OSIPP Discussion Paper: DP-2004-E-004

Shared Literacy and Employment in Non-farm Sector

August 27, 2004

北條　雅一　(Masakazu Hojo) *

大阪大学大学院国際公共政策研究科（ＯＳＩＰＰ）　助手
Research Assistant, Osaka School of International Public Policy (OSIPP)

【キーワード】Literacy; Externalities; Earnings; Non-farm sector; South Africa
【JEL Classification】D10; J24; J31
【要約】This article empirically tests the effects of household literacy on an illiterate person's job selection and wage earnings, using household data from South Africa. Recent empirical studies find that literacy is shared within households on one hand, while non-farm activities are important in determining the welfare of households in developing countries on the other. This article finds that an illiterate person who lives with a literate member in the household is employed in non-farm sector with higher probability than an illiterate person who lives with no literate person. In other words, household literacy affects the illiterate person's wage earnings through raising probability of finding better wage employment in non-farm sector.

---

* I am grateful to Akira Kohsaka and Colin McKenzie for helpful comments.

Address: 1-31 Machikaneyama, Toyonaka, Osaka 560-0043 Japan, Tel/Fax: +81-6-6850-5839, E-mail address: mhoujou@osipp.osaka-u.ac.jp

1 Introduction

Economists have paid attention to the importance of literacy for economic development. While it is widely recognized that increasing the literacy rate in developing countries is an important end in itself, economists have investigated the impacts of literacy on various economic activities such as farm production (summarized in Lockheed et al., 1980), adoption of agricultural innovations (Feder et al., 1985), and children's nutritional status (Strauss and Thomas, 1995).

Literacy, or education more generally, is attained at an individual level. However, recent empirical studies show that there are significant external benefits created by literate individuals. Green et al. (1985) find that a farmer's propensity to adopt modern agricultural practice is enhanced by his own and his family's literacy, indicating that literacy is a shared skill as well as an individual phenomenon. Basu and Foster (1998) provide a concept that distinguishes two types of illiterate individuals: "proximate illiterate" is defined as an illiterate person who lives with at least one literate member in his/her household; "isolated illiterate" is defined as an illiterate person who lives with no literate person in his/her household. The proximate illiterate might benefit from easy access to a literate person. Using data from Papua New Guinea, Gibson (2001) shows the evidence for shared literacy and finds that effects of shared literacy within households are larger than the values supposed in Basu and Foster (1998). An important study on this topic is Basu et al. (2002). They examine whether intra-household externality of literacy and education exists. Using household survey data from Bangladesh, they tested existence of the external effects of household literacy by regressing illiterate person's wage earnings on an indicator of family's literacy and worker's characteristics. Basu et al. (2002) find positive and significant estimates of the indicator of household literacy, that is, they show that the proximate illiterate earns significantly more than the isolated illiterate.

On the other hand, development economists are paying more attention to the importance of rural non-farm sector in developing countries. The rural non-farm sector has been traditionally thought as a low-productivity sector which produces low quality goods, however, it is recently recognized that the rural non-farm sector are becoming more important in determining the welfare of households (Lanjouw and Lanjouw, 2001; Kurosaki and Khan, 2004). Considering the importance of non-farm sector in developing countries, this article extends the empirical analysis of Basu et al. (2002). I hypothesize that the proximate illiterate can benefit from the literate person in his/her household by obtaining more information on employment opportunities than the isolated illiterate. Therefore, household literacy benefits the proximate illiterate *through*

increasing the probability of finding a better job in non-farm sector. In order to investigate whether this hypothesis can be supported empirically, this article conducts a careful empirical analysis using household survey data from South Africa.

This article is organized as follows. The next section describes the characteristics of data used. Section 3 proposes empirical models that controls for endogenous selection of the activities. Section 4 shows estimation results for the selection of activities and wage equations. It is found that household literacy increases the probability that the proximate illiterate can find a job in non-farm sector. Section 5 concludes the article.

2 Data

The data for this article are obtained from the 1993 Project for Statistics on Living Standards and Development (PSLSD) collected in collaboration with the World Bank by the Southern African Labor and Development Research Unit (SALDRU) in the School of Economics at the University of Cape Town. This project is a national survey of about 9,000 households in South Africa, drawn from 360 sample clusters, comprising 43,687 individuals. The survey is an adaptation of the Living Standards Measurement Surveys (LSMS) of the World Bank. The data collection instruments and the two-stage sampling methods used are described in SALDRU (1994).

In order to investigate the effect of shared literacy within households on employment and wage of the illiterates, I construct a sample comprised of 3,534 non-white illiterate individuals who are in working age (ages 15 to 65) and are not going to school.[1] In this sub-sample, 1,260 individuals are regularly employed, 209 individuals are casually employed, and only four individuals are self-employed. Of the regularly employed individuals, 559 individuals engage in farm-related activities and 701 individuals engage in non-farm activities. The drawback of this dataset is, however, that there is no data which directly measures an individual's ability to read and write. For this reason, I define an illiterate person as an individual who have less than two years of formal schooling and a literate person as an individual who have more than five years of formal schooling. The validity of these definitions will be checked in the section 4.[2]

Descriptive statistics of the whole sample and the constructed sample are shown in Table 1. It is obvious that average wage income from non-farm activities is higher than that from farm-

[1] Individuals who are in ill health, already retired, and do not report years of education are excluded from the sub-sample. Whites are dropped from the sub-sample because wage structure for whites differs significantly from that of Africans, Indians, and Colored (Mwabu and Schultz, 2000).

[2] Using our measure of illiteracy, adult illiteracy rate of whole sample is about 23%. World Development Indicators (available from the World Bank's web site) reports adult illiteracy rate of South Africa in 1995 is 15%.

related activities. The importance of income from non-farm activities is similar to the findings of Lanjouw (1999) and Lanjouw and Lanjouw (2001). Average wage income of the illiterate workers from farm-related activities is about half of the overall average. Average wage income of the illiterate workers from non-farm activities is about two-thirds of the overall average. Wage income of male workers are more than 1.5 times than that of female workers for both farm-related and non-farm activities. More than half of the illiterate have a literate person in their households. Illiterate person tends to live in an area labeled "rural" and with less paved roads. Table 2 reports differences between the proximate illiterate and the isolated illiterate. On average, The proximate illiterate receives 28% higher wages than the isolated illiterate. Other variables do not show any significant differences between the two types of the illiterates.

3 Testing for effects of shared literacy on earnings

Descriptive analysis in the previous section suggests that (a) non-farm activities are associated with higher income (Table 1) and (b) earnings of the proximate illiterate is higher than those of the isolated illiterates (Table 2). If a literate person has more information on better employment opportunities than an illiterate person, the proximate illiterate might benefit from easy access to a literate person when he/she seeks for a job. Thus, this article tests the following hypothesis: household literacy affects the earnings of the proximate illiterate through increasing the probability of finding better wage employment, i.e., non-farm activities. In order to investigate whether this hypothesis is supported empirically, this section proposes an empirical model that can control for endogenous selection of the activities.

Consider a situation where an illiterate person is seeking for a job. We assume that he/she chooses an activity j from three alternatives: household work ($j = 0$), employee in farm-related activity ($j = 1$), and employee in non-farm activity ($j = 2$). We ignore the possibility of being self-employed because only four illiterate individuals in the survey are self-employed. We also assume that place of residence is given in the short run and thus the illiterates do not migrate to other areas to find a job. Given that wage labor market is exogenous to the choice of activity by illiterate individuals, we assume that the labor earnings of illiterate worker i employed in activity j, $W_{i,j}$, are determined by

$$\ln W_{i,j} = \alpha_j LIT_i + X_i\beta_j + \varepsilon_{i,j}, \quad j = 1, 2, \tag{1}$$

where LIT_i is a dummy variable which takes the value 1 if at least one person is literate in the illiterate individual's household and 0 otherwise, X_i is a vector of worker attributes such as

education and age (including constant), and $\varepsilon_{i,j}$ is an error term. The parameter α_j measures the effect of family's literacy on illiterate worker's earnings from activity j (Basu et al., 2002), and β_j is a vector of parameters on worker attributes.

A serious econometric problem arises when estimating the wage equation (1). Because $W_{i,j}$ can be observed only when an illiterate individual i engages in farm-related activity ($j = 1$) or non-farm activity ($j = 2$), the error term $\varepsilon_{i,j}$ conditional on this selection has non-zero mean. To control for this, we follow procedures suggested by Lee (1983). By assuming that an illiterate person's utility associated with choosing activity j has a non-stochastic component and a stochastic term with type-I extreme-value distribution, the choice of activity can be characterized by a multinomial logit model (McFadden, 1974). We specify the multinomial logit model as

$$\text{Prob}(d_i = j) = \frac{\exp(\gamma_{j1} LIT_i + X_i\gamma_{j2} + Z_i\gamma_{j3})}{\sum_{k=0,1,2} \exp(\gamma_{k1} LIT_i + X_i\gamma_{k2} + Z_i\gamma_{k3})}, \quad j = 0, 1, 2, \tag{2}$$

and estimate it in the first stage of our empirical analysis, where d_i is an indicator variable denoting the choice of illiterate individual i with respect to j, X_i is a vector of individual attributes defined above, Z_i is a vector of attributes of the household where illiterate individual i resides, γ_{j1} measures the effect of the family's literacy on the probability of choosing activity j, and γ_{j2} and γ_{j3} are vectors of parameters to be estimated, associated with choice j. The multinomial logit model can be estimated by a maximum likelihood method.[3] The fitted probability of illiterate individual i being employed in activity j, $\hat{\text{Prob}}(d_i = j)$, is given by

$$\hat{\text{Prob}}(d_i = j) = \frac{\exp(\hat{\gamma}_{j1} LIT_i + X_i\hat{\gamma}_{j2} + Z_i\hat{\gamma}_{j3})}{\sum_{k=0,1,2} \exp(\hat{\gamma}_{k1} LIT_i + X_i\hat{\gamma}_{k2} + Z_i\hat{\gamma}_{k3})}, \quad j = 0, 1, 2. \tag{3}$$

Then, the correction term is defined as

$$\hat{\lambda}_{i,j} = \frac{\phi(\Phi[\hat{\text{Prob}}(d_i = j)])}{\hat{\text{Prob}}(d_i = j)}, \quad j = 0, 1, 2, \tag{4}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are density and distribution functions for a standard normal variable. Assuming the error term $\varepsilon_{i,j}$ in equation (1) to be normally distributed, estimation of the wage function with correction term $\hat{\lambda}_{i,j}$ produces consistent estimates of parameters α_j and β_j:

$$\ln W_{i,j} = \alpha_j LIT_i + X_i\beta_j + \delta_j\hat{\lambda}_{i,j} + \varepsilon_{i,j}, \quad j = 1, 2. \tag{5}$$

As defined above, LIT_i is a dummy variable which takes the value 1 if individual i is the proximate illiterate and 0 if he/she is the isolated illiterate. The parameter α_j measures the

[3] Stata version 8.2 is used to estimate all the models in this article.

effect of family's literacy on illiterate worker's earnings from activity j. The variables in vector X include years of schooling of the worker, age and age squared, dummy variable that takes the value 1 if the worker's household is located in the cluster labeled "urban", and dummy variable that takes the value 1 if the worker's household is located in the cluster with paved roads. Dummy variable for paved roads are included because local community transportation infrastructure may affect opportunities to commute to wage employment in neighboring areas (Mwabu and Schultz, 2000). The variables in vector Z include dummy variable that takes the value 1 if the worker's household is female-headed and dummy variable that takes the value 1 if the worker's household owns land for cropping or grazing. If at least one variable in Z does not affect earnings directly but affects it indirectly through the choice of activity, the wage function is identified.

4 Results

Table 3 reports estimation results for the first-stage multinomial logit model (2). The assumption of IIA (Independence from Irrelevant Alternatives) is not rejected at 5% significance level for both women and men, and thus the specification of the multinomial logit model is supported. Marginal effects of the family's literacy (LIT) suggest a pattern with increasing the probability of being employed in non-farm sector and decreasing the probability of being employed in farm-related sector, for both women and men. In other words, the proximate illiterate tends to engage in non-farm activities while the isolated illiterate tends to engage in farm-related activities. Also, an illiterate person with more schooling years has a higher probability of finding jobs in non-farm sector.[4] Similarly, an illiterate person who resides in urban area has higher probability of engaging in non-farm activity. Note that these results are obtained as the first-stage regression. The sample selection term $\hat{\lambda}_{i,j}$ is calculated using equation (4).

Table 4 reports estimation results of the second-stage wage equation (5), using sample selection term $\hat{\lambda}_{i,j}$ calculated from the regression of first-stage multinomial logit model. The left-hand side of Table 4 (Model A) reports the results. Family's literacy seems to have a positive and significant effect only on the wage of non-farm activities of illiterate women. A proximate illiterate woman engaging in non-farm sector is expected to be paid 86% ($\simeq e^{0.619} - 1$) higher than an isolated illiterate woman. On the other hand, the effects of family's literacy are positive but insignificant on the wages of farm activities of women and the wages of both farm and non-

[4]Note that years of schooling are at most two years by definition.

farm activities of men. In order to investigate whether or not the wage structure for Africans differs significantly from that of Indians and Colored (Mwabu and Schultz, 2000), I re-estimate the first-stage multinomial logit model and the wage equation using sample comprised of only Africans. The results are shown in the right-hand side of Table 4 (Model B). In effect, the result are similar to that of Model A. Household literacy has a significant and positive effect only on the illiterate women's wages from non-farm activities.

As pointed out in section 2, there is no data in the dataset that directly measures an individual's ability to read or write. Therefore, literacy may be measured with errors. This measurement errors may have caused biases on the above estimates. In order to check whether or not the above results are robust to the measurement errors, three alternative definitions of literacy are examined. The results are reported in Table 5. In Model A, an illiterate person is defined as one who has zero years of schooling. For both women and men, the probability of choosing non-farm activities is increased by the family's literacy (although insignificant for women). An illiterate woman who lives with at least one literate person in her household is expected to be paid 59% higher ($\simeq e^{0.463} - 1$) from non-farm activities than an illiterate woman who lives with no literate person in her household. On the other hand, a literate person is defined as one who has more than three or seven years of schooling in Model B and C, respectively. In effect, the results are similar to that of Model A: A proximate illiterate, both women and men, finds a job in non-farm sector with higher probability, but the effects of family's literacy on wages are significant only for women engaging in non-farm activities. These results are similar to those reported in Table 4, and thus the results of Model A of Table 4 is robust to the measurement errors in literacy.

Variables in X and Z seem to have expected signs. In Table 4, years of education have positive effects on the wages of illiterate workers in many cases. The wage of non-farm activities are higher in urban areas than in rural areas. An illiterate person residing in an area with paved roads receive higher wages from non-farm activities. This indicates that local community transportation infrastructure may increase opportunities to commute to better wage employment in neighboring areas.

5 Conclusions

This paper investigates the external effects of household literacy on the wages and employment of illiterate persons. Taking into account the importance of income from non-farm activities

in developing countries, this paper extends the empirical analysis of Basu et al. (2002) to control for endogenous selection of activities. Specifically, I hypothesized that household literacy affects the earnings of the proximate illiterate through increasing the probability of finding a better job in non-farm sector.

Empirical investigations suggest that a proximate illiterate person finds a job in non-farm sector with higher probability than an isolated illiterate person. Descriptive analysis in Section 2 suggests that, on average, income from non-farm activities are twice of income from farm-related activities. Thus, it is found that the proximate illiterate can earn higher income than the isolated illiterate by engaging in non-farm activities with higher probability. The difference in the probability of finding jobs in non-farm sector between the proximate illiterate and the isolated illiterate can be explained by the amount of information: the proximate illiterate can obtain more information on employment opportunities from easy access to the literate person. On the other hand, the effect of family's literacy on the wages of illiterate workers seems limited: family's literacy only affects women's wages from non-farm activities, and the wages of illiterate men are not affected. This may suggest that women may well be better able to absorb the benefits of household literacy than men (Basu et al., 2002).

At this point, three qualifications are necessary. First, as pointed out in Section 2, the dataset used in this article does not have any direct measures of literacy. Therefore, literacy is defined according to the years of schooling, and robustness of the results are checked. Second, the number of illiterate workers are somewhat small, especially for women. This is a difficult problem because illiterate persons tend to engage in household work. Third, the process of selecting activities may be more complex. For example, decision making might be sequential: an illiterate person decides to work for wages or not in the first stage and then choose farm or non-farm sector in the second stage conditional on the decision made in the first stage. These points are left for a future research.

References

Basu, K., A. Narayan, and M. Ravallion, "Is literacy shared within households? Theory and evidence for Bangladesh," *Labour Economics*, 2002, *8* (6), 649–665.

_ **and J. Foster**, "On Measuring Literacy," *Economic Journal*, 1998, *108*, 1733–1749.

Feder, G., R. E. Just, and D. Zilberman, "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, 1985, *31*, 255–298.

Gibson, J., "Literacy and Intrahousehold Externalities," *World Development*, 2001, *29*, 155–166.

Green, S.E., T.A. Rich, and E.G. Nesman, "Beyond Individual Literacy: The Role of Shared Literacy for Innovation in Guatemala," *Human Organization*, 1985, *44* (4), 313–321.

Kurosaki, T. and H. Khan, "Human Capital, Productivity, and Stratification in Rural Pakistan," 2004. mimeo.

Lanjouw, J.O. and P. Lanjouw, "The Rural Non-farm Sector: Issues and Evidence from Developing Countries," *Agricultural Economics*, 2001, *26* (1), 1–23.

Lanjouw, P., "Rural Non-agricultural Employment and Poverty in Ecuador," *Economic Development and Cultural Change*, 1999, *48* (1), 91–122.

Lee, L.F., "Generalized Econometric Models with Selectivity," *Econometrica*, 1983, *51* (2), 507–512.

Lockheed, M. E., D. T. Jamison, and L. J. Lau, "Farmer Education and Farm Efficiency: A Survey," *Economic Development and Cultural Change*, 1980, *29* (1), 37–76.

McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers in Econometrics*, New York: Academic Press, 1974, pp. 105–142.

Mwabu, G. and T.P. Schultz, "Wage Premiums for Education and Location of South African Workers, by Gender and Race," *Economic Development and Cultural Change*, 2000, *48* (2), 307–334.

SALDRU, *Project for Statistics on Living Standards and Development: South Africans rich and poor: baseline household statistics*, University of Cape Town: Cape Town: South African Labour and Development Research Unit, 1994.

Strauss, J. and D. Thomas, "Human Resources: Empirical Modeling of Household and Family Decisions," in J. Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, Vol. 3, Amsterdam: North-Holland, 1995.

Table 1
Descriptive Statistics

	Whole Sample Ages 16-65					Sub-Sample of Illiterates Ages 16-65				
	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
wage (farm-related jobs)	1288	462.679	444.846	0	3870	559	314.303	289.112	0	1800
women	262	225.749	224.422	0	1300	127	192.324	173.712	0	800
men	1026	523.181	466.511	0	3870	432	350.163	306.138	0	1800
wage (non-farm jobs)	4356	894.751	777.858	0	14620	701	603.406	512.604	0	3500
women	1995	715.761	636.717	0	6000	325	429.383	475.671	0	2600
men	2361	1045.993	850.907	0	14620	376	753.824	496.052	0	3500
family's literacy	14120	0.842	0.365	0	1	3534	0.555	0.497	0	1
years of education	14120	5.261	3.386	0	14	3534	0.662	0.817	0	2
age	14120	34.703	11.373	16	65	3534	39.170	12.309	16	65
living in urban area	14120	0.474	0.499	0	1	3534	0.264	0.441	0	1
with paved roads	14120	0.203	0.402	0	1	3534	0.088	0.283	0	1
sex of the household head	14120	0.261	0.439	0	1	3534	0.233	0.423	0	1
land ownership	14120	0.038	0.192	0	1	3534	0.048	0.215	0	1

Note:
Wage is measured in Rand (1 US$ = 3.55 Rand in 1994).

Table 2
Differences between the Proximate Illiterate and the Isolated Illiterate

	Obs.	Mean	Std. Dev.	Min	Max
Proximate illiterate					
wage	594	537.377	493.204	0	3500
years of education	1962	0.692	0.824	0	2
age	1962	40.920	12.240	16	65
living in urban area	1962	0.298	0.458	0	1
with paved roads	1962	0.085	0.278	0	1
sex of household head	1962	0.248	0.432	0	1
land ownership	1962	0.064	0.245	0	1
Isolated illiterate					
wage	667	419.162	402.826	0	2485.4
years of education	1572	0.624	0.808	0	2
age	1572	36.986	12.046	16	65
living in urban area	1572	0.222	0.416	0	1
with paved roads	1572	0.092	0.289	0	1
sex of household head	1572	0.214	0.410	0	1
land ownership	1572	0.029	0.167	0	1

Note:
Wage is measured in Rand (1 US$ = 3.55 Rand in 1994). The proximate illiterate is defined as an individual who lives with at least one literate member in the household. The isolated illiterate is defined as an individual who lives with no literate member in the household.

Table 3
Estimation Results of the Multinomial Logit Model

Dependent variable: Prob($d_i = j$)				
		Marginal effects on the probability of choosing activity j		
		$j = 0$ (household work)	$j = 1$ (farm wage)	$j = 2$ (non-farm wage)
Women	family's literacy (*LIT*)	0.037	-0.038	0.001
	years of education	-0.020	-0.005	0.025
	age	-0.021	-0.001	0.022
	age squared/100	0.025	0.001	-0.025
	urban (dummy)	-0.128	-0.061	0.189
	paved roads (dummy)	-0.069	-0.022	0.091
	sex of household head	-0.084	0.005	0.079
	land owner (dummy)	0.070	-0.034	-0.037
	number of observations	1564	124	313
Men	family's literacy (*LIT*)	0.146	-0.180	0.034
	years of education	0.001	-0.022	0.021
	age	-0.038	0.021	0.017
	age squared/100	0.039	-0.023	-0.016
	urban (dummy)	-0.015	-0.301	0.316
	paved roads (dummy)	-0.174	0.201	-0.027
	female headed (dummy)	0.304	-0.187	-0.118
	land owner (dummy)	0.079	-0.149	0.070
	number of observations	710	418	352

Note:
Marginal effects on the probability of choosing activity j are shown. Only those explanatory variables whose γ_j is statistically significant at least 10% level for some j are included.

Table 4
Estimation Results of the Wage Equation

Dependent variable: ln W					
		Model A: Full Sample		Model B: African Sample	
		Farm-related activity	Non-farm activity	Farm-related activity	Non-farm activity
Women	family's literacy (LIT)	0.013 (0.221)	0.619 (0.103)	-0.085 (0.185)	0.502 (0.116)
	years of education	0.125 (0.095)	0.105 (0.060)	0.124 (0.099)	0.081 (0.063)
	age	-0.023 (0.036)	0.001 (0.032)	-0.039 (0.038)	0.004 (0.038)
	age squared/100	0.032 (0.043)	-0.014 (0.038)	0.055 (0.045)	-0.013 (0.043)
	urban (dummy)	-1.567 (2.024)	0.396 (0.159)	-1.149 (2.535)	0.325 (0.151)
	paved roads (dummy)	-0.122 (0.313)	0.342 (0.143)	-0.208 (0.514)	0.131 (0.155)
	selection correction term ($\hat{\lambda}$)	0.034 (0.034)	0.029 (0.062)	0.022 (0.032)	0.002 (0.052)
	constant	5.141 (0.648)	5.096 (0.800)	5.422 (0.645)	5.131 (0.913)
	number of observations	124	313	111	284
Men	family's literacy (LIT)	0.179 (0.122)	0.031 (0.081)	0.156 (0.126)	-0.014 (0.083)
	years of education	0.116 (0.057)	-0.022 (0.039)	0.103 (0.058)	-0.003 (0.042)
	age	0.059 (0.026)	0.048 (0.030)	0.056 (0.027)	0.058 (0.033)
	age squared/100	-0.065 (0.030)	-0.051 (0.031)	-0.060 (0.031)	-0.061 (0.034)
	urban (dummy)	0.821 (0.299)	0.212 (0.209)	0.825 (0.296)	0.225 (0.185)
	paved roads (dummy)	0.952 (0.237)	0.202 (0.137)	0.966 (0.238)	0.216 (0.165)
	selection correction term ($\hat{\lambda}$)	-0.034 (0.020)	-0.041 (0.117)	-0.036 (0.022)	-0.030 (0.101)
	constant	4.055 (0.503)	5.364 (0.947)	4.114 (0.530)	5.095 (0.978)
	number of observations	418	352	404	324

Note:
Figures in parentheses are standard errors robust to clustering. Selection correction term ($\hat{\lambda}$) is calculated using equation (4) after first-stage multinomial logit estimation reported in Table 3.

Table 5
Robustness Checks: Effects of Family's Literacy on the Choice of Activities and Wages

		Model A: Illiterate defined as 0 years of schooling		Model B: Literate defined as years of schooling ≥ 3		Model C: Literate defined as years of schooling ≥ 7	
		Farm	Non-farm	Farm	Non-farm	Farm	Non-farm
Women	Choice	-0.005	0.009	-0.032	-0.029	-0.028	0.029
		(0.002)	(0.021)	(0.015)	(0.020)	(0.012)	(0.017)
	Wage	-0.369	0.463	-0.090	0.499	0.380	0.722
		(0.190)	(0.142)	(0.179)	(0.107)	(0.220)	(0.104)
	number of observations	75	150	124	313	124	313
Men	Choice	-0.161	0.060	-0.214	0.030	-0.194	0.077
		(0.050)	(0.039)	(0.053)	(0.029)	(0.035)	(0.031)
	Wage	0.279	0.046	-0.032	-0.023	0.399	0.118
		(0.128)	(0.100)	(0.131)	(0.096)	(0.125)	(0.103)
	number of observations	240	167	418	352	418	352

Note:
Estimates of family's literacy (*LIT*) are shown. Figures in parentheses are standard errors robust to clustering. Estimates of the first-stage multinomial logit model are shown in rows labeled "Choice," and estimates of the second-stage wage equation are shown in rows labeled "Wage." Variables in *X* and *Z* are included in the first- and second-stage regressions, but results are not shown.