



ELSEVIER

Labour Economics 8 (2002) 649–665

LABOUR  
ECONOMICS

www.elsevier.com/locate/econbase

# Is literacy shared within households? Theory and evidence for Bangladesh

Kaushik Basu<sup>a,b,\*</sup>, Ambar Narayan<sup>c</sup>, Martin Ravallion<sup>c</sup>

<sup>a</sup>*Department of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142-1347, USA*

<sup>b</sup>*Department of Economics, Cornell University, Ithaca, NY 14853, USA*

<sup>c</sup>*World Bank, 1818 H. Street NW, Washington, DC, USA*

Received 30 January 2001; received in revised form 19 July 2001; accepted 2 October 2001

---

## Abstract

A member of a collective-action households may or may not share the benefits of literacy with others in that household; the shared gains from doing so may well be offset by a shift in the balance of power within the family. Using household survey data for Bangladesh, we find strong external effects of education on individual earnings. Holding a range of personal attributes constant, an illiterate adult earns significantly more in the non-farm economy when living in a family with at least one literate member. These effects are strongest, and most robust, for women. Omitted-variable bias cannot be ruled out but would also be consistent with an intra-household externality of literacy. © 2002 Elsevier Science B.V. All rights reserved.

*JEL classification:* D10; J24; J31

*Keywords:* Household behavior; Literacy; Externalities; Earnings; Bangladesh

---

## 1. Introduction

This paper asks whether an illiterate worker's labor earnings are affected by the educational attainment of *other* members of that worker's household. The large literature on literacy says very little about intra-household externalities of literacy and education. Yet

---

\* Corresponding author. Department of Economics (E52-383A), Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142-1347, USA. Tel.: +1-617-452-3385; fax: +1-617-253-1330.

*E-mail address:* kbasu@mit.edu (K. Basu).

one might well expect that one person's literacy will have spillover benefits for others within the same household. In their study of the propensity to innovate among Guatemalan farmers, Green et al. (1985) observe how having a literate family member conveys many of the benefits of being literate oneself. Likewise, Drèze and Saran (1995) note how the advantages of literacy can spread to others in the household by virtue of certain kinds of decision-making on behalf of the household shifting toward the literate.

It can be argued that an illiterate person who lives in a household in which at least one person is literate will be better off than an illiterate person who lives in a household of illiterates (Basu and Foster, 1998). For certain activities, all one needs is access to a literate person; one does not become literate this way, but it eases the handicap. For instance, an illiterate farmer who occasionally receives pamphlets from extension workers or needs to read the label on a new fertilizer packet could benefit enormously from easy access to a literate person. An urban worker whose work requires some knowledge of what is happening in the world at large or just plenty of 'commonsense'—which comes from interacting with knowledgeable people—might benefit a lot from having a literate person at home. Of course, employers who are aware of this, will be willing to pay more for someone coming from a literate household.

What evidence is there that literacy, or education more generally, has an external effect within the family? One external effect of education within families has been well researched empirically; namely, the effect of parental education on child welfare. There is a large body of evidence indicating that children's health, nutritional status and educational attainments are enhanced by having better-educated parents, particularly the mother.<sup>1</sup> However, sharing knowledge with children is more obvious than among adults, which will also depend on how sharing knowledge affects other aspects of decision-making within the family.

There is also evidence that a farm-household's total income depends on the highest education level reached, rather than the mean, or the education level of the household head (Foster and Rosenzweig, 1996; Yang, 1997a,b). This does not, however, imply an external effect of one person's human capital on the productivity of another person within that same household—it may simply reflect gains from a division of labor within the family.

More direct tests can be found by looking at individual labor market earnings. In empirical work on earnings, it is commonly assumed that own wages depend on own characteristics and not family characteristics directly. Indeed, the exclusion restriction that such external effects on earnings do not exist has been used to test for the endogeneity of human capital in wage regressions (Alderman et al., 1996).<sup>2</sup> However, in the US, there is evidence that men's earnings respond positively to the wife's schooling (Benham, 1974),

---

<sup>1</sup> An important early contribution was Behrman and Wolfe (1984) who demonstrated the importance of family background (including parental schooling, notably of the mother) to children's schooling in Nicaragua. In a recent example, Gibson (2001) finds strong effects of adult literacy on children's nutritional status in Papua New Guinea. In the same Bangladesh data set that is analyzed here, there is evidence that children of better-educated parents are more likely to attend school (Ravallion and Wodon, 2000). For a survey of the evidence on these effects see Strauss and Thomas (1995).

<sup>2</sup> If there are indeed external effects on earnings, then family characteristics are not valid instruments for endogenizing individual human capital.

and in Brazil, to the schooling of the wife, parents and parents-in-law (Lam and Schoeni, 1993). It is unclear whether this is a direct external effect or the outcome of marital sorting. Krishnan (1996) finds strong family background effects on public sector wages in Ethiopia, which she attributes to a selection effect, whereby family background influences access to public sector jobs; correcting for selection bias, the wage effect vanishes.

We aim here to test for external effects for both men and women in rural Bangladesh. We turn deliberately to a developing country setting in which there are large disparities in educational attainments, in the expectation that this will make it easier to identify any external effects of education. We examine the efficacy of men versus women as recipients of the externality of literacy. It is believed that women are better at generating positive externalities within the family, and there is supportive evidence.<sup>3</sup> However, there does not appear to have been any research on whether women are better or worse as the *recipients* of externalities. Does the presence of a literate member in the household yield significantly greater externalities on illiterate women than on illiterate men?

Before turning to the formal analysis, we should clarify that we are staying away from the larger debate on the inter-relationship between the concepts of literacy, schooling and knowledge. These are of course distinct concepts, though sharing a common core. Once one goes into the distinctiveness of these concepts, one could argue that a minimal level of education enables a person to receive externalities better. However, these are complexities that lie beyond the scope of this paper. Here, literacy is equated with the ability to read, though we will look at the effect of different years of schooling.

We begin by examining a key theoretical question: given that a literate member can, if she so chooses, prevent others from benefiting from her literacy, why does she not do so? Moreover, could it be that men and women have different propensities to share? The following section constructs a theoretical model to answer these questions. The paper then tests for literacy sharing within households. Section 3 describes our data for Bangladesh. Section 4 describes the wage regressions that we estimate across illiterate sub-samples of the workforce, correcting for sample selectivity bias by standard means. In addition to reasonably standard variables for earnings regressions, we include measures of the highest education level in the worker's household. We stratify the regressions by gender and employment sector (subject to data limitations) in the expectation that external effects will be stronger for certain types of jobs. Section 5 presents the results, and Section 6 consists of concluding remarks. A detailed Statistical Addendum is available from the authors.

## 2. Why might there be external effects of education within households?

It is not obvious on a priori grounds that a literate person will share the advantages of literacy with the illiterate members of the same household. The benefits of literacy are, by and large, non-rival but excludable. They are non-rival because for a literate member to read a brochure for an illiterate or to give the illiterate a piece of in-

---

<sup>3</sup> A number of studies have suggested that women share their education and other advantages with others in the household more than do men; for further discussion and references to the literature see Strauss and Thomas (1995).

formation is virtually costless. In a household, one can, for instance, exchange such information over dinner or while doing household chores. They are excludable because a literate member can, if he so chooses, refuse to part with any information. It is true that some of the benefits can be acquired by osmosis—by watching, for instance, what he does with the new fertilizer. Nevertheless, the literate person can block a lot of the benefits if he wants to. So why would we expect him to share knowledge with other members of the household?

In answering this, we could directly appeal to altruism and treat that as the end of the matter. Interestingly, it is possible to go deeper. The new literature on household decision-making (see, for instance, Udry, 1996; and Quisumbing and Maluccio, 1999) has lots of disputes, but one proposition generally agreed upon is that there is some pooling of resources and earnings in most households. There are the household public goods like the common kitchen, which benefit from anybody's rise in income. Then there are the shared tasks, such as one person belonging to the household or hired by the household going to the market to shop for everybody. This means that if individual  $i$  can raise individual  $j$ 's income by conferring some of his literacy advantage to  $j$ , a part of the gain will come back to  $i$  through the sharing process. However, one has to balance against this the fact that a higher income will empower the illiterate members and may alter the expenditure pattern of the household. If preferences differ sufficiently between literate and illiterate people, then a shift in power within the household associated with shared literacy advantage can work against the interests of the literate family member, who will then naturally be disinclined to share the advantages of literacy. To get a grip on these conflicting effects, it is useful to construct a model.

In our model, the household collectively chooses its consumptions, but literate members are free to restrict illiterate members from enjoying the advantages of literacy. Whether literate members choose to help illiterate members of the same household depends on their personal utility gains, which in turn depend on how sharing the literacy advantage alters the level and composition of household consumption. This depends on two potentially opposing effects. The first is a positive income effect of sharing literacy on consumption, through higher earning potential of the illiterate members. The second is a potentially negative effect (from the point of view of the literate members) on the composition of consumption, via effects on the balance of power within the household, given differences in preferences. How these two effects balance out determines whether there will be an external effect of literacy within the household.

Let us, for simplicity, assume that a household has two members, one literate and one illiterate, whose utility functions are  $U_L(\cdot)$  and  $U_O(\cdot)$ , respectively. We assume that there are two goods. If  $x$  is the vector of goods consumed by the household, then the utilities of the literate and illiterate members are given by  $U_L(x)$  and  $U_O(x)$ . It is assumed that both utility functions are strictly increasing and quasi-concave.

The two goods can interpreted in more than one way. We may think of the two goods as tobacco and skirts. Or the two goods may be the same commodity differentiated by the person who consumes it. Thus, we could think of apples consumed by the man of the household as good 1, and apples consumed by the woman as good 2. Either way, both goods are assumed to be private goods. The literature on household decision-making often allows for public goods within the household. In the presence of household public goods,

it is possible to place a kind of transferability restriction on individual utility functions under which the household provision of public goods ceases to be sensitive to the balance of power (see Bergstrom and Comès, 1983). For reasons of simplicity, we stay away from public goods.

We shall take this to be a ‘collective-action household,’ both because this has been modeled and analyzed in recent times (see, for example, Bourguignon and Chiappori, 1991, 1994; Moehling, 1995; Basu, 1999) and because of its flexibility. The latter enables the model to capture households that meet the hard requirements of the unitary model, as one polar extreme, and also of households which are deeply divisive.

Our model assumes that the household consumption maximizes a weighted average of  $U_L$  and  $U_O$ , the weight  $\lambda$  being a function of the illiterate person’s income,  $y_O$ , and the literate person’s income,  $y_L$ . Since the latter is assumed to be constant throughout the analysis, we can think of  $\lambda$  as a function of  $y_O$ . It will be assumed that  $\lambda(y_O) \in [0, 1]$  and  $\lambda'(y_O) \geq 0$ .

Hence, the household’s problem is to choose  $x$  so as to maximize:

$$\lambda(y_O)U_O(x) + [1 - \lambda(y_O)]U_L(x) \quad (1)$$

subject to the budget constraint  $p \cdot x \leq y_L + y_O$ , where  $p = [p_1, p_2]$  is the price vector for the two goods. Let us denote the solution of this problem by:

$$x = x(y_O) \quad (2)$$

(while  $x$  depends also on  $p$  and  $y_L$ , these are assumed to be constant and so are ignored).

The interesting question is: what will be the value of  $y_O$ ? This will depend in part on the literate member’s decision whether or not to exclude the illiterate member from the externalities of his education. Let  $y_p$  and  $y_l$  be the income of the illiterate if she, respectively, gets the benefit of literacy and if she does not get it. Of course,  $y_p > y_l$ .

When will the literate not exclude the illiterate? In other words, when will the literate set  $y_O = y_p$ ? One way of answering this is to think of the household’s problem described above to be the period-2 problem. In period 1, the literate member has to choose between setting  $y_O = y_p$  or  $y_l$ . Clearly, he will set  $y_O = y_p$  if and only if  $U_L(x(y_p)) \geq U_L(x(y_l))$ , where  $x(y_p)$  and  $x(y_l)$  are given by Eq. (2).

One can identify special cases of this problem such that it will be in the interest of the literate member to share the benefits of literacy. It is evident that if preferences do not differ, then the literate member will not exclude the illiterate member from the benefits. If  $U_O(x) = U_L(x)$ , then the household’s consumption choice problem reduces to maximizing  $U_L(x)$  subject to  $p \cdot x \leq y_L + y_O$ , and the maximum level of  $U_L(x)$  will be non-decreasing in  $y_O$ . Therefore,  $U_L(x(y_p)) \geq U_L(x(y_l))$ . Another special case in which the literate member will not exclude the illiterate member from the benefits is when both goods are normal and  $\lambda(y_O) = \lambda$ , for all  $y_O$ . Then the consumption possibilities of the literate family member will be unambiguously higher when knowledge is shared within the household.

Intuitively, what is necessary for members to share their literacy is that they have *similar* preferences, or that the bargaining power (captured by  $\lambda$ ) does not change *too* much in favor of the illiterate family member when literacy benefits are shared. To see this more formally, let us consider a special case:  $U_L(x) = x_1$  and  $U_O(x) = x_1 + \alpha x_2^\alpha$ , with  $\alpha < 1$ .

Note that as  $A$  goes towards 0, the two utility functions tend to converge. Further, and without loss of generality, assume  $P_2=1$ . Thus the household’s (period-2) problem is to maximize

$$\lambda(y_0)(x_1 + Ax_2^\alpha) + [1 - \lambda(y_0)]x_1$$

subject to

$$p_1x_1 + x_2 \leq y_L + y_0.$$

This is equivalent to the problem of maximizing:

$$x_1 + \lambda(y_0)A(y_L + y_0 - p_1x_1)^\alpha.$$

From the first-order condition we have:

$$x_1 = \frac{y_L + y_0}{p_1} - \frac{[p_1\lambda(y_0)\alpha A] \left(\frac{1}{1-\alpha}\right)}{p_1}. \tag{3}$$

Inspecting Eq. (3), it is obvious that changing  $y_0$  from  $y_1$  to  $y_p$  will increase  $x_1$  as long as  $\lambda(y_p) - \lambda(y_1)$  is sufficiently small or  $A$  is sufficiently small. Of course, if either of these terms is zero,  $x_1$  must rise.

Note now that if  $\lambda(y_p) - \lambda(y_1)$  is small, it means that the structure of decision-making changes at most a little in response to changes in the earnings of different members. And if  $A$  is small, it means that both members have similar preferences; preferences being identical if  $A=0$ . From these interpretations, it is clear that these are conditions that one would expect in a normal household. It may not be ‘unitary’ but it is not too divisive either. And when that holds, what we have just shown is that proximate illiterates will be better off than isolated illiterates. This is the hypothesis we are about to test. From the above discussion, it should be evident that the hypothesis meets the criterion of falsifiability. In a sense, what we are testing for is a certain view of the household. If that view is right, we would expect to see external effects of education within households, that is, for an illiterate person to earn more if there exists a literate person at home. Our theory cannot, however, tell how large this effect will be. Therefore, an important purpose of our empirical exercise is to estimate its magnitude.

It is easy to adapt the above model and alter the sharing of literacy from being a 0–1 variable to a variable  $s \in [0,1]$ , where  $s=0$  means not to share at all and  $s=1$  means to fully share. Then we would have the income of an illiterate person in a household where there is someone literate be a function of  $s$ . That is,  $y_0 = y_0(s)$ , with  $y_0(0) = y_l$ ,  $y_0(1) = y_p$  and  $y'_0(s) > 0$ .

From our model, it is easy to see that a literate person’s decision on how much to share, that is, the exact value of  $s$ , will depend on the nature of the utility function of the person with whom literacy can be shared. It is reasonable to expect basic differences between the utility functions of men and women. In that case,  $s$ , and therefore,  $y_0(s)$ , will be different depending on whether the benefits of literacy are to be shared with an illiterate man or an illiterate woman belonging to the household. In anticipation of our empirical finding that women are better “recipients” of the externality of literacy, note that if there is a third

person in the household—a child, for instance—there can be large natural differences between men and women as recipients of externality. Suppose the woman uses increases in her bargaining power in the household to spend in a way that benefits the child, but the man in the same place uses the power to his own advantage. However, if the man (woman) does not get benefits for himself, he (herself, she) prefers that the benefits go to the child rather than the other adults in the household. Given these plausible assumptions, it is natural to expect that a literate person will be more willing to let the external benefits of his or her literacy flow to a woman rather than a man, since that helps direct more advantage towards the child.

For similar reasons,  $s$  and  $y_O(s)$  can also depend on other factors that determine the utility functions of the individuals. These factors include whether the person works in the agricultural or the non-agricultural sector, and whether he or she belongs to the urban or the rural sector. What we mean is that there may be systematic differences in the utility functions that prevail in rural and urban households or households in different sectors of the economy. In that case, we may find that the externality of literacy depends not only on the gender of the participants but also on the sector to which they belong professionally. Theory of course cannot tell us the exact nature of the differences in  $s$  and  $y_O(s)$ . All theory says is that this is potentially an interesting question. The definitive answer has to come from empirical investigation.

The aim of this theoretical section was not to derive some testable hypothesis, but simply to show why the sharing of literacy is neither a foregone conclusion (as the unitary model of the household suggests) nor obviously false (as in a completely individualistic model of decision-making). Hence, there is indeed an open empirical question of how much “literacy sharing” occurs within households.

### 3. Data

Our data are from the Household Expenditure Survey for Bangladesh, 1995–1996 (Bangladesh Bureau of Statistics, 1998). This is a reasonably comprehensive nationally representative socio-economic survey of 7420 randomly sampled households (comprising 39,043 individuals) in 22 regions comprising of 63 districts, including rural, urban metropolitan and urban municipal areas. The survey results indicate that 61.1% of people in Bangladesh over 7 years of age are illiterate (55.7% of males; 66.6% of females). In rural areas, the percentage is higher, namely 65.4% (60.2% of males; 70.8% of females). Bangladesh Bureau of Statistics (1998) provides a detailed description of the survey and summary tabulations of the results. The Statistical Addendum to this paper contains complete tabulations of descriptive statistics on the variables used below for all sub-samples.

As a first step in the analysis, it is important to consider the kind of earnings we are interested in, given that the data reveal incomes from a wide range of sources. For obvious reasons, we are interested in only that part of income that is earned from some form of productive activity. So we exclude incomes derived from personal and individual wealth and assets from the categories of incomes to be considered. For reasons that are less obvious, we also exclude income from own farms and businesses. The reasons for these

exclusion have to do with our objectives as well as with the nature of data. Income from own farms and businesses, we feel, in most cases should be seen as accruing to the entire household rather than to individual members of the household. In such a scenario, education of one or more members of the household is likely to have a direct effect on household earnings and should not be considered a part of the externality effect that we are interested in. Also, due to the nature of these earnings, it is often not possible to apportion them between individual members of the household in a satisfactory way, which makes it very difficult to measure the effect of household education on individual earnings. From eyeballing the data, we get the impression that farm incomes have been apportioned among various household members somewhat arbitrarily. For instance, it has often been added to the income of the head of the household.

For these reasons, we focus on the following types of earnings: salary and wages received from regular and/or irregular/seasonal employment, in both cash and kind, and professional remuneration. Given our data set, quarterly figures for these categories of earnings are used.

To anticipate our results, we found that the external effects of literacy are confined to *non-farm* labor earnings, and so this is what we will focus on in the following discussion. Given the type of work involved, we do not find it implausible that the productivity of an illiterate agricultural workers will be unaffected by whether there is a literate person in their household. Returns to family literacy will presumably appear in less manual non-farm activities. For completeness, however, we also give results for male agricultural laborers in the Statistical Addendum (there are too few female agricultural laborers in the sample for this to be possible).

To separate the sub-samples or groups we are interested in, we identify all illiterate individuals, where illiteracy is defined as being unable to read, irrespective of the level of schooling. The next step is to identify the household each illiterate person belongs to and construct variables that measure (a) whether anybody in the household is literate and (b) the maximum education level of the household. The variable for (a) is constructed by defining a dummy variable that takes the value 1 if at least one member of the household is literate. Variables for (b) are constructed in the following manner: using the available figures on individual educational attainment, we define the maximum education level of a household as the educational level attained by the most educated member of the household. Once we derive the maximum education level of each household, we are in a position to define the dummy variables to proxy the educational attainment of each household.

Accordingly, the four educational levels we consider are being able to read, completed Class I and above, Class IV and above, and Class VII and above. The dummy variable takes the value 1 if that is the highest education attainment in the household. The entire group of illiterate persons is then sub-divided into males and females. The male group is further divided into rural and urban sub-groups, with the latter group being comprised of males who live in urban municipal or metropolitan areas. However, the female group is not sub-divided further because of concerns about the sample-sizes.

To summarize, among illiterate wage earners whose principal work is in the non-farm sector, we consider three sub-groups: (i) urban and rural females; (ii) rural males with primary earnings from the non-agricultural sector, and (iii) urban males.

#### 4. Testing for external effects of education on earnings

There are many ways that literacy might be shared with an illiterate worker, and there could well be contingent factors. For example, it might be conjectured that the extent of literacy sharing depends on the precise relationship between the illiterate worker and the literate family member. We will test whether the gender or age of the literate family member matters. For the main analysis, however, we focus on whether the presence of literacy in the household is associated with higher earnings for illiterate workers. We assume that the labor earnings,  $W_j$ , of an illiterate worker are determined by:

$$\log W_j = \alpha \text{LIT}_j + X_j \beta + \varepsilon_j \quad (4)$$

where  $\text{LIT}_j$  is a dummy variable taking the value 1 if at least one other person in the household is literate,  $X_j$  is a  $(1 \times K)$  vector of control variables for person  $j$  comprising other factors relevant to the worker's productivity and hence wages, including experience. (The first element of  $X_j$  is 1 to allow for an intercept.) The parameter  $\alpha$  measures the external effect of the family's literacy on the worker's earnings, and  $\beta$  is a  $(K \times 1)$  vector of parameters on worker characteristics. The error term  $\varepsilon_j$  is assumed to be normally distributed with mean zero and standard deviation  $\sigma_\varepsilon$ .

There are three potential problems in estimating the test regression in Eq. (4). The first is the classic problem of estimating wage equations, namely that one only observes wages for those individuals who have chosen to enter the (paid) work force. Here we follow standard practice. We assume that wage earnings for person  $j$  are observed if and only if

$$\delta \text{LIT}_j + Z_j \gamma + u_j > 0 \quad (5)$$

where  $Z_j$  is a  $(1 \times M)$  vector of regressors for person  $j$ , whose first element is 1 and other elements are  $M-1$  regressors and  $\gamma$  is a  $(M \times 1)$  vector of parameters. Eq. (5) is interpreted as a model of the individual's labor force participation decision. The error term  $u_j$  is standard normal with a non-zero correlation coefficient ( $\rho$ ) with  $\varepsilon_j$ . Then we can use the maximum-likelihood Heckman selection model to correct for any selection bias. The expected value of observed log earnings is then  $\alpha \text{LIT}_j + X_j \beta + \rho \sigma_\varepsilon \lambda (\delta \text{LIT}_j + Z_j \gamma)$ , where  $\lambda$  is the usual inverse Mills ratio. We can separately identify  $\alpha$  from  $\delta$  even if  $X_j = Z_j$  given that  $\lambda$  is a non-linear function. One of our tests will exploit this fact. However, our main results will also make a widely used identifying assumption in estimating earnings regressions with selectivity, namely that certain household characteristics influence participation but do not influence earnings given participation.

In the wage equation, the variables in the vector  $X$  include age of the earner, square of the age and a dummy variable that takes the value 1 if the earner has been to school.<sup>4</sup> We also test sensitivity to adding extra controls for the maximum level of education in each household (discussed further below). To allow for geographic effects, a series of regional dummies are also included that take the value 1 depending on the region a particular earner belongs to completes the list of independent variables in the wage equation.

<sup>4</sup> A small number of persons in our sample have been to school but remain illiterate.

In the selection equation, the vector  $Z$  includes all the independent variables in the wage equation, along with variables that represent household characteristics that are deemed to determine participation in the workforce. This latter group of variables includes dummies for land holdings, and demographic characteristics including characteristics of the household head. In the case of females with primary earnings from the non-agricultural sector, dummies for whether the household is rural or urban are also included as explanatory variables; in the case of urban males, a dummy for whether the household is from a metropolitan area or not is included as an explanatory variable. However, we also test robustness of our results to including in  $X$  all the variables in  $Z$ , and exploiting the non-linearity of the inverse Mills ratio for identification.

The second potential problem is that the intra-household sharing of literacy may be so complete in some cases that previously illiterate workers become literate. Sample selection bias could arise from the fact that we select illiterate workers. On a priori grounds, we do not think it is common to become literate this way; literacy is normally acquired through schooling. And, of course, a necessary condition for this bias is that there is (very) effective sharing of literacy; there is no bias under the null hypothesis of no sharing. We choose to ignore the possibility of this source of selection bias in the following analysis.

The third problem in estimating our test Eq. (4) is bias due to possible endogeneity of literacy within the family. We believe this to be a more serious concern. In the present context, a key question is whether there are omitted determinants of an individual's earnings—observed by the employer but not in our data—which are correlated with the presence of a literate person in the household. We can suggest two possible reasons why this might happen. The first reason recognizes that workers can make a choice about the extent of schooling in their family, such as through the choice of marriage partner or the schooling choices parents make for their children. This will bias our results if the choices made are correlated with the omitted determinants of earnings. For example, an illiterate person with high latent ability (known to the employer) might be better able to attract a literate spouse, or may choose to educate her children more.

Note that, while such sorting effects will lead us to overestimate  $\alpha$ , they can also be interpreted as literacy externalities in their own right. The choices made reveal that a high-wage illiterate worker derives utility from having a literate spouse or child, even if this effect does not yield higher wages for that worker.

Another reason to suspect endogeneity bias is that the presence of literacy may be correlated with other measures of the extent of schooling in the family, which may also matter to the earnings of an illiterate worker. Possibly it is not literacy that gives the external benefit, but schooling more generally. This would be consistent with the proposition that schooling has external effects on earnings, though it would bias our estimates of the effects of literacy per se.

There are obvious limits to how far one can address such concerns with cross-sectional data. In principle, panel data might help, by permitting a latent time-invariant individual effect in the error term of the earnings regression, correlated with family literacy. Taking deviations from the time means, our estimate of the wage effect of changes over time in family literacy would then be robust to the existence of this individual effect under standard assumptions. In practice, however, family characteristics such as the presence of

a literate member, are also likely to be slow moving, and even time-invariant given the relatively short duration of most household panel data sets available. Standard panel data sets may have little power for detecting external effects of family literacy. It is also unclear whether the short-term fluctuations in family literacy that might be found in panel data—when a literate uncle visits briefly, for example, or simply time varying measurement error—would entail any real scope for passing on the external effect, even when there are large long-term external benefits in having a literate family member over the longer term. There could then be a large attenuation bias in the within estimator on panel data.

We can, however, perform two tests of the robustness of our estimates of  $\alpha$  to endogeneity of LIT<sub>*t*</sub> in our cross-sectional data set. In the first, we restrict the sample to unmarried workers, on the assumption that marriage is the main way in which a high-wage illiterate worker can influence the extent of literacy in her or his family. The wage and probit equations in each case have the same specifications as before, with the difference that the estimates are now based only on the unmarried people belonging to each group. However, this test will not be robust to other forms of selective family formation. For example, an unmarried illiterate worker may choose to continue living with a literate parent or sibling so as to assure that the external gains from literacy are captured; higher earning ability may help assure that this is possible.

In the second test, we add extra control variables measuring the extent of schooling in the family; we use the highest years of schooling reached.

## 5. Results

We estimate Eqs. (4) and (5) for the following sub-groups:

(i) *Women with primary earnings from the non-agricultural sector.* Here the dependent variable in the wage equation is the log of earnings of individuals in this group. In the probit estimated on the full sample of illiterate females, the dependent variable takes the value 1 if an individual's earnings are observed and if her primary earnings are from the non-agricultural sector, and 0 otherwise.

(ii) *Rural males with primary earnings from the non-agricultural sector.* The dependent variable in the wage equation is the log of earnings of individuals in this group. In the probit estimated on the entire sample of rural males, the dependent variable takes the value 1 if an individual's earnings are observed and if his primary earnings are from the non-agricultural sector, and 0 otherwise.

(iii) *Urban males.* The dependent variable in the wage equation is the log of earnings of individuals in this group. In the probit estimated on the entire sample of urban males, the dependent variable takes the value 1 if an individual's earnings is observed.<sup>5</sup>

<sup>5</sup> The Addendum also gives results for rural males with primary earnings from the agricultural sector; in this case, the dependent variable in the wage equation is the log of earnings of individuals belonging to this group. In the probit which is estimated on the entire sample of rural males, the dependent variable takes the value 1 if an individual's earnings are observed and if his primary earnings are from the agricultural sector, and 0 otherwise.

In each of these cases, three models are estimated—one with literacy as the only household-education variable, a second in which the sample is restricted to unmarried workers, and a third with controls for other levels of education.

Table 1 summarizes the results (the Addendum gives full details on the wage regressions and participation probits used to correct for selectivity bias).

We find a large and significant effect of family literacy on the wages of illiterate working women. Living in a household with a literate member adds 0.50 to log earnings of illiterate women, and the effect is highly significant. This implies that illiterate women can expect their earnings to be 65% higher if someone is literate in their family. This effect is robust to restricting the sample to unmarried women; indeed, the gain in log earnings rises to 0.74. This implies that an illiterate female worker from a family with at least one literate member has about double the earnings as one from a family where everyone else is also illiterate.

The fact that we find the earnings gain is even greater in the sample of unmarried women is not suggestive of a bias due to selective sorting through marriage. It is not clear, however, why the effect would be stronger for illiterate unmarried women. Our theoretical model is suggestive of one possible explanation. Illiterate unmarried women may well have very little bargaining power in the household even when they bring extra wage earnings into the family. Thus, only the positive income effect remains and so the literate family member has a strong incentive to share the gains from literacy with the illiterate unmarried woman.

Women’s labor force participation, by contrast, is significantly lower when an illiterate worker lives in a literate household, and the effect is similar for married and unmarried women.

We find significant gains to illiterate males working in both the urban and rural non-farm sectors of living in a household where at least one other person is literate. The gain in

Table 1  
Effects of family literacy on the earnings and labor force participation of illiterate non-farm workers

		Full sample	Unmarried sample only	With land effects on earnings
Women	earnings ( $\alpha$ )	0.504 (4.577)	0.737 (4.477)	0.487 (4.368)
	participation ( $\delta$ )	-0.256 (-3.512)	-0.290 (-2.353)	-0.253 (-3.527)
	number of observations	295	139	295
Men (rural)	earnings ( $\alpha$ )	0.146 (2.769)	0.143 (1.127)	0.143 (2.636)
	participation ( $\delta$ )	0.110 (1.980)	0.113 (1.125)	0.111 (1.991)
	number of observations	582	157	582
Men (urban)	earnings ( $\alpha$ )	0.144 (2.274)	-0.031 (-0.251)	0.136 (2.405)
	participation ( $\delta$ )	-0.023 (-0.153)	-0.023 (-0.153)	-0.151 (1.800)
	number of observations	550	132	550

Note: gain in log earnings of an illiterate worker attributed to literacy within the household. Log earnings regressions include correction for sample selectivity bias as well as controls for worker characteristics (age, age squared, if any school attendance, and regional effects). *t*-ratios in parentheses. We have used the STATA 6 command for the MLE version of the Heckman selection model, with standard errors corrected for heteroskedasticity and clustering.

log earnings is about 0.14–0.15, implying that earnings of illiterate men are on average 15% higher if they come from a family where at least one person is literate. However, this effect is not robust to restricting the sample to unmarried workers; the effect is no longer statistically significant. For urban males, the effect vanishes.

Labor force participation of illiterate rural men is significantly higher if they live in a literate household, although the effect is barely significant at the 5% level. Nor is this effect found for unmarried men. The labor force participation of urban illiterate men is not changed significantly by living in a literate family.

The above tests do not distinguish the external effect of having a male versus female literate family member. To test the possibility that different genders generate different external gains from literacy, we re-ran the tests including two extra dummy variables. One of these variables was for whether the literates are only male, and one was for whether they are only female (the left-out category was the case in which there are both male and female literate family members). The coefficients on these dummy variables were small and (individually and jointly) insignificant in almost all the wage regressions. The one exception was the dummy variable for “male-only” literates in the women’s wage regression, for which there the coefficient was  $-0.361$ , suggesting an attenuation of the external gain to women workers in households where only men are literate; however, the coefficient was only significant at the 8% level ( $t=1.76$ ).

Similarly, we tested for whether age of the literate member matters. We added dummy variables for whether all literates were under 16, or over 16 (again the mixed case was omitted). These were individually and jointly insignificant in all of the wage regressions.

We also tested how robust these results are to the specification of the variables to be included as controls in the wage regression (comprising the vector  $X$  in Eq. (4)). Possibly the most contentious restriction we have imposed is that household landholding only influences earnings via its effect on the correction term for labor force participation, i.e., landholding does not have a direct effect on labor earnings. The final column in Table 1 gives the results when we include dummy variables for landholding class in the wage regression. The estimated effects of family literacy are similar to the first model.

Table 2 presents the main results from the wage and participation regressions which also included controls for other schooling in the household (otherwise the estimation method is the same as used for Table 1, the Addendum also gives full details on the regressions underlying Table 2). The literacy effect on earnings is still strong for women.<sup>6</sup> And there is a very strong effect of high education levels. The significant negative effect of family literacy on female labor force participation is found to be due to higher schooling rather than literacy per se.

For both rural and urban male workers, the positive effect of family literacy on earnings is still present when one adds controls for other family members’ schooling, but it is not significant. Instead, we get a significant positive effect of family schooling. For rural males, participation in non-farm work is not changed significantly by family

---

<sup>6</sup> Similar to Table 1, the effect was also stronger for unmarried women; details available in the Addendum from the authors.

Table 2  
Effects of family literacy and schooling on the earnings and labor force participation of illiterate non-farm workers

		There is another literate person in the household	Highest schooling in the household		
			Completed class I+	Completed class IV+	Completed class VII+
Women	earnings	0.462 (2.578)	0.186 (1.510)	−0.311 (−1.628)	0.757 (4.154)
	participation	0.102 (0.866)	−0.249 (−2.678)	−0.223 (−1.820)	−0.196 (1.705)
Men (rural)	earnings	0.104 (1.252)	0.152 (2.225)	−0.042 (−0.495)	−0.053 (−0.613)
	participation	0.016 (0.189)	−0.069 (−0.949)	0.171 (1.953)	0.012 (0.145)
Men (urban)	earnings	0.048 (0.493)	0.056 (0.743)	−0.020 (−0.205)	0.271 (2.736)
	participation	0.005 (0.004)	0.001 (0.008)	0.026 (0.187)	−0.460 (−4.075)

Note: log earnings regressions include correction for sample selectivity bias as well as controls for worker characteristics (age, age squared, if any school attendance, and regional effects). *t*-ratios in parentheses. Also see notes to Table 1.

literacy; for urban men, higher levels of family schooling attenuate their labor force participation.

So, while we find some signs that literacy in the family increases an illiterate man's earnings, this is not robust to adding controls for higher levels of education in the family. This, of course, is not inconsistent with external effects of education. However, it does lead us to question whether it is the fact of literacy that is driving the results for males, or something else.

The external effects of family literacy and schooling on the earnings of illiterate women that we find are very strong indeed, and much stronger than for men. Why might this be? We can suggest the following explanations.

One might conjecture that women have less power to influence the literacy of their family than do men. Social sanctions on women's behavior might limit their ability to choose a partner, while men are less constrained. Or possibly men have greater influence over whether the children are sent to school. If there are in fact pecuniary gains to an illiterate worker from having at least one literate person in the family, then men may well be better able to capture the external effects of literacy through family formation than are women. We would then find less variance in family schooling for males than females.

This explanation is not, however, easily reconciled with our finding that the external effect of family literacy for illiterate males drops out when we confine attention to unmarried males. If this explanation were right, then we would surely expect a stronger effect to emerge for unmarried males than married ones.

We believe that a more plausible explanation is that illiterate women in Bangladesh are much more dependent on their family for assistance with tasks that require reading and writing than are illiterate men. There are sources of help in reading and writing for illiterate workers, including paying for the services of a scribe. However, cultural and religious sanctions on women's social activities may well entail that they have a harder time than men in accessing these sources of help outside the family. This effect would be no weaker we suspect, and possibly stronger, for unmarried than married men.

Another possible interpretation of our results is that having a literate worker in the family allows illiterate workers to avoid exploitation (in sectors where there is some monopsonistic power), or simply allows a better matching of the worker's other skills to the jobs available (even if all workers receive their marginal product, this may depend on the kind of firm that a person joins).

Finally, women may well be better able to absorb the benefits of household literacy. They may, for instance, be unencumbered by pride or other defenses which could render one unable to learn easily from others.

## **6. Conclusions**

Intuition and piecemeal evidence suggest that literacy has positive externalities on earnings of others in the household. That is, a literate person is likely to share some of the benefits of his or her literacy with other members of the household. This means that it is better to be an illiterate in a household where there is someone literate than in a household of only illiterates.

We have tested this hypothesis in two ways. Firstly, we have asked whether it can be derived from a consistent set of assumptions about household behavior. Our theoretical analysis of the household, based on the collectivist model of the household, confirms that it can make sense for agents to share their education with others in the household. Secondly, we have tested if the hypothesis stands up to systematic scrutiny in earnings data for Bangladesh. We found that it not only stands up, it does so with a considerable amount of robustness. Controlling for observed personal attributes, an illiterate worker earns appreciably more in the non-farm sector when living in a family with at least one literate member. This finding is robust to the tests we perform for selection bias and marital sorting.

The empirical investigation has yielded insights deeper than what we had anticipated on the basis of theory and intuition. It is, for instance, widely noted that a literate mother confers greater benefits on the children than a literate father. But what about differences between male and female *recipients*? Is there any reason to believe that a female is better able to absorb the benefits of literacy in the household than a male illiterate? While theory suggests that they can have differences, it gives no hint as to which way this may go. Our empirical results suggest that women are more efficient recipients.

## Acknowledgements

These are the views of the authors, and should not be attributed to their employers, including the World Bank. The paper is an output of the World Bank research project, ‘Literacy and Child Labor’ (RPO 683-07). We are grateful to two anonymous referees and the editor for helpful comments.

## References

- Aldeman, H., Behrman, J.R., Ross, D.R., Sabot, R., 1996. The returns to endogenous human capital in Pakistan’s rural wage labor market. *Oxford Bulletin of Economics and Statistics* 58 (1), 29–55.
- Bangladesh Bureau of Statistics, 1998. Household expenditure survey 1995–96, Bangladesh Bureau of Statistics, Statistics Division, Dhaka, Bangladesh.
- Basu, K., 1999. Child labor: cause, consequence and cure with remarks on international labor standards. *Journal of Economic Literature* 37, 1083–1119.
- Basu, K., Foster, J., 1998. On measuring literacy. *Economic Journal* 108, 1733–1749.
- Behrman, J.R., Wolfe, B.L., 1984. The socioeconomic impact of schooling in a developing country. *Review of Economics and Statistics* 66, 296–303.
- Benham, L., 1974. Benefits of women’s education within marriage. *Journal of Political Economy* 82 (1), S57–S71.
- Bergstrom, T.C., Comès, R.C., 1983. Independence of allocative efficiency from distribution in the theory of public goods. *Econometrica* 51, 1753–1765.
- Bourguignon, F., Chiappori, P.A., 1991. Collective models of household behavior: an introduction. *European Economic Review* 36, 355–364.
- Bourguignon, F., Chiappori, P.A., 1994. The collective approach to household behavior. In: Blundell, R., Preston, I., Walker, I. (Eds.), *The Measurement of Household Welfare*. Cambridge Univ. Press, Cambridge.
- Drèze, J., Saran, M., 1995. Primary education and economic development in China and India: overview and two case studies. In: Basu, K., Pattanaik, P.K., Suzumura, K. (Eds.), *Choice, Welfare and Development*. Clarendon Press, Oxford.

- Foster, A.D., Rosenzweig, M.R., 1996. Technical change and human-capital returns and investments: evidence from the green revolution. *American Economic Review* 86, 931–953.
- Gibson, J., 2001. Literacy and intrahousehold externalities. *World Development* 29, 155–166.
- Green, S.E., Rich, T.A., Nesman, E.G., 1985. Beyond individual literacy: the role of shared literacy for innovation in Guatemala. *Human Organization* 44, 313–321.
- Krishnan, P., 1996. Family background, education and employment in Urban Ethiopia. *Oxford Bulletin of Economics and Statistics* 58, 167–183.
- Lam, D., Schoeni, R.F., 1993. Effects of family background on earnings and returns to schooling: evidence for Brazil. *Journal of Political Economy* 101 (4), 710–740.
- Moehling, C.M., 1995. The intrahousehold allocation of resources and participation of children, in: *Household Decision-Making: Evidence from Early Twentieth Century America*, mimeo: Northwestern University.
- Quisumbing, A.R., Maluccio, J.A., 1999. Intrahousehold allocation and gender relations: new empirical evidence, mimeo: International Food Policy Research Institute, Washington, DC.
- Ravallion, M., Wodon, Q., 2000. Does child labor displace schooling? Evidence on behavioral responses to an enrollment subsidy in Bangladesh. *Economic Journal* 110, 158–176.
- Strauss, J., Thomas, D., 1995. Human resources: empirical modeling of household and family decisions. In: Behrman, J., Srinivasan, T.N. (Eds.), 1995. *Handbook of Development Economics*, vol. 3. North-Holland, Amsterdam.
- Udry, C., 1996. Gender, agricultural production and the theory of the household. *Journal of Political Economy* 104 (5), 1010–1046.
- Yang, D.T., 1997a. Education in production: measuring labor quality and management. *American Journal of Agricultural Economics* 79, 764–772.
- Yang, D.T., 1997b. Education and off-farm work. *Economic Development and Cultural Change* 45 (3), 613–632.