School Attendance and Child Labor – A Model of Collective Behavior

Holger Strulik*

Leibniz Universitat Hannover, Discussion Paper No. 441

ISSN 0949-9962

November 2009. This Version: February 2010

Abstract. This paper theoretically investigates how community approval or disapproval affects school attendance and child labor and how aggregate behavior of the community feeds back towards the formation and persistence of an anti- (or pro-) schooling norm. The proposed community-model continues to take aggregate and idiosyncratic poverty into account as an important driver of low school attendance and child labor. But it provides also an explanation for why equally poor villages or regions can display different attitudes towards schooling. Distinguishing between three different modes of child time allocation, school attendance, work, and leisure, the paper shows how the time costs of schooling and child labor productivity contribute to the existence of a locally stable anti-schooling norm. It proposes policies that effectively exploit the social dynamics and initiate a permanent escape from the anti-schooling equilibrium. An extension of the model explores how an education contingent subsidy paid to the poorest families of a community manages to initiate a bandwagon effect towards "education for all". The optimal mechanism design of such a targeted transfer program is investigated.

Keywords: School Attendance, Child Labor, Social Norms, Targeted Transfers

JEL: I20 I29 J13 O12

---

*Department of Economics, Brown University, Providence, RI 02912, USA, email: holger_strulik@brown.edu and University of Hannover, Wirtschaftswissenschaftliche Fakultaet, 30167 Hannover, Germany; email: strulik@vwl.uni-hannover.de. I would like to thank Arnab Basu, Arjun Bedi, Karl Dietrich, Katja Michaelowa, Jean-Philippe Platteau, and participants at the ESPE conference, London, at the VfS Conference on Development Economics and Policy, Zurich, and at the EUDN Conference, Paris for their useful comments.

Suddenly, the door slammed shut with a bang. It was her father, and he was angry.
Her mother rushed past. "What's the matter, habibi?"
Baba stamped his feet in fury. "What's the matter? What's the matter? She's the matter!" he roared, pointing a shaking finger at Maha as she emerged from her bedroom, book still in hand.
"I can't walk ten feet in this village without somebody making some rude remark about my daughter and the disgrace she's bringing on this family. The women are talking. The men are talking. The village elders are talking. They came to me today to say they don't approve of Maha going to school alone. Like I don't know! They've been saying nothing else since all this school foolishness began. She brings shame on the entire village, they say! We can't live with that kind of disgrace. No one does business with me anymore. It's like we're outcasts in our own community."
His voice dropped. "Maha, I know what I said, but you can't go to school anymore."
(Rania Al-Abdullah, 2009)

1. Introduction

The observation that many children in developing countries are not attending school and many of them instead spend at least some of their time working is explained in the economics literature primarily by aggregate and individual poverty (Basu, 1998, 1999, Edmonds, 2007). While cross-country comparisons strongly confirm the association between poverty and child labor, within-country studies provide sometimes surprisingly little support (Psacharopolous, 1997, Ray, 2000, Bhalotra, 2006). Similarly the empirical literature on schooling in developing countries, while generally supporting the poverty association, has us left with the questions of why, for example, a Peruvian child is ceteris paribus (i.e. in particular controlling for poverty) more likely to experience schooling than a Pakistani child (Ray, 2000), or why in India 98 percent of the girls are enrolled in middle school in Kerala but only 36 percent in Rajasthan, a state of similar poverty (Kingdon, 2005).

In order to address these questions the present paper sets up a dynamic model of a community where the individual households' decision about the allocation of child time on schooling, work, and leisure is – besides the usual economic determinants – also dependent on the community's approval or disapproval of schooling. The actually observed aggregate attendance at school then feeds back on the evolution of community (dis-) approval of schooling. Households are considered to be heterogenous with respect to poverty and with respect to susceptibility to approval of their actions by the community.

Because of the twofold heterogeneity of households the model supports an equilibrium where it remains to be true that both aggregate and individual poverty are important drivers of low schooling but where the association between the two variables becomes less clear cut than in the standard economic approach. At equilibrium, children of some relatively poor families may attend school while other children of relatively rich parents are absent. More importantly the twofold heterogeneity of households creates social dynamics according to which two similarly poor communities can converge towards very different locally stable equilibria: at one equilibrium every child goes to school and the community sustains a pro-schooling norm whereas at the other equilibrium a majority of children does not go to school and the community sustains an anti-schooling norm.

There exists a small economic literature investigating the evolution of norms. Problems addressed so far include the growing welfare state (Lindbeck, et al., 1999), out-of-wedlock childbearing (Nechyba, 2001), family size (Palivos, 2001), child labor supply (Lopez-Calva, 2002), occupational choice (Mani and Mullin, 2004), and contraceptive use (Munshi and Myaux, 2006). Among these the work of Lopez-Calva is naturally mostly related to the present paper. Lopez-Calva assumes that the strength of an anti-child labor norm depends inversely on the aggregate supply of child labor. Given a static model with homogenous households he derives conditions for which there exist two equilibria, one without child labor, a strong anti-child-labor norm, and a high equilibrium wage and another one where all households supply child labor and both anti-child-labor norm and wages are low. For a given set of parameters both equilibria are equally likely attained such that the child labor decision is reduced to a coordination problem. Schooling attendance, income heterogeneity, and social dynamics are not investigated.

The present paper extends the above literature by investigating how a social norm affects school attendance, taking child labor, and an important third mode of child time allocation, leisure and play, explicitly into account. Another novelty of the model is that it allows, depending on the specification of parameters, that schooling affects child labor either on the intensive margin or on the extensive margin. On the intensive margin school attending children continue to supply labor, yet they spend less time working. In other words, schooling is partly "financed" by a reduction of child leisure and play. On the extensive margin school attendance and child labor are mutually exclusive because parents prefer to leave their children some hours of leisure and play after school. This distinction of interior solution and corner solution and the potential "state transition" of children caused by policy leads to some non-obvious conclusions. For example, increasing the length of the school day is helpful to reduce child labor on the intensive margin. On the extensive margin, however, if schooling becomes too time-intensive, parents may prefer to withdraw their children from school, in particular when schooling gets little approval from the community.

Because social approval and, in equilibrium, the community norm, are assumed to be attached to schooling, social interaction affects child labor only indirectly through the schooling decision, either on the intensive or on the extensive margin. Ideally one may want to investigate two separate social norms, one attached to schooling, the other one to child labor. This idea, however, complicates the model beyond analytical tractability. Anyway, in an equilibrium where schooling and child labor turn out to be mutually exclusive it is, at least qualitatively, unimportant whether social approval is attached to schooling or child labor. This is, of course, no longer true in an interior equilibrium where school attending children continue to work. In that case one can argue that a second best approximation of reality is to attach the social norm to the schooling decision. A schooling norm can be more easily developed and sustained because schooling is a binary decision (sent the girl to school or not) whereas child labor is a continuous variable to which it is harder to attach a specific normative value. Furthermore, the schooling decision is very visible within the community. Children show up at school, wear uniforms and satchels, wait with fellow pupils at the bus stop etc. Whether and how long children work cannot so easily be monitored by others, in particular if they work at home.

Empirically, norms, traditions, and beliefs are hard to measure quantitatively and econometric exercises trying to provide evidence for true community effects are frequently plagued by the self-selection problem (Manski, 1993). In support of the proposed theory there exists, however, ample narrative evidence, in the spirit of the introductory quote from Rania Al-Abdullah's (2009) story and ample indirect evidence. One general conclusion is that the perceived anti-schooling norm is strongest with respect girls. For example, the UNICEF (2008) report on schooling in Sudan concludes:

Cultural barriers to schooling might present the biggest challenge especially in trying to achieve gender parity in education, but yet they have to be addressed if progress is to be achieved in this crucial area. Most cultural barriers manifest themselves as attitudes and practices that have gained *acceptance over long periods of repeated practice and approval....The main cultural obligation for women in most Sudanese communities is home making as wives, caregivers and child bearers. Girls from their early ages are socialized to be just that and sadly most of them do not break off from this bondage. The problem is that both the community and the girls view themselves in this light and any reason to change this "natural" course of things is viewed negatively even by the victims themselves.*

On a more general level, neighborhood peer effects on schooling have been shown be Zelizer (1985), Crane (1991), Case and Katz (1991), and Kling et al. (2007), and with focus on developing countries by Binder (1999) and Chamarbagwala (2007). Particularly interesting for the present paper are two recent evaluations of the Progresa targeted transfer program. Both Lalive and Cattaneo (2009) and Bobonis and Finan (2009) find that schooling-contingent subsidies did not only increase school attendance of targeted children but also that of ineligible children when the program was introduce in their local community. Besides demonstrating the existence of such a "social multiplier" (Glaeser et al., 2003) for the schooling decision the studies provide a couple of other interesting results, indirectly supporting the theory proposed in the present paper. It was found that poorer families are more strongly affected by the behavior of peers, that the social interaction runs through the parents (rather than the children) and that the strength of the response of ineligible families is increasing in the share of eligible families.

The remainder of the paper is organized as follows. The next section sets up the model and analyzes its static solution, i.e. the determinants of schooling and child labor for given strength of social approval. Section 3 investigates dynamics, i.e. how current schooling activities feed back on the evolution and persistence of an anti- (or pro-) schooling norm. It identifies the conditions under which a locally stable anti-schooling equilibrium can exist and discusses the comparative statics of such an equilibrium, in particular with respect to child labor productivity and the time costs of attending school. The section concludes by discussing how policy can achieve the dual goal of escaping from the anti-schooling equilibrium and eliminating child labor.

The twofold heterogeneity of households with respect to poverty and susceptibility to approval makes the model particularly suitable for a theoretical analysis of targeted transfer programs, i.e. policies offering an income subsidy to the poorest families of a community contingent on their children attending school. Section 4 extends the model for such an analysis and provides and explanation for the frequently found large success of schooling-contingent subsidies. Finally the paper investigates the optimal mechanism design of such targeted transfer programs, i.e. it asks how many families should be addressed and how high should subsidies be in order to initiate convergence towards education for all at minimum costs. Derivations of equations and non-obvious proofs or lemmata and propositions can be found in the Appendix.

2. The Model

2.1. **Setup of Society and Family Constraints.** Consider a community populated by a continuum of families indexed by i. Suppose that each family consists of one parent and one child. Suppose that the daily time budget constraint of the child is normalized to one and that the parent decides how this time is spent. At any day t child time may be spent on attending school $h_t(i)$, work $\ell_t(i)$, or leisure and play. Attending school can be modeled conveniently as a binary choice, yes ($h_t(i) = 1$) or no ($h_t(i) = 0$). Child labor, with contrast, is better conceptualized as a continuous variable. Parents decide not only whether at all but also how long their children have to work. Let the time cost of attending school including walking or driving to school be given by τ. The implied child time budget constraint for family i at day t is then given by

$$1 - \ell_t(i) - \tau h_t(i) \geq 0. \tag{1}$$

Suppose that adults allocate all of their time – net of household duties – to work. Without child labor household i earns a family income $w(i)$. A working child contributes to family income a fraction γ of adult income per unit of labor supplied. Although the model is more general and gender is nowhere explicitly stated it may be helpful for the intuition to focus on the time allocation problem for girls. They are probably not supplying any wage work at the labor market but are helping in the household. Girls working ℓ time units at home set free $\gamma\ell$ units of extra time of their parents, which is used by their parents to supply additional adult labor and earn wage income. Suppose that all income is used to finance adult consumption $c_t(i)$ so that family i's budget constraint reads[1]

$$w(i) + \gamma w(i)\ell_t(i) = c_t(i). \tag{2}$$

[1] An alternative assumption providing similar results with more notational clutter would be that income is spent on consumption of adults and children and that adults take utility from child consumption (imperfectly) into account. Also, we could alternatively assume that adults spend the time saved through helping children on leisure. The solution of the adult's own time allocation problem would then provide a value of adult leisure determined by the opportunity cost of the adult wage so that nothing of the model mechanics would be changed. To keep the model simple we thus neglect child consumption and the time allocation problem of adults. Finally, with the focus on daughters, we could allow parents to have one or several sons without affecting any results if preference parameters are such that parents want their sons to attend school.

2.2. **Preferences.** Utility of any adult i consists of a "normal" or autonomous component $u_t^a(i)$ and a socially-dependent component $u_t^s(i)$. In solving the time allocation problem parents consider the trade-off between the empathetically perceived utility from child leisure and play and the children's contribution to family income. They take also into account the empathetically perceived utility from letting their children attend school and social approval or disapproval of such a decision.

Suppose the autonomous part of utility can be captured by the following CES function.

$$u_t^a(i) = c_t^{1-\alpha} \cdot (1 - \ell - \tau h)^{\alpha} \cdot (1 + ah). \tag{3}$$

Here the parameter $\alpha \in (0, 1)$ measures empathy, i.e. the relative weight that parents put on child leisure and play relative to own consumption. For $\alpha \to 0$ we have totally insensitive parents caring only about consumption (and the evaluation of their behavior by the community). As α rises parents increasingly suffer from letting their children work. The parameter a measures how parents evaluate schooling versus leisure time of their children. Since the value for school attendance h_t is either 0 or 1, the associated exponent in the utility function is normalized to one without loss of generality. The crucial feature of this utility function is that attending school reduces the marginal utility perceived from child leisure and play. It will generate the empirically relevant result that higher time costs of schooling substitute partly but not fully child labor, i.e. a part of the schooling activity is "financed" by less leisure and play.

As motivated in the introduction we assume that social approval or disapproval is attached to the schooling decision. Whether sending one's child/daughter to school is socially approved or disapproved depends on the share of community members whose children/daughters are not attending school. A formally equivalent assumption would be that the value of a pro-schooling decision (in terms of parental utility) is learned from observing school attendance of other children/daughters. To formalize this notion we assume that the maximum approval of a pro-schooling decision, which is realized when all children of the community are attending school, is given by $\phi \geq 0$. Social disapproval S_t of a pro-schooling decision is assumed to increase in the relative number of families whose children/daughters were not going school in the past. At day t, S_t is thus endogenous but pre-determined. The complete social evaluation of a schooling decision is given by difference of between maximum approval and disapproval $(\phi - S_t)$.

People are to different degrees susceptible to the evaluation of their behavior by others. We capture this fact by assuming that susceptibility to approval is uniformly and independently distributed within the unit interval. A parent i has susceptibility to approval $\sigma(i) \in [0, 1]$. The most self-assured parent is assigned with value 0 and the most indeterminate one with value 1. Summarizing, the social component of utility experienced by parent i for a "pro schooling" decision is given by

$$u_t^s(i) = \sigma(i) \cdot (S_t - \phi). \tag{4}$$

Note that the "character trait" of susceptibility to approval $\sigma(i)$ is heterogenous across the community but individually given and constant over time while the social evaluation of an individual action is the same for all community members but endogenously explained and time-variant. In simple words, σ reflects how easily people "change their mind" when the community changes the evaluation of their behavior.

The model does not provide an endogenous, deeper explanation for why the social norm existed in the first place. Formally S_0 at time $t = 0$ is given by history. For example, in history there was no school within daytime reach ($\tau > 1$), or there was no school for girls, or no separate toilet for girls at school. The fact that no girl attended school in this initial state then "explains" the initial strength of the social attitude that girls do not belong to school.[2]

2.3. **Child Labor and Schooling Decisions.** Putting the autonomous and social elements of utility together and assuming separability, i.e. $u_t(i) = u_t^a(i) + u_t^s(i)$, parent i at day t solves the problem

$$\max_{c_t(i), \ell_t(i), h_t(i)} u_t(i) = c_t(i)^{1-\alpha}(1 - \ell_t(i) - \tau \cdot h_t(i))^\alpha (1 + a \cdot h_t(i)) + h_t(i) \cdot \sigma(i) \cdot (\phi - S_t). \tag{5}$$

subject to the budget constraint (2).

The solution of the problem with respect to child labor is obtained as

$$\ell_t(i) = \max \left\{ 0,\ (1 - \alpha)(1 - \tau h_t(i)) - \frac{\alpha}{\gamma} \right\}. \tag{6}$$

Child labor depends positively on child productivity γ (in setting free extra adult time) and negatively on preference for child leisure α. At the interior solution, schooling *partly* substitutes

[2]In the spirit of North (1990) it can be imagined that the "institution of daughters not attending school" was once functional in the history of a traditional society, at a time when the skill premium was small and daughters were seen as a household duties performing asset, sold at marriage. With the skill premium close to zero, any "school foolishness" would just delay the matchmaking. See Dreze and Sen (1995), Kingdon (2009).

child labor. The substitution effect is the smaller the higher the weight on child leisure and play, $\partial\ell/\partial\tau = -(1-\alpha)$. Note also, that child labor is always smaller than one. Only totally insensitive parents would let their children work all of the time ($\ell \to 1$ for $\alpha \to 0$ and $h_t = 0$). Empathetic parents leave their children some spare time.

For children to work at all, child labor productivity must be sufficiently high in comparison to preference for child leisure. Some of a child's daily time is used for (household) work, if

$$\gamma(1 - \tau h_t) > \frac{\alpha}{1-\alpha}. \tag{7}$$

Note that the condition may hold for $h_t = 0$ but not for $h_t = 1$. In this case, attending school effectively eradicates child labor.

Because attending school may cause a state transition from working to not-working, we have to distinguish three cases:

WW: the child is working no matter whether she is attending school or not (condition (7) is fulfilled for both $h_t = 0$ and $h_t = 1$).

WN: the child is working if she is not attending school but she not working if she is attending school (condition (7) is fulfilled for $h_t = 0$ but not for $h_t = 1$).

NN: the child is never working (condition (7) is – irrespective of the schooling choice – never fulfilled).

Clearly the WW and WN cases are the more interesting ones from the developing countries' viewpoint. But there are also incidences reported supporting the NN case of completely idle children (Bigeri et al., 2003, Bacolod and Ranjan, 2008). Besides the natural explanation of low productivity, possibly due to insufficient health, captured by low γ in (7), the present paper will provide another explanation for the "puzzle of idle children": a persistent anti-schooling norm.

Next insert (6) into (2) to get consumption and then insert both $\ell_t(i)$ and $c_t(i)$ into (5) to get the indirect utility function $\tilde{u}_t(i)$ for given schooling decision $h_t(i)$. The child of family i is not attending school if $\tilde{u}_t(i)|h_t(i) = 1 < \tilde{u}_t(i)|h_t(i) = 0$. Evaluate and compare utilities for the three cases from above to find that the child of family i is not attending school if

$$E_{ij} \cdot w(i)^{1-\alpha} \leq \sigma(i) \cdot (S_t - \phi), \tag{8}$$

$$i, j = W, N, \quad \text{where}$$

$$E_{WW} = \{[1 + \gamma(1-\tau)]\, a - \gamma\tau\}\, (1-\alpha)^{1-\alpha} \left(\frac{\alpha}{\gamma}\right)^{\alpha}$$

$$E_{WN} = (1-\tau)^{\alpha}(1+a) - (1-\alpha)^{1-\alpha}(1+\gamma) \left(\frac{\alpha}{\gamma}\right)^{\alpha}$$

$$E_{NN} = (1-\tau)^{\alpha}(1+a) - 1.$$

The left hand side of (8) captures the autonomous, socially-independent, i.e. "ordinary", net utility experienced from a pro-schooling decision while the right hand side summarizes utility experienced from social disapproval of schooling. It consists of the social evaluation of a pro-schooling decision by the community, $S_t - \phi$, weighted by parent i's susceptibility to approval. Ceteris paribus the size of the autonomous part of net utility (at the left hand side) is lower for poorer families. The constant E_{ij}, summarizes all decision-relevant parameters. It will be called the "schooling propensity" because it reflects the net utility increase experienced from a pro-schooling decision that an additional unit of weighted income $w(i)^{1-\alpha}$ brings about.

The overall size of the schooling propensity differs, depending on child's time allocation with and without attending school. In particular one can show the following.

Lemma 1. *$E_{WN} \leq E_{WW}$. $E_{WN} < E_{NN}$.*

The result is intuitive. The schooling propensity is lowest when school attendance causes a state transition from an interior solution with child labor towards a corner solution without child labor (E_{WN}), i.e. when when school attendance eliminates child labor but child labor is the preferred state without schooling. The schooling propensity is higher if the school attending child continues to contribute to family income (E_{WW}) or if the school attending child was anyway not working even without attending school (E_{NN}).

In order to avoid uninteresting case differentiation we assume that parameters are such that $E_{WN} > 0$, and thus $E_{WW} > 0$ and $E_{NN} > 0$. In this case the left hand side of (8) is positive and we know that, if a child is not attending school, the cause is a prevailing anti-schooling norm in the community (the right hand side of (8) is positive and larger than the left hand side). For $E_{WN} > 0$ the parent has to evaluate schooling sufficiently strongly in comparison to child productivity, time costs of schooling, and preference for child leisure such that schooling is indeed a feasible option if community norms were absent.

Assumption 1 (Schooling Feasibility).

$$a > (1+\gamma)\left(\frac{\alpha}{\gamma(1-\tau)}\right)^{\alpha} - 1. \tag{9}$$

Note that the assumption of schooling feasibility in (8) continues to produce the result that, ceteris paribus, children of poor families (of low $w(i)$) are more likely not attending school. The mechanism at work, however, is different from the poverty channel established in the literature. Here, poor parents, given their low income, achieve lower autonomous utility such that the social evaluation of their action plays are relatively larger role in their decision making than for rich parents.

Taking the first derivative of the E_{ij}'s with respect to the relevant parameter provides the following comparative statics with a straightforward intuition.

Lemma 2. *Assume schooling is feasible. The schooling propensity E_{ij}, $i, j \in \{W, N\}$ is high, i.e. for given social evaluation $S_t - \phi$ and given income $w(i)$, a child is more likely attending school, if the parent's individual evaluation of schooling a is high and schooling is not time intensive (τ is low).*

An immediate corollary is that making schooling more time consuming, for example, by increasing the school hours per day, is a bad idea in order to eradicate or reduce child labor (unless school attendance is enforceable). Some parents may indeed let the children work less, as indicated by (6). Other parents, however, will react with withdrawing their children from school, as indicated by (8), an effect that will amplify a prevailing anti-schooling and may cause other parents to withdraw their children from school as well because of the initiated community effects (see below).

The comparative statics with respect to child labor productivity reveals a non-monotony.

Lemma 3. *Assume schooling is feasible. A lower child labor contribution to family income γ always increases the schooling propensity E_{WN}. It increases the schooling propensity E_{WW} if γ is sufficiently low, i.e. if*

$$\gamma < \bar{\gamma} \equiv \frac{\alpha a}{(1-\tau)(1+a) - 1}. \tag{10}$$

In these cases, for given social evaluation $S_t - \phi$ and given income $w(i)$, a less productive child is more likely attending school.

If, however, $\gamma > \bar{\gamma}$, a larger γ increases the schooling propensity E_{WW}. For given social evaluation $S_t - \phi$ and given income $w(i)$, a more productive child is more likely attending school.

The result with respect to E_{WW} reflects two countervailing forces of child labor productivity on schooling. First, higher child labor productivity increases the opportunity cost of schooling and will lead for itself to lower schooling propensity. This effect is reflected by the term $(\alpha/\gamma)^{\alpha}$ of E_{WW} of E_{WW} in (8). Second, higher child labor productivity increases family income and consumption, thereby reducing the marginal utility from consumption, making consumption less essential. This in turn reduces the need for child labor and increases the schooling propensity. This effect is reflected by the first term of E_{WW} in (8). Ceteris paribus, the positive effect through the first term is more likely to dominate if child labor productivity is high. Inspecting (10) one also sees that the positive effect on schooling propensity is more likely to dominate at any level of γ if schooling is not very time-consuming, i.e. if τ is low.

While the model thus supports the conclusion that discriminating child labor income (at the extreme through an effective ban of child labor) increases the schooling propensity, it point also to one case when just the opposite holds true. Discriminating child labor deters schooling when school attending children continue to work and their productivity is so high (in setting free adult labor if they work at home) that their parents react on further increasing γ with leaving their children more time for leisure and school. Reducing γ under these circumstances, would lower the schooling propensity and would possibly induce some parents to withdraw their children from school.

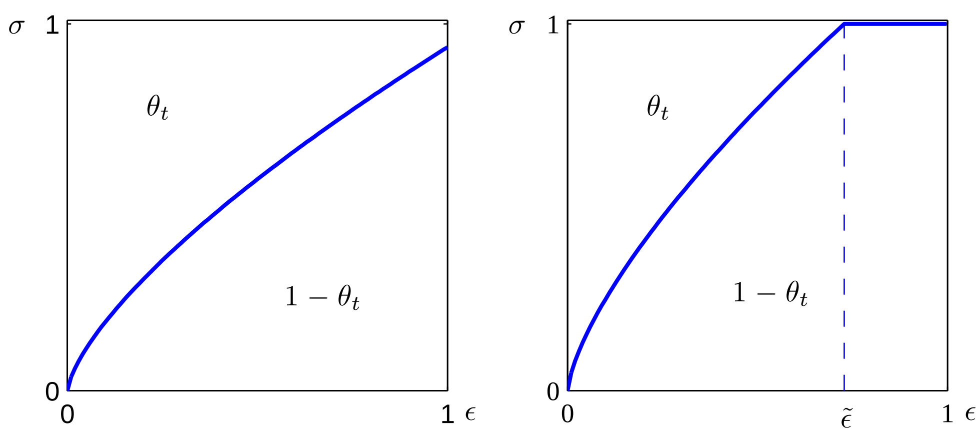
2.4. **How Many and Whose Children Are Not Going to School?** In this subsection we solve for the share of children not attending school when a given schooling norm S_t prevails in the community. The next next section will investigate social dynamics and long-run equilibria.

In order to distinguish between aggregate poverty (of the whole community or region) and idiosyncratic poverty (of a specific family i) we assume that wages are given by $w(i) = \epsilon(i)A$. Here, A is understood as an index of general productivity of the community and $\epsilon(i)$ is the poverty index of family i. In order to allow for an analytical solution we assume that the idiosyncratic part is uniformly and independently distributed within the unit interval, $\epsilon(i) \in (0, 1)$. Employing the new notation we can restate condition (8). The child of family i is not attending school if $A^{1-\alpha}E_{ij}/(S_t - \phi)\epsilon(i)^{1-\alpha} < \sigma(i)$. Diagrammatically speaking, the child of family i is not attending school if the family's (ϵ, σ) tuple lies above the following threshold.

$$\sigma(\epsilon) = \frac{A^{1-\alpha} E_{ij}}{S_t - \phi} \cdot \epsilon^{1-\alpha}, \quad i, j = W, N. \tag{11}$$

Because the right hand side is an increasing function of the poverty index ϵ, the model of collective behavior supports, in principle, the general assertion and frequently observed result that poverty affects education and that it are the children of the poorest families who are not attending school. However, the schooling decision depends also on the general level of community approval and the individual susceptibility to the evaluation by others. In particular threshold (11) embeds cases where some children of relatively rich but highly approval-dependent parents are not attending school if school participation is against the social norm of their neighborhood. On the other hand, it also embeds cases where children of relatively poor yet independently-minded parents are attending school because their parents do not care much about the prevalent anti-schooling attitude in their community.

Figure 1: School Attendance Threshold



Families are distinguished by poverty index $\epsilon \in [0, 1]$, with $\epsilon = 1$ identifying the richest parent, and by susceptibility of the parent to community approval $\sigma \in [0, 1]$, with $\sigma = 1$ reflecting the highest susceptibility. The share of children not attending school in period t is t is θ_t. If disapproval S_t gets sufficiently low all parents above a critical poverty level $\tilde{\epsilon}$ let their children attend school (right hand side, see text for details).

Figure 1 visualizes these result by showing the concave threshold (11) in an ϵ– σ–diagram. The area below the threshold gives the share of children who are attending school at day t. Likewise the area above the threshold given the share of children who are not attending school, denoted by θ_t. For a *given* value of σ it are always the children of poor parents who are not participating in school. But because σ is distributed independently from poverty, there may be also some rich parents who leave their children at home. If the anti-schooling attitude in the community S_t is sufficiently weak, however, the richer parents let their children attend school no matter how large their σ is. Formally, there exists a $\tilde{\epsilon}$ for which $\sigma(\tilde{\epsilon}) = 1$, i.e. $\tilde{\epsilon} = \left[(S_t - \phi)/E_{ij}\right]^{1/(1-\alpha)}/A$. For $\tilde{\epsilon}$ to be in the relevant range, i.e. $\tilde{\epsilon} < 1$, social disapproval of schooling has to be sufficiently small, $A^{1-\alpha}E_{ij} + \phi > S_t > \phi$, see the right hand side of Figure 1. If S_t is even lower and falls short of ϕ, disapproval turns into approval. Consequently all children are attending school and the threshold collapses to the origin.

The simple structure of the model now pays off because we can compute analytically the area below the threshold, and taking its complement to unity we get the share of children who are not attending school as a function of the prevailing social evaluation of schooling S_t for given schooling propensity E_{ij} and aggregate poverty A. Taking the piece-wise definition of the threshold into account we arrive at the following.

$$\theta_t = \theta(S_t) = \begin{cases} 1 - \frac{1}{2-\alpha}\frac{A^{1-\alpha}E_{ij}}{S_t-\phi} & \text{for } 1 \geq S_t \geq A^{1-\alpha}E_{ij} + \phi \\ \frac{1}{A}\left(\frac{S_t-\phi}{E_{ij}}\right)^{\frac{1}{1-\alpha}} & \text{for } A^{1-\alpha}E_{ij} + \phi \geq S_t \geq \phi \\ 0 & \text{for } S_t \geq \phi. \end{cases} \tag{12}$$

Inspecting (12) and using Lemma 2 and 3 proves the following comparative statics.

Proposition 1. *Assume schooling is feasible. The share of children not attending school in a community is increasing in social disapproval of a pro-schooling decision S_t. It is decreasing in stigma costs ϕ, in the time-cost of schooling τ, and in average productivity (income) of the community A.*

The share of children not attending school in a community is increasing in child labor productivity γ if γ is low ($\gamma < \bar{\gamma}$) and decreasing in child labor productivity if γ is high ($\gamma > \bar{\gamma}$).

The general state of children affects the slope of the threshold. Using Lemma 1 we verify that the slope of the threshold is the flattest for E_{WN}, which proves the following result

Proposition 2. *Assume schooling is feasible. For any given community evaluation of schooling S_t, the share of children not attending school is largest, if school attendance eliminates child labor.*

Recall that the child labor decision is not enforced. This means that if schooling is sufficiently time consuming, for example, because the next secondary school for girls is far away, such that

(7) does not hold for $h_t = 1$, parents are sufficiently empathetic that they do not demand any child labor if their daughter is attending school. Instead they prefer to leave her some spare time. However, in this case parents are also, ceteris paribus, more reluctant to let their daughter attend school in the first place. The schooling propensity is relatively low and thus a relatively large share of the families in the community prefers that their daughter does not attend school.

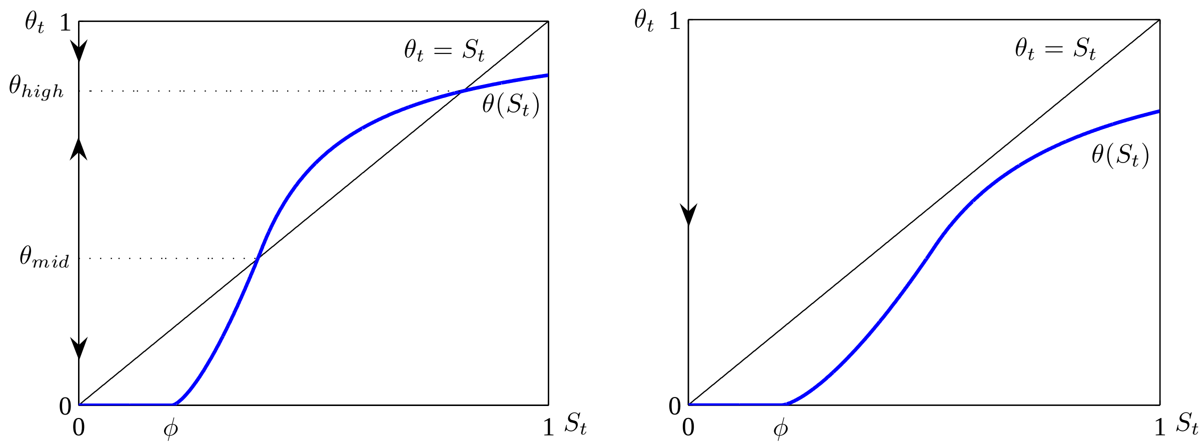
3. The Evolution and Persistence of Community Norms

3.1. **The Evolution of Community Approval.** A community's evaluation of behavior is not a given constant but evolves as a lagged endogenous variable depending on the observation of actual community behavior. The results obtained so far were just providing a snapshot of the socio-economy for given S_t. In order to proceed we assume that the strength of community disapproval of schooling depends positively on the share of children who were not attending school in the past. Let δ denote the time preference rate or rate of oblivion by which these observations are depreciated in mind so that approval is given by $S_t = (1-\delta)\sum_{i=0}^{\infty} \delta^i \theta_{t-1-i}$. Alternatively, this can be written as the period-by-period evolution of approval,

$$S_t = (1-\delta) \cdot \theta_{t-1} + \delta \cdot S_{t-1}. \tag{13}$$

A social equilibrium is obtained where the social norm, i.e. the level of community approval, stays constant, which requires $\theta_t = S_t$. Inspect (12) to see that $\theta(S_t)$ is continuous in $(0,1)$, equal to zero for $S_t \leq \phi$, convex in S_t in its middle part, concave in S_t in its upper part, and converging towards an upper bound less or equal to unity as S_t goes to unity. This implies that in a $S_t - \theta_t$ diagram, the $\theta(S_t)$ curve intersects the identity line either twice (as shown on the left hand side of Figure 2) or not at all (as shown on the right hand side). If the curve does not intersect at all, the only long-run equilibrium is at $S_t = 0$. A possibly observable anti-schooling attitude in the community is not persistent. Even if the community starts out at a situation of high disapproval and low attendance, there are always less children absent from school than needed to sustain the currently prevailing anti-schooling attitude ($\theta(S_t)$ is always below S_t). Thus over time, as the community travels down the $\theta(S_t)$ curve, more families jump on the schooling bandwagon, and the anti-schooling attitude in the community vanishes, eventually turning into a pro-schooling attitude.

Figure 2: Community Dynamics of Schooling and Schooling Norm



S_t denotes community disapproval of schooling at day t evolving according to (13). The share children not attending school at day t is given by $\theta(S_t)$ according to (12). The left-hand side identifies three equilibria, the outer ones being locally stable and separated by θ_{mid}. The right hand side shows the case where $\theta = 0$, i.e. school attendance of all is the only equilibrium.

If the $\theta(S_t)$ curve intersects twice it identifies three social equilibria. The equilibrium in the middle, θ_{mid} is unstable. If a slightly larger share of children is not attending school, i.e. if θ_t is to its right, then the $\theta(S_t)$ curve lies above the identity line implying that the prevailing anti-schooling attitude in the community leads to further non-attendance, raising the anti-schooling attitude even further etc. Likewise for slightly higher school attendance, i.e. for θ_t slightly below θ_{mid}, social dynamics drive θ down until all children are attending school. Employing analogous arguments θ_{high} and θ_{low} (at zero) are identified as locally stable equilibria.

3.2. **Persistence of an Anti-Schooling Norm.** An anti-schooling norm exists, if there are three equilibria and the economy is at or converges towards θ_{high}. Since θ_{high} – if it exists – is located at the upper intersection of the concave branch with the identity line, we insert $S_t = \theta_t$ into the upper part of (12) and solve for the largest root, which is, if it exists, at

$$\theta = \theta_{high} \equiv \frac{1}{2}(1+\phi) + \sqrt{\frac{1}{4}(1-\phi)^2 - \frac{E_{ij}A^{1-\alpha}}{(2-\alpha)}}. \tag{14}$$

Inspect (14) to conclude that θ_{high}, if it exists, is always larger than one half. Since a majority of the girls in the community is not attending school and since this situation is self-sustaining by the prevalent anti-schooling attitude, we can indeed speak of a persistent anti-schooling norm. For such a norm to be possible to exist the radicand in (14) has to be positive implying that θ_{high}

exists if[3]

$$E_{ij} A^{1-\alpha} < \frac{1}{4}(1-\phi)^2(2-\alpha), \quad i,j = W,N. \tag{15}$$

Note that existence of θ_{high} is not sufficient to explain the actual existence of an anti-schooling norm. Actual existence additionally requires that the observed share of children not attending school is above θ_{mid}. If "somehow" school attendance has become sufficiently high ($\theta < \theta_{mid}$) social dynamics lead the community towards $\theta_{low} = 0$ irrespective of the fact that poverty and preferences would, in principle, also support the θ_{high} equilibrium. This feature of *local* stability will be exploited in the subsection on two-step policies.

Inspecting condition (15) and employing Lemma 2 provides the following result.

PROPOSITION 3. *A social equilibrium supporting an anti-schooling norm exists if average poverty in the community is sufficiently high (A sufficiently low), if parents' autonomous evaluation of education is sufficiently low (a is sufficiently low), if maximum approval that can be generated by society (ϕ) is sufficiently small, and if schooling is sufficiently time intensive (τ is sufficiently high).*

Applying Lemma 1 on condition (15) verifies the following result.

PROPOSITION 4. *Ceteris paribus, an anti-schooling norm is more likely to persist, if school attendance – if it occurs – eliminates child labor.*

Using Lemma 3 we prove the non-monotonous effect of productivity on norm prevalence.

PROPOSITION 5. *Assume an anti-child labor norm prevails. a) A sufficiently large decrease of child labor productivity γ can initiate a transition towards a pro-schooling norm and mass school attendance if school attending children are not working or if they are working and are sufficiently unproductive, $\gamma < \bar{\gamma}$. b) If, however, $\gamma > \bar{\gamma}$, decreasing child labor productivity is detrimental to school attendance and amplifies the prevailing anti-child labor norm.*

Diagrammatically, a lower γ increases the schooling propensity E_{ij} and shift the $\theta(S_t)$–curve downwards in case a) but reduces the schooling propensity in case b) and shifts the $\theta(S_t)$–curve

[3]We neglect the degenerate case where the radicand in (14) is zero. In that case θ_{high} collapses with θ_{mid}. Linguistically we distinguish between an anti-schooling attitude and an anti-schooling norm. For example, if the community starts out at high S_t on the right hand side of Figure 2, we would say there exists an anti-schooling attitude. This attitude however does not classify as a norm because it is not persistent and transforms endogenously into social approval of schooling. On the contrary, if the the community starts out at high S_t on the left hand side of Figure 2, it converges towards a persistently strong disapproval of schooling. This feature of local stability of the attitude justifies to characterize it as a norm.

even more upwards. Economically, income from child labor is more essential in case b). The negative effect of decreasing γ on lower family income dominates the positive effect on lower opportunity cost of time spent at school. The policy implication is that if child labor productivity is not too high, a discrimination of child work, for example, a (partly) successful child labor ban that effectively drives down child labor productivity, is helpful to increase attendance at school. However, if child labor productivity is high and an anti-schooling norm prevails although school attending children would continue to work and although schooling is not very time intensive, then reducing the child's contribution to family income is detrimental to schooling via the poverty channel and will aggravate the anti-schooling norm.

More generally, anything that shifts down the $\theta(S_t)$–curve sufficiently far such that the intersection θ_{high} ceases to exist, initiates social dynamics towards high attendance at school. This could, in principle, be brought about by an increase of average productivity of the community (A), for example through widespread introduction of fertilizer or irrigation, or by lowering the time cost of schooling (τ), for example, by providing bus service or establishing a local school, Yet, if such a policy shock is not sufficiently strong, the effect on school attendance will be only marginal since θ_{high} continues to exist and the force of social dynamics is not positively exploited.

3.3. **A Two Step Policy to Eliminate Child Labor.** We next investigate how a two-step policy can successfully address both high attendance at school and eliminating child labor. The policy exploits the fact that once a pro-schooling norm has been established by a sufficiently strong downward shift of the $\theta(S_t)$–curve (from left to right hand side in Figure 2) and the entailed convergence towards θ_{low}, a second upward shift of the $\theta(S_t)$–curve (back from right to left hand side of Figure 2) does *not* trigger a convergence back towards θ_{high} and a relapse of the community into the anti-schooling norm.[4] Sufficient to sustain the pro-schooling norm is an arbitrarily small $\phi > 0$, which, diagrammatically, produces a positive intersection of the $\theta(S_t)$–curve with the abscissa. In words, robustness of a once attained pro-schooling norm against parameter change requires a (arbitrarily small) stigma cost suffered by the first parent whose daughter is not allowed to attend school when everybody else's daughters are attending school. This appears to be a relatively mild requirement.

[4]The only way to relapse into θ_{high} would be a spontaneous non attendance at school of a share of $\theta > \theta_{mid}$ of children. But the theory developed here provides no clue for why such a massive spontaneous move should occur. While "stability" characterizes the impact of the changes of the endogenous variable θ_t on the equilibrium, the term "robustness" refers to the impact of parameter changes on an attained equilibrium.

Proposition 6. *If $\phi > 0$, and the equilibrium of schooling of all ($\theta_{low} = 0$) is attained, the socio-economy remains there irrespective of any change of the model's parameter values.*

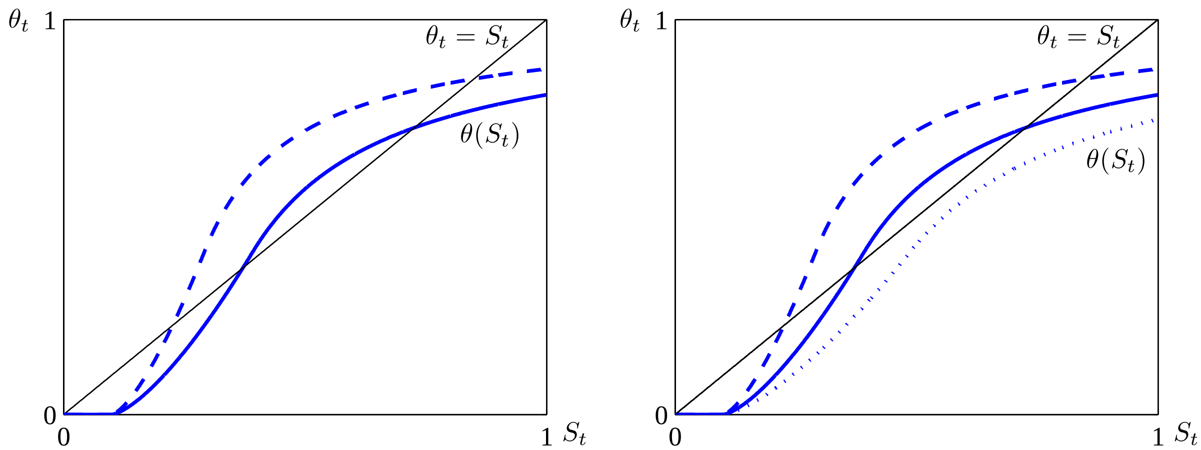
As confirmed in practice, once a pro-schooling norm is established in a society, it is very robust. Robustness can be exploited by policy in a two-step way, reminiscent of the Chinese proverb "To reach your goal take a detour". So far we have established the result that increasing the time requirement for schooling τ reduces child labor and eventually eliminates it, if children are attending school. It could thus be tempting for policymakers to increase τ in order to reduce child labor. The "if", however, is rather big, since we have also established the result that increasing τ reduces the schooling propensity and amplifies through this channel a prevailing anti-schooling norm and leads to less school attendance.

The detrimental effect of prolonging time at school is exemplarily shown at the left hand side of Figure 3 for a parameterized model. Given the parameter values specified below the Figure, school attendance takes away 35 percent of the child's day and, according to condition (7), school attending children continue to work (though less than children not attending school). However, if a strong anti-schooling norm prevails, most parents let their child/daughter not attend school. The $\theta(S_t)$ curve (represented by the solid line) supports an equilibrium θ_{high}. Assume that the community is initially at or close to θ_{high}.

Making the school day longer, now taking away 50 percent of the child's day would indeed eliminate child labor if children were sent to school. Parents are sufficiently empathetic to leave their children the remainder of the day for leisure and play. However, given an anti-schooling norm prevailing under the $\tau = 0.35$ policy, the higher time cost of schooling actually causes some further parents not to let the girl attend school. The anti-schooling situation aggravates and the $\theta(S_t)$–curve shifts upwards (shown by dashed lines), eventually leading to a higher share of girls not attending school and therewith – since girls not attending school work more – actually leading to *more* child labor than at the original equilibrium.

The right hand side panel of Figure 3 visualizes the two-step policy. In the first step the time cost of schooling is sufficiently strongly *reduced*, for example, by introducing a shuttle service to school or by implementing a local school. Attending school now takes only away 20 percent of the girls' days, causing some parents to let their daughters attend school. The $\theta(S_t)$–curve shifts downwards (shown by the dotted line). While not particularly large as such the additional schooling is sufficiently strong to eliminate the θ_{high} equilibrium. The current school attendance

Figure 3: One-Step and Two-Step Attempts to Eliminate Child Labor through Schooling



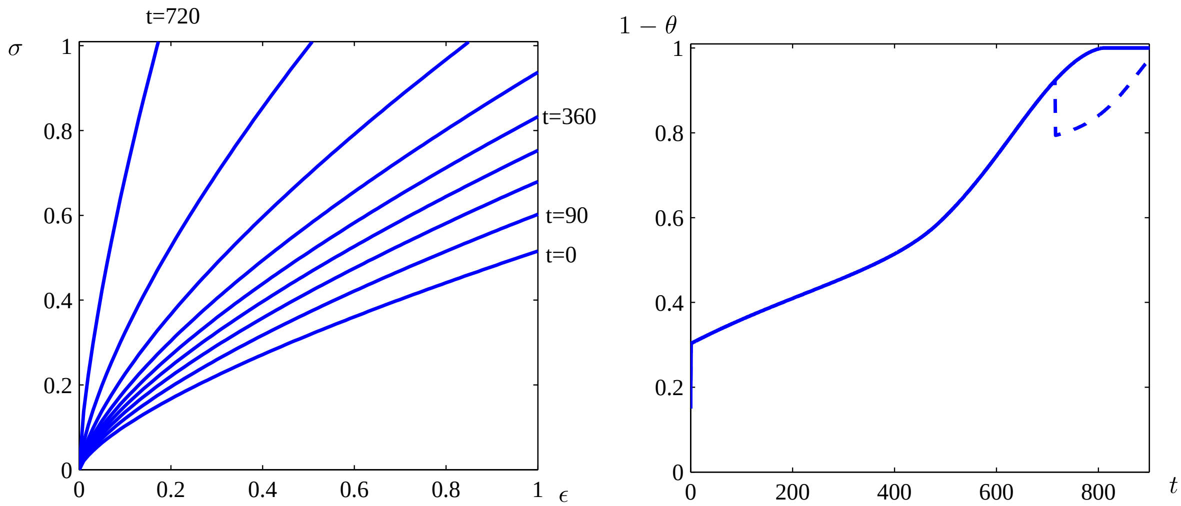
Parameters: $A = 1$, $\alpha = 0.3$, $\gamma = 0.7$, $a = 0.5$, $\phi = 0.1$. Initially $\tau = 0.35$ (solid lines). Left hand side panel: unsuccessful attempt to eliminate child labor by implementing $\tau = 0.5$. Right hand side panel: successful policy implements $\tau = 0.2$ first (dotted line) and then for $\theta < \theta_{mid}$ it implements $\tau = 0.5$ (dashed line).

rate does not support the anti-schooling equilibrium any longer. At the next day, community disapproval of schooling will be somewhat lower. Social dynamics are set in motion and the community moves along the dotted curve towards θ_{low}. With time passing, more and more parents are caused to let their girls attend school thereby further mitigating the anti-schooling attitude and turning it eventually into a pro-schooling norm.

At the time when school attendance has become sufficiently high (precisely when less than θ_{mid} children are not attending school), the movement toward "education for all" has become robust even against high time intensity of schooling. The $\tau = 0.5$ policy can now safely be implemented, for example, by extending the curriculum and/or by having lessons in the afternoon. Diagrammatically, the upward shift of the $\theta(S_t)$–curve (dashed line again) does not prevent that social dynamics are driving the community towards $\theta_{low} = 0$. Even under the increased time costs, more and more parents let their children attend school, which is the behavior that becomes increasingly supported by community approval. But, facing the higher time cost of schooling, parents also no longer prefer their children to work; condition (7) is not fulfilled anymore. Instead parents prefer to leave their children some spare time. Child labor is effectively abandoned.

Figure 4 shows the ongoing social dynamics. The right hand side shows the threshold from Figure 1 for parameters from the above example, and $\delta = 0.98$. The threshold is evaluated at different points in time, every 90 days. Recall that the area below the threshold identifies the share of girls who are attending school. Initially, at $t = 0$, this area is quite small, a majority of

Figure 4: Dynamics of School Attendance



Parameters as for Figure 3 and $\delta = 0.99$. The community is initially stuck at θ_{high}, where 72 percent of the girls were not attending school. Then, at $t = 0$, time cost of schooling is reduced from 0.35 to 0.2. Dotted line: at $t = 700$ time cost of schooling is raised to 0.5, which effectively elimates child labor.

72 percent of girls is not attending school, and the community is stuck at θ_{high}. When the school day becomes less time consuming, some parents (mostly richers ones with high ϵ–index) decide to let the girl attend school. The next snapshot of parameterized Figure 1 is taken after 90 days, i.e. approximately after a quarter of a year. Since the community observes that in its recent past more children were attending school, the anti-schooling attitude has become smaller and as a consequence even more parents are allowing their girls to attend school. After about a year and half S_t has moved down to the convex part of the $\theta(S_t)$–curve and $\sigma(\epsilon)$-threshold intersects the unity line. The richer families in the community now let their girls attend school irrespective of social approval such that social dynamics get even more momentum. More and more of the poorer families let their girls attend school until finally after about two years (precisely 800 days) all girls are attending school.

The solid line on the right hand side of Figure 4 shows the implied evolution of school attendance over time. Dotted lines show the same experiment but now τ is increased to 0.5 after S_t has fallen below $\theta_{mid} \approx 0.22$ (which is reached after about 700 days). School attending children are now no longer working. Because of the increased opportunity cost of schooling the propensity of schooling is lower and some members of the community, mostly poorer ones, are initially reacting by withdrawing their daughters from school and attendance declines. Now, however, social dynamics towards θ_{low} are just too strong for these parents to keep their girls out of school permanently. With more and more girls attending school the community evolves towards an equilibrium with education for all and without child labor.

4. Targeted Transfers

Because the proposed model takes heterogeneity of households with respect to poverty and school attendance into account it is well suited for a theoretical analysis of targeted transfer programs, like Progresa, Bolsa Escola, or Food for Education. These programs are characterized by two salient features. Subsidies are paid contingent on a certain behavior, usually children showing up at school, and they are addressed to only a subset of a community, usually the poorest families. Particularly interesting here is it to provide a theoretical foundation for the success of targeted transfers programs through social norms and collective behavior.[5]

To this end it will be explained how transfers received by only the poorest families of community manage to affect the schooling attitude of the community at large through the spillover of behavior from targeted families to others. From a theoretical viewpoint we investigate a problem of optimal mechanism design. We investigate how many families should receive subsidies and how generous the payment should be in order to initiate a successful transition towards "education for all" at the lowest possible costs.

Suppose that the poorest families of a community are selected to receive an income subsidy of size s contingent on their children going to school. Let p denote the number, i.e. community share, of eligible families. Thus, if the program is entirely effective and triggers school attendance of all children, total costs are given by $\int_0^p s \mathrm{d}w = ps$. To relate the model better to the actually existing transfer programs it is useful to think of the model period as a month. Then, $1-\theta_t$ is the school attendance rate in month t and s is the monthly subsidy received by an eligible family when their daughter has attended school.

In order to reduce complexity and case differentiation we focus on the E_{WN} case, i.e. parameters are such that girls who are not attending school are working but girls who are attending school are not working. While the "optimistic " assumption that schooling replaces child work is indeed supported by some studies it is clearly not generally true.[6] The assumption is thus rather justified

[5]Empirical studies accounting for the success of targeted tranfer programs are, among others, Ravallion and Wodon, 2000, Skoufias and Parker, 2001, Cardoso and Souza, 2004, Schultz, 2004, Bobonis and Finan, 2009, and Lalive and Cattaneo, 2009.

[6]Skoufias and Parker (2001) find for Progresa that schooling was largely replacing child work (on the market for boys, in the household for girls). Ravallion and Wodon (2000) found a much smaller impact of schooling on child labor for the Food for Education Program.

on theoretical grounds. The major objective of this section is to explain the large success of targeted transfer programs through community effects. Since it has been shown (Lemma 1) that the schooling propensity is *lowest* in E_{WN} case, we expect – for given expenditure – a targeted transfer program to be the least effective with respect to schooling in the E_{WN} case. This means that we are investigating with respect to induced schooling a "pessimistic" scenario and expect that the model predicts even more success if school attending children continue to work (the E_{WW} case) or if not school attending children are not working (the E_{NN} case). For convenience we also normalize average productivity ($A = 1$) such that $w(i) = \epsilon(i)$.

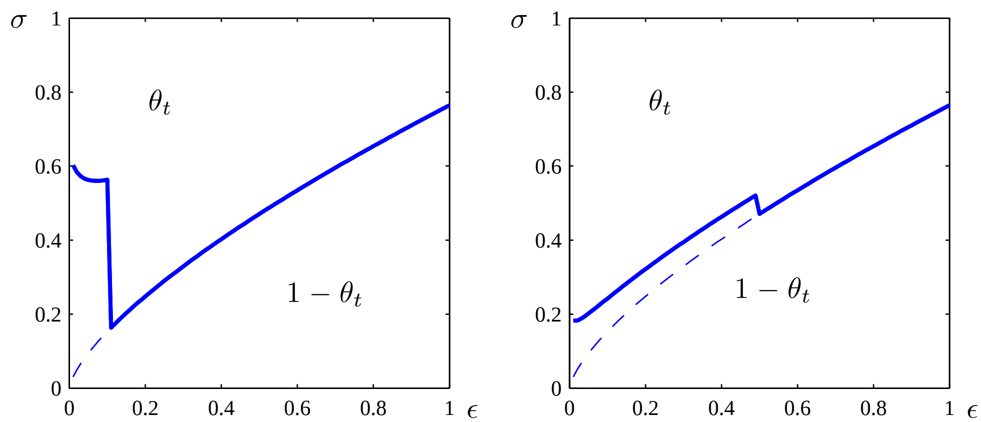
Keeping everything else from the basic model, autonomous utility of a parent whose girl is attending school is given by $u_t^a(i) = (\epsilon(i) + s(i))^{1-\alpha}(1-\tau)^\alpha(1+a)$. Here $s(i) = s$ if $\epsilon(i) < p$ and $s(i) = 0$ otherwise, reflecting the fact that only the poorest families with income below p are eligible to receive transfers. If the parent is not eligible, or if the parent is eligible but the daughter is nevertheless not attending school, utility is, as before, $u_t^a = (1-\alpha)^{1-\alpha}(1+\gamma)(\alpha/\gamma)^\alpha\epsilon(i)$. Comparing utilities, the girl of parent i is not attending school if

$$(\epsilon(i) + s(i))^{1-\alpha}(1-\tau)(1+a)^\alpha - \lambda\epsilon(i)^{1-\alpha} \leq \sigma(i)\cdot(S_t - \phi), \quad \lambda \equiv (1-\alpha)^{1-\alpha}(1+\gamma)\left(\frac{\alpha}{\gamma}\right)^\alpha. \quad (16)$$

Because of the piece-wise definition of eligibility to subsidies, the threshold of Figure 1 becomes also piecewise continuously defined with a kink where $\epsilon = p$, i.e. where eligibility ends. The size of the subsidy determines the size of the kink at p. Figure 5 exemplarily shows the threshold for two quite different designs of the targeted transfer program which both entail the same cost if all eligible parents would sent their girls to school. In both panels dotted lines re-iterate the threshold for a community without targeted transfers. The left hand side of Figure 5 shows a design that targets a relatively small part of the community ($p = 0.1$, i.e. ten percent of all families) with a relatively generous subsidy ($s = 0.1$). The right hand side shows the opposite design. A relatively large part of the community ($p = 0.5$) is addressed with a relatively small subsidy ($s = 0.02$).

The area between the solid line and the dashed line captures the share of children who are caused to attend school by the targeted transfer program. While the design on the left hand side is relatively successful among the relatively few addressed poorest families, the design on the right hand side has relatively little success at each addressed income level but the range of income levels addressed is much larger. Eyeballing, it seems to be impossible to decide which design is

Figure 5: School Attendance: Two Alternative Targeted Transfer Programs



Model parameters are $A = 1$, $a = 1.5$ and $\alpha = 0.3$, $\gamma = 0.4$, $\phi = 0.2$, and $S_t = 0.4$. Left panel: $s = 0.1$ and $p = 0.1$. Right panel: $s = 0.02$ and $p = 0.5$. The dotted line shows the threshold without targeted transfers.

more successful, i.e. which has created a larger area between the solid and dashed threshold. More importantly, both Figures show just a snapshot for one particular anti-schooling attitude ($S_t = 0.4$) and the thresholds and therewith the impact of the subsidy moves with changing S_t.

For a definite judgement of success we have to consider the long-run social equilibrium where $S_t = \theta_t$. There, it is particular interesting whether a transfer program creates enough behavioral change to initiate a transition towards the "education for all" equilibrium. For that purpose we begin with step-wise integrating the area below the threshold to obtain the $\theta(S_t)$–curve.

$$\theta_t = \theta(S_t) = 1 - \frac{1}{2-\alpha}\frac{E_S}{S_t - \phi} - (1 - \tilde{\epsilon}) \quad \text{where} \tag{17}$$

$$E_S \equiv (1-\tau)^\alpha (1+a) \left[(p+s)^{2-\alpha} - s^{2-\alpha} - p^{2-\alpha} + \tilde{\epsilon}^{2-\alpha}\right] - \lambda \tilde{\epsilon}^{2-\alpha}$$

$$\tilde{\epsilon} \equiv \min\left\{ \left[\frac{S_t - \phi}{(1+a)(1-\tau)^\alpha - \lambda}\right]^{\frac{1}{1-\alpha}}, \ 1 \right\}.$$

The compound parameter E_S inherits the function of the schooling propensity from the basic model. Differentiating it with respect to the new policy parameters s and p and using the result in (17) proves the following insight.

Lemma 4. *The schooling propensity E_S is increasing in the generosity of subsidies s and in the share of targeted recipients p. This implies that for any given anti-schooling attitude in the community S_t the share of children attending school $1 - \theta_t$ is increasing in the generosity of subsidies s and the share of targeted recipients p.*

As expected, and as suggested by Figure 5, a larger scale of targeted transfers, $p \cdot s$, always increases school participation on the margin. But which transfer design is successful in the extensive sense, i.e. in managing a transition towards "education for all"?

Inserting the equilibrium condition $S_t = \theta_t = \theta$ into (17), noting that θ_{high}, if it exists, is assumed along the concave branch of the $\theta(S_t)$ curve where $\tilde{e}psilon = 1$, and solving for θ provides, analogously to (14), $\theta = \theta_{high} \equiv \frac{1}{2}(1+\phi) + \sqrt{1/4(1-\phi)^2 - E_S/(2-\alpha)}$. From this follows that a locally stable anti-schooling norm θ_{high}exists if

$$E_S < \frac{1}{4}(1-\phi)^2(2-\alpha). \tag{18}$$

Applying Lemma 4 we conclude that a larger p or s reduces the set of parameters for preferences and child contributions to family income (a, α, and γ) for which θ_{high} exists, which proves the following result.

Proposition 7. *If transfers s are sufficiently high and the share of targeted recipients p is sufficiently large, a prevailing anti-schooling equilibrium ceases to exist.*

We can now address the initial question for the best design of targeted transfer policies, i.e. which combination of p and s moves a socio-economy out of θ_{high} at minimum costs. Formally, we are interested in the solution of the following problem.

$$\begin{aligned} &\min_{p,s} p \cdot s \\ s.t. \quad &E_S = \frac{1}{4}(1-\phi)^2(2-\alpha) \end{aligned} \tag{19}$$

The solution provides a tangency point of the $\theta(S_t)$ curve with the identity line implying that an arbitrarily small further increase of s (or p) initiates a transition towards "education for all". From the first order conditions we get the following result.

Proposition 8. *If a targeted transfer program is optimally designed, the ratio of eligible recipients relative to generosity, p/s, is equal to the inverse ratio of changes of the education propensity that these policies bring about, i.e.*

$$\frac{p}{s} = \frac{\partial E_S/\partial s}{\partial E_S/\partial p} = \frac{(p+s)^{1-\alpha} - s^{1-\alpha}}{(p+s)^{1-\alpha} - p^{1-\alpha}} \quad \Rightarrow s = p. \tag{20}$$

Generosity and eligibility affect the average value of education symmetrically so that the optimal design requires equality of p and s. In other words, the model suggests that a schooling project that targets a large share of the population with small transfers (or vice versa) is ill-designed. Either it does not move the economy out of the bad equilibrium or, if it is successful, there exists a less expensive transfer design with the same effect. The symmetry result originates from the idiosyncratic distribution of σ across households. Intuitively, a small share of parents is already "convinced" by a small subsidy to let their girls attend school. But if the program aspires to change behavior of many, including those with high σ, i.e. those whose decision is highly dependent on the prevailing anti-schooling norm, the subsidy has to be larger. Thus s and p have to move "in sync". If σ were not uniformly distributed the symmetry result would not appear in such a drastic form. But as long as some people depend stronger on social approval than others we expect the main message to continue hold: for an efficient program the size of the subsidy is increasing in the population share of targeted families.

Inserting (20) into E_S in (18) and then into (19) and solving for s we obtain the optimal size of the subsidy.

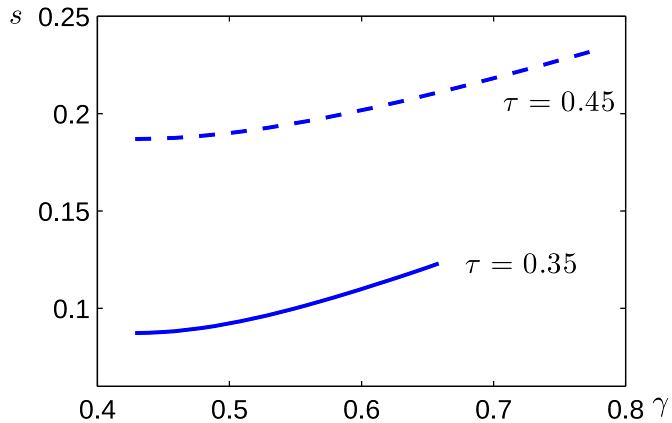
$$s^* = \left\{ \left[\frac{(1-\phi)^2(2-\alpha)+4\lambda}{4(1-\tau)^\alpha(1+a)} - 1 \right] \cdot \frac{1}{2(2^{1-\alpha}-1)} \right\}^{\frac{1}{2-\alpha}} \tag{21}$$

Taking first derivatives proves the following result on comparative statics.

Proposition 9. *The optimal subsidy (that initiates an escape from an anti-schooling equilibrium at minimum costs) is increasing in the time costs of schooling τ and in child labor productivity γ.*

An immediate corollary for policy is that a successful targeted transfer program can be accomplished at lower costs when it is accompanied by a reduction of the time costs of schooling, e.g. by providing bus transfer to school. The two-step policy suggested above can be modified towards a two-step-two-instruments policy. In the first step a reduction of time costs of schooling and a targeted transfer program initiates the escape from θ_{high}. Once the anti-schooling attitude in the community has become sufficiently low, subsidies can be withdrawn and time costs of schooling could be raised without negative repercussions. The social dynamics are strong enough to draw the community towards the "education for all" equilibrium.

Figure 6: The Power of Targeted Transfers



$A = 1$, $a = 0.5$, $\phi = 0.1$, $\alpha = 0.3$.

The proposed theory supports the empirical finding that targeted transfers are not only changing behavior of targeted families but also of ineligible families. We finally demonstrate the potential power of these collective dynamics with a parameterized example. Parameters are chosen such that they support a persistent anti-schooling norm and such that attending school and working are mutually exclusive. Figure 6 shows for alternative time costs of schooling τ and alternative child labor productivity γ the minimum subsidy needed to initiate a permanent transition towards a pro-schooling attitude and education for all. Condition (7) defines the feasible range of γ under the assumptions made. At the lower bound, $\gamma = \alpha/(1-\alpha)$. If γ were lower, girls not attending school would stop working. At the upper bound, $\gamma = \alpha/[(1-\alpha)(1-\tau)]$, if γ were higher, school attending girls would start working.

The most important impression from the Figure is how small the efficient transfer program is in respect to both size and scope. First consider size of the subsidy. For the feasible range of γ, working children contribute between 40 and 80 percent of their parents' income to family income (potentially by setting free extra time of their parent). A value of s between 0.1 and 0.2 means that the efficient subsidy that motivates parents to let the girl attend school and to forego the child contribution to family income is between 10 and 20 percent of average adult income in the community. On average, subsidies are thus much smaller than foregone family income.

Second consider scope, recalling that $p = s$ for the efficient program. The efficient program thus addresses between 10 and 20 percent of the families in the community. We know from theoretical analysis that at θ_{high} a majority of children are not attending school. In the example, it are about

70 percent for $\gamma = 0.6$ when $\tau = 0.35$ and about 80 percent when $\tau = 0.45$. It is thus sufficient (and efficient) to target only a minority of the community in order to change the behavior of a majority. Of course this change of behavior of a majority will not occur spontaneously after introduction of the subsidies. Depending on the size of δ it takes time until people update their believes about social approval and initiate the required collective dynamics. Finally, the example demonstrate the power of short school days. A reduction of the time costs of schooling by 10 percentage points halves the size of the required subsidy and (since $p = s$) quarters total cost of the transfer program.

5. Final Remarks

In this paper we have investigated the effect of community dynamics and social approval in interaction with aggregate and idiosyncratic poverty on school attendance and child labor. While maintaining to deliver the well-known association between schooling and individual and aggregate poverty the theory is additionally able to resolve the puzzle from the empirical literature why communities of the same poverty can display very different schooling behavior and why targeted transfer programs are so successful. It is explained how policy can exploit the endogenous formation of norms and can initiate a bandwagon effect toward an "education for all" equilibrium. The theory strongly supports targeted transfer programs by showing that relatively small behavior-contingent transfers to a minority can change the schooling behavior of a community at large.

The theory is not claiming to replace poverty-based and other available approaches towards child labor and schooling. Instead, credit constraints (Baland and Robinson, 2000), bonded child labor (Basu and Chau, 2004), political economy elements (Doepke and Zilibotti, 2005), fertility (Hazan and Berdugo, 2002), and child mortality (Strulik, 2004) constitute conceivable future extensions of the community-norms based model of schooling and child labor.

Appendix

Appendix

Derivation of (6) and (8). Insert (2) into (5).

$$u_t(i) = [(1+\gamma\ell_t(i))w_t(i)]^{1-\alpha}\,(1-\ell_t(i)-\tau\cdot h_t(i))^{\alpha}(1+a\cdot h_t(i)) + h_t(i)\cdot\sigma(i)\cdot(\phi - S_t). \quad \text{(A.1)}$$

The first order condition with respect to $\ell_t(i)$ is

$$(1-\alpha)(1-\ell_t(i)-\tau h_t(i))\gamma w_t(i) - \alpha(1+\gamma\ell_t(i))w_t(i) = 0.$$

Solving for $\ell_t(i)$ provides

$$\ell_t(i) = (1-\alpha)(1-\tau h_t(i)) - \frac{\alpha}{\gamma},$$

which is the interior solution in (6). Plugging it into (2) provides consumption

$$c_t(i) = (1-\alpha)\left[1+\gamma(1-\tau h_t(i))\right] w_t(i).$$

Plugging $\ell_t(i)$ and $c_t(i)$ into $u_t^a(i)$ and simplifying provides

$$u_t^a(i) = (1-\alpha)\left(\frac{\alpha}{\gamma}\right)^{\alpha}\left[1+\gamma(1-\tau h_t(i)\right](1+a)w_t(i)^{1-\alpha}.$$

Plugging this into $u_t(i)$ and evaluating $u_t(i)|h_t(i) = 1 - u_t(i)|h_t(i) \leq 0$ provides

$$(1-\alpha)^{1-\alpha}\left(\frac{\alpha}{\alpha}\right)^{\alpha}\{[1+\gamma(1-\tau)a-\gamma\tau]\}\,w_t^{1-\alpha} - \sigma(i)(S_t-\phi) \leq 0,$$

which is (8) for $E_{ij} = E_{WW}$. For $\ell_t(i) =$ we get $c_t(i) = w_t(i)$. Plugging these into $u_t^a(i)$ provides

$$u_t^a(i) = w_t(i)^{1-\alpha}(1-\tau h_t(i))^{\alpha}(1+ah_t(i)).$$

Computing the above utility difference for the case where $\ell_t(i)$ irrespective of $h_t(i)$ provides

$$\left[(1-\tau)^{\alpha}(1+a)-1\right]w_t(i)^{1-\alpha} - \sigma(i)(S_t-\phi) \leq 0$$

which is (8) for $E_{ij} = E_{NN}$. Finally computing the utility difference when $\ell_t(i) > 0$ for $h_t(i) = 0$ and $\ell_t(i) = 0$ for $h_t(i) = 1$ provides (8) for $E_{ij} = E_{WN}$.

Lemma 1. Begin with showing $E_{WN} < E_{NN}$. For E_{WN} to materialize according to condition (7), $(1-\alpha)\gamma/\alpha > 1$. Raising both sides of the expression to the power of $-\alpha$ we get

$$(1-\alpha)^{-\alpha}\left(\frac{\alpha}{\gamma}\right)^{\alpha} > 1. \tag{A.2}$$

Condition (7) also implies $\gamma(1-\alpha)-\alpha > 0$. Expanding both sides of the expression with 1 we get

$$(1-\alpha)(1+\gamma) > 1 \tag{A.3}$$

Computing the product of the left hand sides of (A.2) and (A.3) we observe

$$(1-\alpha)^{1-\alpha}\left(\frac{\alpha}{\gamma}\right)^{\alpha}(1+\gamma) > 1.$$

Now, observe that the left hand side of the above expression is the second term in E_{WN}. Thus the second term of $E_{WN} > 1$ while the second term of $E_{NN} = 1$. Since the second term is subtracted from the first term and the first term is identical for both schooling propensities, $E_{WN} < E_{NN}$.

In order to show that $E_{WW} \geq E_{WN}$ let τ_a denote the value of τ that has to hold for E_{WW} to materialize, i.e. according to condition (7)

$$1 - \tau_a > \frac{\alpha}{(1-\alpha)\gamma}. \tag{A.4}$$

Likewise, let τ_b denote the value of τ that has to hold for E_{WN} to materialize, i.e. according to condition (7)

$$1 - \tau_b < \frac{\alpha}{(1-\alpha)\gamma}. \tag{A.5}$$

Taking the difference of the schooling propensities we see that for $E_{WW} \geq E_{WN}$

$$(1-\alpha)^{1-\alpha}\left(\frac{a}{\gamma}\right)^{\alpha}[1+\gamma(1-\tau_a)] \geq (1-\tau_b)^{\alpha}.$$

Because of condition (A.5) a sufficient condition for this to be true is

$$(1-\alpha)^{1-\alpha}\left(\frac{a}{\gamma}\right)^{\alpha}[1+\gamma(1-\tau_a)] \geq \left(\frac{\alpha}{\gamma}\right)^{\alpha}(1-\alpha)^{-\alpha} \quad \Rightarrow \quad (1-\alpha)\left[1+\gamma(1-\tau_a)\right] \geq 1.$$

Because of condition (**??**) a sufficient condition for this to be true is

$$(1-\alpha)\left[1+\gamma\frac{\alpha}{(1-\alpha)\gamma}\right] \geq 1 \quad \Rightarrow \quad 1 \geq 1,$$

which is true.

Lemma 3. Take the derivative to see that the sign of $\partial E_{WN}/\partial\gamma$ is equal to the sign of

$$-\left\{-\alpha\gamma^{-\alpha-1}(1+\gamma)+\gamma^{-\alpha}\right\} = \gamma^{-\alpha-1}\left[\alpha(1+\gamma)-\gamma\right]$$

Yet for E_{WN} to materialize, condition (7) has to hold when $h_t = 0$ demanding that

$$\gamma > \frac{\alpha}{1-\alpha} \quad \Rightarrow \quad \alpha-(1-\alpha)\gamma < 0 \quad \Rightarrow \quad \alpha(1+\gamma)-\gamma < 0,$$

implying $\partial E_{WN}/\partial\gamma < 0$.

Take the derivative to see that the sign of $\partial E_{WW}/\partial\gamma$ is equal to the sign of

$$\gamma^{\alpha-1}\left\{-\alpha\left[1+\gamma(1-\tau)\right](1+a)+\alpha(1+\gamma)+\gamma(1-\tau)(1+a)-\gamma\right\}$$

Simplifying one see that the sign of the above expression equals the sign of

$$(1-\alpha)\gamma\left[(1-\tau)(1+a)-1\right]-\alpha a,$$

which is positive if $\gamma > \bar{\gamma}$ and negative if $\gamma < \bar{\gamma}$.

Derivation of (12). Start with noting that $\theta_t = 0$ for $S_t < \phi$. Then Integrating the area below the threshold (11) for $S_t > \phi$.

$$1-\theta = \int_0^{\tilde{\epsilon}} \epsilon^{1-\alpha}A^{1-\alpha}\frac{E_{ij}}{S_t-\phi}\mathrm{d}\epsilon + (1-\tilde{\epsilon}) = \frac{A^{1-\alpha}}{S_t-\phi}\left|\frac{1}{2-\alpha}\epsilon^{2-\alpha}\right|_0^{\tilde{\epsilon}} + (1-\tilde{\epsilon}).$$

For $S_t > A^{1-\alpha}E_{ij} + \phi$, $\tilde{\epsilon} = 1$ and the integral is obtained as

$$1 - \theta = \frac{A^{1-\alpha}E_{ij}}{(S_t - \phi)(2-\alpha)}.$$

For $\phi < S_t < A^{1-\alpha}E_{ij} + \phi$, $\tilde{\epsilon} = [(S_t - \phi)/E_{ij}]^{1/(1-\alpha)}/A$ and the integral is obtained as

$$1-\theta = \frac{A^{1-\alpha}E_{ij}}{(S_t - \phi)(2-\alpha)} \frac{1}{A^{2-\alpha}} \left(\frac{S_t - \phi}{E_{ij}}\right)^{\frac{2-\alpha}{1-\alpha}} + 1 - \frac{1}{A}\left(\frac{S_t - \phi}{E_{ij}}\right)^{\frac{1}{1-\alpha}} = 1 - \frac{1-\alpha}{2-\alpha} \cdot \frac{1}{A}\left(\frac{S_t - \phi}{E_{ij}}\right)^{\frac{1}{1-\alpha}}.$$

The last equality sign follows from noting that $1 - (2-\alpha)/(1-\alpha) = -(1-\alpha)$. Altogether this results in (12).

Derivation of (14). Start with noting that upper concave branch of $\theta_(S_t)$ cut the identity line at most twice and that θ_{high}, if it exists, has to be at the upper intersection. Next, plug in $S_t = \theta_t = \theta$ into the upper branch in (12).

$$\theta = 1 - \frac{x}{\theta - \phi}, \quad x \equiv \frac{A^{1-\alpha}E_{ij}}{2-\alpha}.$$

Sorting terms provides $\theta^2 - \theta(1+\phi) + x + \phi = 0$. The largest root of the quadratic supports the equilibrium θ_{high}, if it exists. It is obtained as

$$\theta = \theta_{high} = \frac{1}{2}(1+\phi) + \sqrt{\frac{1}{4}(1+\phi)^2 - \phi - x}.$$

Noting that $\frac{1}{4}(1+\phi)^2 - \phi = \frac{1}{4}(1-\phi)^2$ provides (14) in the text.

Derivation of (17). As for the basic model let the level of poverty where $\sigma(\epsilon)$ intersects the unity line if such an intersection exists. If no such intersection exists the upper limit of integration is unity, taking piece-wise continuity into account, the area below the threshold is

$$\begin{aligned} 1 - \theta_t &= \int_0^p \frac{1}{S_t - \phi}(\epsilon + s)^{1-\alpha}(1-\tau)^\alpha(1+a) - \lambda\epsilon^{1-\alpha}\mathrm{d}\epsilon + \int_p^1 \frac{1}{S_t - \phi}\epsilon^{1-\alpha}\left[(1-\tau)^\alpha(1+a) - \lambda\right]\mathrm{d}\epsilon \\ &= \frac{1}{(S_t - \phi)(2-\alpha)}\left\{(1-\tau)^\alpha(1+a)\left[(p+s)^{2-\alpha} - p^{2-\alpha} - s^{2-\alpha} + 1\right] - \lambda\right\} \end{aligned}$$

Thus

$$\theta = 1 - \frac{1}{(S_t - \phi)(2-\alpha)}E_S$$

with E_S as defined in (17) for $\tilde{\epsilon} = 1$. If $\tilde{\epsilon}$ is smaller than unity it is found where $\tilde{\epsilon} = \left[\frac{S_t - \phi}{(1+a)(1-\tau)^\alpha - \lambda}\right]^{\frac{1}{1-\alpha}}$. Integrating

$$\begin{aligned} 1 - \theta_t = &\int_0^p \frac{1}{S_t - \phi}(\epsilon + s)^{1-\alpha}(1-\tau)^\alpha(1+a) - \lambda\epsilon^{1-\alpha}\mathrm{d}\epsilon \\ &+ \int_p^{\tilde{\epsilon}} \frac{1}{S_t - \phi}\epsilon^{1-\alpha}\left[(1-\tau)^\alpha(1+a) - \lambda\right]\mathrm{d}\epsilon + (1 - \tilde{\epsilon}) \end{aligned}$$

we arrive at

$$\theta = 1 - \frac{1}{(S_t - \phi)(2-\alpha)}\left\{(1-\tau)^\alpha(1+a)\left[(p+s)^{2-\alpha} - p^{2-\alpha} - s^{2-\alpha} + \tilde{\epsilon}^{2-\alpha}\right] - \lambda\tilde{\epsilon}^{2-\alpha}\right\} - (1 - \tilde{\epsilon})$$

which is (17) for interior $\tilde{\epsilon}$.

5.1. **Lemma 4.** Inspect the definition of E_S to note that the sign of $\partial E_S/\partial s$ equals the sign of $\partial x/\partial s$, $x \equiv (p+s)^{2-\alpha} - p^{2-\alpha} - s^{2-\alpha}$ and

$$\frac{\partial x}{\partial s} = (2-\alpha)\left[(p+s)^{1-\alpha} - s^{1-\alpha}\right] > 0$$

because $p + s > s$. The proof with respect to p proceeds analogously. The sign of the second derivative $\partial^2 E_S/\partial s^2$ equals the sign of $\partial^2 x/\partial s^2$ and

$$\frac{\partial^2 x}{\partial s^2} = (1-\alpha)(2-\alpha)\left[(p+s)^{-\alpha} - s^{-\alpha}\right] < 0$$

because $p + s > s$. The extremum is a maximum.

Derivation of (20). Let $L = ps + \mu\left[(1-\phi)^2(2-\alpha)/4 - E_S\right]$ be the Lagrangian associated with problem (19). The first order conditions are

$$\frac{\partial L}{\partial p} = s - \mu\frac{\partial E_S}{\partial p} = 0, \quad \frac{\partial L}{\partial s} = p - \mu\frac{\partial E_S}{\partial s} = 0 \quad \Rightarrow \quad \frac{p}{s} = \frac{\partial E_S/\partial s}{\partial E_S/\partial p}.$$

Insert the results from the proof of Lemma 4 above to obtain

$$\frac{p}{s} = \frac{(p+s)^{1-\alpha} - s^{1-\alpha}}{(p+s)^{1-\alpha} - p^{1-\alpha}}.$$

The unique solution (in the $(0, 1)$ interval) is $s = p$.

Derivation of (??). Since the community is initially situated at θ_{high}, where the concave branch of the $\theta(S_t)$-curve applies, $\tilde{\epsilon} = 1$. Taking this information into account insert $p = s$ from (20) into E_S from (17) to obtain

$$E_S = (1-\tau)^\alpha(1+a)\left[1 + 2s^{2-\alpha}(2^{1-\alpha} - 1)\right] - \lambda.$$

Insert this into (19) and solve for s to get (**??**).

Proposition 9. Note that $2^{1-\alpha} > 1$. The result with respect to τ is then obvious from inspecting (**??**). For the result with respect to γ recall the definition of λ from (18) and note that the sign of $\partial\lambda/\partial\gamma$ equals the sign of

$$\gamma^{\alpha-1}\left[\gamma - \alpha(1+\gamma)\right]$$

This expression is positive because $\gamma > \alpha/(1-\alpha)$ according to condition (7) since children are working when not attending school. Thus $\partial\lambda/\partial\gamma > 0$ which together with noting that $\partial s^*/\partial\lambda > 0$ completes the proof.

References

Al-Abdullah, R., 2009, Maha of the Mountains, in: *The Big Read: Stories in Support of Education*, Unesco, Paris.

Bacolod, M.P. and Ranjan, P., 2008, Why children work, attend school, or stay idle, *Economic Development and Cultural Change* 56, 791-828.

Baland, J.-M. and Robinson, J.A., 2000, Is child labor inefficient?, *Journal of Political Economy* 108, 663-679.

Basu, A.K. and Chau, N.H., 2004, Exploitation of child labor and the dynamics of debt bondage, *Journal of Economic Growth* 9, 209-238.

Basu, K. and Van, P.H., 1998, The economics of child labor, *American Economic Review* 88, 412-427.

Basu, K., 1999, Child labor: cause, consequence, and cure, with remarks on international labor standards, *Journal of Economic Literature* 37, 1083-1119.

Bhalotra, S., 2007, Is child work necessary, *Oxford Bulletin of Economics and Statistics* 69, 29-55.

Bigeri, M., and Guarcello, L. and Lyon, S. and Rosati, F., 2003, The puzzle of 'idle' children: neither in school nor performing economic activity: evidence from six Countries, UCW Working Paper.

Bobonis, G.J. and Finan, F., 2009, Neighborhood peer effects in secondary school enrollment decisions, *Review of Economics and Statistics*, forthcoming.

Cardoso, E. and Souza, A.P., 2004, The impact of cash transfers on child labor and school attendance in Brazil, Working Paper 04-W07, Vanderbilt University.

Chamarbagwala, R., 2007, Social norms, spatial dependence, and children's activities: evidence from India, *Journal of Developing Areas*, forthcoming.

Case, A.C. and Katz, L.F., 1991, The company you keep: the effect of family and neighborhood on disadvantaged youth, NBER Working Paper 3708.

Doepke, M. and Zilibotti, F., 2005, The Macroeconomics of child labor regulation, *American Economic Review* 95, 1492-1524.

Dreze, J. and Sen, A., 1995 Basic education as a political issue, *Journal of Educational Planning and Administration* 9, 1-26.

Edmonds, E.V., 2007, Child labor, in: Strauss, J. and Schultz, T.P. (eds), *Handbook of Development Economics* Vol. 4, forthcoming.

Glaeser, E.L., Sacerdote, B., and Scheinkman, J.A., 2003, The social multiplier, *Journal of the European Economic Association* 1, 345-353.

Granovetter, M., 1978, Threshold models of collective behavior, *American Journal of Sociology* 83, 1420-1443.

Hazan, M. and Berdugo, B., 2002, Child labor, fertility, and economic growth, *Economic Journal* 112, 810-828.

Jensen, P. and Nielsen, H.S., 1997, Child labor or school attendance? Evidence from Zambia, *Journal of Population Economics* 10, 407-424.

Kingdon, G.G., 2005, Where has all the bias gone? Detecting gender bias in the household allocation of educational expenditure in India, *Economic Development and Cultural Change* 53, 409-451.

Kingdon, G.G., 2009, Education of females in India: determinants and economic consequences, a case study of Uttar Pradesh, Discussion Paper, Economics Department, Oxford University.

Kling, J.R., Liebman, J.B. and Katz, L.F., 2007, Experimental analysis of neighborhood effects, *Econometrica* 75, 83-119.

Lalive, R. and Cattaneo, A., 2009, Social interactions and schooling decisions, *Review of Economics and Statistics* 91, 457-477.

Lindbeck, A., Nyberg, S., and Weibull, W., 1999, Social norms and economic incentives in the welfare state, *Quarterly Journal of Economics* 114, 1-35.

Lopez-Calva, L.F., 2002, A social stigma model of child labor, *Estudios Economicos* 17, 193-217.

Mani, A. and Mullin, C.H., 2004, Choosing the right pond: social approval and occupational choice, *Journal of Labor Economics* 22, 835-861.

Manski, C., 1993, Identification of endogenous social effects: the reflection problem, *Review of Economic Studies* 60, 531-542.

Munshi, K. and Myaux, J., 2006, Social norms and the fertility transition, *Journal of Development Economics* 80, 1-38.

Nechyba, T.J., 2001, Social approval, values, and AFDC: A reexamination of the illegitimacy debate, *Journal of Political Economy* 109, 637-672.

North, D.C., 1990, *Institutions, Institutional Change, and Economic Performance*, Cambridge University Press, Cambridge.

Palivos, T., 2001, Social norms, fertility and economic development, *Journal of Economic Dynamics and Control* 25, 1919-1934.

Psacharopoulos, G., 1997, Child labor versus educational attainment: some evidence from Latin America, *Journal of Population Economics* 10, 377-386.

Ray, R., 2000, Analysis of child labor in Peru and Pakistan: a comparative study, *Journal of Population Economics* 13, 3-19.

Ravallion, M. and Wodon, Q., 2000, Does child labor displace schooling? Evidence on behavioral responses to an enrollment subsidy, *Economic Journal* 110, C158-C175.

Schultz, T.P., 2004, School subsidies for the poor: evaluating the Mexican Progresa poverty program, *Journal of Development Economics* 74, 199-250.

Skoufias, E. and S.W. Parker, 2001, Conditional cash transfers and their impact on child work and schooling: evidence from the PROGRESA program in Mexico, *Economica* Fall 2001, 45-96.

Strulik, H., 2004, Child mortality, child labour, and economic development, *Economic Journal* 114, 547-568.

UNICEF, 2008, *Socio-economic and Cultural Barriers to Schooling in Southern Sudan*, New York, United Nations Children's Fund.

Zelizer, V.A., 1985, *Pricing the Priceless Child: The Changing Social Value of Children*, New York, Basic Books.