods for own use. Child labor consists of all children under age 15 years who are economically active excluding (1) those who are under age 5 years and (2) those ages 12-14 years who spend less than 14 hours a week on their jobs, unless their activities or occupations are hazardous by nature or circumstance. Added to this are children ages 15-17 years in the worst forms of child labor (ILO 2002).

[11] The worst forms of child labor refers to child labor in the context of slavery or slave-like conditions, in prostitution or pornography, in illicit activities such as drug trafficking, or in conditions that are likely to affect the health or safety of the children involved. Unconditional worst forms of child labor exclude the last type.

about the extent and nature of the potential harm they are facing, and whether they have the market power to extract compensation for experiencing harmful child labor.

The theoretical literature on harmful child labor has proceeded in two distinct directions. In the first, labor markets are characterized along largely standard textbook lines where poor households possess some information about the nature of potential harm associated with different employment opportunities for their children and maintain the right to exit from an employment relationship (see, e.g., Chapter 1; Dessy and Pallage 2005). In the second, households effectively relinquish the child's right to exit from an employment relationship such as under child servitude or child trafficking (see, e.g., Rogers and Swinnerton 2003; Dessy and Pallage 2003). The predictions regarding the relationship between child earnings and harmful child labor between the two characterizations of the labor market differ. Under the first characterization, positive compensating wages arise for harmful child labor, while under the second characterization, *negative* compensating wages arise (i.e., poor working conditions are accompanied by low earnings).

To the best of my knowledge, to date, there has been no research on estimating compensating wage differentials for harmful child labor. In fact, there has been very little work on estimating earnings equations for child workers in general. The bulk of the empirical literature on child labor has focused on estimating child labor force participation equations in order to identify the key determinants of child labor and, to a lesser extent, the effects of child labor force participation and earnings on household and child socioeconomic outcomes such as nutrition, fertility, and education (see, e.g., Brown et al. 2003 for an extensive survey of the empirical literature on child labor).

In recent years, there has been some empirical research on harmful child labor. However, the focus of this strand of research has been on estimating the adverse health effects of child labor both in the short-term and long-term (see, e.g., Beegle et al.

2005; O'Donnell et al. 2005; Rosati and Straub 2007; Rogerro et al. 2008). The rest of this research essentially consists of descriptive statistics on child work-related injury or illness rates. For example, Ashagrie (1998) provides information on the incidence of workplace injuries and illnesses among child workers disaggregated by industry and gender from household survey data collected in four countries.

The empirical research on compensating wage differentials for occupational harm and hazards for adult workers is extensive (see, e.g., Viscusi 1993 and Viscusi and Aldy 2003 for surveys of the literature). In general, in this literature, harmful labor is defined by aggregating work-related injury or illness data to certain level, typically either at the detailed industry or occupation level, and assigning all individual workers in a given occupation-industry cell the average work-related injury/illness rate in that occupation-industry cell. Compensating wage differentials are then estimated for an occupation-industry level dataset.

This approach is not feasible in this study for two reasons. First, the occupation and industry classifications for child workers are not detailed enough to permit a reasonable sample size once the data are collapsed to an occupation-industry level dataset. Second, labor force (wage) labor market participation rates among children are significantly lower than for adults; further child workers are more concentrated in certain industries and occupations than adult workers—consequently, even if sufficient detail were present in the occupation and industry classifications, it is likely that a significant share of occupation-industry cells would be lack (sufficient numbers of) observations. Notwithstanding, individual-level injury and illness data are available in my data; using these data at this level would however be a flawed approach. Given that these data represent the overt and immediate manifestation of harmful child labor, they would, as a result, likely grossly understate the true extent of harmful child labor across child workers. That is, there are likely to be large number of cases where child workers would not have suffered a work-related injury or illness in the survey reference period but may nevertheless be employed under harmful employment conditions.

In this chapter, I define harmful child labor as child labor in certain workplaces that is *likely* to result in adverse health effects in the short-term and/or over the longer-term. Specifically, given the available data, harmful child labor is defined as child labor in activities in which children self-report to be physically strenuous, psychologically stressful, or hazardous. This is in contrast to Chapter 1, where I define harmful child labor as child labor in certain workplaces which results in adverse, irreversible health effects later in life (as an adult) in a deterministic way. This stylization of harmful child labor facilitated the modeling of household decisionmaking over child employment choices without having to resort to the use of expectations. In reality however harm is typically a stochastic event. Irrespective of whether harmful child labor is treated as a deterministic or stochastic event, what is common across the definitions is that workplaces are viewed to systematically differ in their capacity to cause harm to the child worker.

This research benefits from the discussions on methodological issues and various empirical approaches undertaken to estimate compensating wage differentials in the literature in general (e.g., Greene 2001; Moretti 2000; Gunderson and Hyatt 2001); these lessons are reflected in the design of the empirical framework. In this chapter, I attempt to estimate the earnings-harmful child labor trade-off using both parametric (ordinary least squares) and semiparametric methods (quantile regression). I first examine the simple bivariate relationship between harmful child labor and earnings. This serves as a starting point for the conditional analysis, where I estimate a log-linear earnings equation via ordinary least squares and quantile regression. The latter method serves as a more robust alternative estimator to the ordinary least square estimator at the center of the conditional distribution of earnings; it also allows me to characterize the earnings-harmful child labor trade-off at different points along the conditional distribution of earnings. Needless to say, the conditional analysis provides the more compelling evidence on the nature of the relationship between earnings and harmful child labor. However, it is important to note that the evidence on an earnings-harmful child labor trade-off is descriptive—the available data do not allow me to address the potential simultaneity and selectivity problems that are generally present in earnings estimations.

The analysis yields six main empirical results. First, I find that children who work in paid employment systematically report higher incidences of harmful child labor than those who work in unpaid employment. Second, I find strong evidence of an unconditional positive relationship between earnings and harmful child labor, irrespective of the type of harmful child labor examined. Third, examining this relationship within a multiple regression framework with a range of sociodemographic, employment, and other controls, I find evidence of positive compensating wages only for hazardous labor and physically strenuous labor when evaluated at the conditional mean of earnings via ordinary least squares; further, the result for hazardous labor is not robust when evaluated at the conditional median of earnings via quantile regession. Fourth, in the cases of psychologically stressful and physically strenuous labor, I do not find any evidence that children receive additional positive compensating wages at the conditional mean or median for higher levels of harmful child labor, as represented by the frequency of harmful child labor. Fifth, I do not find evidence that compensating wages for harmful child labor at the conditional mean or median systematically vary between girls and boys or between urban and rural children. Sixth and last, I find that the estimated earnings premia for physically strenuous and hazardous labor at the conditional mean appear to be largely driven by substantial premia in the bottom half of the conditional distribution of earnings; the premia in the upper half of the conditional distribution of earnings are relatively modest and not significantly different from zero.

The remaining sections of the chapter are organized as follows. Section 2.2 presents the data and the sample. Section 2.3 presents the various empirical strategies employed to determine the regression relationship between harmful child labor and child earnings. Section 2.4 presents the main findings from the unconditional and conditional analyses of the earnings-harmful child labor trade-off. Section 2.5 concludes by summarizing the main results and briefly discussing them.

2.2. Data and sample

2.2.1. Data

The chapter uses observational individual- and household-level data from the Filipino *2001 Survey on Children: 5 to 17 Years Old* (2001 SOC). The 2001 SOC is a nationally-representative sample survey conducted by the National Statistics Office (NSO), the Philippines, in collaboration with the International Labor Organization's International Program on the Elimination of Child Labor (ILO-IPEC). Its purpose was to gather a wide range of information on child work activities.

The 2001 SOC adopted the same three-stage sample design as the 2001 NSO Labor Force Survey, using listings from the 1995 Population Census. The first stage consisted of the systematic selection of *barangys*, the smallest administrative unit in the country, with probability proportional to size. In order to ensure broad geographic coverage, prior to their selection, the barangys were stratified explicitly along urban-rural lines as well as implicitly by, inter alia, municipal district affiliation and groupings based on accessibility and similarities in socioeconomic characteristics and religious composition.[12] The second stage consisted of the systematic selection of enumeration areas (EAs), physical divisions of barangys, with probability proportional to size. These EAs serve as the primary sampling units (PSUs) in the sample. The master sample consisted of 3,416 PSUs, out of which, a subsample of 2,247 PSUs, designated as the core sample, was used for the 2001 SOC. In the third stage, from each PSU in the core sample, a total of twelve households were selected systematically, providing a total of 26,964 private households for the 2001 SOC.[13]

Two separate survey questionnaires were fielded as part of the 2001 SOC: SOC Form 1 which largely collected information on the socioeconomic characteristics of households with children aged 5-17 years and SOC Form 2 which collected detailed information on the employment characteristics of working children. The respondent for the SOC Form 1 was either the parent or guardian of the child or children aged 5-17 years; the main purpose of this survey was to identify eligible child respondents for administering the SOC Form 2. The respondents for the SOC Form 2 were children aged 5-17 years who engaged in any economic activity for at least one hour in the twelve months preceding the interview date (September 2000-October 2001). Information for both survey questionnaires was collected through personal interviews conducted by field interviewers.

Out of the sample of 26,964 households, 17,454 households (64.7%) had members aged 5-17 years; 17,444 of these households (or 99.9%) were successfully interviewed using the SOC Form 1. Among the interviewed households, 6,523 children indicated that they worked during the reference period, of which, 6,365 children (97.6%) were successfully interviewed using the SOC Form 2. Consequently,

[12] Information for the explicit and implicit stratification was obtained from the 1990 Census of Population and Housing as well as other administrative reports produced by the NSO.
[13] Individuals who reside in institutions or establishments were not covered as part of the 2001 SOC.

survey nonresponse by households and children appears to be a negligible source of selection bias.

Several steps were taken to ensure the reliability and interpersonal comparability of the data. For example, survey interviewers were expected to read the questions exactly as worded in the questionnaires (either in English or the local language), and maintain a professional, dispassionate demeanor through the entire interview. In addition, to the extent possible, interviews were to be one-on-one and conducted in private. Castro et al. (2005) however cautions that the reliability of data from parents and children may be undermined if child labor is a sensitive subject (this is plausible as children below ages 15 and 18 are legally prohibited from general and hazardous work in the Philippines, respectively). They also point out that children may be less-reliable respondents than adults. For example, they are more likely to be (1) unaware of salient characteristics of their work (such as workplace hazards) and/or (2) unable to accurately convey them to the interviewer due to recall and spoken language command problems; further, these problems are likely to be exacerbated the younger the child. The degree to which these issues affect the reliability of these data is however unknown. Notwithstanding, in the SOC Form 2, the interviewer was asked to assess the levels of interest and sincerity of the child respondent during the interview (based on predefined scales) as well as note down sections and questions where the child had any doubts, difficulties, or apprehensions. While the notes are unavailable, I include the interviewers' categorical assessments as controls in all regression estimations.

2.2.2. Sample

In order to arrive at the appropriate sample for the study, I pared down the survey sample in stages. A problem with data collection was responsible for the first round of eliminations. The SOC Form 2 asks the first few questions on the two longest employment activities undertaken by the child by referring to them separately. However, the subsequent majority of questions, including the key questions on earnings and harmful child labor, were asked without explicitly specifying the employment activity of concern. Clearly, this issue poses a problem only in the cases in which children reported two employment activities during the reference period. Fortunately, this occurs only in 9% of the sample. These cases were excluded, yielding a sample of 5,791 children.

The second round of eliminations targeted certain categories of child workers either due to their negligible shares in the sample or their incompatibility with the research question. Given that the aim of the chapter is to estimate the earnings-harmful child labor trade-off, all child workers in the sample who did not work for pay were excluded. These cases accounted for roughly 60% of the sample. Almost all of these children were employed in own household-operated enterprises. Children who reported that they were self-employed, whom accounted for about 7% of the sample, were excluded as these children, at least in principle, likely determine working conditions themselves. Children who reported working as paid workers in own household-operated enterprises, which accounted for slightly over 1% of the sample, were also excluded as the wage determination process is likely not (fully) subject to market forces. Finally, two other categories of workers, namely home-based workers and public sector or parastatal workers, were excluded due to their small sample size: they jointly accounted for less than 1% of the sample. Collectively, these exclusions resulted in a sample of 1,677 child workers with labor earnings, which is roughly 29% of the sample of children with only one employment activity in the reference period. These 1,677 child workers were employed in either private households (31%) or private enterprises (69%).

The third and final round of eliminations was conducted in order to restrict a subset of the variables to be included in the regression analysis to certain values or ranges. Specifically, observations with rare values were omitted. Restricting the sample to children (1) ages 8-17 years, (2) who were either paid on a time-rate or piece-rate basis, and (3) who received monetary earnings (instead of in-kind earnings), yielded in an eventual sample size of 1,479 children. This sample of child workers is hereafter referred to as the *wage-earners sample*.

In addition to the wage-earners sample, in order to examine how the incidence and composition of harmful child labor differ between children who work in paid employment and children who work in unpaid employment, a corresponding sample of non-earners is constructed. Restricting the sample to children (1) aged 8-17 years, (2) who engaged in only one employment activity during the reference period, and (3) who are unpaid workers in an own household-operated enterprise, yielding a final sample size of 3,407 children. This sample of child workers is hereafter referred to as the *non-earners sample*.

2.2.3. Construction of harmful child labor variables

The variables that are used in the chapter to represent harmful child labor were constructed from responses to the following three questions in the SOC Form 2 (provided verbatim): (1) "Did you or do you find your work mentally or emotionally stressful?"; (2) "Did you or do you perform heavy physical labor?"; and (3) "Did you or do you consider some aspects of your work risky or dangerous?". The available response options for questions 1 and 2 were "frequently", "sometimes", "seldom", and "never", while the available response options for question 3 were "yes" and "no".

Survey interviewers are offered some guidance on what the above questions on harmful child labor are asking. For example, illustrative examples of heavy physical work provided in the survey interviewer manual include transporting heavy loads, lifting heavy items, using heavy tools and machinery, digging, and quarrying. For stressful work, the manual provides examples of work situations which could cause stress such as working long hours, poor posture/vision due to lack of appropriate furniture/equipment, and harassment from work supervisors. For risky or dangerous work, the manual lists the following risks as examples: vehicular accidents for delivery workers; extreme light from welders which can cause loss of sight; and loud noise of machines which can cause loss of hearing. The manual also provided guidance on the meaning of the response categories. For the questions on heavy physical and stressful work, "frequently" was defined as occurring daily or 3-6 times per week, while "sometimes" was defined as occurring 1-2 times per week or 2-3 times per month (NSO 2001).

Five separate dichotomous variables, denoted by *stressful*, *strenuous*, *hazardous*, *often stressful*, and *often strenuous* were constructed based on the responses to the above questions regarding harmful child labor. The first three variables essentially reflect whether or not the child reported experiencing harmful child labor, while the latter two variables reflect whether or not the child reported experiencing *frequent* harmful child labor. First, with respect to the occurrence-related harmful child labor variables, the variable *stressful* was set equal to one if the respondent indicated that the work activity was frequently or sometimes psychologically stressful, and zero otherwise. The variable *strenuous* was constructed analogously. The variable *hazardous* was set equal to one if the respondent indicated the work activity was hazardous, and zero otherwise. Next, with respect to the magnitude-related harmful child labor variables, the variable *often stressful* was set equal to one if the respondent indicated to have engaged in work that was frequently psychologically stressful, and zero otherwise. Similarly, *often strenuous* was set equal to one if the respondent indicated to have engaged in work that involved frequent physically strenuous work, and zero otherwise.

2.3. Empirical methodology

In order to test for the presence of compensating wage differentials for harmful child labor, I estimate a semi-logarithmic child earnings equation of the form

$$\ln y_i = h_i'\beta + d_i'\gamma + e_i'\lambda + t_i'\phi + \varepsilon_i, \tag{16}$$

where y_i denotes the average weekly gross cash earnings reported by child worker *i*; h_i a vector of harmful child labor responses by the child; d_i a vector of sociodemographic characteristics comprising of the child's age (in quadratic form), gender, urban/rural, region, highest level of formal schooling, and current school attendance status; e_i a vector of employment characteristics comprising of the age when the child first started working (in quadratic form), location of work, contract type, work benefits, and sector of employment; and t_i a vector comprised of controls for usual hours of work per day, usual days of work per week (in quadratic form), and the sincerity and interest levels during the interview as assessed by the interviewer. The vectors β, γ, λ, and ϕ are model parameters to be estimated, and ε_i is a stochastic error term associated with child worker *i*.[14]

I specify the vector h_i in three alternative ways in order to permit a richer description of the conditional relationship between child earnings and harmful child labor. In the first specification, each occurrence-related measure of harmful child labor, namely *hazardous*, *strenuous*, and *stressful*, is included separately in the

[14] A large portion of the literature on estimating compensating wage differentials uses hourly wages as the dependent variable in hedonic wage regressions. However, given the available data, only a crude measure of hourly wage can be constructed as the information on usual hours of work per day was collected in categorical rather than continuous form. Consequently, reported weekly earnings are used as the dependent variable in the estimations but usual days of work per week and hours of work per day are included as covariates. This alternative approach is not uncommon in the literature (see, e.g., Viscusi and Aldy 2003).

earnings regression. I refer to the specification as the *individual inclusion* specification. This particular specification allows me to estimate the partial effect for each type of harmful child labor, *without* controlling for the other types of harmful child labor. While the regression specification includes all the sociodemographic and work time controls, the specification is estimated both with and without the employment controls, thereby allowing me to investigate how much of the partial effect of harmful child labor on earnings is absorbed by these particular controls. Note that I perform this investigation only with the *individual inclusion* specification; the objective is to ascertain whether the included employment attributes partially account for harmful child labor.

In the second specification, I include all three occurrence-related measures of harmful child labor in additive fashion in the regression, along with the full set of sociodemographic, employment, and work time control variables. I refer to this specification as the *additive inclusion* specification. This specification allows me to examine the partial effect of a given type of harmful child labor on earnings, controlling for, among other things, the other types of harmful child labor.

In the third specification, for physically strenuous labor and psychologically stressful labor only, I include the occurrence-related and the magnitude-related measures of harmful child labor in additive fashion in the regression, along with the full set of sociodemographic, employment, and work time controls. I refer to this specification as the *frequency* specification. This specification allows me to examine the partial effect of frequent harmful child labor, controlling for, among other things, the occurrence of harmful child labor.

I first estimate the alternative model specifications of the child earnings equation using ordinary least squares (OLS). However, given the use of survey sample data, the classical assumption of identically and independently distributed errors is likely to be violated due to the use of stratification and clustering in the sample design. To explain, children found within a specific cluster are likely to possess characteristics that are more similar to each other than children found in other clusters. Thus, the amount of intra-cluster variation in the residuals is likely to be significantly different from the amount of inter-cluster variation. Breusch-Pagan tests of the residuals from estimating the various specifications of the earnings regression via OLS strongly suggest the presence of heteroskedasticity. Consequently, I estimate the standard errors by using a formula which corrects for survey design effects (see Deaton 1997 for details).

I also estimate the alternative model specifications of the child earnings equation via quantile regression. The quantile regression estimator has several attractive features (see Koenker and Hallock 2001 for a fuller description). First, quantile regression, specifically the least absolute deviations (LAD) or median regression, provides a robust measure of location as it minimizes a weighted sum of absolute deviations of errors. Thus, the estimation results are less susceptible to *y*-outliers as compared to those from OLS. Second, when the error term is nonnormally distributed, LAD regression may be more efficient than OLS. Diagnostic tests after estimating the alternative model specifications of the earnings equation via OLS strongly suggest the presence of *y*-outliers and nonnormal residuals, making LAD regression a particularly suitable alternative estimation method.[15] Third, as elaborated further below, quantile regression, unlike OLS, allows the regression parameters to vary across the entire distribution of the dependent variable.

[15] The Kolmogorov-Smirnov and Shapiro-Wilks tests were implemented to test whether the OLS residuals were distributed normally or not. Both tests strongly rejected the null hypothesis of normally-distributed residuals.

The basic quantile regression model specifies the conditional quantile in linear form. Adopting the notation in Buckinsky (1998), for the θth quantile, the child earnings regression model is given by

$$\begin{aligned} &\ln y_i = h_i'\beta_\theta + d_i'\gamma_\theta + e_i'\lambda_\theta + t_i'\phi_\theta + \varepsilon_{\theta i}, \\ &Quant(\ln y_i \mid h_i, d_i, e_i, t_i) = h_i'\beta_\theta + d_i'\gamma_\theta + e_i'\lambda_\theta + t_i'\phi_\theta,\ \theta \in (0,1), \end{aligned} \tag{17}$$

where $Quant\left(\ln y_i \mid h_i, d_i, e_i, t_i\right) = h_i'\beta_\theta + d_i'\gamma_\theta + e_i'\lambda_\theta + t_i'\phi_\theta$ denotes the quantile of $\ln y_i$ conditional on the vectors of covariates h_i, d_i, e_i, and t_i. The distribution of the error term, $\varepsilon_{\theta i}$, is left unspecified. The only assumption made is that $\varepsilon_{\theta i}$ satisfies the quantile restriction $Quant\left(\varepsilon_{\theta i} \mid h_i, d_i, e_i, t_i\right) = 0$. The θth sample quantile $\left(0 < \theta < 1\right)$ of $\ln y$ solves

$$\min_{\beta,\gamma,\lambda,\phi} \frac{1}{n} \left\{ \begin{array}{l} \displaystyle\sum_{i:\ln y_i \geq h_i'\beta + d_i'\gamma + e_i'\lambda + t_i'\phi} \theta \left| \ln y_i - h_i'\beta - d_i'\gamma - e_i'\lambda - t_i'\phi \right| + \\ \displaystyle\sum_{i:\ln y_i < h_i'\beta + d_i'\gamma + e_i'\lambda + t_i'\phi} (1-\theta) \left| \ln y_i - h_i'\beta - d_i'\gamma - e_i'\lambda - t_i'\phi \right| \end{array} \right\}. \tag{18}$$

An important feature of the framework is that the partial effects of the covariates, given by β_θ, γ_θ, λ_θ, and ϕ_θ, may differ across quantiles (different values of θ). As is well known, under OLS, the partial effects of the covariates are only estimated at the conditional mean of $\ln y$, and hence are constant across the distribution of $\ln y$. Quantile regression relaxes this restriction, permitting me to characterize the partial effects of covariates along the entire conditional distribution of $\ln y$.

Apart from estimating the partial effect of harmful child labor on earnings at the conditional mean and median via OLS and LAD regressions, respectively, I am also interested in estimating the partial effects of the three forms of harmful child labor on earnings at various distinct quantiles, and to investigate whether there are systematic differences in the estimated effects across the conditional distribution of earnings. Specifically, I estimate the *individual inclusion* specification of the earnings regression at quantiles $\theta \in [0.25, 0.75]$.[16] Quantile regressions for $\theta < 0.25$ or $\theta > 0.75$ are not estimated due to concerns that there may be insufficient observations above and below the selected quantile in the tails of the distribution to permit a robust fit of the data; consequently statistical inference might be impaired (Chernozhukov 2000).

For the purpose of statistical inference, in implementing the quantile regression, I estimate the standard errors by bootstrapping as the formula for obtaining analytical standard errors proposed by Bassett and Koenker (1982) underestimates the standard errors in the presence of heteroskedasticity. Bootstrapping is viewed as a valid alternative for estimating the variance-covariance matrix (see Deaton 1997). However, the standard bootstrap method corrects for heteroskedasticity but fails to address the problem of correlated errors. This is because the standard method resamples observations from the original sample by assuming that each observation has an equal probability of being selected into the bootstrap sample (i.e., it assumes a simple random sample). Given that my sample is instead generated from a complex survey design comprised of both stratification and clustering, the standard bootstrap method needs to be modified to reflect these survey design elements.

I implement my bootstrap method in two steps. In the first step, clusters are randomly selected with replacement from each stratum. The number of clusters selected from each stratum is equal to the total number of clusters found in that particular stratum. In the second step, observations are randomly selected with replacement from each of the selected clusters from the first step. The number of observations selected from each cluster is equal to the total number of observations

[16] The different quantiles are estimated by weighting the residuals differently. For the LAD or median regression ($\theta = 0.5$), all residuals receive equal weight. However, when estimating say the 75th percentile, negative residuals are weighted by 0.25 and positive residuals by 0.75. The criterion is minimized when 75% of the residuals are negative. This is set up as a linear programming model and solved.

found in that particular cluster. In this way, each observation in the original sample does not have an equal probability of being selected into the bootstrap sample. Rather, the probability of an observation being selected into the bootstrap sample increases if it already includes a observation from the same cluster. Thus, the bootstrapped standard errors from the two-step method are robust to violations in both homoskedasticity and independence.

Finally, in order to select the optimal number of bootstrap replications *B*, I apply the three-step method proposed by Andrews and Buchinsky (2000). In principle, the larger the number of replications, the more accurate the bootstrapped statistic is likely to be. However, this benefit has to be traded off against a practical constraint: computational time. Denote the standard error of a parameter of interest estimated using a finite and infinite number of bootstrap replications as $\hat{\sigma}_B$ and $\hat{\sigma}_\infty$, respectively. The object is to choose *B* such that $\hat{\sigma}_B$ is "close" $\hat{\sigma}_\infty$, where "closeness" is defined as

$$\Pr\left(100 \times \frac{\left|\hat{\sigma}_B - \hat{\sigma}_\infty\right|}{\hat{\sigma}_\infty} \le pdb\right) = 1 - \tau,$$

and where *pdb* denotes the maximum desired percentage difference between $\hat{\sigma}_B$ and $\hat{\sigma}_\infty$ and τ the probability that $\hat{\sigma}_B$ and $\hat{\sigma}_\infty$ differ by more than *pdb* percent. For this study, I choose *pdb* and τ to be 5% and 0.05, respectively.[17]

2.4. Findings

2.4.1. Harmful child labor profile

Before discussing the results from examining the conditional relationship between child earnings and harmful child labor, I construct a profile of the incidence and nature of harmful child labor in the wage-earners sample, which I then contrast against that of

[17] These are default values in `bssize`, the Stata ado program that automates the procedure (see Poi 2004).

the non-earners sample. This investigation is useful as it casts light on the relative incidence and composition of harmful child labor experienced by child workers in paid employment. Evidence of this kind is currently missing in the literature.

Table 2.1 reports the incidence of harmful child labor separately by the three types of harmful child labor: hazardous, physically strenuous, and psychologically stressful labor. Column 1 refers to the wage-earners sample and Column 2 to the non-earners sample. The evidence suggests that the incidence of harmful child labor is high among working children in general. Further, the evidence suggests that the incidence of harmful child labor is systematically higher among wage-earners than non-earners. Focusing first on the wage-earners sample, roughly 61% of children reported that they experienced at least one of the three forms of harmful child labor. Psychologically stressful labor is the most commonly reported form of harmful child labor (48%) followed, in turn, by physically strenuous labor (38%) and hazardous labor (27%). In contrast, in the non-earners sample, only 43% of children reported that they experienced at least one form of harmful child labor. What is more, while the ranking in terms of the incidence of harmful child labor by type is identical across the two samples, the incidence of each type of harmful child labor is considerably lower in the non-earners sample. The differences in the incidences of harmful child labor between the two samples are significant.

Table 2.1: Incidence of reported harmful child labor, by sample

Harmful child labor type	(1) Wage-earners sample (in %)	(2) Non-earners sample (in %)	(3) Difference in proportions (1)-(2)
Hazardous	27.0	14.4	12.6***
Stressful	48.1	29.6	18.5***
Strenuous	37.6	20.0	17.6***
N	1,479	3,407	--

Notes: The harmful child labor categories are not mutually exclusive. *denotes statistically significance at the 10% level; ** at the 5% level; and *** at the 1% level.

Table 2.2 provides a more detailed disaggregation of the data by distinguishing between various distinct combinations of the three types of harmful child labor. Again, Column 1 refers to the wage-earners sample and Column 2 to the non-earners sample. Focusing on the wage-earners sample, the most commonly reported combinations of harmful child labor are all three types (15%), psychologically stressful labor only (15%), and psychologically stressful and physically strenuous labor (14%); the remaining combinations all have incidences in the single digits. In contrast, the most commonly reported combination of harmful child labor in the non-earners sample is psychologically stressful labor only (14%); the remaining combinations all have incidences in the single digits. The statistics presented in the table also show clearly that (while a *chi*-squared test indicates that the distributions of the various combinations of harmful child labor across the two samples are significantly different) the lower overall incidence of harmful child labor in the non-earners sample is principally driven by the lower incidence of non-earners reporting experiencing multiple types of harmful child labor; the incidences of children reporting experiencing only one type of harmful child labor are not significantly different between the two samples.

Table 2.3 reports the incidence of a higher magnitude of physically strenuous and psychologically stressful labor as represented by its frequency, *conditional* on having reported physically strenuous or psychologically stressful labor. Column 1 refers to the wage-earners sample and Column 2 to the non-earners sample.

Similar to the evidence on the incidence of harmful child labor presented earlier, the incidence of a higher magnitude of harmful child labor is systematically higher among wage-earners than non-earners. Specifically, in the wage-earners sample, the incidences of frequent strenuous and stressful labor are 28% and 17%, respectively. In contrast, the corresponding incidences in the non-earners sample are

14% and 10%, respectively. Further, the differences in the incidences between the two samples are significantly different.

Table 2.2: Incidence of distinct combinations of reported harmful child labor, by sample

Harmful child labor combination	(1) Wage-earners sample (in %)	(2) Non-earners sample (in %)	(3) Difference in proportions (1)-(2)
None	38.7	57.3	-21.4***
Hazardous only	5.2	5.3	-0.1
Stressful only	14.5	14.2	0.3
Strenuous only	5.9	5.9	0.0
Stressful and *Hazardous* only	4.1	3.1	1.0
Stressful and *Strenuous* only	14.0	8.1	5.9***
Hazardous and *Strenuous* only	3.3	1.9	1.4**
All three types	14.5	4.2	10.3***
Total	100.0	100.0	--
N	1,479	3,407	--

Chi-squared test of the equality of distributions: Test statistic = 267.9; *p*-value = 0.0000.

Notes: * denotes statistically significant at the 10% level; ** at the 5% level; and *** at the 1% level.

As mentioned before, the available response options to the hazardous labor question in the survey precluded the construction of a magnitude-related measure similar to those for physically strenuous and psychologically stressful labor. However, a follow-up question to children who reported their work to be hazardous provided information on the types of work-related risks they encountered.

Table 2.3: Incidence of *frequent* harmful child labor of type *x*, conditional on reporting type *x*, by sample

Frequent type	(1) *Wage-earners sample* (in %)	(2) *Non-earners sample* (in %)	(3) Difference in proportions (1)-(2)
Often strenuous	27.9	13.9	14.0***
Often stressful	16.8	10.1	6.7**
N	1,479	3,407	--

Notes: * denotes statistically significant at the 10% level; ** at the 5% level; and *** at the 1% level.

Table 2.4 presents statistics on the incidences of specific types of work-related risks reported by child in hazardous labor, again separately by the two samples of interest (wage-earners and non-earners). Among wage-earners, the most frequently reported risks are vehicular accidents (23%), disease and sickness (22%), falls and physical mutilation (16% each). Among non-earners, they are disease and sickness (38%), other (17%), and physical mutilation (17%). While I did not conduct any formal statistical tests, it appears that the major differences in the incidences of specific risks between wage-earners and non-earners are the higher incidence of the risk of vehicular accident and the lower incidence of the risk of disease and sickness in the former sample.

Table 2.4: Distribution of specific occupational hazards cited by children in self-reported hazardous workplaces, by sample

Type of occupational hazard	(1) Wage- earners sample (in %)	(2) Non-earners sample (in %)
Vehicular accident	23.1	8.5
Burns	5.8	4.7
Fall	15.5	12.0
Hearing impairment	1.8	0.6
Visual impairment	1.0	1.8
Physical mutilation	15.5	16.5
Disease/sickness	21.8	38.4
Mental/psychological torture	1.0	0.2
Other (unknown)	14.5	17.3
Total	100.0	100.0
N	399	492

2.4.2. Unconditional analysis of the earnings-harmful child labor trade-off

In this subsection, mainly to set the stage for the conditional results to follow, I report the findings on the unconditional bivariate relationship between child earnings and harmful child labor. I begin by comparing the full earnings distributions of children who reported experiencing harmful child labor and those who did not. Figure 2.1 depicts the cumulative distribution functions of earnings separately by whether the child reported engaging in hazardous labor or not. Average weekly earnings is reported on the *x*-axis and the proportion of child workers earning less than average weekly earnings *x* on the *y*-axis. Figures 2.2 and 2.3 depict analogous cumulative distribution functions for physically strenuous and psychologically stressful labor, respectively. A simple visual examination reveals that, for each of the three harmful child labor variables, until at least about the 80th percentile of earnings, the cumulative distribution function of earnings for child workers who reported experiencing harmful child labor generally lies below the cumulative distribution function of earnings for children who reported that they did not. This suggests that, by and large, children who reported harmful child labor earn more than children who did not. More formally, Kolmogorov-Smirnov and Wilcoxon-Mann-Whitney tests strongly reject the null hypothesis of equality of distributions (the *p*-values for all test statistics are 0.000).[18]

Next, focusing on the locations of the earnings distributions, I estimate means and medians of earnings separately for children who reported experiencing harmful child labor and those who did not. The evidence suggests that, in general, the differences in mean and median earnings between child workers who reported harmful child labor and those who did not are highly statistically significant for each type of harmful child labor. The only exception is the difference in mean earnings for psychologically stressful labor which is borderline insignificant (the test statistic is 1.64 and the *p*-value 0.101).

To summarize, the unconditional results strongly suggest that child workers tend to earn more when they engage in work they report to be hazardous, psychologically stressful, or physically strenuous. The next subsection examines

[18] The Kolgomorov-Smirnov and Wilcoxon-Mann-Whitney tests are both nonparametric tests that assess whether the two samples (in my case, earners who reported harmful child labor and earners who did not) are drawn from the same underlying distribution.

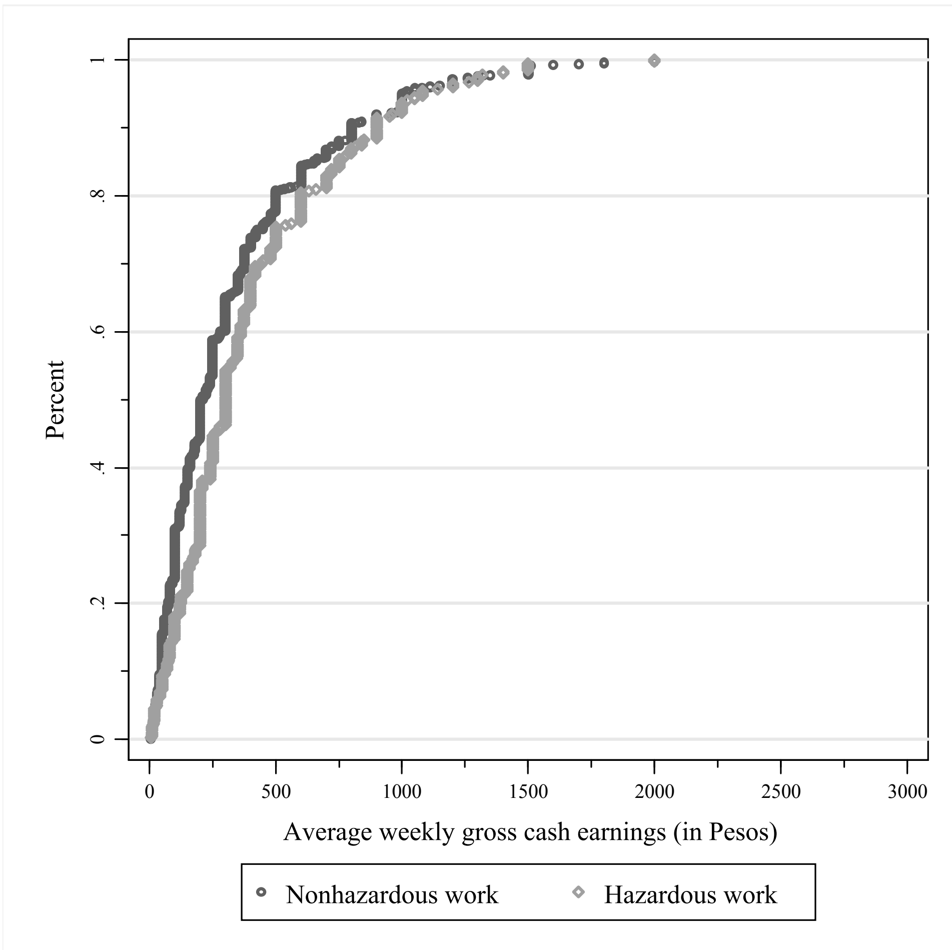

Figure 2.1: Cumulative distribution functions of weekly earnings, by hazardous labor status

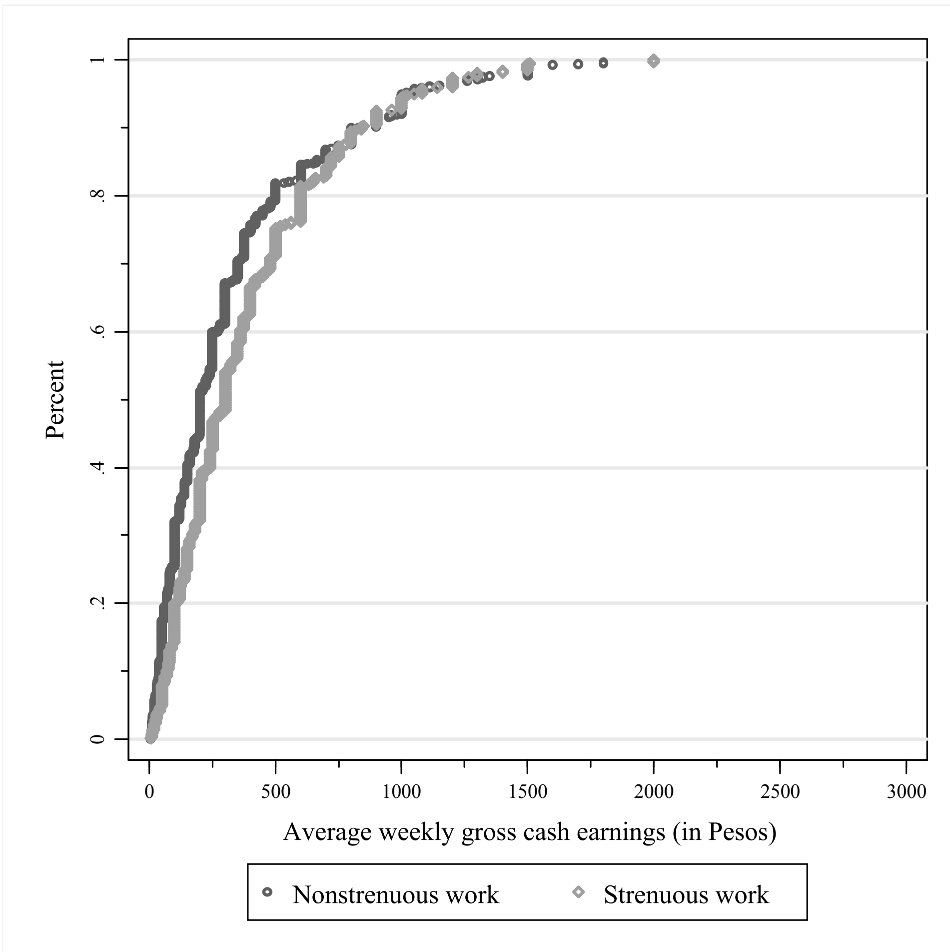

Figure 2.2: Cumulative distribution functions of weekly earnings, by strenuous labor status

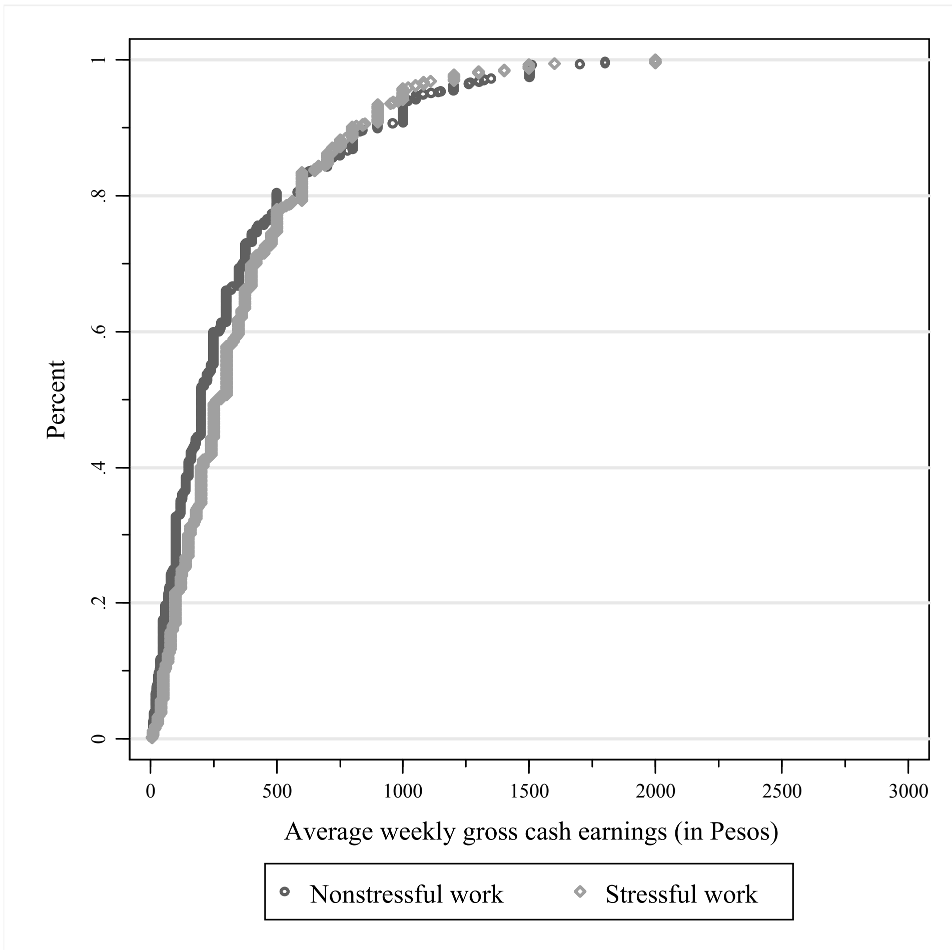

Figure 2.3: Cumulative distribution functions of weekly earnings, by stressful labor status

whether the findings from the unconditional analysis remain intact qualitatively when the relationship between earnings and harmful child labor is investigated within a multiple regression framework. The results that follow are considered the more compelling evidence on the earnings-harmful child labor trade-off.

2.4.3. Conditional analysis of the earnings-harmful child labor trade-off

In this subsection, I present the results of estimating compensating wage differentials for harmful child labor via ordinary least squares as well as quantile regressions for the alternative model specifications of harmful child labor. Table 2.5 presents descriptive statistics for the wage-earners sample for the regressions. Mean average weekly cash earnings is roughly 365 Pesos, with a maximum of 2000 Pesos and a minimum of 6 Pesos. In 2001 US dollars, these values translate to $7.2, $39.2, and $0.12, respectively. In terms of selected socio-demographic characteristics, the mean age of wage-earners is 15 years; 61% are male; 42% reside in urban areas; and 40% are currently attending school. In terms of selected employment characteristics, the mean age when wage-earners first started working is 13 years; 60% work in non-agriculture; 70% work in a private establishment as opposed to a private household; 30% work at night; 22% receive piece-rate pay as opposed to time-rate pay; and 23% have long-term or steady employment as opposed to short-term or casual employment. The incidences of harmful child labor among wage-earners are not reported here as they were discussed earlier.

Table 2.6 reports the ordinary least squares (OLS) estimates of the earnings-harmful child labor trade-off, separately by each type of harmful child labor: hazardous, physically strenuous, and psychologically stressful labor. The results are for the *individual inclusion* specification, where each harmful child labor variable is included separately. For each harmful child labor variable, the earnings regression is estimated first without and then with employment controls. However, all earnings regressions are estimated with the full set of socio-demographic controls as well as controls for hours worked per day, days worked per week, and indicators for the child's sincerity and interest during the interview as assessed by the interviewer. The full set of estimation results from fitting these regressions to the data is presented in Table 2.A1. Note the results in Table 2.A1 are arranged analogously to those in Table 2.6.

Table 2.5: Summary statistics for wage earners sample

Variable	Median	Mean	SD	Max.	Min.
Avg. weekly gross cash earnings (in Pesos)	250.00	350.13	346.89	2000.00	6.00
Natural log avg. weekly gross cash earnings	5.52	5.34	1.12	7.60	1.79
Urban	1.00	0.58	0.49	1.00	0.00
Male	1.00	0.60	0.49	1.00	0.00
Age	15.00	15.00	2.00	17.00	8.00
Age-squared	225.00	228.99	55.73	289.00	64.00
Secondary incomplete	0.00	0.38	0.49	1.00	0.00
Secondary complete	0.00	0.12	0.33	1.00	0.00
Currently attending school	0.00	0.42	0.49	1.00	0.00
Age first started working	14.00	13.12	2.53	17.00	5.00
Age first started working squared	196.00	178.43	63.44	289.00	25.00
Work hours/day: 5-8 hours	1.00	0.56	0.50	1.00	0.00
Work hours/day: 9+ hours	0.00	0.24	0.43	1.00	0.00
Work days/week	5.00	4.42	2.18	7.00	1.00
Work days/week squared	25.00	24.29	18.16	49.00	1.00
Hazardous	0.00	0.27	0.44	1.00	0.00
Strenuous	0.00	0.38	0.48	1.00	0.00
Stressful	0.00	0.47	0.50	1.00	0.00
Often strenuous	0.00	0.10	0.31	1.00	0.00
Often stressful	0.00	0.08	0.27	1.00	0.00
Night work	0.00	0.31	0.46	1.00	0.00
Meals at work	0.00	0.22	0.42	1.00	0.00
Work location: Farm	0.00	0.31	0.46	1.00	0.00
Work location: Non-house and non-farm	0.00	0.30	0.46	1.00	0.00
Worker in private establishment	1.00	0.70	0.46	1.00	0.00
Piece rate pay	0.00	0.22	0.42	1.00	0.00
Long-term/steady employment	0.00	0.24	0.43	1.00	0.00
Non-agriculture	1.00	0.65	0.48	1.00	0.00
Sincere during interview	1.00	0.61	0.49	1.00	0.00
Interested during interview	1.00	0.80	0.40	1.00	0.00

Notes: $N = 1{,}479$. SD denotes standard deviation. Summary statistics for region dummies are not reported.

Given that there are presently very limited examples of earnings estimations in the child labor literature, I take advantage of the analysis to present the earnings estimation results for the other covariates before turning to the main results on the earnings-harmful child labor trade-off (see Table 2.A1).

Table 2.6: OLS regression estimates of earnings premia for harmful child labor
Individual inclusion specification

Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Hazardous	0.104*	0.100*	--	--	--	--
	(0.058)	(0.056)				
Strenuous	--	--	0.140***	0.135***	--	--
			(0.052)	(0.050)		
Stressful	--	--	--	--	0.068	0.052
					(0.049)	(0.049)
R-squared	0.469	0.481	0.471	0.483	0.469	0.481
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	Yes	No	Yes	No	Yes

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. Demographic controls include indicators for urban/rural, region, gender, highest level of formal schooling, and current school attendance status; it also includes age in quadratic form. Employment controls include indicators for night work, meals at work, location of work, payment type, contract type, and sector of activity; it also includes the age the child started working in quadratic form. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Standard errors are reported in parentheses. *** denotes $p<0.01$, ** $p<0.05$, and * $p<0.10$.

To begin, it appears that the earnings regression model fitted to the data explains a significant share of the variation in child earnings; across the various *individual inclusion* specifications, the R-squared statistic suggests that slightly less than 50% of the variation in earnings is explained by the variation in the covariates. This fit is markedly better than what is often found with earnings regressions for adults. Turning to the individual model covariates, with respect to the included socio-demographic factors, as expected a priori, I find robust evidence that (1) earnings are increasing in child age; (2) boys earn more than girls; and (3) those who have completed secondary schooling earn more than those who have primary schooling or less. Second, with respect to the employment factors, again, by and large, as expected a priori, I find consistent evidence that (1) children who receive meals at work earn less than those who do not; (2) those who receive piece-rate pay earn less than those who receive time-rate pay; and (3) those who work in a private establishment earn more than those who work in a private household.[19] The evidence also shows that children who work on average 1-4 hours per day earn less than those who work 5-8 hours per day, and that earnings are increasing in days worked per week, but at a decreasing rate peaking outside the range of the data. There is however no evidence of earnings differences between urban and rural children; children who work at night and those who do not; and children who work in long-term or steady employment and those who work in short-term or casual employment. There is also no evidence that the age at which the child first started working matters for earnings.

Turning next to the OLS estimates of the earnings-harmful child labor trade-off at the conditional mean (see Table 2.6), I find that children receive earnings premia for engaging in hazardous or physically strenuous labor, but do not for psychologically stressful labor. Further, the inference results hold even when employment controls are additionally included in the earnings regressions. Specifically, at the conditional mean, children receive roughly a 10% premium for engaging in hazardous work while they receive roughly a 14% premium for engaging in physically strenuous work. Children who engage in psychologically stressful labor obtain an earnings premium of 7% or 5% at the conditional mean, estimated without and with the employment controls, respectively. These results are however not statistically significant at standard significance levels.

[19] The finding regarding piece-rate pay is surprising as standard theory suggests that piece-rate pay should be associated with higher earnings as it promotes the self-selection of more productive workers into work activities with piece-rate pay systems and/or induces greater work effort.

Table 2.7 presents the quantile regression estimates of the earnings-harmful child labor trade-off at the conditional median. As mentioned before, this particular estimator is referred to as the least absolute deviations (LAD) estimator.

Table 2.7: LAD regression estimates of earnings premia for harmful child labor
Individual inclusion specification

Covariate	(1)	(2)	(3)	(4)	(5)	(6)
Hazardous	0.086	0.027	--	--	--	--
	(0.055)	(0.053)				
Strenuous	--	--	0.141***	0.106**	--	--
			(0.049)	(0.048)		
Stressful	--	--	--	--	0.098**	0.045
					(0.049)	(0.051)
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	Yes	No	Yes	No	Yes

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. Demographic controls include indicators for urban/rural, region, gender, highest level of formal schooling, and current school attendance status; it also includes age in quadratic form. Employment controls include indicators for night work, meals at work, location of work, payment type, contract type, and sector of activity; it also includes the age the child started working in quadratic form. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Bootstrapped standard errors are reported in parentheses. *** denotes $p<0.01$, ** $p<0.05$, and * $p<0.10$.

Analogous to the OLS estimates presented in Table 2.6, the LAD estimates are for the *individual inclusion* specification, where each harmful child labor variable is included separately in the earnings regression along with the socio-demographic and work time controls, first without and then with the employment controls. I find that there are notable differences between the LAD and OLS results. While, like the OLS results, the LAD estimates suggest earnings premia for physically strenuous labor, unlike the OLS results, the LAD earnings premia estimates for hazardous labor are smaller and lose significance. In addition, the LAD results suggest a significant earnings premium for psychologically stressful labor; however the estimated earnings premium becomes smaller and loses significance when the employment controls are additionally included in the regression. Thus, in sum, the LAD results provide consistent evidence of earnings premia for harmful child labor only in the case of one type of harmful child labor: physically strenuous labor.

The above findings answer the question: do children receive positive compensating wages for a given type of harmful child labor treated in isolation? Here we attempt to answer the following questions: (1) do children receive positive compensating wages for a given type of harmful child labor controlling for other types of harmful child labor and (2) do children receive *additional* positive compensating wages for a higher magnitude of harmful child labor? Table 2.8 presents OLS and LAD estimates of the earnings-harmful child labor trade-off for the *additive inclusion* and *severity* specifications.

Table 2.8: OLS and LAD regression estimates of earnings premia for harmful child labor
Additive inclusion and severity specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Covariate	Additive inclusion		Frequency: Frequent strenuous work		Frequency: Frequent stressful work	
	OLS	LAD	OLS	LAD	OLS	LAD
Hazardous	0.064	0.008	--	--	--	--
	(0.060)	(0.055)				
Strenuous	0.122**	0.100*	0.126**	0.121**	--	--
	(0.054)	(0.055)	(0.055)	(0.055)		
Often strenuous	--	--	0.034	-0.038	--	--
			(0.072)	(0.067)		
Stressful	-0.007	0.020	--	--	0.052	0.037
	(0.052)	(0.055)			(0.051)	(0.050)
Often stressful	--	--	--	--	0.001	0.060
					(0.081)	(0.081)
R-squared	0.484	--	0.483	--	0.481	--

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. All regressions include the full set of demographic and employment controls. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Analytical standard errors are reported in parentheses for OLS regressions. Bootstrapped standard errors are reported in parentheses for LAD regressions. *** denotes p<0.01, ** p<0.05, and * p<0.10.

Columns 1 and 2 present the OLS and LAD estimates when the harmful child labor variables are included together in additive fashion in the regression, respectively. Columns 3 and 4 present the OLS and LAD estimates when physically strenuous work and *frequent* physically strenuous work are included together in additive fashion in the regression, respectively. Lastly, Columns 5 and 6 present the OLS and LAD estimates when psychologically stressful labor and *frequent* stressful labor are included together in additive fashion in the regression, respectively. All regressions include the full set of controls.

I find that children receive an earnings premium for engaging in physically strenuous labor, *controlling* for hazardous and stressful labor, among other things; the estimated compensating wages are 12% and 10% at the conditional mean and median, respectively. However, I do not find evidence that the same is true for hazardous or stressful labor. In addition, I find that children receive an earnings premium for physically strenuous labor but no additional premium for frequent strenuous labor. On the other hand, there is no evidence of a significant earnings premium for either the occurrence or a higher magnitude of psychologically stressful labor. These results are robust to the choice of estimator (OLS or LAD).

Additionally, I investigate whether the presence and size of positive compensating wages for harmful child labor differ systematically across selected socio-demographic groups. Tables 2.9-2.12 present estimates of the earnings-harmful child labor trade-off interacted separately with the child's gender and residence location (urban vs. rural). All earnings regressions estimate the *individual inclusion* specification with the full set of controls. Table 2.9 and 2.11 present the OLS estimation results for the trade-offs interacted with boys and urban children at the conditional mean, respectively. Tables 2.10 and 2.12 present the corresponding LAD estimation results at the conditional median. I find that compensating wages for harmful child labor do not systematically differ between urban and rural children or between girls and boys; this finding holds irrespective of the type of harmful child labor. The main, interaction, and joint effects are not significant for psychologically stressful or hazardous labor. In the case of physically strenuous work, while the main and interaction effects are not individually significant, consistent with the earlier findings, the joint effects are significant across the LAD and OLS estimators.

Thus far, I have used quantile regression only as a robust alternative to OLS, estimating the earnings-harmful child labor trade-off at the conditional median. Quantile regression can also be used to estimate the earnings premia of harmful child labor at other quantiles along the conditional distribution of earnings. Table 2.13 reports the estimated earnings premia for the three types of harmful child labor at the 25th, 50th and 75th percentiles of the conditional earnings distribution. The table is accompanied by Figures 2.4, 2.5, and 2.6 which depict the estimated earnings premia between the 25th and 75th percentiles of conditional earnings (in 2.5 percentile increments) along with their respective 95% confidence intervals for hazardous, physically strenuous, and psychologically stressful labor, respectively. The dashed and dotted lines in the figures represent the OLS earnings premia estimates (evaluated at the conditional mean) and their 95% confidence intervals, respectively. All estimates are obtained from fitting the *individual inclusion* regression specification with all controls to the data.

First, with respect to psychologically stressful labor, as shown in Table 2.13 and Figure 2.6, I find that the earnings premia for stressful labor generally decline as one moves up the conditional distribution of earnings—for example, the estimated earnings premium at the 25th, 50th, and 75th percentiles are 10%, 5%, and 2%, respectively. However, the size differences in the earnings premia at different points in the conditional earnings distribution are not significant; further, by and large, none of the estimated earnings premia are significantly different from zero. Thus, the OLS evidence at the conditional mean and the quantile regression evidence over a large range of conditional quantiles are consistent.

Table 2.9: OLS regression estimates of earnings premia for harmful child labor, by gender
Individual inclusion specification

Variables	(1) *Hazardous*	(2) *Strenuous*	(3) *Stressful*
Male	0.098	0.104	0.180**
	(0.062)	(0.066)	(0.075)
Hazardous	-0.035	--	--
	(0.113)		
Male × *Hazardous*	0.176	--	--
	(0.126)		
Strenuous	--	0.112	--
		(0.096)	
Male × *Strenuous*	--	0.034	--
		(0.111)	
Stressful	--	--	0.101
			(0.079)
Male × *Stressful*	--	--	-0.080
			(0.095)
R-squared	0.482	0.483	0.481

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. All regressions include the full set of demographic and employment controls. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Standard errors are reported in parentheses. *** denotes p<0.01, ** p<0.05, and * p<0.10.

Table 2.10: LAD regression estimates of earnings premia for harmful child labor, by gender
Individual inclusion specification

Variables	(1) *Hazardous*	(2) *Strenuous*	(3) *Stressful*
Male	0.218***	0.017**	0.250***
	(0.067)	(0.068)	(0.080)
Hazardous	-0.007	--	--
	(0.114)		
Male × *Hazardous*	0.038	--	--
	(0.130)		
Strenuous	--	0.069	--
		(0.085)	
Male × *Strenuous*	--	0.048	--
		(0.104)	
Stressful	--	--	0.071
			(0.088)
Male × *Stressful*	--	--	-0.049
			(0.107)

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. All regressions include the full set of demographic and employment controls. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Bootstrapped standard errors are reported in parentheses. *** denotes p<0.01, ** p<0.05, and * p<0.10.

Table 2.11: OLS regression estimates of earnings premia for harmful child labor, by urban/rural

Individual inclusion specification

Variables	(1) *Hazardous*	(2) *Strenuous*	(3) *Stressful*
Urban	-0.006	-0.018	0.008
	(0.065)	(0.071)	(0.083)
Hazardous	0.104	--	--
	(0.081)		
Urban × *Hazardous*	-0.008	--	--
	(0.100)		
Strenuous	--	0.112	--
		(0.075)	
Urban × *Strenuous*	--	0.041	--
		(0.096)	
Stressful	--	--	0.068
			(0.078)
Urban × *Stressful*	--	--	-0.027
			(0.099)
R-squared	0.481	0.483	0.481

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. All regressions include the full set of demographic and employment controls. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Standard errors are reported in parentheses. *** denotes p<0.01, ** p<0.05, and * p<0.10.

Table 2.12: LAD regression estimates of earnings premia for harmful child labor, by urban/rural

Individual inclusion specification

Variables	(1) *Hazardous*	(2) *Strenuous*	(3) *Stressful*
Urban	0.026	0.053	0.035
	(0.068)	(0.076)	(0.078)
Hazardous	0.025	--	--
	(0.075)		
Urban × *Hazardous*	0.003	--	--
	(0.106)		
Strenuous	--	0.143**	--
		0.073)	
Urban × *Strenuous*	--	-0.073	--
		(0.097)	
Stressful	--	--	0.049
			(0.078)
Urban × *Stressful*	--	--	-0.004
			(0.094)

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. All regressions include the full set of demographic and employment controls. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Bootstrapped standard errors are reported in parentheses. *** denotes p<0.01, ** p<0.05, and * p<0.10.

Table 2.13: OLS and quantile regression estimates of earnings premia for harmful child labor at selected conditional quantiles
Individual inclusion specification

Variable	Statistic	Conditional mean	Conditional quantile		
			0.25	0.50	0.75
	b	0.100*	0.177***	0.027	0.038
Hazardous	$100 \times e^b - 1$	10.5%	19.4%	2.7%	4.0%
	s.e.	(0.056)	(0.065)	(0.053)	(0.058)
	b	0.135***	0.253***	0.106***	0.067
Strenuous	$100 \times e^b - 1$	14.5%	28.8%	11.2%	6.9%
	s.e.	(0.050)	(0.057)	(0.048)	(0.053)
	b	0.052	0.098	0.045	0.020
Stressful	$100 \times e^b - 1$	5.3%	10.3%	4.6%	2.0%
	s.e.	(0.049)	(0.060)	(0.051)	(0.052)

Notes: N = 1,479 for all regressions. Dependent variable is the natural log of average weekly gross cash earnings. All regressions include the full set of demographic and employment controls. All regressions include controls for hours worked per day and days worked per week as well as interviewer-assessed respondent sincerity and interest during the interview. Analytical standard errors are reported in parentheses for OLS regressions. Bootstrapped standard errors are reported in parentheses for quantile regressions. *** denotes p<0.01, ** p<0.05, and * p<0.10.

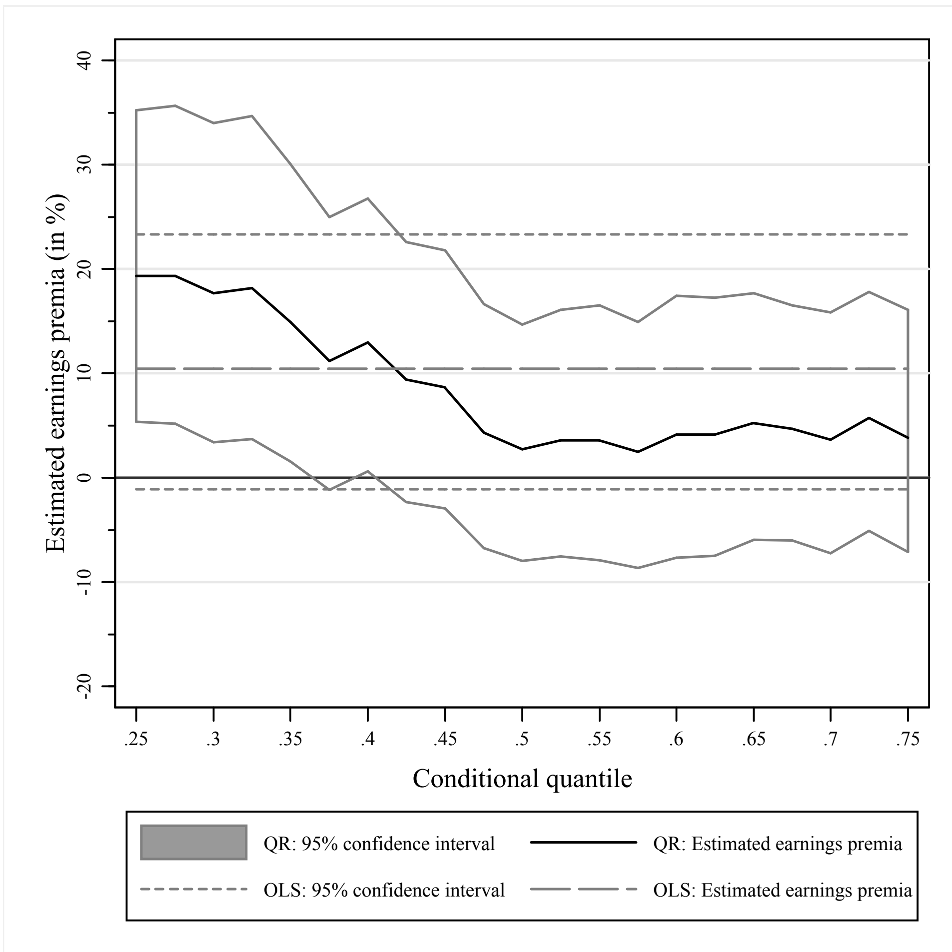

Figure 2.4: Quantile regression estimates of earnings premia for hazardous labor, by conditional quantile

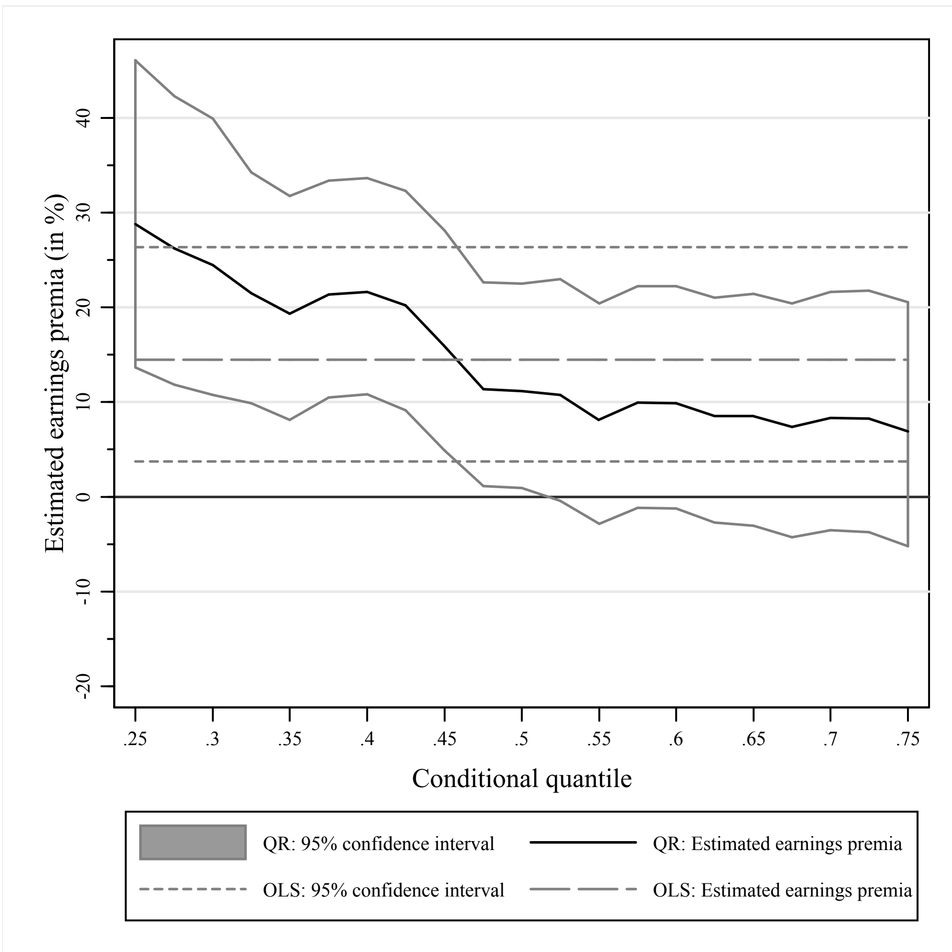

Figure 2.5: Quantile regression estimates of earnings premia for strenuous labor, by conditional quantile

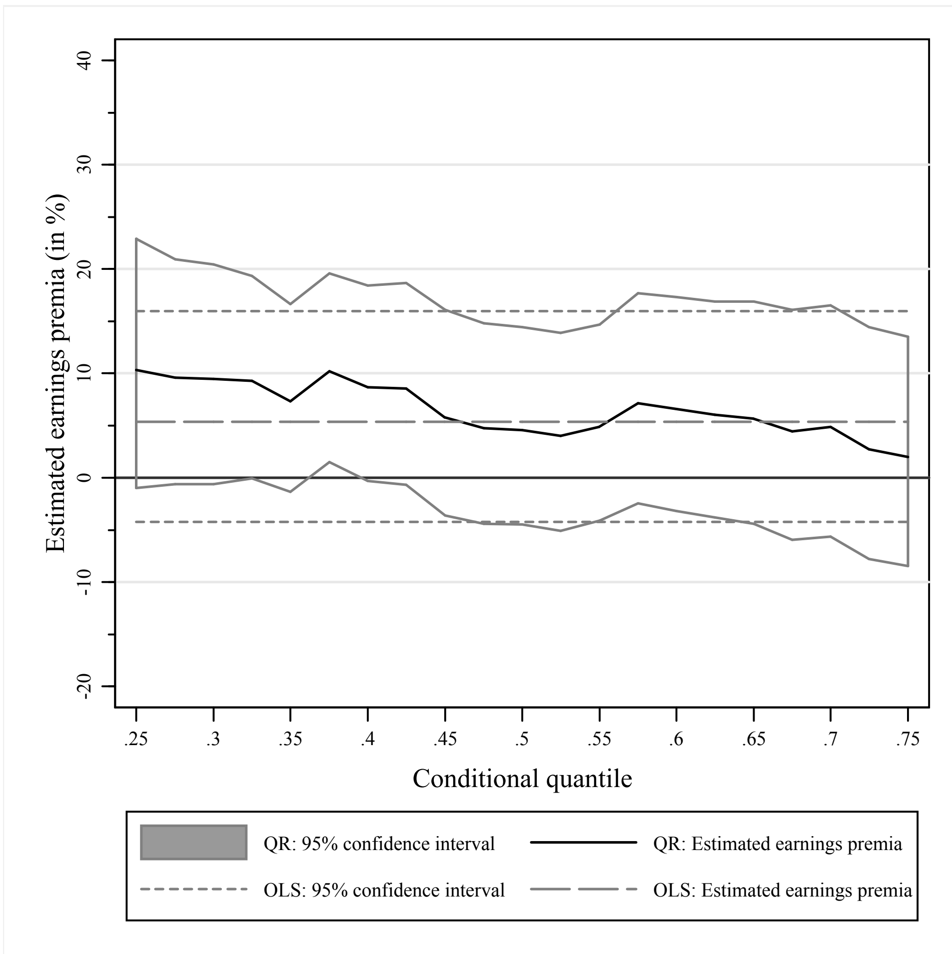

Figure 2.6: Quantile regression estimates of earnings premia for stressful labor, by conditional quantile

In contrast, I find substantial earnings premia for hazardous and physically strenuous labor in the lower half of the conditional distribution of earnings. As Table 2.13 and Figures 2.4 and 2.5 show, for both types of harmful child labor, the size of the estimated earnings premia falls rapidly before stabilizing somewhat as one moves up the conditional distribution of earnings. For example, the estimated earnings premia for hazardous labor at the 25th, 50th, and 75th percentiles are 19%, 3%, and 4%, respectively. Likewise, the estimated earnings premia for physically strenuous labor at the 25%, 50th, and 75th percentiles are 29%, 11%, and 7%, respectively. Further, formal tests suggest that the estimated earnings premia at the 25th percentile for both types of harmful child labor are significantly different from those at the 50th and 75th percentiles; the earnings premia between the latter two percentiles are not. More generally, moving up from the bottom, by roughly the 40th percentile of conditional earnings for hazardous labor (roughly the 50th percentile of conditional earnings for strenuous labor), the estimated earnings premia for harmful chid labor cease to be significant. To summarize the evidence, the OLS and LAD estimates of significant earnings premia at the conditional mean and median for hazardous and physically strenuous labor appear to be driven by the substantial earnings premia in the lower half of the conditional distribution of earnings.

2.5. Conclusion

To summarize, in this chapter, in order to gain some insight into the reasons behind harmful child labor and the socioeconomic conditions underlying its existence, I examine whether child workers in harmful employment settings are compensated monetarily in the form of higher labor market earnings. I define harmful child labor as child labor in certain workplaces that is *likely* to result in adverse health effects in the short-term and/or over the longer-term. Specifically, given the available data, harmful child labor is defined as child labor in activities in which children self-report to be physically strenuous, psychologically stressful, or hazardous.

In terms of the empirical strategy, I first examine the simple bivariate relationship between harmful child labor and earnings. This serves as a starting point for the conditional analysis, where I estimate a log-linear earnings equation via ordinary least squares and quantile regression. The latter method serves as a more robust alternative estimator to the ordinary least square estimator at the center of the conditional distribution of earnings; it also allows me to characterize the earnings-harmful child labor trade-off at different points along the conditional distribution of earnings. Needless to say, the conditional analysis provides the more compelling evidence on the nature of the relationship between earnings and harmful child labor. However, it is important to note that the evidence on the earnings-harmful child labor trade-off is descriptive—the available data do not allow me to address the potential simultaneity and selectivity problems that are generally present in earnings estimations.

The analysis yields six main empirical results. First, I find that children who work in paid employment systematically report higher incidences of harmful child labor than those who work in unpaid employment. Second, I find strong evidence of an unconditional positive relationship between earnings and harmful child labor, irrespective of the type of harmful child labor examined. Third, examining this relationship within a multiple regression framework with a range of sociodemographic, employment, and other controls, I find evidence of positive compensating wages only for hazardous labor and physically strenuous labor when evaluated at the conditional mean of earnings via ordinary least squares; further, the result for hazardous labor is not robust when evaluated at the conditional median of earnings via quantile regession. Fourth, in the cases of psychologically stressful and physically strenuous labor, I do not find any evidence that children receive additional positive compensating wages at the conditional mean or median for higher levels of harmful child labor, as represented by the frequency of harmful child labor. Fifth, I do not find evidence that compensating wages for harmful child labor at the conditional mean or median systematically vary between girls and boys or between urban and rural children. Sixth and last, I find that the estimated earnings premia for physically strenuous and hazardous labor at the conditional mean appear to be largely driven by substantial premia in the bottom half of the conditional distribution of earnings; the premia in the upper half of the conditional distribution of earnings are relatively modest and not significantly different from zero.

A straightforward (and clearly simplistic) interpretation of the above findings is that there appears to be compensating wages for immediate, transparent, and certain harmful child labor as represented by physically strenuous labor or probabilistic physical harm as represented by hazardous labor, especially in relatively lower-wage employment activities (conditional on other factors). The findings however suggest the absence of compensating wages for harm that takes the relatively intangible form of psychological stress. The differing results across the types of harmful child labor could indicate perceptional differences over the types of harm. For example, households may tend to discount harm when it is psychological. This behavior may be exacerbated when the physical manifestations of this harm may occur later in life or are difficult to attribute to work and working conditions.

APPENDIX

Table 2.A1: OLS earnings regression estimates, all model parameters
Individual inclusion specification

Covariate	(1)	(2)	(3)	(4)	(5)	(6)
	Hazardous		*Strenuous*		*Stressful*	
Urban	-0.009	-0.008	-0.001	-0.003	-0.003	-0.005
	(0.058)	(0.059)	(0.058)	(0.059)	(0.058)	(0.059)
Male	0.177***	0.132**	0.159***	0.113*	0.193***	0.146**
	(0.056)	(0.060)	(0.057)	(0.060)	(0.055)	(0.059)
Age	0.326**	0.376**	0.328**	0.375**	0.314**	0.362**
	(0.161)	(0.148)	(0.158)	(0.148)	(0.159)	(0.147)
Age-squared	-0.008	-0.010**	-0.008	-0.010**	-0.007	-0.010*
	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Secondary incomplete	0.040	0.034	0.038	0.031	0.040	0.034
	(0.058)	(0.059)	(0.058)	(0.059)	(0.058)	(0.059)
Secondary complete	0.219***	0.154*	0.215***	0.148*	0.212***	0.145*
	(0.079)	(0.085)	(0.079)	(0.084)	(0.079)	(0.084)
Attending school	-0.263***	-0.236***	-0.261***	-0.234***	-0.266***	-0.239***
	(0.075)	(0.073)	(0.076)	(0.074)	(0.075)	(0.073)
Hazardous	0.104*	0.100*	--	--	--	--
	(0.058)	(0.056)				
Strenuous	--	--	0.140***	0.135***	--	--
			(0.052)	(0.050)		
Stressful	--	--	--	--	0.068	0.052
					(0.049)	(0.049)
Work hours/day: 1-4	-0.510***	-0.440***	-0.498***	-0.435***	-0.508***	-0.445***
	(0.075)	(0.077)	(0.075)	(0.077)	(0.075)	(0.077)
Work hours/day: 9+	0.063	0.092	0.064	0.091	0.064	0.094
	(0.066)	(0.067)	(0.066)	(0.067)	(0.066)	(0.067)
Work days/week	0.432***	0.408***	0.423***	0.399***	0.433***	0.410***
	(0.080)	(0.080)	(0.080)	(0.080)	(0.080)	(0.080)
Work days/week-squared	-0.033***	-0.029***	-0.032***	-0.028***	-0.033***	-0.030***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Age started work	--	-0.076	--	-0.073	--	-0.072
		(0.099)		(0.099)		(0.100)
Age started work-squared	--	0.004	--	0.004	--	0.004
		(0.004)		(0.004)		(0.004)
Night work	--	0.035	--	0.043	--	0.036
		(0.059)		(0.059)		(0.059)
Meals at work	--	-0.142**	--	-0.142**	--	-0.140**
		(0.059)		(0.059)		(0.058)
Farm	--	-0.019	--	-0.020	--	-0.015
		(0.109)		(0.109)		(0.109)
Non-farm, non-house	--	0.042	--	0.051	--	0.057
		(0.074)		(0.074)		(0.074)

Table 2.A1 (Continued)

Private establishment	--	0.170**	--	0.174**	--	0.174**
		(0.074)		(0.074)		(0.074)
Piece rate pay	--	-0.157**	--	-0.146**	--	-0.145**
		(0.068)		(0.067)		(0.068)
Long-term/steady work	--	0.064	--	0.060	--	0.062
		(0.059)		(0.058)		(0.058)
Non-agriculture	--	-0.118	--	-0.113	--	-0.111
		(0.094)		(0.095)		(0.094)
Sincere in responses	-0.104**	-0.108**	-0.100*	-0.104**	-0.102*	-0.107**
	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.053)
Interested in interview	0.051	0.052	0.040	0.041	0.045	0.047
	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)
Constant	0.893	1.008	0.888	1.000	0.960	1.053
	(1.194)	(1.137)	(1.172)	(1.118)	(1.179)	(1.128)
R-squared	0.469	0.481	0.471	0.483	0.469	0.481

Notes: N = 1,479 for all regressions. Standard errors reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Estimated coefficients for region dummies are not reported.

REFERENCES

Andrews, Donald W. K., and Moshe Buchinsky. 2000. A Three-Step Method for Choosing the Number of Bootstrap Repetitions. *Econometrica*. v68 (n1):23-51.

Ashagrie, Kebebew. 1998. *Statistics on Child Labor and Hazardous Child Labor in Brief*. Geneva: International Labor Organization. http://www.ilo.org/public/english/standards/ipec/simpoc/stats/child/stats.htm

Bassett, Gilbert, Jr., and Roger Koenker. 1982. An Empirical Qauntile Function for Linear Models with iid Errors. *Journal of the American Statistical Association* 77(378): 407-415.

Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti. 2005. *Why should we care about child labor? The education, labor market, and health consequences of child labor*. World Bank Policy Research Working Paper Series No. 3479. Washington, D.C.: World Bank.

Brown, Drusilla K., Alan V. Deardorff, and Robert Stern. 2003. Child Labor: Theory, Evidence, Policy. In Kaushik Basu, Henrik Horn, Lisa Roman, and Judith Shapiro (eds.), *International Labour Standards- Issues, Theories and Policy*, Basil Blackwell, 2003.

Buchinsky, Moshe. 1998. Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *Journal of Human Resources*. v33 (n1):88-126.

Castro, Charita L., Sarah Gormly, and Amy R. Ritualo. 2005. The SIMPOC Philippine Survey of Children 2001: A Data Source for Analyzing Occupational Injuries to Children. *Public Health Reports*. v120 (n6):631-641.

Chernozhukov, Victor. 2000. *Conditional Extremes and Near-Extremes.* MIT Department of Economics Working Paper No. 01-21.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore: Johns Hopkins University Press.

Dessy, Sylvain, and Stephane Pallage. 2003. *The economics of child trafficking.* CIRPEE Working Paper No. 03-23.

Dessy, Sylvain E., and Stephane Pallage. 2005. A Theory of the Worst Forms of Child Labour. *Economic Journal*. v115 (n500):68-87.

Greene, William H. 1993. *Econometric Analysis*. 2nd edition. New York: Macmillan Publishing Company.

Guarcello, Lorenzo, Scott Lyon, and Furio C. Rosati. 2004. *Impact of Working Time on Children's Health*. Understanding Children's Work (UCW) Working Paper.

Gunderson, Morley, and Douglas Hyatt. 2001. Workplace Risks and Wages: Canadian Evidence from Alternative Models. *Canadian Journal of Economics*. v34 (n2):377-95.

Hagemann, Frank, Yacouba Diallo, Alex Etienne, and Farhad Mehran. 2006. *Global Child Labour Trends 2000 to 2004.* Geneva: International Labour Organization.

ILO 2002. *Every Child Counts: New Global Estimates of Child Labour*. Geneva: International Labour Office.

Koenker, Roger, and Kevin F. Hallock. 2001. Quantile Regression. *Journal of Economic Perspectives*. v15 (n4):143-56.

Koenker, Roger W., and Gilbert, Jr Bassett. 1978. Regression Quantiles. *Econometrica*. v46 (n1):33-50.

Moretti, Enrico. 2000. Do Wages Compensate for Risk of Unemployment? Parametric and Semiparametric Evidence from Seasonal Jobs. *Journal of Risk and Uncertainty* v20(n1):45-66.

National Statistics Office. 2001. *2001 Survey of Children: Enumerator's Manual*. Philippines: National Statistics Office.

O'Donnell, Owen, Furio C. Rosati, and Eddy van Doorslaer. 2005. Health Effects of Child Work: Evidence from Rural Vietnam. *Journal of Population Economics*. v18 (n3):437-67.

Poi, Brian P. 2004. From the help desk: Some bootstrapping techniques. *The Stata Journal* v4(n3):312-328.

Rogerro, Paola, Viviana Mangiaterra, Flavia Bustreo, and Furio Rosati. The Health Impact of Child Labor in Developing Countries: Evidence from Cross-Country Data. *American Journal of Public Health*. v97 (n2):271-75.

Rogers, Carol Ann, and Kenneth A. Swinnerton. 2008. A Theory of Exploitative Child Labor. *Oxford Economic Papers*. v60 (n1):20-41.

Rosati, Furio, and Roland Straub. 2007. Does Work during Childhood Affect the Health of Guatemalan Adults? *Review of Economics of the Household*. v5 (n1):83-94.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World. *Journal of Risk and Uncertainty* v27 (n1):5-76.

Viscusi, W. Kip. 1993. The Value of Risks to Life and Health. *Journal of Economic Literature*. v31 (n4):1912-46.

CHAPTER 3

DO CONFLICTING PARENT-CHILD RESPONSES MATTER? INSIGHTS FROM INVESTIGATING THE HARMFUL CHILD LABOR DECISION

3.1. Introduction

Child labor is a mass and persistent phenomenon in much of the developing world, particularly in its poorer parts. According to the International Labor Organization (ILO), 191 million children ages 5-14 years were classified as economically active in 2004, of which, 166 million children were classified as child laborers (Hagemann et al. 2006).[20] What is more, 74 million children (or 39% of economically-active children) were classified as engaged in hazardous work, that is, in forms of employment likely to endanger their physiological, psychological, and/or moral health and development. A number of recent empirical studies find that child labor has adverse effects on important child outcomes such as school attainment and cognitive achievement (e.g., Akabayashi and Psacharopoulos 1999; Rosati and Rossi 2003; Gunnarsson et al. 2006), health (e.g., O'Donnell et al. 2005; Rosati and Straub 2007; Rogerro et al. 2008), future labor market earnings (e.g., Ilahi et al. 2005), and the likelihood of child labor among the child's future sons and daughters (e.g., Emerson and Souza 2003).[21]

[20]Child labor is a more stringent definition than child economic activity or work—it takes into account the age of the child, the type of work, and the numbers of hours the child works. Specifically, the ILO classifies a child as a child laborer if (1) the child is aged 5-11 years and is economic active regardless of the number of hours or (2) the child is aged 12-14 years and is economically active for 14 hours or more per week, or works in a hazardous activity or occupation or in the "worst forms of child labor" regardless of the number of hours.

[21] There is also some evidence that not all child labor is necessarily potentially harmful and some forms may actually provide valuable apprenticeship skills or on-the-job training which raises the future labor productivity and earnings of the child. For example, Beegle et al. (2005) find that child labor increases the probability of wage employment and higher daily labor and farm earnings in the future, fully offsetting the forgone earnings attributable to schooling lost due to child labor.

Understanding the underlying reasons for why children engage in arduous and/or hazardous work is instrumental for devising policy and program solutions to effectively and sustainably extricate children from such undesirable situations. Much of the available research on what explains the existence of child labor, whether theoretical or empirical, is typically based on a unitary household model (Becker 1981), where all household members are assumed to have identical preferences. In addition, parents are assumed to behave altruistically towards their children. Under such a model, the prevalence of child labor is typically explained by household characteristics such as chronic poverty (e.g., Basu and Van 1998; Basu 1999; Emerson and Souza 2003), risk-coping in the face of adverse income shocks (e.g., Grootaert and Kanbur 1995; Jacoby and Skoufias 1997; Beegle et al. 2005) or supply-side issues such as the low quality of or lack of access to schools (e.g., Grootaert and Patrinos 1999), imperfect markets for credit, land, and labor (e.g., Baland and Robinson 2000; Ranjan 2001; Basu et al. 2007), or changes in commodity prices and labor market returns (e.g., Edmonds and Pavcnik 2005; Kruger 2007).

To be sure, the same factors considered to influence households' decision to supply child labor can be straightforwardly shown to also influence households' decision on whether to send their children to work in harmful settings. The theoretical literature which focuses specifically on explaining harmful child labor is however presently limited. For example, it comprises of explanations such as household sorting into different child employment choices on the basis of depth of poverty in the presence of compensating wage differentials for potential harm in the wage labor market (see Chapter 1) and households' seeking immediate returns in an environment of limited labor market opportunities (Dessy and Pallage 2005). However, with a few exceptions (e.g., Chapter 2), little empirical evidence is available on what specific factors influence the harmful child labor decision.

Given growing evidence that there may be differing preferences among members within the household, recent studies have examined how children's outcomes, including children's labor market outcomes, are affected by conceptualizing the problem using an intrahousehold model (Manser and Brown 1980; Horney 1981; Chiappori and others 1982; 1992; 1998). However, in using the intrahousehold model, by and large, studies have examined how (changes in) the intrahousehold position of mothers vis-à-vis fathers have affected child investments and outcomes. Virtually none of these studies model the interaction between parents and children or treat children as potential decisionmakers within the households who might have distinct preferences from their parents.

In this chapter, the child is viewed as having differing preferences from her parent over choices such as child education and labor. Further, in the presence of asymmetries in decisionmaking power/leverage between the parent and the child, specifically, in the presence of greater decisionmaking power in the hands of the parent, child outcomes such as whether to work or not as well as what type of setting the child works in in terms of harm, will reflect more the parent's valuation of the costs and benefits of (harmful) child work than the child's. Given this, one can conceive of a scenario where the resulting net welfare position at the household level is suboptimal—in other words, changing the power dynamic between the parent and the child may yield a higher welfare level for the household, though obviously the welfare gains for the household will have to be traded off against the welfare losses of some individuals within the household (those that initially held greater power).

In the intrahousehold literature in general, differing preferences between household members are usually inferred from differences in observed household behavior as a function of some variable that captures divergence in decisionmaking power between members. For example, the conditional relationship between child outcomes and (changes in the) mother-father gap in, say, income/assets is used to infer if households with lower mother-father income/asset gaps have systematically different child outcomes from those with higher mother-father income/asset gaps. In the household survey dataset I use, I have information on whether a working child suffered a work-related injury or illness as reported by the child *as well as* the parent. I use this information to construct a similar gap measure to the above example. Namely a gap exists if the parent reports the working child did not suffer a work-related injury/illness whereas the child reports that she did (I refer to this specific gap as a parent-child injury report mismatch) and a gap does not exist if the child work-related injury/illness reports by the parent and the working child are congruent. Note that while the reverse gap—that is, the parent reports that the child suffered a work-related injury/illness but the child reports that she did not—is, in principle, possible, given my data, it appears to be a highly rare occurrence: roughly 1-2% of all working children had responses of this nature.

Using this information, this chapter examines whether parent-child injury report mismatches have an impact on the probability of harmful child labor, where harmful child labor is measured using both subjective and objective indicators of harmful child labor constructed from information provided by the working child. Given the interest in the intrahousehold literature in investigating potential gender differences, the chapter also examines whether the estimated impacts systematically differ between working boys and girls as well as by whether the respondent parent is the father or the mother. In attempting to answer these questions, the chapter contributes to the presently limited literature on the determinants of the harmful child labor decision, albeit from a different angle than is typically found in the child labor literature. While the aim of this chapter is not to challenge the assumption that parents behave altruistically towards their children, the constraints that poverty imposes on households may well include parents valuing child welfare lower (i.e., reduced altruism). In other words, parental altruism may itself be a choice variable rather than a fixed parameter in parents' utility functions, with altruism increasing in household income or wealth.

In this chapter, I define harmful child labor as child labor in certain workplaces that is *likely* to result in adverse health effects in the short-term and/or over the longer-term. Specifically, given the data, harmful child labor is defined as child labor in activities in which children self-report to be physically strenuous or hazardous. It also includes child labor in workplaces where the child reports exposure to at least one identified physical or chemical hazard. This is in contrast to Chapter 1, where I define harmful child labor as child labor in certain workplaces which results in adverse, irreversible health effects later in life (as an adult) in a deterministic way. This stylization of harmful child labor facilitated the modeling of household decisionmaking over child employment choices without having to resort to the use of expectations. In reality however harm is typically a stochastic event. Irrespective of whether harmful child labor is treated as a deterministic or stochastic event, what's common across the definitions is that workplaces are viewed to systematically differ in their capacity to cause harm to child.

The impacts of parent-child injury report mismatches on the probability of harmful child labor are estimated using three alternative approaches for establishing a valid empirical counterfactual (under certain assumptions) for working children with parent-child injury report mismatches. The first is standard multiple regression with controls for a range of child, respondent parent, and household characteristics which also potentially explain the harmful child labor decision. The second is multiple regression augmented by household fixed effects to additionally control for unobservables that may vary between households. The third is propensity score matching (matching for short), which explicitly models the parent-child injury report mismatch outcome and generates a presumably stronger counterfactual than is possible via regression by only considering similar observations. It is important to note that all three methods essentially assume that the selection process which determines the parent-child injury report mismatch outcome can be modeled using the available data (i.e., selection is on observables). This assumption is plausible given data on an extensive set of covariates collected similarly across all sample observations (Heckman et al. 1997; Heckman et al. 1998)

While matching is typically used in program evaluations where exposure to or participation in a program (referred to in the evaluation literature as treatment) is binary, the method can be used more generally to estimate the impact of any binary variable such as the parent-child injury report mismatch variable in this chapter (Wooldridge 2002). For example, previous studies in the labor economics literature that have interpreted treatment broadly and applied matching methods include studies of the impact of union membership on wages (e.g., Bryson 2002) and the impact of migration on wages (e.g., Ham et al. 2005), both of which represent research questions where the estimation issues of endogeneity and selection are well-known in the literature (Cahuc and Zylberberg 2004).[22]

All three estimators suggest that parent-child injury report mismatches have a large and significant positive impact on the probability of harmful child labor. This finding is largely robust to the choice of harmful child labor measure which comprises of both self-reported measures as well as measures constructed using child information on specific workplace hazards. While the assumption of selection on observables on which this empirical analysis hinges on is untestable, the results from applying a

[22] Caliendo (2008) also lists other nonevaluation research which uses matching methods to estimate the effects of binary variables.

standard bounding test suggests that the significance of the matching-based results appears to be robust to the potential influence of varying degrees of arbitrary unobserved heterogeneity in explaining the presence of parent-child injury report mismatches among households. Finally, the evidence suggests that the impact of parent-child injury report mismatches on the probability of harmful child labor does not vary systematically between working boys and girls, and between children where the adult respondent to the question on child work-related injury is the mother or the father.

The remaining sections of the chapter are organized as follows. Section 3.2 motivates how parent-child injury report mismatches may matter for the probability of harmful child labor by using insights from the theoretical literature on intrahousehold interactions and individual and household outcomes. Section 3.3 discusses the identification and estimation strategies for determining the impact of parent-child injury report mismatches on the harmful child labor decision. Section 3.4 describes the data and sample for this study. Section 3.5 presents the findings. Finally, Section 3.6 summarizes the main findings and provides some concluding remarks.

3.2 Theoretical discussion

Following Becker (1981), the standard economic model of the household assumes that a household behaves like a single individual; that is, all household members have identical preferences and household behavior reflects the solution to an optimization problem involving a single objective function subject to a household budget constraint. However, in the last two decades, the unitary model of the household has been challenged empirically. There is growing evidence that the distribution of resources and decisionmaking power between household members affect outcomes. In particular, greater control over resources by mothers has been consistently documented to result in positive impacts on children's human capital investments and outcomes (see, e.g., Thomas 1990; Pitt and Khandker 1998; Duflo 2003; Pitt et al. 2003; Duflo and Udry 2004; Schady and Rosero 2007; Rubalcava et al. 2009).

The findings from these studies support the intrahousehold allocation model, which allows for preferences to differ between household members. There are two main classes within intrahousehold models. The first is the bargaining model which was developed by Manser and Brown (1980), McElroy and Horney (1981), and Lundberg and Pollak (1993). In this model, household decisions are treated as the outcome of a bargaining game between household members. The solution to this game is sensitive to the threat point of each member (i.e., his or her utility if cooperation fails) as well as the equilibrium concept assumed, where the threat point reflects the "bargaining power" of the household member. The second is the collective model which was developed by Chiappori (1988, 1992), Bourguignon and Chiappori (1992), and Browning and Chiappori (1998). This model does not specify the underlying process within the household but assumes that households make Pareto efficient decisions. This assumption implies that household behavior can be modeled as the solution to maximizing the weighted sum of the individual utility functions of the household decisionmakers, where the weights reflect the bargaining powers of the relevant decisionmakers.

To date, the existing theoretical and empirical child labor literature based on the intrahousehold model has focused on the interactions between spouses. The literature typically assumes a household to comprise of two heads (mother and father) and some number of children (who can be male or female). The male and female heads are assumed to have different preferences over the levels of child labor and schooling of their children; these preferences may also be sensitive to the gender of the child. The realized household outcomes are seen to be the result of the resolution to a bargaining game between male and female heads, and reflect the underlying preferences and the bargaining powers of the two agents. The children are assumed to not have any say in the final outcome. This is a reasonable assumption in many contexts as children may have little autonomy over their fate and parents have the final decisionmaking authority.

One exception to existing child labor studies is Moehling's (2005) study of child labor as a decisionmaking game where the agents involved are parents and children. Using data from the United States from the early 20th century, she finds that children's bargaining power within the household increases as their share of income contributed to total household income increases. She concludes that the implicit threat that children may cease to contribute to household income gives them the ability to influence parents' spending patterns in their favor, demonstrating that children appear to have some agency in the child labor decision.

Following Moehling, I model household decisionmaking as a bargaining process between two agents—parents and children. I posit that parents are altruistic and obtain utility from the human capital development of their children and disutility from child labor in general and harmful child labor in particular. Parents also obtain utility from household consumption. In a developing country setting, children can go to school or work (which includes working in harmful settings).[23] Given that parents place a positive value on their children's human capital development, they desire to send their children to school. However, confronted by poverty, parents decide to send their children to work to supplement household income. Given compensating wage differentials for harmful child labor, certain parents may prefer to send their children

[23] They can also be engaged in both activities, in which case the time spent in school is the inverse of the time spent in harmful child labor. The discussion of the intrahousehold decision model is similar to that above, with parents placing a smaller negative weight on work-related injuries than their children. The outcome of the bargaining process would then be in terms of the number of hours in school versus work.

to work in harmful employment. Further, like parents, children too obtain utility from attending school as well as from household consumption. Further, like parents, they obtain disutility from child labor in general and harmful child labor in particular.

Within this framework, a first potential explanation for parent-child injury report mismatches is simply that parents are not perfectly altruistic and thus discount the event and extent of work-related injury/illness experienced by the child. A second (and more benign) potential explanation is that harmful child labor is an "experience" good, whose attributes and effects cannot be perfectly communicated to parents, thus introducing information asymmetries within the household which can bias household choices. A third potential explanation is that even in the presence of full information as well as perfect sharing of information within the household, parents might engage in what is referred to in the social psychology literature as defensive distortion and denial, namely, phenomena in which individuals deliberately deny or minimize the magnitude of negative events in order to tolerate certain difficult choices.[24] In a collective household model where the household maximizes a weighted sum of the utility functions of the parent and the child, and where the weights reflect the bargaining power of the two actors, the higher bargaining power of the parent vis-à-vis the child's will result in outcomes which reflect more the parent's preferences. Given the above explanations, one such outcome could well be the increased likelihood of the child working in a harmful setting.

More formally, let $u_i : R_+^n \to R$ be agent i's utility function, where R is the set of real numbers, and u_i a well-behaved utility function. The argument $x \in R_+^n$ is a vector of n goods consumed by the household, including the child's schooling. The vector x also includes "bads" that provide disutility as consumption increases.[25] Child

[24] Defensive distortion or denial is also discussed in happiness research by economists where these phenomena can make subjective measures of well-being noisy (see, e.g., Veenhoven 2004).
[25] See, e.g., Felkey (2006) on modeling household "goods" and "bads".

labor, including harmful child labor, can be thought of as such a "bad". In the collective model approach, the household maximizes the objective function

$$\Omega = \theta u_p(x) + (1-\theta)u_c(x),$$

where u_p and u_c are the utility functions of the parent and the child, respectively, and the parameter $\theta \in [0,1]$ reflects the balance of power between the two agents. If θ represents the index of parental power in the household, then θ closer to one implies that the parent has a higher bargaining power relative to the child's. The alternative explanations posited above suggest that the manifestation of parent-child injury report mismatches can be incorporated into the model to reflect either differences in the utility functions *u* between the parent and the child or differences in the perceived attributes of harmful child labor (one of the choice variables in the *x* vector) that enter into the utility functions.

Note that the above theoretical sketch is far from a structural model of the relationship between parent-child injury report mismatches and harmful child labor and does not isolate a specific channel through or set of conditions under which the impact of injury report mismatches emerges. The above discussion however motivates a reduced form approach to estimating the relationship, in which, to the extent possible given the data, other factors considered to be associated with the harmful child labor decision are controlled for.

3.3 Empirical methodology

In this section, I discuss the identification and estimation strategies used to determine the average partial effect of parent-child work injury report mismatches on the probability of harmful child labor using observational data from a single cross-sectional household sample survey dataset, which is all that is available to me (details on the data are discussed in the following section). As a reminder, it is important to keep in mind that parent-child injury report mismatch is defined as a mismatch in a particular direction: the parent reports that the working child did not suffer a work-related injury/illness whereas the child does.

It is plausible that there may be systematic differences between working children with parent-child injury report mismatches and those without which influence harmful child labor outcomes. I use three different empirical strategies to address this potential selection bias and to construct well-defined counterfactuals for the sample of working children with parent-child injury report mismatches. These strategies comprise of comparing outcomes of working children with parent-child injury report mismatches against those without via (1) multiple regression with a rich set of relevant child, adult respondent and household covariates (a standard regression estimator); (2) multiple regression accounting for unobserved heterogeneity that varies at the household level or higher (a fixed-effects regression estimator); and (3) propensity score matching.

Counterfactual framework: Given that my interest lies in understanding the relationship between parent-child injury report mismatches and the probability of harmful child labor, not unlike understanding the relationship between program participation (where selection via endogenous program placement and participation can play important roles) and relevant outcomes of interest, I adopt a program evaluation or counterfactual framework (Rubin 1974; Rosenbaum and Rubin 1983) to present the problem. While uncommon, the adoption of a program evaluation framework for a nonevaluation problem is not an ill fit: for example, Wooldridge (2002) argues that consistent estimators derived under the assumptions made in the program evaluation framework typically reduce to familiar standard estimators such as OLS regression which are used for ceteris-paribus analysis and are traditionally derived very differently, at least prima facie.

Following the exposition in Todd (2008) and Ravallion (2008), suppose there are two states of the world for each working child i $(=1,...,N)$: mismatch $(m_i = 1)$, that is, where the parent does not report a work-related injury suffered by the child whereas the child does, or match $(m_i = 0)$, that is, where the child injury reports by the parent and the child are congruent. Let y_{0i} and y_{1i} denote the potential harmful child labor outcomes in the match and mismatch states for working child *i*, respectively. The observed outcome, y_i, is then given by

$$y_i = m_i y_{1i} + (1 - m_i) y_{0i}. \tag{19}$$

Suppose the outcomes in the match and mismatch states are written as additive-separable functions of a vector of observables *x* and an unobservable *u* as follows:

$$y_{0i} = x_i \beta_0 + u_{0i} \quad i = 1,..., N \tag{20}$$

$$y_{1i} = x_i \beta_1 + u_{1i} \quad i = 1,..., N. \tag{21}$$

Incorporating (2) and (3) and rearranging terms, (1) can be rewritten as

$$y_i = x_i \beta_0 + m_i \left(x_i \left(\beta_1 - \beta_0 \right) \right) + \left\{ u_{0i} + m_i \left(u_{1i} - u_{0i} \right) \right\} \quad i = 1,...,N. \tag{22}$$

Assuming that $E(u_1 \mid x) = E(u_0 \mid x) = 0$ (i.e., the *x* vector is exogenous), the change in the probability of harmful child labor from moving from the match to the mismatch state for working child *i* is given by

$$\Delta_i = m_i \left(x_i \left(\beta_1 - \beta_0 \right) \right) + m_i \left(u_{1i} - u_{0i} \right) \quad i = 1,..., N. \tag{23}$$

It follows that the conditional mean impact on the probability of harmful child labor for children with mismatches is given by

$$\alpha \equiv E(\Delta \mid x, m = 1) = E(y_1 \mid x, m = 1) - E(y_0 \mid x, m = 1) = x(\beta_1 - \beta_0) + E((u_1 - u_0) \mid x, m = 1). \tag{24}$$

This is my parameter of interest, commonly referred to in the program evaluation literature as the *average treatment effect on the treated (ATT)*. Estimating (6) requires essentially solving a missing data problem as information on y_1 is missing for working children with matches $(m=0)$, and, likewise, information on y_0 is missing for working children with mismatches $(m=1)$. Thus, the impact Δ is not directly observed for any working child. Consequently, (6) is unidentified. Importantly, information on the counterfactual conditional expectation $E(y_0 \mid x, m=1)$ required for obtaining the impact parameter α is not directly estimable from the data. Given this, a common way to impute the missing counterfactual conditional expectation is to assume that

$$E(y_0 \mid x, m=1) = E(y_0 \mid x, m=0) = E(y_0 \mid x), \tag{25}$$

that is, that the counterfactual conditional expectations do not vary with mismatch status. Under this assumption, the impact parameter α is identified.

3.3.1. Standard regression estimator

The impact parameter α can be, for example, consistently estimated via standard parametric regressions of (2) and (3) for the match and mismatch samples, respectively. Alternatively, the parameter can be consistently estimated by pooling the two samples together and estimating a regression of (4). Taking (4), and assuming that the value of the unobservable is the same in both states, $u_{1i} = u_{0i}$ for all *i*, and that the value of $x_i(\beta_1 - \beta_0)$ is a constant for all *i*, yields the "constant effects" model

$$y_i = \alpha m_i + x_i \beta_0 + u_{0i} \quad i = 1, \ldots N, \tag{26}$$

where α^* is defined as the solution to α. Assuming (7) under this model is equivalent to assuming that

$$E(u_0 \mid x, m=1) = E(u_0 \mid x, m=0) = E(u_0 \mid x), \tag{27}$$

or that selection into mismatch status is entirely conditional on observables.

Given that the harmful child labor outcome measures of interest are dummy variables, (8) is formulated to represent an underlying latent variable structure. Let y_i^*

denote a latent continuous random variable, reflecting the unobserved tendency of working child *i* towards harmful child labor, determined by the regression function

$$y_i^* = \alpha m_i + x_i \beta + u_i \quad i = 1, \ldots N, \tag{28}$$

where m_i is a dummy variable denoting parent-child injury report mismatch, x_i a vector of child, adult respondent, and household covariates, α and the vector β parameters to be estimated, and *u* the stochastic error term. I estimate (10) by fitting a linear probability model (LPM) to the data.[26] Consistency of the standard LPM estimator of α requires $E\left(u_i \mid x_i, m_i\right) = 0$.

3.3.2. Fixed effects regression estimator

I also attempt to identify and parametrically estimate the impact of parent-child injury report mismatches by exploiting the variation in harmful child labor outcomes among working children from *within* the same household, thus accounting for unobserved heterogeneity that varies between households—this heterogeneity can arise from differences at the household level, village level, or higher. Accounting for this source of potential selection bias is likely to lead to a reduction in total selection bias but may not completely eliminate it. The LPM is now characterized as

[26] It is important to recognize that the LPM is not a typical choice for binary choice estimation; more common choices in applied econometric work are the binomial probit and logit. An important advantage of the LPM over these alternative estimators is that the estimated parameters can be readily interpreted as the marginal effects of the covariates on the conditional probability of harmful child labor, $P\left(y = 1 \mid x, m\right)$. On the other hand, the econometric literature points out two general weaknesses of the LPM relative to its nonlinear competitors. First, some of the predicted probabilities from the LPM can fall outside the $\left[0,1\right]$ interval. Second, a unit change in a given covariate always changes the conditional probability of the outcome variable by the same magnitude, irrespective of the initial values of the covariates. Thus, continually increasing the value of the covariate can eventually result in the conditional probability of the outcome falling outside the unit interval. As Wooldridge (2002) notes, if the objective is to estimate the marginal effects on the conditional probability of the outcome at the center of the distribution of the covariates or averaged over the distribution of covariates rather than at extreme values of covariates, then the LPM typically performs well. Nonetheless, I test the sensitivity of the findings to the choice of estimator; specifically, I also estimate (11) by fitting a binomial logit model to the data, and calculating the marginal effects averaged across the distribution of the covariates (the average marginal effects).

$$y_{ih}^{*} = \alpha m_{ih} + I_h \lambda + x_{ih}\beta + u_{ih}, \quad i = 1,...N, \; h = 1,...H, \qquad (29)$$

where y_{ih} reflects the unobserved tendency of the working child towards harmful child labor and where $y_{ih} = 1\left[y_{ih}^{*} > 0\right]$,[27] m_{ih} a dummy variable denoting parent-child injury report mismatch, x_{ih} a vector of household-varying child covariates (note that adult respondent and household characteristics are not included here as they are household-invariant), I_h a vector of household-specific fixed effects, α and the vectors β and λ parameters to be estimated, and u_{ih} the error term. The error term in (11) can be decomposed as $u_{ih} = \theta_h + \upsilon_{ih}$, where θ_h denotes the effects of household-invariant unobservables and υ_{ih} the standard stochastic error term. Consistency of the household fixed effects LPM regression estimator of α requires that $E\left(\left(\upsilon_{ih} - \upsilon_{jh}\right) | \, x_{ih}, x_{jh}, m_{ih}, m_{jh}\right) = 0, \;\; i \neq j$, that is, that the injury report mismatch status of working children within the household is uncorrelated with the error term υ_{ih}.

3.3.3. Propensity score matching estimator

As an alternative to regression, I also estimate the impact parameter α using matching. Matching is similar to regression but essentially compares the average outcomes of working children with parent-child injury report mismatches to the average outcomes of similar matched working children without parent-child injury report mismatches. To avoid confusion, in the subsection, in line with the program evaluation terminology, children with parent-child injury report mismatches are referred to as treated observations, children without parent-child injury report mismatches as untreated observations, and mismatch status as treatment status; for the remainder of this section, the word "match" is reserved for the method.

To begin, matching has several important advantages over standard regression (Ravallion 2008). First, matching does not require specifying a parametric functional

[27] $y_{ih} = 1\left[y_{ih}^{*} > 0\right]$ is an abbreviated form that indicates that $y_{ih} = 1$ if $y_{ih}^{*} > 0$, and 0 otherwise.

form for relating outcomes to covariates, and thus is less exposed to potential misspecification bias. In contrast, following common practice, the LPM regression specifications discussed earlier specify the outcomes as linear functions of the observables. Second, matching restricts the treated and untreated observations to those with similar propensity scores or in the common support (these terms will be discussed shortly). In contrast, regression uses the full (unmatched) samples of treated and untreated observations, regardless of whether they are indeed comparable. Simulations show that impact estimates based on unmatched samples are more biased and are less robust to misspecification of the regression function than those based on matched samples (Rubin and Thomas 2000). Third, matching looks for covariates that explain treatment status (as well as outcomes of interest), while regression looks for covariates that explain the outcomes of interest, treating these covariates as exogenous. If the covariates only weakly explain outcomes, this poses a problem for regression. In contrast, these same covariates can still help attenuate potential bias in estimating the impact parameter under matching (Rubin and Thomas 2000). Fourth and final, in regression, all untreated observations receive equal weight in determining the counterfactual for each treated observation. On the other hand, in matching, only untreated observations similar to a given treated observation receive positive weight; further, the weights decrease in size as the distance in propensity scores between the positively-weighted untreated observations and the treated observation expands. In this way, potential bias is further attenuated.

In order to identify and estimate the impact parameter using matching, following Heckman et al. (1998), I assume that there exists a set of child, adult respondent, and household covariates x such that

$$E\left(y_0 \mid x, m=1\right) = E\left(y_0 \mid x, m=0\right) = E(y_0 \mid x), \tag{30}$$

that is, the treatment status *m* does not help predict the counterfactual outcome y_0 conditional on the covariates *x*. Further, I assume that

$$P\left(m=1 \mid x\right)<1; \tag{31}$$

this assumption guarantees the possibility of an untreated analogue for each treated observation. Under assumptions (12) and (13), the impact parameter is given by

$$\begin{aligned} \alpha &= E\left(y_1 - y_0 \mid m=1\right) \\ &= E\left(y_1 \mid m=1\right) - E_{x|m=1}\left\{E_y\left(y \mid m=1 \mid x\right)\right\} \\ &= E\left(y_1 \mid m=1\right) - E_{x|m=1}\left\{E_y\left(y \mid m=0 \mid x\right)\right\}, \end{aligned} \tag{32}$$

where the second term can be estimated from the average outcomes of the matched untreated group. It might be the case that there does not exist a set of covariates *x* such that (12) and (13) hold, that is, the support of *x* does not overlap for the treated and untreated groups; in which case, matching is not valid. Matching can only be performed over the "region of common support", and the impact parameter is in fact additionally defined conditional on the region of overlap.

Matching is generally not practicable if the set of covariates *x* is large, as one has to then deal with what is commonly referred to as the "curse of dimensionality".[28] In this context, Rosenbaum and Rubin (1983) provide a solution to reduce the dimensionality of the problem and, thus, greatly increase the tractability of matching, namely

$$E\left(y_0 \mid m,x\right) = E\left(y_0 \mid P(x)\right), \tag{33}$$

that is, if the counterfactual outcomes y_0 are independent of treatment status *m* conditional on the covariates *x*, then they are also independent of treatment status conditional on the probability of treatment $P\left(x\right)=\Pr\left(m=1 \mid x\right)$. In other words, if I can match on *x*, I can also match on $P\left(x\right)$, the propensity score. Thus, I can reduce the dimensionality of the matching problem to that of a univariate matching problem.

[28] As Todd (2008) notes, if *x* is discrete then it is possible to have cells without matches. On the other hand, if *x* is continuous, then nonparametric estimations suffer from slow convergence rates.

The matching problem is then solved using a two-step process. In the first step, the propensity score is estimated using a binary choice model (following common practice, I use binomial logit regression). In the second step, observations are matched on the predicted probabilities of treatment obtained from the first stage using the selected matching estimator described next.

Choice of matching estimator: Several alternative propensity score matching estimators have been proposed in the literature. A typical matching estimator for the impact parameter α takes the form

$$\alpha_M = \frac{1}{n_1} \sum_{i \in I_1 \cap S_p} \left[y_{1i} - \sum_{j \in I_0} W(i,j)\, y_{0i} \right], \tag{34}$$

where I_1 denotes the set of treated observations, I_0 the set of untreated observations, S_p the region of common support, and n_1 the number of observations in the subset $I_1 \cap S_p$. The match for each treated observation $i \in I_1 \cap S_p$ is constructed as a weighted average of the outcomes of untreated observations, where the weights $W(i,j)$ depend on the distance between the propensity scores P_i and P_j. Further, the untreated observations $j \in I_0$ matched to treated observation $i \in I_1 \cap S_p$ are those untreated observations in set $A_i = \{ j \in I_0 \mid P_j \in C(P_i) \}$, where $C(P_i)$ denotes the neighborhood for the treated observation *i*. The alternative matching estimators basically differ in how the neighborhood $C(p_i)$ is defined and how the weights $W(i,j)$ are constructed (see, e.g., Caliendo and Kopeinig 2005, Smith and Todd 2005, and Todd 2008 for an overview of alternative estimators).

From among the set of available estimators, I choose to implement local linear matching to estimate the impact parameter α. Proposed by Heckman et al. (1997), local linear matching is a generalized version of kernel matching. The local linear weighting function is given by

$$W(i,j) = \frac{K_{ij}\sum_{k\in I_0} K_{ik}(P_k - P_i) - \left[K_{ij}(P_j - P_i)\right]\left[\sum_{k\in I_0} K_{ik}(P_k - P_i)\right]}{\sum_{j\in I_0} K_{ij}\sum_{k\in I_0} K_{ij}(P_k - P_i)^2 - \left(\sum_{k\in I_0} K_{ik}(P_k - P_i)\right)^2}, \quad (35)$$

where $K_{ij} = K\left((P_j - P_i)/h\right)$ denotes a kernel function that integrates to one and has mean zero and *h* the parameter determining the kernel bandwidth. For a kernel bounded by -1 and 1, the neighborhood $C(P_i)$ equals $\left\{\left|\frac{P_i - P_j}{h}\right| \le 1\right\}, j \in I_0.$ Compared to standard kernel estimation, local linear estimation has been shown to have faster convergence rates at boundary points. This is a potentially useful property when (1) untreated observations are distributed asymmetrically around treated observations in terms of *P*; (2) *P* for some observations are near zero or one; and (3) there are gaps in the distribution of *P* (Galdo et al. 2007; Smith and Todd 2005). In addition, the performance of local linear regression is more robust to different data design densities (Fan 1992).

Choice of kernel function: Two commonly used kernels in the applied matching literature are the Epanechnikov and Gaussian kernels. Black and Smith (2004) test both kernels in a study of the effects of college quality using data from the National Longitudinal Survey of Youth (NLSY), and find that the Epanechnikov kernel performs slightly better irrespective of bandwidth choice. Specifically, they find that their nonparametric estimations using the Epanechnikov kernel have faster convergence rates and implicitly impose the common support condition through the choice of the kernel bandwidth. In light of this evidence, I use the Epanechnikov kernel, which is a popular choice in applied studies using matching. Notwithstanding, simulation evidence suggests that nonparametric estimates are much less sensitive to the choice of kernel than the choice of bandwidth (Cameron and Trivedi 2005; Imbens and Wooldridge 2008).

Choice of kernel bandwidth: The kernel bandwidth or smoothing parameter *h* has important implications for the tradeoff between variance and bias: a smaller bandwidth yields a smaller bias but a larger variance (due to undersmoothing), while a larger bandwidth yields a larger bias but a smaller variance (due to oversmoothing). In the specific case of matching, too large a bandwidth size means that untreated observations which may be quite different from the treated observation are included in the estimation of the expected counterfactual outcome and too small a bandwidth size means that too few untreated observations are used to estimate the expected counterfactual outcome, yielding a noisy estimate of the impact parameter. Following Galdo et al. (2007), I determine the optimal bandwidth size by minimizing the mean squared error (MSE)—which is equal to the sum of the variance and the square of the bias— of the local linear matching estimator.

Choice of variance estimator: Standard errors for impact parameter estimates are typically generated using bootstrap resampling methods in the applied matching literature. Abadie and Imbens (2006) find that standard bootstrapping does not yield valid standard errors in the case of the nearest-*k* neighbor matching estimator and arrive at a consistent analytical estimator for the standard error; this result does not apply to kernel or local linear matching estimators. In addition, Heckman et al. (1997) derive the asymptotic properties for the kernel and local linear matching estimators. In light of these results, I report both analytical and bootstrapped standard errors (resampling 1,000 times from the data) for the matching-based parameter estimates.

3.4 Data and sample

3.4.1. Data

The data for this study come from the Filipino 2001 *Survey of Children, 5-17 Year Olds* (SOC). This survey was administered as a rider to the October round of the 2001

Labor Force Survey by the National Statistics Office in collaboration with the Bureau of Labour Employment Statistics of the Department of Labour and Employment, with technical and financial assistance from the ILO's International Program on the Elimination of Child Labour (IPEC).

The survey adopted a multi-stage clustered sampling design, yielding data that are representative at the national as well as regional levels.[29] In the first stage, sample *barangays* (barangy is the smallest administrative unit in the country) were selected systematically with probability proportional to size from a frame of barangays stratified on the basis of province, urban/rural, and other dimensions to ensure nearly complete geographic coverage. In the second stage, sample enumeration areas (EAs, physical divisions of barangays) were selected systematically with probability proportional to size in the sample barangays. Finally, in the third stage, private households were selected systematically in the sample EAs (twelve households per EA) based on the 1995 population census frame.[30] These complex survey design features of stratification, clustering, and sampling weights given the unequal probability of selection are incorporated into all estimations.

The original survey sample was 2,247 primary sampling units (EAs) and 26,964 private households. Out of this number, 17,454 households (64.7%) had children aged 5-17 years; 17,444 out of the 17,454 households (99.9%) were successfully interviewed. The total number of children residing in the interviewed households was 41,924; out of this number, 6,523 children (15.6%) reported that they worked in the twelve months preceding the survey (September 2000-October 2001); 6,365 out of the 6,523 working children (97.6%) were successfully interviewed.

---

[29] At the time of the survey, there were 16 regions in the Philippines. A region is a grouping of provinces (province is the largest administrative unit in the country), based on similarities in geographical, cultural, and ethnographic characteristics.
[30] Individuals residing in institutions or establishments were not covered in the survey.

Consequently, survey nonresponse by households and children appears to be a negligible source of selection bias. The definition of work used in the survey follows ILO guidelines: a child was considered to be working or economically active if at any time in the reference period the child was engaged in an economic activity for at least one hour (ILO 2009).

Two survey questionnaires were fielded as part of the 2001 SOC: SOC Form 1 which largely collected information on the socioeconomic characteristics of households with children aged 5-17 years and SOC Form 2 which collected detailed information on the work and working characteristics of working children. The respondent for the SOC Form 1 is the parent or the guardian of the child or children aged 5-17 years; the main purpose of this survey was to identify eligible child respondents for administering the SOC Form 2. The respondent for the SOC Form 2 is the working child residing in the household.[31]

Several steps were taken to ensure the reliability and interpersonal comparability of the data. For example, survey interviewers were expected to read the questions exactly as worded in the questionnaires (either in English or the local language), and maintain a professional, dispassionate demeanor through the entire interview. In addition, to the extent possible, interviews were to be one-on-one and conducted in private. Castro et al. (2005) however cautions that the reliability of data from parents and children may be undermined if child labor is a sensitive subject (this is plausible as children below ages 15 and 18 are legally prohibited from general and hazardous work in the Philippines, respectively). They also point out that children may

[31] A source of potential selection bias arises from the fact that comparable work data on children residing outside the household were not collected in the survey. The survey however collected limited basic data on unmarried children aged 5-17 years residing outside the household. The total number of such children is very small: 462. Among these, 265 children (or 57.4%) were reported to have worked in the reference period. Given these small absolute numbers, selection due to sample truncation is likely to be minor—if comparable data were available for these children, they would increase the total child and working child samples by 1% and 4.1%, respectively.

be less-reliable respondents than adults. For example, they are more likely to be (1) unaware of salient characteristics of their work (such as workplace hazards) and/or (2) unable to accurately convey them to the interviewer due to recall and spoken language command problems; further, these problems are likely to be exacerbated the younger the child. The degree to which these issues affect the reliability of these data is however unknown. Notwithstanding, in the SOC Form 2, the interviewer was asked to assess the levels of interest and sincerity of the child respondent (based on predefined scales) as well as note down sections and questions where the child had any doubts, difficulties, or apprehensions. While the notes are unavailable, I include the interviewers' categorical assessments as controls in the outcome and propensity score regression estimations.

Construction of the mismatch variable: The SOC Form 1 collected limited work-related information on children from the parent; as part of this information, the parent was asked the question "Did you child suffer from a work-related injuries/illnesses?" This was followed by a related question on the type of injury or illness, which was postcoded by the interviewer using a scheme provided in the survey interviewer manual—the scheme lists different types of injuries and illnesses and provides definitions. Multiple responses were allowed. If the response could not be categorized using the given code list, the interviewer coded it as "other" and wrote in the injury or illness. Similarly, in the SOC Form 2, the working child was asked the question "Have you ever experienced any injuries/illnesses while working?" This is followed by a series of related questions on the different types of injuries or illnesses suffered using a precoded list (with multiple responses allowed), the severity of the injury or illness measured in terms of work cessation, and treatment.[32] The equivalent questions posed

[32] As noted by Castro et al. (2005), the questions on work-related injuries/illnesses do not explicitly specify a reference period (in fact, this applies to all the work-related questions); only the question used to screen-in respondents for the SOC Form 2 explicitly states the reference period. However, it is

to the parent and child on the event of injury or illness are used to construct the measure of parent-child injury report *mismatch*, where one denotes a mismatch $(m = 1)$, that is, where the parent reports that the child did not suffer a work-related injury/illness but the child does, and zero a match $(m = 0)$, that is, where the child injury reports by the parent and the child are congruent. However, as explained later, this indicator is constructed for a marginally smaller sample than the full sample of working children due to missing data and child work-related injury reports by adult respondents other than parents.

Construction of the harmful child labor variables: Data on harmful child labor are obtained from responses of working children in the SOC Form 2. Specifically, I use data from the following questions: (1) "Did/do you perform heavy physical work?"; (2) "Did/do you consider some aspects of your work risky or dangerous?"; (3) "Did/do your work often involve exposure to the following physical environmental characteristics?"; and (4) "Did/do your work often involve exposure to the following chemicals?" The response options for question (1) are "always", "sometimes", "seldom", or "never"; for question (2), "yes" or "no"; and for questions (3) and (4), "yes" or "no" to a precoded list of hazards (with multiple responses allowed). The distributions of responses to these questions are presented in Table 3.1.

Survey interviewers are offered some guidance on what the above questions on harmful child labor are asking. For example, illustrative examples of heavy physical work provided in the survey interviewer manual include transporting heavy loads, lifting heavy items, using heavy tools and machinery, digging, and quarrying. For risky or dangerous work, the manual lists the following risks as examples: vehicular

---

possible that both parents and children might report work-related injuries and illnesses that date from farther back in time than the last twelve months, potentially upwardly biasing injury and illness incidence rates in the sample.

accidents for delivery workers; extreme light from welders which can cause loss of sight; and loud noise of machines which can cause loss of hearing.

Table 3.1. Sample distribution of responses to original harmful child labor questions

	Frequency	Share (%)
Did/do you perform heavy physical work?		
Always	352	5.53
Sometimes	1,376	21.62
Seldom/rarely	986	15.49
Never	3,651	57.36
Did/do you consider some aspects of your work risky or dangerous?		
Yes	1,288	20.24
No	5,077	79.76
Did/does your work often involve exposure to the following? Physical hazards (multiple responses permitted)		
Noise	257	4.04
Temperature/humidity	1,698	26.68
Pressure	147	2.31
Inadequate lighting	47	0.74
Slip/trip/fall hazards	435	6.83
Insufficient exit for prompt escape	19	0.3
Congested layout	31	0.49
Radiation/ultraviolet/microwave	302	4.74
Other	307	4.82
Not exposed to physical elements	3,122	49.05
Chemical hazards (multiple responses permitted)		
Dust (e.g., silica dust, saw dust, sanding dust)	762	11.97
Liquid (e.g., oil, gasoline, mercury)	90	1.41
Mist/fumes/vapors (e.g., paint, insecticides, pesticides)	404	6.35
Gas (e.g., oxygen, ammonia)	75	1.18
Other	107	1.68
Not exposed to chemicals	4,927	77.41

For the questions on exposure to specified physical and chemical hazards, the interviewers are explicitly directed to read each listed type of hazard and explain the meaning by providing examples included in the manual. The manual also provides guidance on the meaning of response categories. For the question on heavy physical work, the frequency level "always" is defined as daily to 3-6 times a week and "sometimes" as 1-2 times per week or 1-3 per month.

Using these data, the following harmful child labor indicator variables are constructed: *physically strenuous work*, where one denotes response options "always" or "sometimes" and zero "seldom" or "never"; *risky work*, where one denotes "yes" and zero "no"; *exposure to physical hazards*, where one denotes an affirmative response to exposure to at least one physical hazard, zero otherwise; and analogously for *exposure to chemical hazards*.

Note that the variables physically strenuous work and risky work can be construed as capturing subjective data on harmful child labor, whereas the variables exposure to physical and chemical hazards can be viewed as capturing objective data, given that the underlying data required the child to specify the precise hazard. While these variables may not necessarily reflect the same latent distribution of harmful child labor (and subjective measures may be noisier due to interpersonal differences in the interpretation of the response scales and options), I examine whether there is consistency in inference results across objective and subjective measures of harmful child labor. Notwithstanding, alternative measures of harmful child labor are also used to gauge the sensitivity of the results to the choice of the harmful child labor outcome measure.

3.4.2. Sample

An inspection of the data indicates that the child injury report was not necessarily provided by the child's parent, though the survey question was worded to suggest that the adult respondent was the parent. As my interest lies in understanding injury report mismatches between filially-connected individuals, the sample is restricted to parent-child pairs. First, in terms of data coverage, 6,283 out of the 6,365 working children (98.7%) have child injury reports by an adult respondent. Unfortunately, data that link the child directly to the adult respondent in terms of relationship are unavailable. However, a workaround solution exists. For 6,271 out of the 6,283 child injury reports (99.8%) by an adult respondent, information on the relationship of the adult respondent to the household head is available. Likewise, for all working children, information on their relationship to the household head is available. Thus, the relationship between the child and the adult respondent can be determined *via* the child's relationship to the household head.

Among the 6,271 working children with child injury reports by adult respondents, 5,520 children (or 88%) were recorded as the sons or daughters of the household head. For these 5,520 children, 1,562 (28.3%) and 3,650 (66.1%) children had child injury reports by the household head or the head's spouse (virtually always female), respectively. I assume that the head's spouse is the mother of the child. This is highly plausible as less than a handful of the surveyed households with working children had household heads with multiple spouses. Further, without accounting for child birth order, the 10th-90th percentile range for the age of the head's spouses at the year of birth of the child was 20-36 years, which corresponds to the high fertility period for women generally. There will certainly be some errors of inclusion but I expect these cases to be low and have a negligible effect on the empirical analysis. Thus, at the end of this process, out of the original sample of 6,365 working children, 5,212 children (81.9%) have child injury reports by one of their two parents.

Two additional restrictions are imposed in order to arrive at the final sample used for inferring the impact of parent-child injury report mismatches on the probability of harmful child labor. First, working children who reported not having experienced a work-related injury or illness but have a conflicting child injury report by one of their parents are excluded. This decision was made as this particular report mismatch type is arguably less salient than a conflicting parental response when the working child reports a work-related injury or illness; notwithstanding, the former mismatch type is found to be a highly rare event and one that can be plausibly explained by simple random reporting error (more details on this are provided in the next section). Second, children who reported holding more than one job in the reference period were excluded. This additional restriction was made due to a data issue: for working children that reported more than one job, data on employment characteristics were collected without specific reference to which particular job they pertain to.[33] These two restrictions result in a reduction of the sample from 5,212 to 4,600 working children (a 12% reduction in sample size).

3.5 Findings

3.5.1. Profile of work-related injuries and illnesses

In this section, I present the main empirical findings from examining the impact of parent-child injury report mismatches on the probability of harmful child labor. However, before presenting these findings, I first provide some basic descriptive statistics on the incidence of work-related injuries and illnesses (hereafter, for the sake of brevity, referred to as injuries) among working children, the distribution of the types of injuries and their treatment, as well as the incidence of parent-child injury report mismatches. This information not only serves to contextualize the impact

[33] This restriction can potentially introduce sample selection bias. Given this, I also redo the analysis on the sample which relaxes this restriction and find that the size and significance of the impact estimates are robust to sample choice along this dimension (these results are available from the author upon request).

findings reported later, but is also useful in its own right, given the limited available information in the literature on this particular aspect of child labor.

Table 3.2 presents the share of working children ages 5-17 years that reported work-related injuries as well as the distribution of injuries and illnesses by type.

Table 3.2. Incidence of work-related injuries and illnesses among children, 5-17 year olds

	(1) All working children (in percent)	(2) Working children with parental second-hand reports (in percent)
Share reported work-related injuries or illnesses	29.1	29.9
Of which, share reported …		
Injuries		
Contusions and bruises	18.3	19.1
Cuts and puncture wounds	71.2	72.3
Crushing injuries	3.1	3.2
Dislocations, fractures, and sprains	4.5	4.7
Burns	6.3	6.2
Other injuries	2.1	1.8
Illnesses		
Itching and skin rashes	24.9	25.8
Body aches and pains	48.0	46.7
Visual or hearing problems	1.3	1.1
Respiratory or gastrointestinal problems	7.6	7.6
Other illnesses	6.8	6.9
N	6,365	5,212

Notes: Children were permitted to report more than one type of injury or illness. Consequently, the distribution of injuries and illnesses can exceed 100%. One category of injury was omitted: loss of body parts and amputation. Statistics are adjusted for sampling weights.

Column 1 of the table reports statistics for the full sample of working children while Column 2 reports the corresponding statistics for those working children who have child injury reports from their respondent parents. First, looking at Column 1, roughly 30% of children reported work-related injuries. Second, the most commonly-reported injury or illness was open wounds (71%), followed, in turn, by body aches and pains (48%) and skin irritation and rashes (25%). The rest of the specified injuries and illnesses were relatively uncommon—each type garnered less than 10% of reporting children. Third, comparing the statistics in Column 1 to Column 2, the shares across the two samples appear to be similar, suggesting that negligible bias was introduced by restricting the sample to those working children with child injury reports from their parents as opposed to any household adult member.

Table 3.3 restricts attention to those working children that reported work-related injuries and/or illnesses and presents statistics on their treatment and the extent to which the injuries resulted in work cessation.

Table 3.3. Treatment and severity of injuries and illnesses suffered by children

	(1) Children reporting injuries or illnesses (in percent)	(2) Children reporting injuries with parental second-hand reports (in percent)
Share received treatment	70.7	70.4
Distribution of treatment by type of provider		
Parent	48.8	51.0
Self	27.4	27.1
Employer	5.1	4.2
Other	3.5	2.6
Share for whom injury/illness resulted in work stoppage	16.4	16.2
N	1,751	1,479

Notes: Children were permitted to report more than one treatment provider; consequently, the distribution of treatment providers can exceed 100%. Shares are adjusted for sampling weights.

Column 1 reports statistics for the full sample of working children that reported work-related injuries while Column 2 reports the corresponding statistics for those working children who reported injuries and have child injury reports from their respondent parents. Focusing first on Column 1, the data suggest that the majority of working children who reported injuries (71%) were treated for their injuries or illnesses. Further, the most-common provider of treatment was their parents (49%), followed by self-treatment (27%). This evidence seems to discount one of the alternative hypotheses I offered in Section 2 for the presence of parent-child injury report mismatches, namely asymmetric information on injuries/illnesses between the child and the parent.

While one source of information which can be used to develop an indicator of the severity of an injury or illness is the breakdown of injuries and illnesses experienced by working children presented in Table 3.2, an alternative indicator is whether the injury or illness resulted in the cessation of work for a period of time. Using the latter indicator, Table 3.3 shows that roughly 16% of children stopped work. Comparing Column 1 to Column 2 in the table, the statistics again appear to be similar across the two samples, suggesting negligible sample selection bias along these lines.

Table 3.4 presents statistics from a cross-tabulation of parents' and children's responses on work-related injuries experienced by the child. I highlight two findings here.

Table 3.4. Work-related injury or illness experienced by the working child: parent vs. child's response

		Parental-report of child work-related injury or illness		
	Key: Observations Cell proportion	No	Yes	Total
Self-report (child) of work-related injury or illness	No	3,660 70.2 (68.6)	73 1.4 (1.5)	3,733 71.6 (70.1)
	Yes	727 14.0 (13.9)	752 14.4 (16.0)	1,479 28.4 (29.9)
	Total	4,387 84.2 (82.5)	825 15.8 (17.5)	5,212 100.0 (100.0)

Notes: Shares corrected using sampling weights are provided in parentheses.

First, the data suggest that a match in the child injury report between the respondent parent and the child is largely the norm in the sample—specifically, for 85% of the sample, there is a parent-child match in responses. Decomposing this finding however reveals that the high incidence of parent-child injury report matches is largely the result of the better confirmatory performance of parents in one direction: that of a match in responses when the child reports *no* injury. If I was to restrict the sample to working children who reported an injury, I find that a striking 49% of respondent parents report that their children did *not*. Second, conditional on a parent-child mismatch in response (the cross-diagonal cells in the table), the response combination of the parent reporting that the child did not suffer a work-related injury when the child reports that she did is roughly ten-fold more likely than the response combination of the parent reporting that the child did suffer a work-related injury when the child reports that she did not; the difference in the incidences of these two combinations is highly statistically significant. Further, the absolute numbers of the latter response combination are so small relative to the full sample size that they can be plausibly explained by random reporting error and thus viewed as unmeaningful.

Table 3.5 provides definitions for all variables used in the empirical analysis. Table 3.6 presents summary statistics for the alternative dichotomous measures of harmful child labor, the parent-child injury report mismatch measure, and the child, respondent parent, and household covariates included in the different regressions. The extent of harmful child labor in the sample varies from a low of 20% when measured using risky work to a high of 51% when measured using exposure to physical hazards. The share of children with parent-child injury report mismatches, the treatment measure, is 16%. In terms of other selected sample characteristics, the majority of children are male (62%), attend school (69%), and work in an own-household operated or owned enterprise (61%).

Table 3.5. Description of variables

Variable	Description
Outcome variables	
Physically strenuous work	Indicator variable based on child's self-report of frequency of arduous work. The variable equals 1 if the frequency is "always" or "sometimes"; 0 zero otherwise.
Risky work	Indicator variable based on child's self-report of risky or dangerous work. The variable equals 1 if yes; 0 otherwise.
Exposed to physical hazards	Indicator variable signifying child's exposure to physical hazards. Variable equals 1 if exposed to at least one of the listed hazards; 0 otherwise.
Exposed to chemical hazards	Indicator variable signifying child's exposure to chemical hazards. Variable equals 1 if exposed to at least one of the listed chemical hazards; 0 otherwise.
Child covariates	
Male	Male indicator
Age/10	Age (in completed years)/ 10
Age squared/100	Square of age (in completed years)/100
Age first worked/10	Age at which the child first started working/10
Age first worked squared /100	Square of age at which the child first started working/100
Daily work hours: 5-8 hours	Indicator variable constructed from question on normal working hours per day. Reference category: 1-4 hours.
Daily work hours: 9+ hours	Same as above.
Weekly work days/10	Days worked per week/10
Weekly work days squared /100	Square of days worked per week/100
Night work	Indicator variable constructed from question on whether the child worked usually in the evening or night. The variable equals 1 if yes; 0 otherwise.
Presently in school	Indicator variable signifying attending school presently.
Ever left school	Indicator variable signifying stopped or dropped out of school in the past.
Secondary or tertiary schooling	Indicator variable constructed from question on highest level of education completed. Variable equals 1 if highest level is "high school undergraduate", "high school graduate", or "college undergraduate"; 0 otherwise.
Location: Worked in own house	Indicator variable constructed from question on place of work in the last twelve months. Reference category: Site other than farm or own home
Location: Worked on farm	Same as above.
Worked in own household enterprise	

Table 3.5 (Continued)

Paid for work	Indicator variable constructed from question on the nature of payment for work. Variable equals 1 if child received any form of payment; 0 otherwise.
Received meals as worker benefit	Indicator variable constructed from question on whether the child received meal allowance either in cash or kind.
Worked for financial reasons	Indicator variable constructed from question on main reason for working or having a job in the reference period. The variable equals 1 if responses are "to help pay own family debts", "to pay own schooling", "to supplement family income", or "to earn money to establish own business"; 0 otherwise.
Gave earnings to parents	Indicator variable constructed from question on share of earnings given to parents. The variable equals 1 if responses are "Yes, wholly" or "Yes, partly"; 0 otherwise.
Parent supervised work	Indicator variable constructed from question on whether the child's work was supervised by an adult and who the adult was. Reference category: Employer supervised work.
Other relative supervised work	Same as above.
Unsupervised	Same as above.
Interview interest	Indicator variable constructed from question to interviewer on the level of interest of the respondent. The variable equals 1 if assessment is "very interested" or "interested"; 0 otherwise.
Interview sincerity	Indicator variable constructed from question to interviewer on the level of sincerity of the respondent. The variable equals 1 if the assessment is "sincere"; 0 otherwise.
Respondent parent covariates	
Male	Male indicator.
Age/10	Age (in completed years)/10
Age squared/100	Square of age (in completed years)/100
Secondary or tertiary education	Indicator variable constructed from question on highest level of education completed. Variable equals 1 if highest level is "high school undergraduate", "high school graduate", or "college undergraduate"; 0 otherwise.
Head of household	Head of household indicator

Table 3.5 (Continued)

Household covariates	
Run household enterprise	Indicator variable constructed form question on whether the household was engaged in a household business/enterprise in the last twelve months.
Own farm land	Indicator variable constructed from question on the types of land the household owns.
Household size/10	Number of household members/10
Household size squared/100	Square of the number of household members/100
Children/10	Number of children (<17 years)/10
Children squared/100	Square of the number of children (<=17 years)/100
Household income: P2000-P2999	Average monthly gross income in the last twelve months. Reference category: Less than P2,000.
Household income: P3000-4999	Same as above
Household income: P5000-9999	Same as above
Household income: P10000-14999	Same as above
Household income:>P15000	Same as above
Urban	Urban indicator.

Table 3.6. Summary statistics for working child sample

Variable	Mean	Standard deviation	Minimum	Maximum
Outcome variables				
Physically strenuous work	0.27	0.44	0.00	1.00
Risky work	0.20	0.40	0.00	1.00
Exposed to physical hazards	0.51	0.50	0.00	1.00
Exposed to chemical hazards	0.23	0.42	0.00	1.00
Child covariates				
Male	0.62	0.48	0.00	1.00
Age/10	1.38	0.25	0.50	1.70
Age squared/100	1.98	0.65	0.25	2.89
Age first worked/10	1.13	0.27	0.40	1.70
Age first worked squared /100	1.35	0.63	0.16	2.89
Daily work hours: 5-8 hours	0.39	0.49	0.00	1.00
Daily work hours: 9+ hours	0.09	0.29	0.00	1.00
Weekly work days/10	0.35	0.22	0.10	0.70
Weekly work days squared /100	0.17	0.18	0.01	0.49
Night work	0.21	0.41	0.00	1.00
Presently in work	0.69	0.46	0.00	1.00
Ever left school	0.35	0.48	0.00	1.00
Secondary or tertiary schooling	0.42	0.49	0.00	1.00
Location: Worked in own house	0.19	0.39	0.00	1.00
Location: Worked on farm	0.47	0.50	0.00	1.00
Worked in own household enterprise	0.61	0.49	0.00	1.00
Paid for work	0.37	0.48	0.00	1.00
Received meals as worker benefit	0.09	0.28	0.00	1.00
Worked for financial reasons	0.40	0.49	0.00	1.00
Gave earnings to parents	0.35	0.48	0.00	1.00

Table 3.6. (Continued)

Parent supervised work	0.52	0.50	0.00	1.00
Other relative supervised work	0.05	0.23	0.00	1.00
Unsupervised	0.25	0.44	0.00	1.00
Interested during interview	0.78	0.42	0.00	1.00
Sincere during interview	0.59	0.49	0.00	1.00
Report mismatch	0.16	0.37	0.00	1.00
Respondent parent covariates				
Male	0.23	0.42	0.00	1.00
Age/10	4.24	0.75	2.40	7.80
Age squared/100	18.49	6.74	5.76	60.84
Secondary or tertiary education	0.41	0.49	0.00	1.00
Head of household	0.30	0.46	0.00	1.00
Household covariates				
Run household enterprise	0.86	0.35	0.00	1.00
Own farm land	0.40	0.49	0.00	1.00
Household size/10	0.67	0.22	0.20	2.30
Household size squared/100	0.50	0.35	0.04	5.29
Children/10	0.39	0.19	0.10	1.40
Children squared/100	0.18	0.17	0.01	1.96
Household income: P2000-P2999	0.16	0.37	0.00	1.00
Household income: P3000-4999	0.29	0.45	0.00	1.00
Household income: P5000-9999	0.28	0.45	0.00	1.00
Household income: P10000-14999	0.10	0.30	0.00	1.00
Household income:>P15000	0.08	0.28	0.00	1.00
Urban	0.43	0.50	0.00	1.00

Notes: Summary statistics for region dummies are omitted. All statistics are adjusted for sampling weights.

Respondent parents are mainly female (77%) and hence non-household heads; 41% have some secondary or tertiary education. Finally, the children come mainly from rural households (57%) and those that run household enterprises (86%).

3.5.2. Effects of other regression covariates on the probability of harmful child labor

Tables 3.7 and 3.8 present the full set of results from estimating the standard and household fixed-effects LPM regressions (see Eqns. (10) and (11)) for the alternative harmful child labor outcome measures, namely physically strenuous work, risky work, exposure to physical hazards, and exposure to chemical hazards. Before turning to the impact of parent-child injury report mismatches on the probability of harmful child labor, I discuss the estimated effects of the other covariates on the probability of harmful child labor given that little is known in the literature on what factors are associated with harmful child labor. Across the four outcome measures and the two regression estimators, the data suggest that only two of the covariates have generally consistent significant effects on the conditional probability of harmful child labor: the child's gender and place of work. Specifically, boys are more likely to be engaged in harmful child labor than girls—for example, in absolute terms, the magnitude of the estimated marginal effects from the standard regression estimator ranges from a low of 3.9 percentage points in the case of exposure to physical hazards to a high of 15.4 percentage points in the case of physically strenuous work. Relative to the mean incidence of harmful child labor measured using the relevant outcome indicator, these effects translate into roughly 8% and 57%, respectively. In addition, children who work in their own homes are less likely than children who work outside their own homes in nonfarm employment—for example, in absolute terms, the estimated marginal effects from the standard regression estimator ranges from a low of

Table 3.7. Determinants of the probability of harmful child labor
Linear probability model (LPM) regression estimates

Independent variables	(1) Physically strenuous work	(2) Risky work	(3) Exposure to physical hazards	(4) Exposure to chemical hazards
Child variables				
Male	0.154***	0.095***	0.039**	0.063***
	(0.016)	(0.015)	(0.019)	(0.015)
Age/10	-0.682**	-0.654*	0.077	-0.111
	(0.320)	(0.336)	(0.368)	(0.325)
Age squared/100	0.337***	0.265**	-0.009	0.059
	(0.123)	(0.129)	(0.138)	(0.125)
Age first worked/10	0.612**	-0.064	-0.947***	-0.166
	(0.251)	(0.281)	(0.326)	(0.259)
Age first worked squared /100	-0.296***	0.026	0.417***	0.101
	(0.108)	(0.121)	(0.138)	(0.114)
Daily work hours: 5-8 hours	0.095***	0.033	0.079***	0.057***
	(0.022)	(0.021)	(0.026)	(0.022)
Daily work hours: 9+ hours	0.099***	0.074**	0.025	0.007
	(0.036)	(0.036)	(0.041)	(0.032)
Weekly work days/10	1.067***	0.599**	0.323	0.279
	(0.209)	(0.237)	(0.267)	(0.225)
Weekly work days squared /100	-1.221***	-0.764***	-0.619**	-0.376
	(0.248)	(0.273)	(0.312)	(0.258)
Night work	0.036*	0.038*	0.012	-0.002
	(0.022)	(0.022)	(0.027)	(0.021)
Presently attending school	-0.022	-0.020	0.016	0.025
	(0.028)	(0.026)	(0.029)	(0.025)
Ever left school	0.043*	-0.006	0.020	-0.001
	(0.023)	(0.020)	(0.024)	(0.021)
Secondary or tertiary schooling	0.000	-0.018	-0.003	-0.008
	(0.021)	(0.019)	(0.022)	(0.019)
Location: Worked in own house	-0.042*	-0.149***	-0.292***	-0.133***
	(0.021)	(0.023)	(0.032)	(0.024)
Location: Worked on farm	0.046*	-0.057**	0.096***	-0.025
	(0.024)	(0.026)	(0.032)	(0.025)

Table 3.7 (Continued)

Worked in own household enterprise	0.041	0.018	0.076*	0.096***
	(0.035)	(0.036)	(0.042)	(0.032)
Paid for work	0.030	0.055	0.024	0.080**
	(0.037)	(0.041)	(0.044)	(0.038)
Received meals as worker benefit	-0.025	-0.114***	-0.081**	-0.020
	(0.033)	(0.032)	(0.038)	(0.035)
Worked for financial reasons	0.034	0.014	0.017	0.010
	(0.021)	(0.020)	(0.025)	(0.021)
Gave earnings to parents	0.034	0.081***	0.043	0.055*
	(0.028)	(0.026)	(0.031)	(0.029)
Parent supervised work	-0.042	0.048	-0.002	0.046
	(0.034)	(0.034)	(0.039)	(0.031)
Other relative supervised work	-0.044	0.065	0.062	0.026
	(0.044)	(0.045)	(0.052)	(0.043)
Unsupervised	-0.018	0.051	0.004	0.034
	(0.032)	(0.033)	(0.036)	(0.029)
Interested in interview	0.033	-0.029	-0.050*	-0.011
	(0.021)	(0.023)	(0.027)	(0.021)
Sincere in interview	-0.041**	0.005	0.012	-0.014
	(0.018)	(0.018)	(0.021)	(0.017)
Injury report mismatch	*0.065****	*0.118****	*0.112****	*0.049***
	(0.023)	*(0.028)*	*(0.029)*	*(0.023)*
Respondent parent variables				
Male	-0.040	-0.005	-0.038	--
	(0.041)	(-0.13)	(0.039)	(0.046)
Age/10	0.072	-0.007	-0.040	0.043
	(0.083)	(-0.070)	(0.105)	(0.078)
Age squared/100	-0.008	0.001	0.003	-0.009
	(0.009)	(0.097)	(0.011)	(0.008)
Secondary or tertiary education	-0.038**	-0.022	-0.013	0.010
	(0.017)	(-1.41)	(0.021)	(0.017)
Head of household	0.031	0.008	0.026	0.026
	(0.039)	(0.21)	(0.036)	(0.040)

Table 3.7 (Continued)

Household variables				
Run household enterprise	-0.084***	0.030	0.043	-0.010
	(0.029)	(1.13)	(0.032)	(0.029)
Own farm land	0.021	-0.011	-0.054**	0.029
	(0.018)	(-0.55)	(0.026)	(0.018)
Household size/10	-0.320	-0.425	-0.001	-0.199
	(0.199)	(-2.18)	(0.246)	(0.193)
Household size squared/100	0.198	0.255	0.068	0.275**
	(0.124)	(2.03)	(0.155)	(0.114)
Children/10	0.380*	0.369	-0.016	-0.201
	(0.212)	(1.85)	(0.243)	(0.207)
Children squared/100	-0.362	-0.369	-0.038	-0.127
	(0.226)	(-1.82)	(0.258)	(0.199)
Household income: P2000-P2999	-0.041	-0.027	0.007	0.010
	(0.033)	(-0.75)	(0.041)	(0.031)
Household income: P3000-4999	-0.053	-0.042	-0.059	-0.001
	(0.035)	(-1.10)	(0.045)	(0.034)
Household income: P5000-9999	-0.041	-0.026	-0.045	-0.007
	(0.038)	(-0.64)	(0.049)	(0.035)
Household income: P10000-14999	-0.048	-0.032	-0.034	0.003
	(0.041)	(-0.72)	(0.054)	(0.041)
Household income:>P15000	-0.047	-0.068	0.062	-0.045
	(0.045)	(-1.42)	(0.059)	(0.046)
Urban	-0.027	-0.013	0.033	0.091***
	(0.020)	(-0.61)	(0.025)	(0.021)
Intercept	-0.167	0.539	0.969***	0.145
	(0.262)	(1.83)	(0.346)	(0.257)
Observations	4600	4600	4600	4600
R-squared	0.20	0.13	0.18	0.09

Notes: Standard errors reported in parentheses. * denotes statistical significance at the 10%; ** at the 5% level; and *** at the 1% level. All estimates are adjusted for complex survey design features. Estimates for region dummies are not reported in table.

Table 3.8. Determinants of the probability of harmful child labor
Linear probability model (LPM) regression estimates, Household fixed effects

Independent variables	(1) Physically strenuous work	(2) Risky work	(3) Exposure to physical hazards	(4) Exposure to chemical hazards
Child variables				
Male	0.088***	0.067***	-0.012	-0.012
	(0.020)	(0.018)	(0.022)	(0.014)
Age/10	-0.437	-0.100	-0.219	0.010
	(0.382)	(0.398)	(0.362)	(0.326)
Age squared/100	0.253*	0.078	0.150	0.017
	(0.149)	(0.153)	(0.139)	(0.125)
Age first worked/10	-0.331	-1.013**	0.369	-0.130
	(0.374)	(0.431)	(0.404)	(0.381)
Age first worked squared /100	0.150	0.405**	-0.210	0.068
	(0.166)	(0.188)	(0.186)	(0.166)
Daily work hours: 5-8 hours	0.097***	0.013	0.146***	-0.001
	(0.031)	(0.030)	(0.033)	(0.025)
Daily work hours: 9+ hours	0.044	0.004	0.009	0.050
	(0.051)	(0.069)	(0.052)	(0.058)
Weekly work days/10	0.946***	-0.077	-0.100	0.308
	(0.318)	(0.310)	(0.315)	(0.294)
Weekly work days squared /100	-1.071***	0.063	0.049	-0.355
	(0.387)	(0.378)	(0.370)	(0.369)
Night work	-0.025	-0.072*	-0.021	-0.039
	(0.037)	(0.041)	(0.037)	(0.033)
Presently attending school	-0.010	-0.051	-0.037	-0.050
	(0.043)	(0.033)	(0.035)	(0.032)
Ever left school	0.055*	0.004	0.012	-0.000
	(0.033)	(0.034)	(0.030)	(0.028)
Secondary or tertiary schooling	0.001	0.019	-0.055***	-0.020
	(0.031)	(0.027)	(0.021)	(0.022)
Location: Worked in own house	-0.133**	-0.183**	-0.483***	-0.073
	(0.061)	(0.073)	(0.086)	(0.063)
Location: Worked on farm	0.039	-0.093	0.022	0.151***
	(0.064)	(0.079)	(0.084)	(0.057)

Table 3.8 (Continued)

Worked in own household enterprise	0.066	0.089	0.125	0.071
	(0.065)	(0.065)	(0.081)	(0.048)
Paid for work	0.106	0.156**	0.041	0.035
	(0.081)	(0.075)	(0.104)	(0.075)
Received meals as worker benefit	-0.048	-0.226***	-0.133	-0.118*
	(0.071)	(0.086)	(0.084)	(0.066)
Worked for financial reasons	0.011	-0.039	-0.023	0.012
	(0.037)	(0.035)	(0.035)	(0.034)
Gave earnings to parents	0.010	0.077*	0.019	0.009
	(0.045)	(0.042)	(0.056)	(0.046)
Parent supervised work	-0.102	-0.072	-0.008	-0.147***
	(0.063)	(0.067)	(0.061)	(0.056)
Other relative supervised work	-0.068	-0.006	0.037	-0.128**
	(0.066)	(0.078)	(0.070)	(0.059)
Unsupervised	-0.069	-0.039	-0.096	-0.172***
	(0.060)	(0.076)	(0.070)	(0.051)
Interested in interview	-0.019	-0.021	-0.079**	0.022
	(0.027)	(0.027)	(0.031)	(0.024)
Sincere in interview	0.007	0.008	0.015	-0.023
	(0.027)	(0.022)	(0.020)	(0.018)
Injury report mismatch	*0.117****	*0.070**	*0.081***	*0.045*
	(0.037)	*(0.041)*	*(0.038)*	*(0.029)*
Intercept	0.290	0.835***	0.481*	0.236
	(0.239)	(0.259)	(0.260)	(0.211)
Observations	4600	4600	4600	4600
R-squared	0.86	0.90	0.93	0.92

Notes: Standard errors reported in parentheses. * denotes statistical significance at the 10%; ** at the 5% level; and *** at the 1% level. All estimates are adjusted for complex survey design features. Estimates for region dummies are not reported in table.

-4.2 percentage points (-16%) for physically strenuous work to a high of -29.2 percentage points (-57%) for exposure to physical hazards. There is also somewhat less robust evidence across the outcome measures and regression estimators that working more hours per day and working more days per week are associated with a higher likelihood of harmful child labor.

The regression-based results presented in Tables 3.7 and 3.8 also suggest that respondent parent and household covariates are generally not significantly associated with the probability of harmful child labor (that is, *conditional* on child labor). Interestingly, the evidence here directly contrasts with those from studies of the determinants of child labor which generally find that parental characteristics such as the level of formal education and household characteristics such as location (urban versus rural), household income and whether the household engages in a household enterprise significantly affect the probability of child labor (see, e.g., Dar et al. 2002 and Sedlacek et al. 2005 who report evidence from a range of countries in Latin America, Africa, and Asia, including the Philippines).

As an aside, the finding on the conditional association between household income and the probability of harmful child labor deserves some attention. An important result derived from modeling the harmful child labor decision in Chapter 1 is that the probability of harmful child labor is decreasing in parental income. Proxying parental income using reported household income, and estimating the unconditional relationship between the probability of harmful child labor and household income, I find that the probability of harmful child labor, as separately measured by physical strenuous work, hazardous work, and exposure to physical hazards, is significantly and negatively associated with household income; the only exception is exposure to chemical hazards.[34] Thus, the unconditional evidence

[34] The estimation results of this unconditional analysis are available from the author upon request.

supports my result from Chapter 1.[35] However, the effect of household income loses its significance once a range of child, parent, and other household characteristics are included in the regression, suggesting that much of its effect is indirect.

3.5.3. Regression-based impacts of parent-child injury report mismatches

I now turn to the regression-based estimates of the impact of parent-child injury report mismatches on the conditional probability of harmful child labor reported in Table 3.7. First, using the standard regression estimator (Eqn. 10), I find that parent-child injury report mismatches have a positive and highly significant impact on the probability of harmful child labor. This finding holds both across the self-reported measures of harmful child labor as well as the harmful child labor measures I constructed based on child reports of the presence of specific workplace hazards. In absolute terms, the magnitude of the estimated impacts range from a low of 4.9 percentage points in the case of exposure to chemical hazards to a high of 11.8 percentage points in the case of risky work. In relative terms (i.e., relative to the incidence of harmful child labor in the sample using these measures), the magnitude of these estimated impacts translate to 21% and 59%, respectively.

As noted earlier, the household fixed-effects regressions of the probability of harmful child labor (Eqn. 11) identify the effects of the covariates by exploiting the variation across working children within the household. Using this estimator, across the harmful child labor measures examined, I find that over 90% of the variation in harmful child labor is explained by this regression model, suggesting that *inter*-household differences (captured through the household fixed effects) explain much of

[35] I find this despite the fact that the strength of the relationship between household income and the probability of harmful child labor will be dampened in the presence of positive compensating wages for children working in harmful settings (which I find some evidence of in Chapter 2) given that these wages will also be included in household income.

the variation in harmful child labor outcomes between working children. In other words, differences in the household characteristics (both observed and unobserved) across working children matter a great deal in explaining whether the child works in a harmful setting or not. Notwithstanding, similar to the standard regression estimator results, with the exception of exposure to chemical hazards, the impact parameter estimates from the fixed-effects regression estimator are positive and statistically significant. In absolute terms, the magnitude of the estimates range from a low of 7 percentage points in the case of risky work to a high of 11.7 percentage points in the case of physically strenuous work. In relative terms, the magnitude of the estimates range from 16% in the case of exposure to physical hazards to a high of 43% in the case of physically strenuous work.

Implementation of the matching estimator

For the discussion of the matching-based results, to avoid any confusion, I revert back to referring to parent-child injury report mismatches as the treatment, and children with and without parent-child injury report mismatches as the treated and untreated groups, respectively. While Section 3.3 discussed the specific matching estimator applied in the study, before presenting its results, I provide some details on the practical implementation of the matching estimator here.

A three-step process is followed in order to arrive at the matching-based impact estimates. First, the model to estimate the probability of treatment (the propensity score model) is specified and estimated. Second, the common support restriction and various tests to assess the quality of the matching procedure (via whether the propensity score model covariates balance across the treated and untreated samples) are implemented. Third, for the matched sample, the impact of the treatment variable on harmful child labor is estimated via local linear matching. The first two steps are discussed in this subsection.

Specification of the propensity score model: In order to match treated children to untreated children, I estimate the conditional probability (or propensity score) of parent-child injury report mismatches by fitting a binomial logit to the data. These results of this estimation are presented in Table 3.9. In specifying the propensity score model, I include a rich set of covariates that potentially predict both the probability of treatment as well as the probability of harmful child labor—these covariates comprise of the full set of child, respondent parent, and household covariates included in the outcome regressions. As discussed by Heckman et al. (1997) and Heckman et al. (1998), including a rich set of covariates greatly increases the likelihood that matching generates valid impact estimates. Note the results of the propensity score estimation are not discussed here as the objective of this step is not statistical inference but rather generating the predicted propensity scores.

Figure 3.1 depicts the densities of the propensity scores for the treated and untreated samples generated from the estimated propensity score model. It is easy to see that the supports of both samples largely overlap, with most observations lying between a propensity score of 0.1 and 0.3. This finding implies that applying the common support restriction will only have a minimal effect in terms of the number of discarded observations. Thus, the resulting sample for the matching exercise is expected to largely mirror the original underlying sample of working children.

Implementation of the common support restriction: Given that the impact parameter is only defined over the region of common support, applying the min-max method discussed in Caliendo and Kopeinig (2008), I discard treated observations whose propensity scores are smaller than the smallest and larger than the largest propensity

Table 3.9. Propensity score model: Determinants of parent-child injury report mismatch
Pseudo-MLE logit regression estimates

Covariates	Estimated coefficients
Child variables	
Male	0.153
	(0.116)
Age/10	-0.372
	(2.341)
Age squared/100	0.259
	(0.897)
Age first worked/10	-3.313*
	(1.904)
Age first worked squared /100	1.323
	(0.810)
Daily work hours: 5-8 hours	-0.025
	(0.138)
Daily work hours: 9+ hours	-0.249
	(0.248)
Weekly work days/10	2.353
	(1.625)
Weekly work days squared /100	-2.398
	(1.925)
Night work	0.114
	(0.148)
Presently in work	-0.270
	(0.165)
Ever left school	-0.062
	(0.143)
Secondary or tertiary schooling	-0.004
	(0.137)
Location: Worked in own house	-0.327
	(0.226)
Location: Worked on farm	0.391**
	(0.172)
Worked in own household enterprise	-0.142
	(0.243)
Paid for work	-0.105
	(0.263)
Received meals as worker benefit	0.083
	(0.222)
Worked for financial reasons	0.259*
	(0.147)
Gave earnings to parents	0.125
	(0.178)
Parent supervised work	0.190
	(0.252)
Other relative supervised work	0.268
	(0.297)
Unsupervised	0.300
	(0.229)
Interested in interview	0.272
	(0.176)

Table 3.9 (Continued)

Sincere in interview	-0.162
	(0.133)
Respondent parent variables	
Male	-0.428
	(0.295)
Age/10	-0.413
	(0.629)
Age squared/100	0.046
	(0.066)
Secondary or tertiary education	0.228*
	(0.129)
Head of household	0.183
	(0.253)
Household variables	
Run household enterprise	-0.005
	(0.191)
Own farm land	0.211
	(0.129)
Household size/10	1.959
	(1.395)
Household size squared/100	-0.820
	(0.812)
Children/10	-2.914**
	(1.475)
Children squared/100	2.754**
	(1.371)
Household income: P2000-P2999	-0.523**
	(0.214)
Household income: P3000-4999	-0.460**
	(0.206)
Household income: P5000-9999	-0.561**
	(0.226)
Household income: P10000-14999	-0.578**
	(0.288)
Household income:>P15000	-0.250
	(0.346)
Urban	-0.174
	(0.142)
Intercept	-0.163
	(2.168)
LR χ-squared	238.89
p-value	0.000
McFadden's Pseudo R-squared	0.0670
N	4600

Notes: Robust standard errors reported in parentheses. * denotes statistical significance at the 10%; ** at the 5% level; and *** at the 1% level. All estimates are adjusted for complex survey design features. Estimates for region dummies are not reported in table.

score for the untreated sample. In principle, this is particularly important in the case of kernel-based matching estimators such as the local linear matching estimator, since all

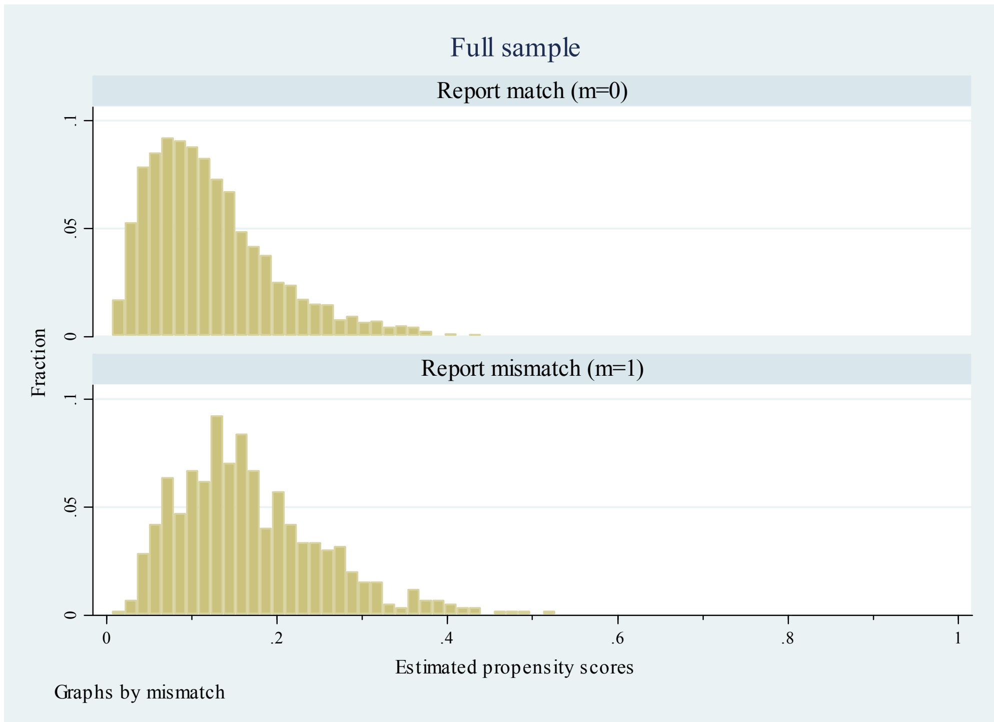

Figure 3.1. Distribution of predicted propensity scores by report mismatch status

available observations are used to estimate the missing counterfactual. However, in practice here, applying this procedure results in only three treated observations (less than 1% of the treated sample) being discarded. This negligible reduction in the sample suggests that any emerging differences in the impact estimates between the outcome regressions (Eqns. 10 and 11) and matching are not driven by the common support restriction. Rather, they will be driven by the weighting process in matching, in which the counterfactual for a given treated observations is formed by weighting more similar untreated observations higher (recall that in regression, the counterfactual is formed by weighting all untreated observations equally).

Implementation of alternative balancing tests: In order to assess the quality of the matching procedure, I implement three alternative common tests to see if the covariates across the treated and matched untreated samples have comparable distributions, or are "balanced" (see Caliendo and Kopeinig 2008 for a review of alternative tests in the literature). This is important since I match on the propensity score (a coarse summary measure of the predictive information contained within the covariates) and not on the values of the individual covariates themselves.

As a first test, I examine the means of included covariates for the matched treated and untreated samples (see Table 3.A3). I find that the difference in means for the two samples is not statistically different for any of the included covariates. As a second test, I examine the pseudo *R*-squared of the propensity score model before and after matching, using observations in the common support region only as well as weights generated from the matching algorithm (see Table 3.A4).[36] The measure provides information on how well the covariates predict the probability of treatment. I find that the pseudo *R*-squared is smaller after matching. I also find that the joint significance of the covariates of the model after matching is rejected; in contrast, the joint significance of the covariates before matching was not rejected. As a third and final test, I examine the standardized difference before and after matching (also see Table 3.A4). The standardized difference for a variable is the difference in means between the treated and matched untreated samples, as a percentage of the square root of the mean variance across the samples. As expected, I find that the standardized difference is smaller after matching and below a value of three, which is well within the acceptable limits per guidance available in the literature. Thus, to summarize, all

[36] The pseudo *R*-squared indicates how well the covariates predict the probability of treatment. If the covariates are balanced after matching, the pseudo *R*-squared should be low since there should be no systematic differences in the distribution of covariates between the matched treated and untreated groups.

three tests strongly suggest that the covariates are comparable or balance across the treated and untreated samples.

3.5.4 Matching-based impact estimates of parent-child injury report mismatches

Table 3.10 presents the local linear matching-based estimates of the impact of parent-child injury report mismatches on the probability of harmful child labor measured variously (see Eqn. 16 for the estimator). The estimator is parameterized by using the Epanechnikov kernel as the weighting function and a kernel bandwidth size of 0.01 (which was found to be optimal by minimizing the MSE).[37] Consistent with the regression-based findings presented earlier, the matching-based estimates suggest that parent-child injury report mismatches have a positive and statistically-significant impact on the probability of harmful child labor, regardless of the outcome measure examined. The magnitudes of the impacts are also generally comparable to those from the regressions. In absolute terms, the estimated impacts vary from a low of 5.9 percentage points in the case of exposure to chemical hazards to a high of 12.3 percentage points in the case of risky work. In relative terms, the estimated impacts vary from a low of 24% in the case of exposure to physical hazards to a high of 62% in the case of risky work. The magnitudes of the impacts are also generally comparable to those from the regressions. In absolute terms, the estimated impacts vary from a low of 5.9 percentage points in the case of exposure to chemical hazards to a high of 12.3 percentage points in the case of risky work. In relative terms, the estimated impacts vary from a low of 24% in the case of exposure to physical hazards to a high of 62% in the case of risky work.

[37] I examine the MSE for a range of bandwidth sizes between 0.01 and 0.21, in 0.05 increments. The bandwidth size used for the final analysis, 0.01, has the lowest MSE for the outcome measures in general.

Table 3.10: Impact of mismatch in injury reports on harmful child labor outcomes
Local linear regression propensity score matching estimates

	Physically strenuous work	Risky work	Exposure to physical hazards	Exposure to chemical hazards
ATT	0.094***	0.123***	0.124***	0.059***
S.E.	(0.022)	(0.021)	(0.022)	(0.020)

Notes: * Statistically significant at the 10% level; ** at the 5% level; and *** at the 1% level. Analytical standard errors in parentheses.

Sensitivity of impact estimates to unobserved heterogeneity: The assumption that treatment status can be predicted based on observables is central to the validity of the matching estimates. While this assumption is untestable, following common practice in the applied matching literature, I apply a bounding test proposed by Rosenbaum (2002) which basically evaluates how strongly an arbitrary unobserved factor must influence selection into treatment to alter inference on any significant impacts. Following Aakvik (2001), the bounding test I apply uses the Mantel and Haenszel (MH) test statistic Q_{MH} (see the appendix for a detailed discussion of the method). This test statistic can be bounded by two known distributions: Q_{MH}^{+} for overestimation of the impact (which would be the case if children who are more likely have a parent-child injury report mismatch are also *more* likely to engage in harmful child labor) and Q_{MH}^{-} for underestimation of the impact (which would be the case if children who are more likely to have a parent-child injury report mismatch are *less* likely to engage in harmful child labor). The parameter e^{γ} is a measure of the extent to which the analysis suffers from bias due to unobserved heterogeneity, which indicates the odds of differential assignment to treatment (i.e., the odds of having an injury report mismatch) due to unobserved factors. By examining the significance level of the Q_{MH} statistic as e^{γ} varies, the sensitivity of the matching estimates to unobserved heterogeneity can be evaluated.

Table 3.11 presents the results of such a sensitivity test, where I vary e^{γ} between 1 and 10 in one unit increments.[38] While $e^{\gamma} = 1$ represents no bias due to unobserved heterogeneity, higher levels of e^{γ} denote increasing bias due to unobservables. For example, if the results lose significance at $e^{\gamma} = 2,$ it implies that two units that are similar based on observables need differ only by a factor of two (or 100%) due to unobservables in order for inference to be altered. The selected range $e^{\gamma} \in [0,10]$ therefore represents a fairly large range of potential bias due to unobserved heterogeneity.

Table 3.11: Sensitivity of treatment effect estimates to unobserved heterogeneity

e^{γ}	Q^{+}_{MH}	p^{+}_{MH}	Q^{-}_{MH}	p^{-}_{MH}
1.0	67.692	0.000	67.692	0.000
2.0	52.535	0.000	89.819	0.000
3.0	45.894	0.000	107.234	0.000
4.0	41.914	0.000	122.126	0.000
5.0	39.167	0.000	135.363	0.000
6.0	37.111	0.000	147.403	0.000
7.0	35.489	0.000	158.524	0.000
8.0	34.162	0.000	168.911	0.000
9.0	33.046	0.000	178.692	0.000
10.0	32.089	0.000	187.963	0.000

Notes: e^{γ} : odds of differential assignment due to unobserved factors. Q^{+}_{MH} : Mantel-Haenszel statistic (assumption: overestimation of treatment effect). Q^{-}_{MH} : Mantel-Haenszel statistic (assumption: underestimation of treatment effect). p^{+}_{MH} : *p*-value (assumption: overestimation of treatment effect). p^{-}_{MH} : *p*-value (assumption: underestimation of treatment effect).

Within the range examined, I find that my results are not sensitive to unobserved heterogeneity. To reiterate, while the test results do not provide conclusive evidence that the probability of treatment is based entirely on observables, it offers a level of confidence that inference is robust to unobservable heterogeneity within the range examined.

[38] Following Becker and Caliendo (2007), I use the mhbounds procedure in Stata to implement this sensitivity test.

3.5.5. Heterogeneous impacts of parent-child injury report mismatches

Regression-based impact estimates: Given the interest in the intrahousehold modeling and empirical literature on how gender differences affect outcomes, Tables 3.12 and 3.13 present regression-based impact estimates for the standard and fixed-effects regression estimators (Eqns. 10 and 11), respectively, where the parent-child injury report mismatch variable is separately interacted with the child's gender and the respondent parent's gender. Across estimators and outcome measures, I find generally consistent evidence that the child's gender and parent-child injury report mismatches individually have significant positive impacts on the conditional probability of harmful child labor. However, I find little consistent evidence that the impact of

Table 3.12. Regression estimates of heterogeneous impacts of parent-child injury report mismatches

Linear probability model (LPM) regression results

Covariate	(1) Physically strenuous work	(2) Risky work	(3) Exposure to physical hazards	(4) Exposure to chemical hazards
Interaction with child's gender				
Male	0.154***	0.096***	0.051***	0.065***
	(0.016)	(0.016)	(0.020)	(0.016)
Injury report mismatch	0.065**	0.119***	0.182***	0.063*
	(0.030)	(0.045)	(0.051)	(0.038)
Male × Injury report mismatch	0	-0.002	-0.104**	-0.02
	(0.040)	(0.047)	(0.052)	(0.044)
Interaction with the respondent parent's gender				
Male	-0.041	-0.008	-0.042	0.004
	(0.043)	(0.039)	(0.040)	(0.047)
Injury report mismatch	0.064**	0.114***	0.109***	0.056**
	(0.027)	(0.031)	(0.031)	(0.026)
Male × Injury report mismatch	0.005	0.017	0.015	-0.034
	(-0.041)	(-0.008)	(-0.042)	0.004

Notes: Standard errors reported in parentheses. * denotes statistical significance at the 10%; ** at the 5% level; *** at the 1% level. All estimates are adjusted for complex survey design features. All regression includes the full set of child, respondent parent, and household covariates.

parent-child injury report mismatches on the probability of harmful child labor differs systematically by the child's gender. Likewise, I do not find evidence that the impact of parent-child injury report mismatches on the probability of harmful child labor differs systematically by the respondent parent's gender.

Table 3.13. Regression estimates of heterogeneous impacts of parent-child injury report mismatches

Linear probability model (LPM) regression results, Household fixed effects

Covariate	(1) Physically strenuous work	(2) Risky work	(3) Exposure to physical hazards	(4) Exposure to chemical hazards
Interaction with child's gender				
Male	0.086***	0.064***	-0.006	-0.006
	(0.021)	(0.018)	(0.019)	(0.016)
Injury report mismatch	0.110**	-0.004	0.096**	0.036
	(0.050)	(0.043)	(0.048)	(0.041)
Male × Injury report mismatch	0.013	0.117**	-0.025	-0.012
	(0.057)	(0.056)	(0.053)	(0.043)

Notes: Standard errors reported in parentheses. * denotes statistical significance at the 10%; ** at the 5% level; *** at the 1% level. All estimates are adjusted for complex survey design features. All regression includes the full set of child, respondent parent, and household covariates.

Matching-based impact results: In order to estimate the differential impacts of parent-child injury report mismatches by the child's and respondent parent's gender, I implement the matching analysis on the relevant subsamples, namely boys only, girls only, male parent respondents only, and female parent respondents only.

I follow the same implementation steps discussed earlier with respect to the full sample. I estimate the conditional probability of treatment as a function of the full set of relevant child, respondent parent, and household covariates by fitting a binomial logit regression to the data (see Table 3.14 for the results from the estimated propensity score models for the subsamples). The density plots of the propensity scores for the treated and untreated groups for each subsample show that the supports largely overlap (see Figures 3.2-3.5). Thus, the application of the common support restriction using the min-max method results in a negligible percentage of observations being discarded. Specifically, in each subsample, less than ten treated observations (less than 1% of the relevant treated subsample) were discarded. Finally, the results from testing differences in covariate means as well as differences in the pseudo R-squared and the standardized difference before and after matching strongly suggest that the covariates are balanced for the matched treated and untreated subsamples (see Tables 3.A5-3.A9).

Table 3.15 presents the impacts of parent-child injury report mismatches on the probability of harmful child labor measured variously in each of the subsamples. As before, the impacts in each subsample are estimated using a local linear matching estimator parameterized using the Epanechnikov kernel and a bandwidth size of 0.01. To allow statistical inference between subsamples, the 95% confidence interval are reported. Comparing the impact estimates for boys versus girls, I find that, with the exception of exposure to physical hazards where the estimated impact was higher for girls than for boys (23 versus 9 percentage points), the impact estimates do not appear to systematically differ between boys and girls. Comparing the impact estimates for children with injury reports provided by fathers versus mothers, I find that the impact estimates do not systematically differ between these children. Applying the Rosenbaum bounding test and varying e^{γ} between 1 and 10, I find that, across the subsamples, inference is not sensitive to unobservable heterogeneity within the range examined (see Table 3.16). Thus, in general, the matching-based findings on differential impacts by gender are consistent with the regression-based findings: the estimated impacts of parent-child injury report mismatches on the probability of

Table 3.14. Propensity score model for estimation of heterogeneous effects: Determinants of parent-child injury report mismatches
Pseudo-MLE binomial logit regression estimates

Variable	Coefficients			
	(1) Boys only	(2) Girls only	(3) Fathers only	(4) Mothers only
Child variables				
Male	--	--	0.182	0.173
			(0.271)	(0.135)
Age/10	0.944	-2.002	-1.486	0.136
	(3.143)	(4.526)	(6.703)	(2.618)
Age squared/100	-0.184	0.871	1.121	0.009
	(1.176)	(1.762)	(2.512)	(1.001)
Age first worked/10	-5.257***	0.076	-5.897	-3.310*
	(2.010)	(3.310)	(3.851)	(1.951)
Age first worked squared /100	2.243**	-0.485	2.726*	1.216
	(0.874)	(1.425)	(1.645)	(0.845)
Daily work hours: 5-8 hours	0.109	-0.174	0.406	-0.123
	(0.172)	(0.239)	(0.343)	(0.149)
Daily work hours: 9+ hours	-0.256	0.308	-0.213	-0.283
	(0.301)	(0.412)	(0.556)	(0.276)
Weekly work days/10	1.823	2.789	6.445*	1.268
	(1.914)	(2.778)	(3.612)	(1.928)
Weekly work days squared /100	-1.753	-3.328	-7.147*	-1.158
	(2.247)	(3.315)	(4.316)	(2.281)
Night work	0.005	0.309	0.131	0.128
	(0.193)	(0.253)	(0.323)	(0.172)
Presently in work	-0.297	-0.188	0.328	-0.411**
	(0.201)	(0.331)	(0.374)	(0.192)
Ever left school	-0.169	0.092	0.067	-0.079
	(0.170)	(0.265)	(0.322)	(0.157)
Secondary or tertiary schooling	-0.035	0.041	-0.051	0.023
	(0.166)	(0.273)	(0.305)	(0.158)
Location: Worked in own house	-0.589*	-0.319	0.048	-0.404*
	(0.332)	(0.324)	(0.479)	(0.244)
Location: Worked on farm	0.563***	-0.016	-0.009	0.499***
	(0.203)	(0.308)	(0.348)	(0.192)
Worked in own household enterprise	-0.155	0.183	-0.031	-0.082
	(0.282)	(0.363)	(0.391)	(0.297)
Paid for work	-0.036	-0.451	-0.218	0.005
	(0.301)	(0.452)	(0.556)	(0.308)
Received meals as worker benefit	0.066	0.252	0.248	0.058
	(0.261)	(0.375)	(0.546)	(0.265)
Worked for financial reasons	0.252	0.356	0.447	0.201
	(0.177)	(0.222)	(0.303)	(0.171)
Gave earnings to parents	0.224	-0.163	-0.376	0.267
	(0.211)	(0.379)	(0.458)	(0.198)
Parent supervised work	0.283	-0.164	0.371	0.220
	(0.268)	(0.461)	(0.516)	(0.285)
Other relative supervised work	0.331	-0.111	0.418	0.252
	(0.332)	(0.527)	(0.597)	(0.332)

Table 3.14 (Continued)

Unsupervised	0.317	0.004	0.241	0.313
	(0.264)	(0.424)	(0.480)	(0.249)
Interested in interview	0.402**	-0.014	0.302	0.251
	(0.197)	(0.280)	(0.356)	(0.185)
Sincere in interview	-0.169	-0.032	0.101	-0.219
	(0.157)	(0.207)	(0.295)	(0.138)
Respondent parent variables				
Male	-0.682**	0.078	--	--
	(0.313)	(0.503)		
Age/10	-1.419*	1.440	-0.016	-0.794
	(0.785)	(1.178)	(1.103)	(0.881)
Age squared/100	0.161*	-0.162	0.004	0.089
	(0.084)	(0.123)	(0.108)	(0.099)
Secondary or tertiary education	0.159	0.407*	-0.295	0.289*
	(0.156)	(0.217)	(0.319)	(0.148)
Head of household	0.371	-0.162	--	0.271
	(0.279)	(0.424)		(0.252)
Household variables				
Run household enterprise	0.212	-0.784**	-0.107	0.034
	(0.208)	(0.381)	(0.425)	(0.230)
Own farm land	0.204	0.171	-0.199	0.259*
	(0.159)	(0.195)	(0.261)	(0.150)
Household size/10	0.774	2.194	-2.600	2.938*
	(1.827)	(2.487)	(3.164)	(1.670)
Household size squared/100	-0.529	-0.079	1.326	-1.267
	(1.069)	(1.512)	(1.824)	(1.018)
Children/10	-1.714	-4.737*	5.526	-4.687***
	(1.911)	(2.501)	(3.830)	(1.649)
Children squared/100	2.406	3.077	-4.254	4.114***
	(1.835)	(2.349)	(3.665)	(1.533)
Household income: P2000-P2999	-0.197	-1.261***	-0.632	-0.449*
	(0.263)	(0.373)	(0.422)	(0.243)
Household income: P3000-4999	-0.163	-0.929***	-1.155***	-0.251
	(0.260)	(0.301)	(0.392)	(0.240)
Household income: P5000-9999	-0.290	-1.186***	-1.227***	-0.362
	(0.276)	(0.341)	(0.384)	(0.263)
Household income: P10000-14999	-0.205	-1.415***	-1.569***	-0.329
	(0.354)	(0.456)	(0.525)	(0.335)
Household income:>P15000	0.271	-1.264***	-0.162	-0.175
	(0.434)	(0.460)	(0.617)	(0.407)
Urban	-0.209	-0.140	0.177	-0.235
	(0.168)	(0.227)	(0.312)	(0.163)
Intercept	1.505	-2.737	-0.515	0.269
	(2.684)	(3.749)	(5.541)	(2.376)
LR χ^2	182.35	141.09	113.85	204.80
p-value	0.000	0.000	0.000	0.000
McFadden's Pseudo-R^2	0.0775	0.118	0.146	0.0737
N	2898	1702	1093	3506

Notes: Robust standard errors reported in parentheses. * denotes statistical significance at the 10%; ** at the 5% level; *** at the 1% level. All estimates are adjusted for complex survey design features. Estimates for region dummies are not reported in table.

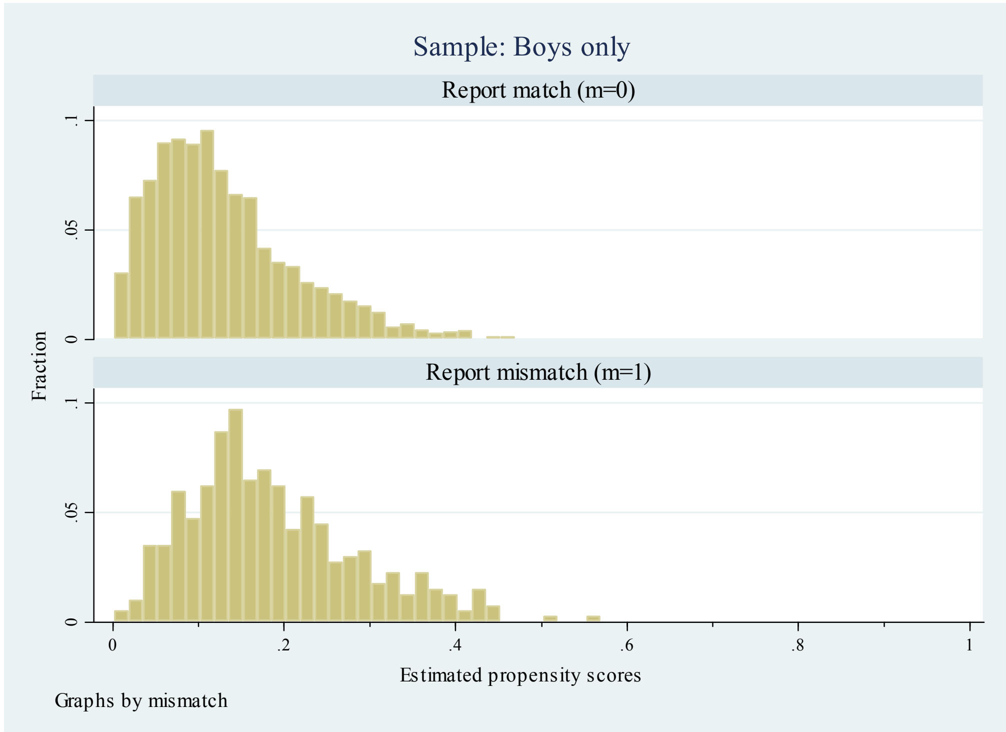

Figure 3.2. Distribution of predicted propensity scores by report mismatch status

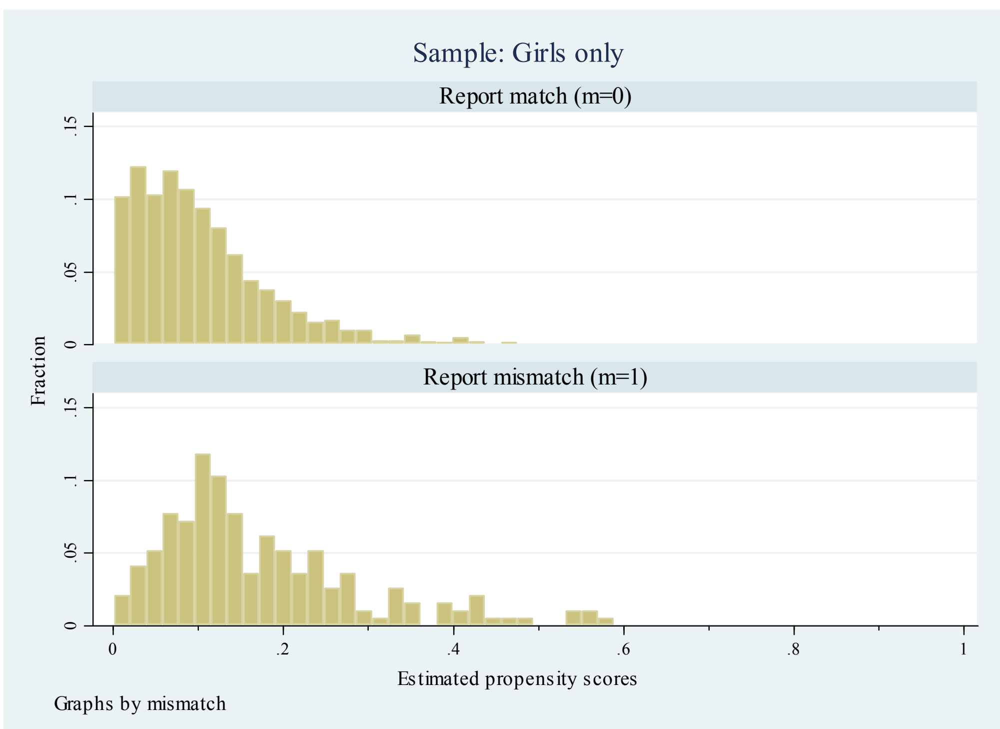

Figure 3.3. Distribution of predicted propensity scores by report mismatch status

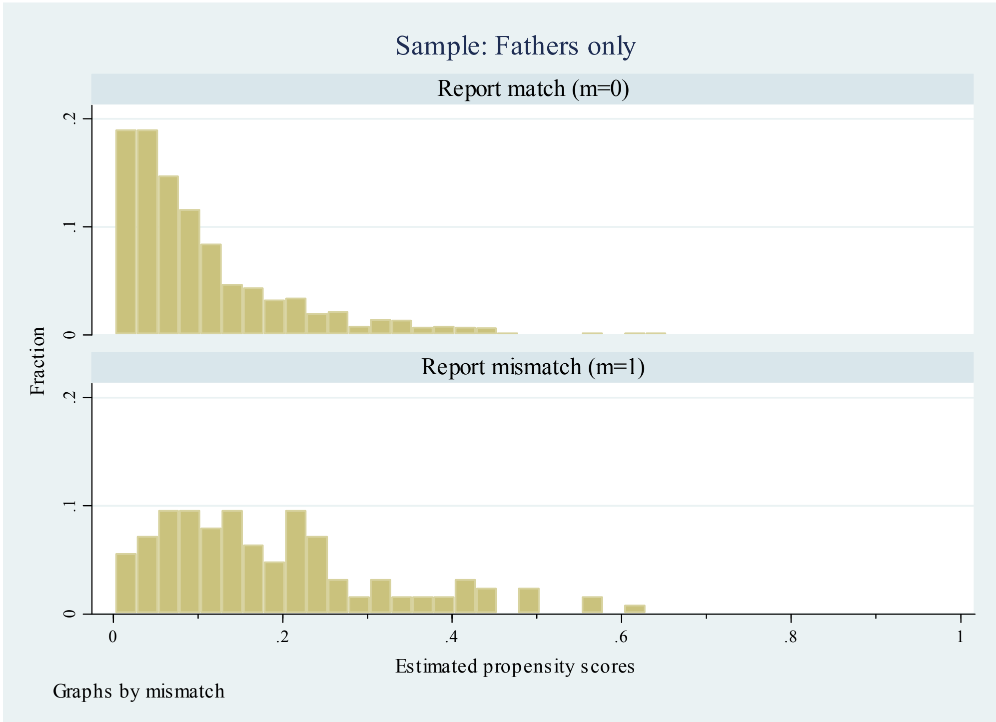

Figure 3.4. Distribution of predicted propensity scores by report mismatch status

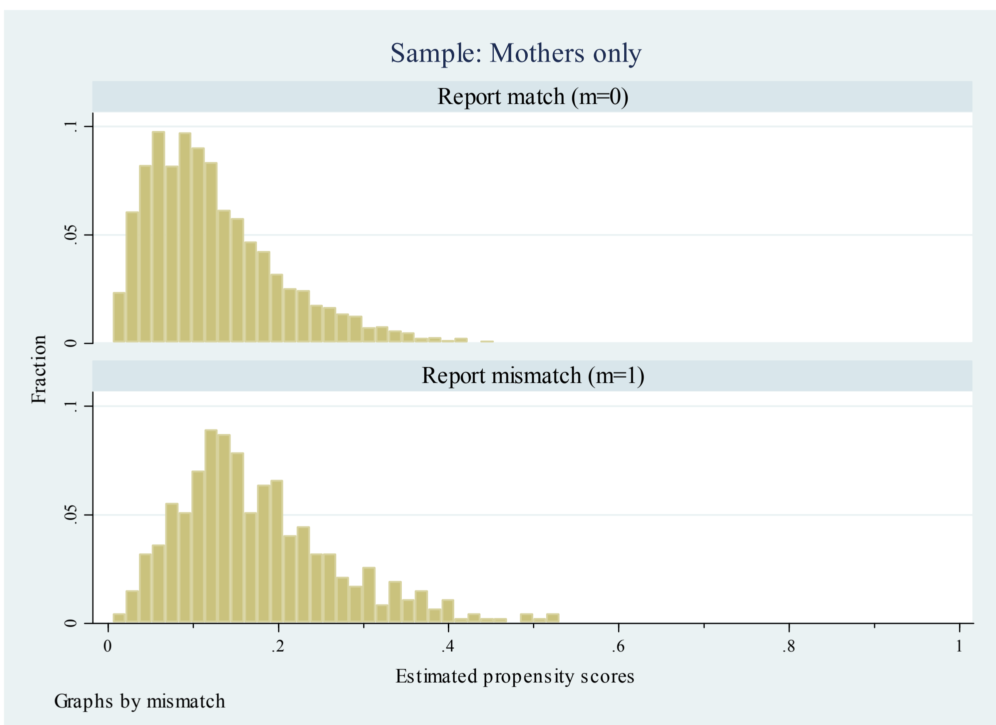

Figure 3.5. Distribution of predicted propensity scores by report mismatch status

Table 3.15: Heterogeneous impacts of injury report mismatches on harmful child labor

Local linear regression propensity score matching estimates

Sample	Physically strenuous work	Risky work	Exposure to physical hazards	Exposure to chemical hazards
Boys only				
ATT	0.100***	0.114***	0.087***	0.036
S.E.	(0.028)	(0.026)	(0.027)	(0.027)
95% C.I.	[0.045,0.155]	[0.063,0.165]	[0.034,0.140]	[-0.017,0.089]
Girls only				
ATT	0.075**	0.136***	0.229***	0.097***
S.E.	(0.033)	(0.034)	(0.042)	(0.035)
95% C.I.	[0.010,0.140]	[0.069,0.203]	[0.147,0.311]	[0.028,0.166]
Fathers only				
ATT	0.094*	0.124***	0.160***	0.080*
S.E.	(0.048)	(0.046)	(0.049)	(0.044)
95% C.I.	[0.000,0.188]	[0.034,0.214]	[0.064,0.256]	[-0.006,0.166]
Mothers only				
ATT	0.088***	0.121***	0.125***	0.056**
S.E.	(0.025)	(0.023)	(0.025)	(0.023)
95% C.I.	[0.039,0.137]	[0.076,0.166]	[0.076,0.174]	[0.011,0.101]

Notes: * Statistically significant at the 10% level; ** at the 5% level; and *** at the 1% level. Analytical standard errors in parentheses.

Table 3.16: Sensitivity of treatment effect estimates to unobserved heterogeneity

	e^{γ}	Q^{+}_{MH}	p^{+}_{MH}	Q^{-}_{MH}	p^{-}_{MH}
Sample: boys only					
	1.0	53.588	0.000	53.588	0.000
	2.0	41.790	0.000	70.787	0.000
	3.0	36.596	0.000	84.327	0.000
	4.0	33.472	0.000	95.912	0.000
	5.0	31.309	0.000	106.217	0.000
	6.0	29.686	0.000	115.595	0.000
	7.0	28.404	0.000	124.260	0.000
	8.0	27.353	0.000	132.355	0.000
	9.0	26.468	0.000	139.980	0.000
	10.0	25.708	0.000	147.209	0.000
Sample: girls only					
	1.0	40.888	0.000	40.888	0.000
	2.0	31.412	0.000	54.770	0.000
	3.0	27.286	0.000	65.671	0.000
	4.0	24.830	0.000	74.976	0.000
	5.0	23.145	0.000	83.237	0.000
	6.0	21.889	0.000	90.744	0.000
	7.0	20.902	0.000	97.672	0.000
	8.0	20.098	0.000	104.139	0.000
	9.0	19.423	0.000	110.226	0.000
	10.0	18.846	0.000	115.993	0.000
Sample: Fathers only					
	1.0	32.650	0.000	32.650	0.000
	2.0	25.109	0.000	43.707	0.000
	3.0	21.820	0.000	52.387	0.000
	4.0	19.861	0.000	59.797	0.000
	5.0	18.516	0.000	66.376	0.000
	6.0	17.512	0.000	72.355	0.000
	7.0	16.724	0.000	77.874	0.000
	8.0	16.081	0.000	83.025	0.000
	9.0	15.541	0.000	87.874	0.000
	10.0	15.079	0.000	92.469	0.000
Sample: Mothers only					
	1.0	59.020	0.000	59.020	0.000
	2.0	45.901	0.000	78.163	0.000
	3.0	40.142	0.000	93.231	0.000
	4.0	36.684	0.000	106.119	0.000
	5.0	34.294	0.000	117.579	0.000
	6.0	32.504	0.000	128.005	0.000
	7.0	31.090	0.000	137.636	0.000
	8.0	29.933	0.000	146.632	0.000
	9.0	28.960	0.000	155.105	0.000
	10.0	28.124	0.000	163.136	0.000

Notes: See Table 3.11 for definitions of e^{γ}, Q^{+}_{MH}, Q^{-}_{MH}, p^{+}_{MH}, p^{-}_{MH}.

harmful child labor do not appear to differ by the child's or respondent parent's gender.

3.6 Summary and concluding remarks

Recent ILO estimates show that child labor in potentially harmful settings constitutes a widespread and significant problem across the developing world. In addition, there is growing evidence that such child labor has adverse short- and long-term health effects which are undesirable in themselves leave alone other negative effects of child labor in general such as reduced educational attainment, cognitive achievement, and future labor market earnings, as well as the intergenerational persistence in child labor. Despite this, few studies attempt to carefully distinguish between different forms of child labor and focus narrowly on the causes of and attributes that typify harmful child labor.

This chapter contributes to the presently limited literature on the specific determinants of the harmful child labor decision, albeit from an unconventional angle. Using household sample survey data from the Philippines, the chapter examines whether parent-child injury report mismatches have an impact on the probability of harmful child labor. Given the interest in the intrahousehold literature of investigating potential gender differences, the chapter also examines whether the impacts systematically vary between working boys and girls as well as by whether the respondent parent is the father or the mother. The above questions are based on a decidedly intrahousehold view of the interactions between the parent and the child, where the harmful child labor decision is framed as the outcome of a potentially adversarial relationship between the parent and the child arising from different perceptions on and preferences over harmful child labor along with asymmetries in the decisionmaking power between the two agents.

Assuming that differences in children with parent-child injury report mismatches and children without are largely explained by selection on observables, I use both multiple regression (with and without household fixed effects) and propensity score matching methods to infer the impact of parent-child injury report mismatches on the conditional probability of harmful child labor. Harmful child labor is measured using variables that capture different types of harm such as an immediate condition (physical strain) and a future stochastic condition (hazardous work). Harmful child labor is also measured using both self-reported and constructed measures to assess whether the findings differ across subjective and objective measures of potential harm.

Across estimators and the harmful child labor measures examined, I find consistent and statistically-significant evidence that parent-child injury report mismatches increase the probability of harmful child labor. I also find that in general the impacts do not systematically vary by the child's gender or the respondent parent's. With respect to the matching estimation, statistical tests based on Rosenbaum bounds suggest that the impact estimates are not sensitive to potential unobserved heterogeneity in explaining the presence of parent-child injury report mismatches, indicating the robustness of the results to potential "hidden bias".

If these findings indeed reflect the outcome of differences in preferences over harmful child labor between the parent and the child, where the parent wields higher bargaining/decisionmaking power than the child, then a policy solution such as effectively banning harmful child labor can potentially lead to a welfare gain for the child, though at the expense of the parent's welfare. In general, the net welfare result for the household is hence ambiguous. However, in the case where harmful child labor is a result of information asymmetries within the household (specifically, the parent's information set is poorer), if banning harmful child labor "reveals" information that leads to the parent's perception and valuation of harmful child labor converging to those of the child, under certain formulations, the intervention can simultaneously yield welfare gains for the child as well as the parent. Thus, banning harmful child labor functions as a signal of the nature of harm in the child labor market that is accurately read by parents.

Notwithstanding a ban on harmful child labor, policymakers can also raise the welfare of both children and parents by expanding the set of decision options to include other potentially desirable ones or increasing the relative attractiveness of existing options. For example, this could entail providing affordable, higher-quality schooling, potentially increasing the opportunity cost of harmful child labor, and, thus, leading to household members to optimally choosing more schooling and less harmful child labor. It could also entail simply improving the occupational health and safety levels of child labor. An intrahousehold view of household decisionmaking would only imply that the welfare gains across household members will vary to the extent that preferences and power vary between the members.

APPENDIX

SENSITIVITY TO UNOBSERVED HETEROGENEITY: USING THE ROSENBAUM BOUNDING APPROACH

In the matching-based program evaluation literature, the assumption that selection into participation can be modeled using observable characteristics is central. In order to ascertain the validity of matching estimates, it is therefore important to examine their sensitivity to any deviation from this identifying assumption. In this context, sensitivity analysis using a bounding approach suggested by Rosenbaum (2002) is increasingly used in matching applications. The objective of this analysis is to examine the extent to which unobserved variables or "hidden bias" may alter inference about treatment effects.

Following the exposition in Caliendo and Kopeinig (2008), and using the language of program evaluation, I denote the probability of participation P_i as follows:

$$\mathrm{P}(\mathrm{m}_i = 1 \mid x_i) = F(\beta x_i + \gamma u_i),$$

where x_i are the observed characteristics for individual i, u_i are unobserved characteristics, and β and γ are the impacts of x_i and u_i on the participation decision. If no unobservable characteristics affect participation, i.e., $\gamma = 0,$ then two individuals with the same set of observable characteristics have the same probability of participation. However, if $\gamma \neq 0,$ i.e., there are unobservable characteristics that affect participation, then two individuals with the same x have differing probabilities of participation. If F is the logistic distribution, the odds that two individuals i and j participate are given by $\mathrm{P}_i / (1 - \mathrm{P}_i)$ and $\mathrm{P}_j / (1 - \mathrm{P}_j)$ respectively. Then, the odds ratio can be written as

$$\frac{P_i/(1-P_i)}{P_j/(1-P_j)} = P_i(1-P_j)/P_j(1-P_i) = \frac{\exp(\beta x_i + \gamma u_i)}{\exp(\beta x_j + \gamma u_j)}.$$

If i and j form a matched pair, then the x vector cancels out and the odds ratio can be simply

written as $\exp[\gamma(u_i - u_j)]$. If there are no differences in unobservables, i.e. $u_i = u_j$, or the unobservable factors do not affect the probability of participating, i.e. $\gamma = 0$, the odds ratio equals 1, implying that the matching estimates do not suffer from unobserved selection bias. However, if this is not the case, then the matching estimates are said to suffer from a "hidden bias". In this context, Rosenbaum (2002) shows that the following bounds can be placed on the odds ratio:

$$\frac{1}{e^\gamma} \le \frac{P_i/(1-P_i)}{P_j/(1-P_j)} \le e^\gamma.$$

As e^γ increases, the bounds move apart reflecting uncertainty due to the presence of unobserved selection bias. Thus, e^γ is a measure of the extent to which the analysis suffers from this bias.

For binary outcomes, Aakvik (2001) suggests using a test based on the Mantel and Haenszel statistic. As discussed by Aakvik (2001), the treatment effect on outcome y (which in this case is the impact on the harmful child labor decision) is said to be significant if it crosses some test statistic $t(m, y)$, where m is a dummy variable denoting program participation (which in this case is response mismatch). Let n_1 and n_0 be the number of households with and without response mismatched reports, where $n = n_1 + n_0$. Let y_1 and y_0 denote households with and without response mismatches, where $y_t = y_1 + y_0$. The test statistic Q_{MH}, which asymptotically follows the normal distribution, is given by

$$Q_{MH} = \frac{|y_1 - E(y_1)| - 0.5}{\sqrt{Var(y_1)}} = \frac{\left|y_1 - \frac{n_1 y_t}{n}\right| - 0.5}{\sqrt{\frac{n_1 n_0 y_t (n - y_t)}{n^2 (n-1)}}}.$$

As shown by Rosenbaum (2002), Q_{MH} can be bounded by two known distributions. Let Q_{MH}^{+} denote the test statistic in the case of overestimation of the treatment effect (which in this case would occur if households that are more likely to have a parent-child response mismatch are also households that are more likely to have a harmful child labor outcome) and Q_{MH}^{-} in the case of underestimation of the treatment effect (which in this case would occur if households that are more likely to have a parent-child response mismatch are households that are less likely to have a harmful child labor outcome). The bounds for overestimation and underestimation of the true treatment effect are given by:

$$Q_{MH}^{+} = \frac{\left|y_1 - \tilde{E}^{+}\right| - 0.5}{\sqrt{Var\left(\tilde{E}^{+}\right)}}$$

and

$$Q_{MH}^{-} = \frac{\left|y_1 - \tilde{E}^{-}\right| - 0.5}{\sqrt{Var\left(\tilde{E}^{-}\right)}},$$

where $\tilde{E}$ and $Var\left(\tilde{E}\right)$ are the estimated large sample expectation and variance of treated units with a successful outcome, for given values of γ.

$e^{\gamma} = 1$ implies no unobservable bias, while higher levels of e^{γ} denote increasing bias due to unobservable characteristics. The *p*-values indicate the level of significance as e^{γ} is varied. For example, if the results lose significance at $e^{\gamma} = 2$, it implies that two units that are similar based on observables need differ by a factor of 2 (or 100%) due to unobservables for inference to be altered.

Table 3.A1. Binomial logit estimates of the determinants of harmful child labor
Estimated coefficients

Covariates	Estimated Coefficients			
	(1) Physically strenuous work	(2) Risky work	(3) Exposure to physical hazards	(4) Exposure to chemical hazards
Child variables				
Male	1.095***	0.720***	0.182**	0.439***
	(0.121)	(0.134)	(0.092)	(0.115)
Age/10	-2.385	-4.354*	0.324	-0.505
	(2.321)	(2.275)	(1.799)	(2.337)
Age squared/100	1.364	1.755**	-0.021	0.315
	(0.850)	(0.858)	(0.674)	(0.878)
Age first worked/10	3.240*	-0.693	-4.816***	-0.143
	(1.834)	(1.950)	(1.622)	(1.896)
Age first worked squared /100	-1.543**	0.244	2.110***	0.255
	(0.759)	(0.816)	(0.685)	(0.795)
Daily work hours: 5-8 hours	0.523***	0.201	0.376***	0.334**
	(0.125)	(0.138)	(0.122)	(0.135)
Daily work hours: 9+ hours	0.607***	0.468**	0.131	0.065
	(0.194)	(0.195)	(0.185)	(0.203)
Weekly work days/10	6.325***	4.259**	1.547	1.907
	(1.375)	(1.673)	(1.267)	(1.520)
Weekly work days squared /100	-7.179***	-5.459***	-3.062**	-2.606
	(1.622)	(1.930)	(1.493)	(1.749)
Night work	0.255*	0.238*	0.067	-0.017
	(0.133)	(0.143)	(0.137)	(0.140)
Presently attending school	-0.046	-0.097	0.065	0.162
	(0.149)	(0.157)	(0.137)	(0.160)
Ever left school	0.241*	-0.044	0.098	0
	(0.126)	(0.129)	(0.114)	(0.135)
Secondary or tertiary schooling	0.096	-0.068	-0.012	-0.06
	(0.120)	(0.126)	(0.104)	(0.120)
Location: Worked in own house	-0.628***	-1.491***	-1.501***	-1.080***
	(0.198)	(0.240)	(0.170)	(0.185)
Location: Worked on farm	0.251*	-0.334**	0.438***	-0.117
	(0.143)	(0.157)	(0.145)	(0.154)

Table 3.A1 (Continued)

Worked in own household enterprise	0.25	0.123	0.360*	0.642***
	(0.217)	(0.228)	(0.201)	(0.233)
Paid for work	0.202	0.318	0.115	0.502**
	(0.218)	(0.239)	(0.209)	(0.235)
Received meals as worker benefit	-0.061	-0.668***	-0.370**	-0.082
	(0.182)	(0.227)	(0.179)	(0.207)
Worked for financial reasons	0.206*	0.073	0.087	0.082
	(0.119)	(0.125)	(0.118)	(0.134)
Gave earnings to parents	0.172	0.476***	0.205	0.339*
	(0.164)	(0.161)	(0.150)	(0.175)
Parent supervised work	-0.226	0.252	-0.002	0.269
	(0.180)	(0.191)	(0.182)	(0.182)
Other relative supervised work	-0.267	0.322	0.297	0.152
	(0.246)	(0.251)	(0.246)	(0.252)
Unsupervised	-0.074	0.269	0.02	0.169
	(0.168)	(0.182)	(0.168)	(0.167)
Interested in interview	0.192	-0.209	-0.248*	-0.077
	(0.134)	(0.143)	(0.130)	(0.134)
Sincere in interview	-0.248**	0.034	0.081	-0.091
	(0.113)	(0.120)	(0.103)	(0.108)
Injury report mismatch	*0.354****	*0.700****	*0.546****	*0.281***
	(0.130)	*(0.159)*	*(0.143)*	*(0.137)*
Respondent parent variables				
Male	-0.279	-0.044	-0.19	-0.019
	(0.245)	(0.259)	(0.194)	(0.290)
Age/10	0.461	0.009	-0.162	0.381
	(0.556)	(0.664)	(0.495)	(0.625)
Age squared/100	-0.05	0.002	0.012	-0.069
	(0.060)	(0.073)	(0.053)	(0.068)
Secondary or tertiary education	-0.253**	-0.148	-0.067	0.07
	(0.107)	(0.110)	(0.101)	(0.110)
Head of household	0.212	0.051	0.128	0.187
	(0.227)	(0.244)	(0.177)	(0.258)

Table 3.A1 (Continued)

Household variables				
Run household enterprise	-0.437***	0.16	0.206	-0.044
	(0.153)	(0.143)	(0.145)	(0.166)
Own farm land	0.116	-0.065	-0.271**	0.194
	(0.117)	(0.139)	(0.125)	(0.122)
Household size/10	-1.632	-2.553**	-0.175	-0.95
	(1.163)	(1.215)	(1.483)	(1.330)
Household size squared/100	1.054	1.477**	0.466	1.517*
	(0.674)	(0.694)	(0.986)	(0.802)
Children/10	2.091	2.334*	0.058	-1.475
	(1.356)	(1.388)	(1.214)	(1.316)
Children squared/100	-1.963	-2.279*	-0.389	-0.623
	(1.420)	(1.350)	(1.302)	(1.293)
Household income: P2000-P2999	-0.281	-0.18	0.039	0.101
	(0.192)	(0.215)	(0.197)	(0.229)
Household income: P3000-4999	-0.336	-0.282	-0.28	0.031
	(0.205)	(0.235)	(0.213)	(0.245)
Household income: P5000-9999	-0.277	-0.194	-0.226	0.013
	(0.218)	(0.254)	(0.232)	(0.254)
Household income: P10000-14999	-0.334	-0.258	-0.168	0.1
	(0.261)	(0.285)	(0.258)	(0.284)
Household income:>P15000	-0.372	-0.559	0.315	-0.266
	(0.312)	(0.371)	(0.297)	(0.338)
Urban	-0.19	-0.068	0.164	0.615***
	(0.133)	(0.145)	(0.126)	(0.134)
Intercept	-5.293***	0.746	2.44	-2.864
	(1.906)	(1.997)	(1.671)	(1.941)
Observations	4600	4600	4600	4600
Log likelihood, intercept only model	-2686	-2370	-3187	-2373
Log likelihood, full model	-2166	-2046	-2731	-2139
Pseudo R-squared	0.194	0.137	0.143	0.0986

Notes: Standard errors reported in parentheses. * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. Estimates are adjusted for complex survey design features. Estimates for region dummies are not reported in table.

Table 3.A2. Binomial logit estimates of the determinants of harmful child labor
Average marginal effects

Covariates	Average marginal effects			
	(1) Physically strenuous work	(2) Risky work	(3) Exposure to physical hazards	(4) Exposure to chemical hazards
Child variables				
Male	0.162***	0.098***	0.038**	0.064***
	(0.021)	(0.020)	(0.019)	(0.018)
Age/10	-0.368	-0.621*	0.066	-0.076
	(0.352)	(0.321)	(0.368)	(0.351)
Age squared/100	0.21	0.251**	-0.004	0.047
	(0.129)	(0.122)	(0.138)	(0.132)
Age first worked/10	0.499*	-0.099	-0.972***	-0.021
	(0.282)	(0.279)	(0.315)	(0.284)
Age first worked squared /100	-0.238**	0.035	0.431***	0.038
	(0.117)	(0.117)	(0.138)	(0.119)
Daily work hours: 5-8 hours	0.084***	0.029	0.077***	0.051**
	(0.022)	(0.021)	(0.025)	(0.022)
Daily work hours: 9+ hours	0.100***	0.073**	0.027	0.01
	(0.034)	(0.033)	(0.037)	(0.031)
Weekly work days/10	0.966***	0.607**	0.316	0.286
	(0.206)	(0.239)	(0.258)	(0.227)
Weekly work days squared /100	-1.108***	-0.780***	-0.625**	-0.39
	(0.250)	(0.276)	(0.304)	(0.261)
Night work	0.040*	0.035	0.014	-0.003
	(0.022)	(0.022)	(0.028)	(0.021)
Presently attending school	-0.007	-0.014	0.013	0.024
	(0.023)	(0.022)	(0.028)	(0.024)
Ever left school	0.038*	-0.006	0.02	0
	(0.021)	(0.018)	(0.023)	(0.020)
Secondary or tertiary schooling	0.015	-0.01	-0.002	-0.009
	(0.019)	(0.018)	(0.021)	(0.018)
Location: Worked in own house	-0.091***	-0.163***	-0.314***	-0.134***
	(0.025)	(0.015)	(0.029)	(0.016)
Location: Worked on farm	0.039*	-0.048**	0.092***	-0.018
	(0.023)	(0.021)	(0.030)	(0.022)

Table 3.A2 (Continued)

Worked in own household enterprise	0.038	0.017	0.073*	0.092**
	(0.034)	(0.033)	(0.040)	(0.038)
Paid for work	0.032	0.047	0.024	0.078*
	(0.035)	(0.038)	(0.043)	(0.041)
Received meals as worker benefit	-0.009	-0.084***	-0.076**	-0.012
	(0.028)	(0.024)	(0.037)	(0.030)
Worked for financial reasons	0.032*	0.011	0.018	0.012
	(0.019)	(0.018)	(0.024)	(0.021)
Gave earnings to parents	0.027	0.071***	0.042	0.052*
	(0.026)	(0.026)	(0.030)	(0.029)
Parent supervised work	-0.035	0.036	0	0.04
	(0.027)	(0.029)	(0.037)	(0.029)
Other relative supervised work	-0.04	0.049	0.06	0.023
	(0.035)	(0.041)	(0.049)	(0.040)
Unsupervised	-0.011	0.04	0.004	0.026
	(0.025)	(0.028)	(0.034)	(0.026)
Interested in interview	0.029	-0.031	-0.050*	-0.012
	(0.021)	(0.020)	(0.027)	(0.020)
Sincere in interview	-0.039**	0.005	0.017	-0.014
	(0.017)	(0.017)	(0.021)	(0.016)
Injury report mismatch	*0.057****	*0.111****	*0.110****	*0.044**
	(0.022)	*(0.028)*	*(0.028)*	*(0.023)*
Respondent parent variables				
Male	-0.042	-0.006	-0.039	-0.003
	(0.035)	(0.037)	(0.040)	(0.043)
Age/10	0.071	0.001	-0.033	0.057
	(0.087)	(0.095)	(0.101)	(0.093)
Age squared/100	-0.008	0	0.002	-0.01
	(0.009)	(0.010)	(0.011)	(0.010)
Secondary or tertiary education	-0.039**	-0.021	-0.014	0.011
	(0.016)	(0.015)	(0.021)	(0.017)
Head of household	0.033	0.007	0.026	0.028
	(0.037)	(0.036)	(0.036)	(0.041)

Table 3.A2 (Continued)

Household variables				
Run household enterprise	-0.071***	0.022	0.042	-0.007
	(0.023)	(0.021)	(0.030)	(0.025)
Own farm land	0.018	-0.009	-0.055**	0.029
	(0.018)	(0.020)	(0.026)	(0.019)
Household size/10	-0.252	-0.365**	-0.036	-0.142
	(0.179)	(0.173)	(0.303)	(0.200)
Household size squared/100	0.163	0.211**	0.095	0.227*
	(0.104)	(0.100)	(0.201)	(0.121)
Children/10	0.322	0.333*	0.012	-0.221
	(0.209)	(0.198)	(0.248)	(0.197)
Children squared/100	-0.303	-0.326*	-0.08	-0.093
	(0.220)	(0.193)	(0.266)	(0.194)
Household income: P2000-P2999	-0.042	-0.025	0.008	0.015
	(0.028)	(0.029)	(0.040)	(0.036)
Household income: P3000-4999	-0.051*	-0.039	-0.057	0.005
	(0.029)	(0.031)	(0.044)	(0.037)
Household income: P5000-9999	-0.042	-0.027	-0.046	0.002
	(0.032)	(0.034)	(0.048)	(0.038)
Household income: P10000-14999	-0.049	-0.035	-0.035	0.015
	(0.037)	(0.037)	(0.053)	(0.044)
Household income:>P15000	-0.055	-0.071*	0.063	-0.038
	(0.043)	(0.041)	(0.059)	(0.045)
Urban	-0.029	-0.01	0.033	0.098***
	(0.020)	(0.020)	(0.025)	(0.024)
Intercept				
Observations	4600	4600	4600	4600

Notes: Standard errors reported in parentheses. * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level. Estimates are adjusted for complex survey design features. Estimates for region dummies are not reported in table.

Table 3.A3: Balancing *t*-tests for propensity score model covariates
Matched treated and untreated groups, based on kernel propensity score matching

Variables	Means ($m=1$) (1)	Means ($m=0$) (2)	Difference in means (1)-(2)	*t*-statistic
Child variables				
Male	0.672	0.682	-0.010	0.721
Age/10	1.396	1.396	0.000	0.999
Age squared/100	2.007	2.008	-0.001	0.980
Age first worked/10	1.124	1.121	0.003	0.835
Age first worked squared /100	1.339	1.333	0.007	0.861
Daily work hours: 5-8 hours	0.439	0.435	0.003	0.912
Daily work hours: 9+ hours	0.077	0.069	0.008	0.579
Weekly work days/10	0.340	0.352	-0.013	0.300
Weekly work days squared /100	0.159	0.170	-0.011	0.272
Night work	0.188	0.200	-0.011	0.622
Presently in work	0.657	0.646	0.011	0.688
Ever left school	0.378	0.374	0.004	0.885
Secondary or tertiary schooling	0.400	0.394	0.006	0.831
Location: Worked in own house	0.113	0.118	-0.005	0.771
Location: Worked on farm	0.593	0.571	0.023	0.428
Worked in own household enterprise	0.624	0.602	0.021	0.448
Paid for work	0.331	0.348	-0.017	0.538
Received meals as worker benefit	0.084	0.076	0.009	0.589
Worked for financial reasons	0.461	0.446	0.015	0.604
Gave earnings to parents	0.355	0.375	-0.020	0.466
Parent supervised work	0.524	0.522	0.002	0.941
Other relative supervised work	0.049	0.052	-0.003	0.803
Unsupervised	0.297	0.284	0.014	0.605
Interested in interview	0.783	0.801	-0.018	0.454
Sincere in interview	0.551	0.564	-0.013	0.652
Respondent parent variables				
Male	0.212	0.210	0.002	0.929
Age/10	4.302	4.282	0.020	0.651
Age squared/100	19.075	18.922	0.153	0.701
Secondary or tertiary education	0.397	0.408	-0.012	0.682
Head of household	0.284	0.280	0.004	0.891

Table 3.A3 (Continued)

Household variables				
Run household enterprise	0.861	0.854	0.007	0.746
Own farm land	0.449	0.430	0.019	0.518
Household size/10	0.699	0.699	0.000	0.993
Household size squared/100	0.543	0.535	0.008	0.729
Children/10	0.409	0.413	-0.004	0.731
Children squared/100	0.208	0.209	-0.001	0.917
Household income: P2000-P2999	0.175	0.169	0.006	0.793
Household income: P3000-4999	0.301	0.305	-0.004	0.88
Household income: P5000-9999	0.261	0.265	-0.005	0.854
Household income: P10000-14999	0.071	0.082	-0.012	0.445
Household income:>P15000	0.069	0.061	0.008	0.593
Urban	0.343	0.360	-0.018	0.527

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Estimates for region dummies not reported in table.

Table 3.A4: Balancing tests using measure of pseudo *R*-squared and standardized bias

Pseudo-R^2		Wald test *p*-value		Standardized bias	
Before matching	After matching	Before matching	After matching	Before matching	After matching
0.061	0.009	0.000	1.000	7.323	2.273

Notes: Tests after matching only on observations in matched sample having common support.
Pseudo *R*-squared from logit estimation of the propensity score (the conditional treatment probability) on all the control variables before and after matching. *p*-values of the likelihood-ratio test of the joint insignificance of all the regressors before and after matching.
Standardized bias is the difference of the sample means in the treated and non-treated sub-samples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups.

Table 3.A5: Balancing *t*-tests for propensity score model covariates (Sample: Boys only)
Matched treated and untreated samples, based on kernel propensity score matching

Variables	Means ($m=1$) (1)	Means ($m=0$) (2)	Difference in means (1)-(2)	*t*-statistic
Child variables				
Age/10	1.411	1.413	-0.002	0.893
Age squared/100	2.047	2.054	-0.007	0.878
Age first worked/10	1.142	1.140	0.003	0.881
Age first worked squared /100	1.381	1.373	0.007	0.869
Daily work hours: 5-8 hours	0.511	0.519	-0.008	0.823
Daily work hours: 9+ hours	0.062	0.057	0.005	0.765
Weekly work days/10	0.354	0.361	-0.008	0.608
Weekly work days squared /100	0.167	0.175	-0.007	0.541
Night work	0.165	0.171	-0.006	0.821
Presently in work	0.596	0.570	0.026	0.461
Ever left school	0.429	0.426	0.003	0.925
Secondary or tertiary schooling	0.384	0.369	0.015	0.664
Location: Worked in own house	0.052	0.063	-0.011	0.507
Location: Worked on farm	0.653	0.643	0.010	0.763
Worked in own household enterprise	0.589	0.562	0.027	0.445
Paid for work	0.369	0.382	-0.013	0.711
Received meals as worker benefit	0.077	0.058	0.019	0.281
Worked for financial reasons	0.491	0.481	0.011	0.765
Gave earnings to parents	0.406	0.433	-0.027	0.443
Parent supervised work	0.526	0.525	0.001	0.977
Other relative supervised work	0.052	0.057	-0.004	0.791
Unsupervised	0.289	0.268	0.022	0.493
Interested in interview	0.781	0.798	-0.018	0.544
Sincere in interview	0.514	0.536	-0.022	0.532
Respondent parent variables				
Male	0.217	0.201	0.016	0.584
Age/10	4.286	4.264	0.022	0.689
Age squared/100	18.945	18.800	0.145	0.771
Secondary or tertiary education	0.379	0.373	0.006	0.859
Head of household	0.289	0.271	0.018	0.566

Table 3.A5 (Continued)

Household variables				
Run household enterprise	0.850	0.849	0.001	0.959
Own farm land	0.446	0.438	0.008	0.811
Household size/10	0.694	0.695	-0.002	0.917
Household size squared/100	0.536	0.529	0.007	0.799
Children/10	0.412	0.420	-0.007	0.598
Children squared/100	0.210	0.214	-0.004	0.761
Household income: P2000-P2999	0.187	0.189	-0.002	0.951
Household income: P3000-4999	0.309	0.297	0.012	0.712
Household income: P5000-9999	0.267	0.276	-0.009	0.773
Household income: P10000-14999	0.072	0.085	-0.013	0.509
Household income:>P15000	0.065	0.061	0.004	0.823
Urban	0.334	0.348	-0.014	0.680

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Estimates for region dummies not reported in table.

Table 3.A6: Balancing *t*-tests for propensity score model covariates (Sample: Girls only)
Matched treated and untreated samples, based on kernel propensity score matching

Variables	Means ($m = 1$) (1)	Means ($m = 0$) (2)	Difference in means (1)-(2)	*t*-statistic
Child variables				
Age/10	1.368	1.387	-0.019	0.446
Age squared/100	1.930	1.984	-0.054	0.425
Age first worked/10	1.093	1.091	0.002	0.944
Age first worked squared /100	1.263	1.263	0.000	0.999
Daily work hours: 5-8 hours	0.296	0.267	0.028	0.547
Daily work hours: 9+ hours	0.108	0.099	0.009	0.777
Weekly work days/10	0.310	0.339	-0.029	0.192
Weekly work days squared /100	0.140	0.163	-0.023	0.205
Night work	0.210	0.240	-0.031	0.482
Presently in work	0.790	0.794	-0.004	0.929
Ever left school	0.269	0.258	0.011	0.811
Secondary or tertiary schooling	0.435	0.476	-0.041	0.430
Location: Worked in own house	0.226	0.243	-0.017	0.695
Location: Worked on farm	0.478	0.429	0.050	0.338
Worked in own household enterprise	0.694	0.702	-0.008	0.868
Paid for work	0.253	0.241	0.012	0.789
Received meals as worker benefit	0.108	0.087	0.020	0.507
Worked for financial reasons	0.403	0.368	0.035	0.488
Gave earnings to parents	0.258	0.267	-0.009	0.851
Parent supervised work	0.527	0.551	-0.024	0.646
Other relative supervised work	0.043	0.049	-0.006	0.776
Unsupervised	0.306	0.283	0.024	0.616
Interested in interview	0.790	0.801	-0.011	0.792
Sincere in interview	0.618	0.618	0.001	0.989
Respondent parent variables				
Male	0.199	0.222	-0.023	0.587
Age/10	4.342	4.350	-0.008	0.916
Age squared/100	19.395	19.405	-0.010	0.988
Secondary or tertiary education	0.441	0.436	0.005	0.924
Head of household	0.274	0.301	-0.027	0.566

Table 3.A6 (Continued)

Household variables				
Run household enterprise	0.887	0.895	-0.008	0.801
Own farm land	0.468	0.406	0.062	0.230
Household size/10	0.706	0.704	0.002	0.937
Household size squared/100	0.552	0.548	0.004	0.917
Children/10	0.400	0.395	0.005	0.805
Children squared/100	0.200	0.196	0.004	0.849
Household income: P2000-P2999	0.156	0.150	0.006	0.870
Household income: P3000-4999	0.296	0.325	-0.030	0.539
Household income: P5000-9999	0.253	0.267	-0.014	0.756
Household income: P10000-14999	0.070	0.082	-0.012	0.662
Household income:>P15000	0.081	0.057	0.024	0.369
Urban	0.371	0.385	-0.014	0.783

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Estimates for region dummies not reported in table.

Table 3.A7: Balancing *t*-tests for propensity score model covariates (Sample: Fathers only)
Matched treated and untreated samples, based on kernel propensity score matching

Variables	Means ($m=1$) (1)	Means ($m=0$) (2)	Difference in means (1)-(2)	*t*-statistic
Child variables				
Male	0.694	0.638	0.057	0.354
Age/10	1.426	1.424	0.002	0.957
Age squared/100	2.087	2.083	0.004	0.962
Age first worked/10	1.152	1.150	0.002	0.956
Age first worked squared /100	1.405	1.396	0.009	0.914
Daily work hours: 5-8 hours	0.512	0.455	0.058	0.371
Daily work hours: 9+ hours	0.066	0.068	-0.002	0.962
Weekly work days/10	0.369	0.366	0.002	0.929
Weekly work days squared /100	0.182	0.179	0.003	0.892
Night work	0.190	0.215	-0.025	0.635
Presently in work	0.620	0.633	-0.013	0.832
Ever left school	0.380	0.367	0.013	0.837
Secondary or tertiary schooling	0.430	0.435	-0.005	0.933
Location: Worked in own house	0.141	0.169	-0.028	0.544
Location: Worked on farm	0.570	0.521	0.049	0.445
Worked in own household enterprise	0.620	0.581	0.039	0.537
Paid for work	0.339	0.359	-0.020	0.75
Received meals as worker benefit	0.066	0.087	-0.021	0.54
Worked for financial reasons	0.471	0.478	-0.007	0.914
Gave earnings to parents	0.355	0.369	-0.014	0.825
Parent supervised work	0.496	0.495	0.001	0.984
Other relative supervised work	0.050	0.040	0.009	0.73
Unsupervised	0.281	0.270	0.011	0.845
Interested in interview	0.810	0.800	0.010	0.841
Sincere in interview	0.620	0.596	0.024	0.701
Respondent parent variables				
Age/10	4.596	4.567	0.029	0.783
Age squared/100	21.745	21.528	0.217	0.832
Secondary or tertiary education	0.413	0.372	0.041	0.514

Table 3.A7 (Continued)

Household variables				
Run household enterprise	0.860	0.867	-0.008	0.863
Own farm land	0.372	0.394	-0.022	0.723
Household size/10	0.697	0.685	0.012	0.71
Household size squared/100	0.558	0.513	0.045	0.501
Children/10	0.409	0.406	0.003	0.888
Children squared/100	0.208	0.196	0.012	0.662
Household income: P2000-P2999	0.174	0.181	-0.008	0.88
Household income: P3000-4999	0.223	0.274	-0.051	0.366
Household income: P5000-9999	0.289	0.273	0.016	0.782
Household income: P10000-14999	0.074	0.055	0.020	0.537
Household income:>P15000	0.099	0.082	0.017	0.649
Urban	0.380	0.398	-0.018	0.781

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Estimates for region dummies not reported in table.

Table 3.A8: Balancing *t*-tests for propensity score model covariates (Sample: Mothers only)
Matched treated and untreated samples, based on kernel propensity score matching

Variables	Means ($m=1$) (1)	Means ($m=0$) (2)	Difference in means (1)-(2)	*t*-statistic
Child variables				
Male	0.666	0.678	-0.012	0.69
Age/10	1.386	1.386	-0.001	0.973
Age squared/100	1.979	1.982	-0.003	0.946
Age first worked/10	1.116	1.110	0.006	0.73
Age first worked squared /100	1.321	1.307	0.013	0.75
Daily work hours: 5-8 hours	0.418	0.428	-0.010	0.746
Daily work hours: 9+ hours	0.081	0.072	0.009	0.604
Weekly work days/10	0.331	0.343	-0.013	0.349
Weekly work days squared /100	0.152	0.163	-0.011	0.297
Night work	0.188	0.197	-0.008	0.753
Presently in work	0.672	0.653	0.020	0.524
Ever left school	0.373	0.368	0.005	0.879
Secondary or tertiary schooling	0.392	0.381	0.011	0.731
Location: Worked in own house	0.107	0.110	-0.003	0.874
Location: Worked on farm	0.597	0.587	0.010	0.754
Worked in own household enterprise	0.627	0.617	0.011	0.739
Paid for work	0.330	0.346	-0.016	0.605
Received meals as worker benefit	0.088	0.071	0.017	0.335
Worked for financial reasons	0.452	0.421	0.030	0.35
Gave earnings to parents	0.353	0.376	-0.023	0.474
Parent supervised work	0.535	0.533	0.002	0.953
Other relative supervised work	0.047	0.052	-0.005	0.717
Unsupervised	0.300	0.282	0.018	0.544
Interested in interview	0.777	0.799	-0.022	0.411
Sincere in interview	0.537	0.560	-0.023	0.49
Respondent parent variables				
Age/10	4.217	4.193	0.024	0.612
Age squared/100	18.293	18.088	0.205	0.617
Secondary or tertiary education	0.398	0.416	-0.018	0.58
Head of household	0.092	0.090	0.002	0.899

Table 3.A8 (Continued)

Household variables				
Run household enterprise	0.863	0.856	0.007	0.767
Own farm land	0.469	0.433	0.036	0.271
Household size/10	0.700	0.700	0.000	0.984
Household size squared/100	0.539	0.537	0.002	0.933
Children/10	0.409	0.414	-0.005	0.721
Children squared/100	0.207	0.212	-0.005	0.699
Household income: P2000-P2999	0.176	0.171	0.005	0.848
Household income: P3000-4999	0.321	0.305	0.016	0.602
Household income: P5000-9999	0.257	0.264	-0.007	0.815
Household income: P10000-14999	0.071	0.092	-0.021	0.237
Household income:>P15000	0.062	0.057	0.005	0.743
Urban	0.338	0.352	-0.014	0.658

Notes: * Statistically significant at the 10% level; ** at the 5% level; *** at the 1% level. Estimates for region dummies not reported in table.

Table 3.A9: Balancing tests using measure of pseudo *R*-squared and standardized bias (by separate samples)

Sample	Pseudo *R*-squared		Wald test *p*-value		Standardized bias	
	Before matching	After matching	Before matching	After matching	Before matching	After matching
Boys only	0.078	0.012	0.000	1.000	6.984	2.156
Girls only	0.118	0.047	0.000	1.000	8.855	3.160
Fathers only	0.146	0.039	0.000	1.000	11.542	3.526
Mothers only	0.074	0.013	0.000	1.000	6.744	2.468

Notes: Tests after matching only on observations in matched sample having common support.
Pseudo-R^2 from logit estimation of the propensity score (the conditional treatment probability) on all the control variables before and after matching. *p*-values of the likelihood-ratio test of the joint insignificance of all the regressors before and after matching.
Standardized bias is the difference of the sample means in the treated and non-treated sub-samples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups.

REFERENCES

Aakvik, Arild. 2001. Bounding a Matching Estimator: The Case of a Norwegian Training Program. *Oxford Bulletin of Economics and Statistics*. v63 (n1):115-43.

Abadie, Alberto and Guido W. Imbens. 2005. On the Failure of the Bootstrap for Matching Estimators. Mimeo, University of California, Berkeley.

Akabayashi, Hideo, and George Psacharopoulos. 1999. The Trade-off between Child Labour and Human Capital Formation: A Tanzanian Case Study. *Journal of Development Studies*. v35 (n5):120-40.

Baland, Jean-Marie, and James A. Robinson. 2000. Is Child Labor Inefficient? *Journal of Political Economy*. v108 (n4):663-79.

Basu, Kaushik, and Pham Hoang Van. 1998. The Economics of Child Labor. *American Economic Review*. v88 (n3):412-27.

Basu, Kaushik. 1999. Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standards. *Journal of Economic Literature*. v37 (n3):1083-1119.

Basu, Kaushik. 2006. Gender and Say: A Model of Household Behaviour with Endogenously Determined Balance of Power. *Economic Journal*. v116 (n511):558-80.

Basu, Kaushik, Sanghamitra Das, and Bhaskar Dutta. Child Labor and Household Wealth: Theory and Empirical Evidence of a Inverted-U. IZA Discussion Paper No. 2736. Bonn: IZA.

Becker, Gary S. 1981. *A treatise on the family*. Cambridge: Harvard University Press.

Becker, Sascha O. and Marco Caliendo. 2007. Sensitivity Analysis for Average Treatment Effects. *Stata Journal.* v7 (n1):71-83.

Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti. 2006. Child Labor and Agricultural Shocks. *Journal of Development Economics*. v81 (n1):80-96.

Beegle, Kathleen, Rajeev H. Dehejia, and Roberta Gatti. 2005. *Why should we care about child labor? The education, labor market, and health consequences of child labor*. World Bank Policy Research Working Paper Series No. 3479. Washington, D.C.: World Bank.

Black, Dan A., and Jeffrey A. Smith. 2004. How Robust Is the Evidence on the Effects of College Quality? Evidence from Matching. *Journal of Econometrics*. v121 (n1-2):99-124.

Bourguignon, Francois, and Pierre-Andre Chiappori. 1992. Collective Models of Household Behavior: An Introduction. *European Economic Review*. v36 (n2-3):355-64.

Browning, Martin, and Pierre-Andre Chiappori. 1998. Efficient Intra-household Allocations: A General Characterization and Empirical Tests. *Econometrica*. v66 (n6):1241-78.

Bryson, Alex. 2002. *The Union Membership Wage Premium: An Analysis Using Propensity Score Matching*. Centre for Economic Performance Discussion Paper No. 530. London: London School of Economics.

Cahuc, Pierre, and Andre Zylberbeg. 2004. *Labor Economics*. Cambridge: MIT Press.

Caliendo, Marco. 2008. *Microeconometric Evaluation of Labor Market Policies*. New York: Springer.

Caliendo, Marco, and Sabine Kopeinig. 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys.* v22 (n1): 31-72.

Cameron, Colin A., and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Castro, Charita L., Sarah Gormly, and Amy R. Ritualo. 2005. The SIMPOC Philippine Survey of Children 2001: A Data Source for Analyzing Occupational Injuries to Children. *Public Health Reports*. v120 (n6):631-641.

Chiappori, Pierre-Andre. 1988. Nash-Bargained Households Decisions: A Comment. *International Economic Review*. v29 (n4):791-96.

Chiappori, Pierre-Andre. 1992. Collective Labor Supply and Welfare. *Journal of Political Economy*. v100 (n3):437-67.

Dar, Amit, Niels-Hugo Blunch, Bona Kim, and Masaru Sasaki. 2002. Participation of Children in Schooling and Labor Activities: A Review of Empirical Studies. Social Protection Discussion Paper No. 0221. Washington D.C.: World Bank.

Dessy, Sylvain E., and Stephane Pallage. 2005. A Theory of the Worst Forms of Child Labour. *Economic Journal*. v115 (n500):68-87.

Duflo, Esther, and Christopher Udry. 2004. *Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices*. National Bureau of Economic Research Working Paper No. 10498.

Duflo, Esther. 2003. Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa. *World Bank Economic Review*. v17 (n1):1-25.

Edmonds, Eric V., and Nina Pavcnik. 2005. The Effect of Trade Liberalization on Child Labor. *Journal of International Economics*. v65 (n2):401-19.

Emerson, Patrick M., and Andre Portela Souza. 2003. Is There a Child Labor Trap? Intergenerational Persistence of Child Labor in Brazil. *Economic Development and Cultural Change*. v51 (n2):375-98.

Emerson, Patrick M., and Andre Portela Souza. 2007. Child Labor, School Attendance, and Intrahousehold Gender Bias in Brazil. *World Bank Economic Review*. v21 (n2):301-16.

Felkey, Amanda J. 2006. "Husbands, wives and the peculiar economics of household public goods". In Essays on the household formation, resource allocation and consequences for the labor market. Ph.D. dissertation, Cornell Univeristy.

Galdo, Jose, Jeffrey Smith, and Dan Black. 2007. *Bandwidth Selection and the Estimation of Treatment Effects with Unbalanced Data.* Institute for the Study of Labor Discussion Paper No. 3095.

Grootaert, Christiaan, and Harry Anthony, Patrinos, eds. 1999. *The policy analysis of child labor: A comparative study*. New York: St. Martin's Press.

Grootaert, Christiaan, and Ravi Kanbur. 1995. Child Labour: An Economic Perspective. *International Labour Review*. v134 (n2):187-203.

Gunnarsson, Victoria, Peter F. Orazem, and Mario A. Sanchez. 2006. Child Labor and School Achievement in Latin America. *World Bank Economic Review*. v20 (n1):31-54.

Hagemann, Frank, Yacouba Diallo, Alex Etienne, and Farhad Mehran. 2006. *Global Child Labour Trends 2000 to 2004.* Geneva: International Labour Organization.

Ham, John C., Xianghong Li, and Patricia B. Reagan. 2005. *Propensity score matching, a distance-based measure of migration, and the wage growth of young men*. Federal Reserve Bank of New York Staff Report No. 212.

Heckman, James, Hidehiko Ichimura, and Petra Todd. 1997. Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies*. v64 (n4): 605-654.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. *Characterizing Selection Bias Using Experimental Data*. National Bureau of Economic Research Working Paper No. 6699.

Ilahi, Nadeem, Peter F. Orazem, and Guilherme Sedlacek. 2005. *How Does Working as a Child Affect Wage, Income and Poverty as an Adult?* World Bank Social Protection Discussion Paper No. 0514.

International Labour Organization. 2009. *Report of the Conference: 18th International Conference of Labour Statisticians: Geneva, 24 November-5 December 2008.* Geneva: International Labour Office.

Imbens, Guido M., and Jeffrey M. Wooldridge. 2008. *Recent Developments in the Econometrics of Program Evaluation*. National Bureau of Economic Research Working Paper No. 14251.

Jacoby, Hanan G., and Emmanuel Skoufias. 1997. Risk, Financial Markets, and Human Capital in a Developing Country. *Review of Economic Studies*. v64 (n3):311-35.

Kruger, Diana I. 2007. Coffee Production Effects on Child Labor and Schooling in Rural Brazil. *Journal of Development Economics*. v82 (n2):448-63.

Lundberg, Shelly, and Robert A. Pollak. 1993. Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy*. v101 (n6):988-1010.

Manser, Marilyn, and Murray Brown. 1980. Marriage and Household Decision-Making: A Bargaining Analysis. *International Economic Review*. v21 (n1):31-44.

McElroy, Marjorie B., and Mary Jean Horney. 1981. Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand. *International Economic Review*. v22 (n2):333-49.

Moehling, Carolyn M. 2005. "She Has Suddenly Become Powerful": Youth Employment and Household Decision Making in the Early Twentieth Century. *Journal of Economic History*. v65 (n2):414-38.

O'Donnell, Owen, Furio C. Rosati, and Eddy van Doorslaer. 2005. Health Effects of Child Work: Evidence from Rural Vietnam. *Journal of Population Economics*. v18 (n3):437-67.

Pitt, Mark M., and Shahidur R. Khandker. 1998. The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter? *Journal of Political Economy.* v106 (n5):958-96.

Pitt, Mark M., Shahidur R. Khandker, Omar H. Choudhury, and Daniel L. Millimet. 2003. Credit Programs for the Poor and the Health Status of Children in Rural Bangladesh. *International Economic Review*. v44 (n1):87-118.

Ranjan, Priya. 2001. Credit Constraints and the Phenomenon of Child Labor. *Journal of Development Economics*. v64 (n1):81-102.

Ravallion, Martin. 2008. Evaluating Anti-Poverty Programs. In Paul Schultz and John Strauss (eds.) *Handbook of Development Economics* vol. 4. Amsterdam: North-Holland.

Rogerro, Paola, Viviana Mangiaterra, Flavia Bustreo, and Furio Rosati. The Health Impact of Child Labor in Developing Countries: Evidence from Cross-Country Data. *American Journal of Public Health*. v97 (n2):271-75.

Rosati, Furio Camillo, and Mariacristina Rossi. 2003. Children's Working Hours and School Enrollment: Evidence from Pakistan and Nicaragua. *World Bank Economic Review*. v17 (n2):283-95.

Rosati, Furio, and Roland Straub. 2007. Does Work during Childhood Affect the Health of Guatemalan Adults? *Review of Economics of the Household*. v5 (n1):83-94.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*. v70 (n1):41-55.

Rosenbaum, Paul R. 2002. *Observational Studies*. New York: Springer.

Rubalcava, Luis, Gracela Teruel, and Duncan Thomas. 2009. Investments, Time Preferences, and Public Transfers Paid to Women. *Economic Development and Cultural Change*. v57 (n3):507-538.

Rubin, Donald B. 1974. Characterizing the Estimation of Parameters in Incomplete-Data Problems. *Journal of the American Statistical Association*. v69 (n346):467-74.

Rubin, Donald and Neal Thomas. 2000. Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates. *Journal of the American Statistical Association*. v95:573-585.

Schady, Norbert, and Jose Rosero. 2008. Are Cash Transfers Made to Women Spent Like Other Sources of Income? *Economics Letters*. v101 (n3):246-48.

Sedlacek, Guilherme, Suzanne Duryea, Nadeem Ilahi, and Masaru Sasaki. 2005. Child Labor, Schooling, and Poverty in Latin America. World Bank Social Protection Discussion Paper No. 0511. Washington, D.C.: World Bank.

Smith, Jeffrey A., and Petra E. Todd. 2005. Does Matching Overcome LaLonde's Critique of Non-experimental Estimators? *Journal of Econometrics*. v125 (n1-2):305-53.

Thomas, Duncan. 1990. Intra-household Resource Allocation: An Inferential Approach. *Journal of Human Resources*. v25 (n4):635-64.

Thomas, Duncan. 1994. Like Father, Like Son; Like Mother, Like Daughter: Parental Resources and Child Height. *Journal of Human Resources*. v29 (n4):950-88.

Todd, Petra. 2008. Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated. In Paul Schultz and John Strauss (eds.) *Handbook of Development Economics* vol. 4. Amsterdam: North-Holland.

Veenhoven, Ruut. 2004. Subjective measures of well-being. UNU-WIDER Discussion Paper No. 2004/07.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.