HITOTSUBASHI UNIVERSITY

Global COE Hi-Stat Discussion Paper Series 080

Research Unit for Statistical and Empirical Analysis in Social Sciences (Hi-Stat)

Education and Its Distributional Impacts on Living Standards: Evidence from Rural India

Takahiro Ito

August 2009
(Revised: March 2010)

Hi-Stat Discussion Paper

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

Education and Its Distributional Impacts on Living Standards: Evidence from Rural India[†]

Takahiro Ito [‡]

This version: March 2010
(First version: August 2009)

Abstract

This paper investigates the determinants of living standards (measured by per capita consumption expenditure) at the household level, addressing heterogeneity in the impact of education and endogeneity of educational attainment. The estimation results obtained through an instrumental variables quantile regression suggest that the endogeneity of education matters in determining the causal effect of education on living standards. On the other hand, no evidence of heterogeneity in the *percentage* impact of education is found. However, the results also provide evidence that the impact of other determinants varies significantly over the outcome (expenditure) distribution, and consequently a simulation based on the results shows that the *level* impact of education on consumption expenditure differs substantially between the instrumental variables quantile regression and standard instrumental variables regression results. The comparison of the two shows that the poverty alleviation impact of education estimated through the instrumental variables quantile regression are much smaller than the impact estimated through the standard instrumental variable regression.

JEL classification: D12, I21, I32, O15

Key words: rural poverty, heterogeneous impact of education, instrumental variables quantile regression

---

[†] I am deeply grateful to Takashi Kurosaki and Daiji Kawaguchi for extensive discussions and suggestions on this paper. I am also indebted to Aparajita Bakshi, Yukinobu Kitamura, Kohei Kubota, Fumiko Oshikawa, Chiaki Moriguchi, Takeshi Sakurai, Masaru Sasaki, Takahiro Sato, Deepa Sinha, Shinsuke Tanaka, Yoshifumi Usami and participants at various seminars at Hitotsubashi University and Osaka University for their helpful comments. Finally, I acknowledge the financial support from the Global COE Program "Research Unit for Statistical and Empirical Analysis in Social Sciences."
[‡] Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047 Japan. E-mail: t_ito@iser.osaka-u.ac.jp.

1. Introduction

Despite the past two decades of significant economic growth, a high incidence of poverty remains an important policy issue in India. One of the latest estimates indicates that about 35% of the population (or 360 million people), which accounts for approximately one-third of the world's poor, still lived on less than one PPP dollar a day (UNDP, 2007). Although the incidence of poverty has been declining gradually and steadily, understanding how poverty can be alleviated significantly remains an issue of considerable concern to policy makers.

This paper investigates the causal relationship between poverty and education at the household level. There seem to be two distinct strands in the literature on the determinants of poverty. The first strand is poverty profiles, which are the commonly-used way of providing information on the characteristics of the poor. Calculating and tabulating poverty measures by communities, social classes, and/or other characteristics enable us to investigate the determinants of poverty. However, they are generally unsuitable for examining the effects of multivariate factors on poverty, since cross-tabulations become increasingly complex as the number of factors increases. The second strand is the regression approach, where living standards measured by household income or consumption are regressed on several factors. Due, in part, to its simplicity and usefulness, there is a vast body of literature on regression analysis of the determinants of living standards or poverty at the micro/household level.

This paper also conducts a multivariate regression analysis of living standards (measured by per capita consumption expenditure) at the household level, paying special attention to the role of education. While the importance of education to poverty reduction appears to be commonly recognized among researchers at the micro level,[1] some studies have suggested that educational returns may vary widely depending on the standard of living. For example, agricultural wages are likely to be less responsive to schooling attainment, and gaps in educational returns between agricultural and non-agricultural sectors are noticeable.[2] In the context of rural economies in developing countries, this sectoral difference in educational returns may not be negligible, since the degree of economic dependence on agriculture is rather high and a large share of the poor are agricultural wage workers. There is also a issue of the quality of education. If there is a sorting of households into different quality schools based on their income levels, educational returns could be heterogeneous according to the level of income. It is frequently found in developing countries like India that poorer households cannot afford to send their children to a high-quality (private) school. These facts suggest that educational returns for low-income households are less than those for high-income households.

Thus, the aim of this paper is to quantify the heterogeneous impacts of education on living standards, and this will be done by adopting a quantile regression model.[3] The quantile regression model that allows the effect of variables to vary according to the quantiles of an outcome distribution enables us to examine whether the heterogeneity in the impact of education is substantial, and to what extent. Especially, the impact of education on living standards for the poor is of great interest to both policy makers and

---

[1] Tilak (2007) also found a negative correlation between education and poverty using semi-macro (state) level data in India. In this regard, however, macro/country level studies have often found no correlation between schooling expansion and per capita GDP growth (See, e.g., Pritchett, 2001; and Easterly, 2001).
[2] See, e.g., Kurosaki and Khan (2006), Dutta (2006), and Ito (2009) for South Asian studies.
[3] There are several theoretical and empirical studies on heterogeneous returns to education. See, for instance, Wooldridge (1997), Heckman and Vytlacil (1998), Arias et. al. (2001), Heckman and Li

academic researchers. In addition, this paper employs an instrumental variables estimation for the quantile regression model. As it is naturally expected that households' living standards simultaneously affect their members' educational status (Behrman and Knowles, 1999), the endogeneity of education may also matter in the empirical analysis of educational returns. Thus, the main contribution of this paper is to investigate the empirical relationship between living standards and education, addressing issues of both the heterogeneity in educational returns and the endogeneity of education simultaneously.

The remainder of this paper is structured as follows. In the next section, the data set used in the analysis is described. The sample consists of rural households in Bihar and Uttar Pradesh in north India. Section 3 briefly discusses econometric issues in relation to the causal model for education and earnings and presents the empirical models adopted in this paper. The empirical results are presented in Section 4. The estimation results obtained through the instrumental quantile regression suggest that the endogeneity of education matters in determining the causal effect of education on living standards. On the other hand, no evidence of the heterogeneity in the *percentage* impact of education is found. However, the results also provide evidence that the impact of other determinants varies significantly over the outcome (per capita expenditure) distribution, and consequently a simulation based on the results shows that the *level* impact of education on consumption expenditure differs substantially between the instrumental variables quantile regression and standard instrumental variables regression results. The comparison of the two shows that the poverty alleviation impact of education estimated through the instrumental variables quantile regression are much smaller than the impact estimated through the standard instrumental variable regression. Section 5 concludes the

(2004), and Carneiro et. al. (2006).

paper.

2. Data

The data employed in this paper are from the *Survey of Living Conditions*, *Uttar Pradesh and Bihar*, which is one of the Living Standard Measurement Study (LSMS) surveys. The survey was conducted in 1997/98 and covered 1,035 households, 57 villages, and 13 districts in Bihar and 1,215 households, 63 villages, and 12 districts in Uttar Pradesh (UP). UP and Bihar are located in the Ganges Plain of north India and are known for their high incidence of poverty. An official estimate in 1999 says that the ratios of rural population living below the poverty line in UP and Bihar are respectively 31.22% and 44.30%, while the ratio in all India is 27.09%.[4]

Table 1 shows the key features of the sample households by per capita expenditure quartiles. Among 2,250 households included in the survey, 2,062 households are used here, after excluding households with missing information on related characteristics. As the table shows, poorer households have less farm land, fewer working-age (and more dependent) members, and fewer educated members. "Working-age" is defined as ages between 15 and 60, and "schooling years" denotes the average number of schooling years among working-age adults. The table suggests that human and productive capitals are very important in determining households' welfare. It is also worth noting that the average level of education in this study region is considerably low: even for the richest group, it is less than 5 years. Considering the fact that Indian compulsory education system consists fundamentally of 5 years of primary school and 3

[4] These figures are based on the government's official estimates (GOI, 2001). There has been an ongoing debate on poverty estimates for India, even among researchers using the same micro data collected by the National Sample Survey Organization. See, for example, Deaton and Kozel (2005).

years of middle school, this figure indicates that the study region falls behind in education level.

The last four rows of the table show the fraction of households belonging to each caste category. “Caste” represents the traditional hereditary class. Although castes consist of thousands of endogamous groups called *jatis* (the word literally means “birth”), the classification in the table is based on the classification in the survey. Note that scheduled castes (*dalits*, or once known as “untouchables”) and scheduled tribes are those who sit at the bottom of the social hierarchy. As can be seen from the table, households belonging to the upper or middle Hindu castes are more likely to be rich, while scheduled castes and tribes are more likely to be poor. This implies that caste-based discrimination is still severe in this study region, despite that several policy efforts against the discrimination have been implemented since independence. This is consistent with the implications of Ito’s (2009) study on caste discrimination in the labor market using the same dataset.

[Table 1]

To investigate the possibility of the heterogeneous impact of education graphically, a boxplot analysis is implemented (Figure 1). The boxes represent the interquartile range of monthly per-capita expenditure for each group, classified by the average years of schooling of adult members (aged 15 to 60 years). With the exception of a dip at the 4th year of schooling, consumption expenditure is steadily increasing with schooling years. In line with previous studies, this implies that education plays an important role in improving living standards in this study region. Moreover, the boxes show that increases in expenditure with schooling years expand at an increasing rate at the 75th percentile, while those are relatively constant at the lower percentiles. This suggests the impact of education on consumption expenditure vary according to income (consumption) level. This will be further investigated using regression techniques by introducing additional covariates and allowing endogeneity of educational status.

[Figure 1]

3. Empirical Specification

3.1. Econometric Issues and the Empirical Model

This section briefly discusses econometric issues related to the causal model for schooling and earnings outcomes[5] and presents the empirical methods adopted in the analysis. For the sake of simplicity, consider the following model where the log of household income ($\ln Y_i$) is a function of education level (E_i) with an intercept:

(1) $$\ln Y_i = a_i + b_i E_i,$$

where a_i and b_i are household i's specific attribute (e.g., *ability* associated with income generation) and the household's marginal rate of returns to education, respectively.[6] Issues related to a measurement of household-level educational attainment (E_i) are discussed later.

If there is *no* household-specific heterogeneity through either the intercept (a_i) or

[5] The discussion below owes much to Card's (1999) excellent survey on this topic.
[6] Usually, when it comes to an estimation of the "rate of returns to education," the analysis is implemented at individual level by regressing individuals' wages on their schooling years. However, in developing countries like India, the majority of the labor force population is engaged in self-employment activities, and this is especially true in rural areas. Together with the fact that income generated from self-employed activities is almost always measured at the household level, analyses need to be implemented at the household level in quantifying the impact of education on living standards. In this connection, the term "rate of returns to education" does not seem appropriate in the household-level analysis, but it is referred to in a broad sense as the "household rate of returns to education."

the schooling coefficient (b_i), the ordinary least squares (OLS) regression of Equation (1) can provide consistent estimates of α and β (as the means of a_i and b_i). However, if heterogeneity does exist, this is not the case. Equation (1) can be rewritten as:

(2) $$\ln Y_i = \alpha + \beta E_i + (a_i - \alpha) + (b_i - \beta)E_i,$$

In this case, the OLS estimator for the average impact of education (β) has bias, as its probability limit is expressed as:

(3) $$\text{plim } \beta_{\text{ols}} = \beta + \text{Cov}(a_i, E_i) / \text{Var}(E_i) + \overline{E}\text{Cov}(b_i, E_i) / \text{Var}(E_i),$$

where $\overline{E}$ denotes the mean of schooling (E_i).[7] The second term in Equation (3) represents the bias owing to the correlation between education (E_i) and the individual-specific attribute (a_i), and the last term results from the correlation between education level (E_i) and its slope (b_i). Hence, the OLS estimator suffers from two possible biases because of unobserved heterogeneity (a_i) and heterogeneous schooling coefficients (b_i).

One common way to account for this situation would be to employ an instrumental variables (IV) estimation. Using instrumental variables Z_i that are uncorrelated with a_i and b_i, the IV estimator for β is able to eliminates bias resulting from unobserved heterogeneity (a_i) and heterogeneous effects (b_i) if $\text{E}[(b_i - \beta)E_i,| Z_i]$ is not a function of Z_i, since

(4) $$\begin{aligned} \text{E}[\ln Y_i \mid Z_i] &= \alpha + \beta Z_i + \text{E}[(a_i - \alpha) \mid Z_i] + \text{E}[(b_i - \beta)\, E_i \mid Z_i] \\ &= \alpha + \beta Z_i + \text{E}[(b_i - \beta)\, E_i \mid Z_i]. \end{aligned}$$

Nevertheless, in the case that $\text{E}[(b_i - \beta)E_i,| Z_i]$ is dependent on Z_i, the standard IV regression cannot rule out the influence of heterogeneous returns (b_i) on the causal effect

[7] Note that Equation (3) can be derived if E_i and b_i have a jointly symmetric distribution. See Appendix A in Card (1999) for the derivation.

of education on living standards. In order to deal with this possibility, therefore, an alternative IV regression technique, control function approach (Garen, 1984) is employed.[8] Under the zero conditional mean assumptions of $E[(a_i - \alpha) \mid Z_i] = 0$ and $E[(b_i - \beta) \mid Z_i] = 0$, the following equation is derived using linear projections of $(a_i - \alpha)$ and $(b_i - \beta)$ on (E_i, Z_i);

$$(5) \qquad E[\ln Y_i \mid E_i, Z_i] = \alpha + \beta E_i + E[(a_i - \alpha) \mid E_i, Z_i] + E[(b_i - \beta) \mid E_i, Z_i]E_i$$

$$= \alpha + \beta E_i + \gamma v_i + \delta v_i E_i,$$

where v_i is the random error from the linear projection of E_i on Z_i, and γ and δ are parameters to be estimated.[9] Thus, inclusion of v_i and v_iE_i purges two possible biases owing to heterogeneity via a_i and b_i.

Furthermore, in order to estimate the heterogeneous effects of intercept and education, a completely different approach, quantile regression (QR) model, is also employed. Heterogeneous effects models with unobserved heterogeneity, as expressed in Equation (1), are an important application field for the QR model (Koenker, 2005). The QR model allows explanatory variables to have different impacts according to the outcome distribution by estimating:

$$(7) \qquad \ln Y = Q_{\ln Y|E}(U) = \alpha(U) + \beta(U)E, \; U \sim \text{Uniform}(0,1),$$

where U is the rank variable that indexes household-specific heterogeneity, and $Q_{\ln Y|E}(U)$ is the quantile function of $\ln Y$ conditional on E, which is increasing in U. Thus, heterogeneity of income for households with the same level of education is characterized by the rank variable U and is captured as quantile treatment effects denoted by $\alpha(\tau)$ and $\beta(\tau)$ ($\tau \in U$). However, as in the case of OLS, when the household-specific heterogeneity

[8] For further discussion on the related topic, see also Wooldridge (1997), and Heckman and Vytlacil (1998).

(via a_i and b_i), which is partly captured through U, is correlated with educational attainment (E_i), running the standard QR model may cause the estimates to be biased. To account for this possibility of bias, following the procedure proposed by Chernozhukov and Hansen (2006),[10] the instrumental variables estimation of the quantile regression (IV-QR) model is employed, and the endogeneity is statistically tested by comparing QR estimates and IV-QR estimates. The choice of valid instrumental variables, which should be independent of the rank variable U, is discussed in the next subsection.

3.2. Empirical Variables

The estimation of Equation (1) is implemented at the household level to investigate the relationship between living standards and education. Out of 2,250 households in the Indian LSMS survey, 2,062 households are used in the analysis, after excluding households with missing information on related characteristics. Regarding the dependent variable $\ln Y$ in Equation (1), the log of monthly per capita consumption expenditure (*Log of expenditure*) is employed. Thus, the analysis focuses on income/consumption poverty, though poverty itself is a multidimensional concept. The reason for using consumption expenditure instead of income is that expenditure is smoothed to some extent, whereas income is subject to various transitory shocks and is likely to be volatile (Walker and Ryan, 1991; Townsend, 1994). Hence, consumption expenditure is considered a better proxy for permanent income.

Explanatory variables are the standard ones commonly used in this type of regression (see, e.g., Datt and Jolliffe, 2005). Regarding a household-level educational

---

[9] See Appendix A in Card (1999) for the derivation.

[10] With regard to necessary conditions for parameter identification, see Chernozhukov and Hansen (2005, 2006).

attainment (*E*), unlike individual-level estimations of earnings functions, there is an issue of how the education level in a household is measured. For instance, based on several hypotheses, Jolliffe (2002) tested three measures of education (the average level, maximum level, and minimum level of education in the household) and found that that the maximum level of education is a significant determinant of total household income using Ghanaian household data.[11] In the analysis in this paper, however, the average schooling years for working-age members (between 15 to 60 years) is used, since a preliminary analysis shows that the average level of education is the only statistically significant determinant of *Log of expenditure*, among the three education measures.[12]

In addition to this *schooling years* variable, various controls for household composition and farming assets are also included to avoid possible omitted variable bias. Controls for household composition are household size (*Household size*), ratio of male members (*Male ratio*), ratio of non-working-age members (*Dependency ratio*), average age for working-age members (*Age*), age of the household head (*Age of head*), and dummy for female-head households (*Female head*). With regard to controls for farming assets, the size of farmland owned by households (*land owned*), the share of irrigated farmland (*Irrigation ratio*), the value of semi-fixed capital in agricultural production (*Agric. capital*) and livestock (*Livestock*) are employed.

Finally, instrumental variables for the first-stage regression of *Schooling years* are discussed. As mentioned earlier, instruments should be independent of the rank

[11] Education is thought to enhance workers' productivity through two distinct channels — improving their ability to produce more with given resources, technology, or information and improving access to resources, technology, or information sources (Welch, 1970; Rosenzweig, 1995). The average level of education is expected to capture mainly the former, "productivity effect", while the maximum level of education is expected to capture the latter, "allocative effect" (Yang, 1997).

[12] The preliminary analysis is conducted by the OLS and standard quantile regression. The result is available on request.

variable U, which characterizes household-specific unobserved heterogeneity (via a_i and b_i). One of the candidates is the accessibility of education, since supply-side variations are more likely to be uncorrelated with household (unobserved) characteristics. In our data set, information about distance to the nearest schools (public primary school, middle school, and secondary school) is available. However, there is a possibility that proximity to schools are correlated with unobserved characteristics. For instance, India has a strong caste-based power structure, and hence public goods like education facilities might be provided in favor of the upper castes (Banerjee et. al., 2005). To see this, Table 2 shows the distances to several facilities by caste.

[Table 2]

As can be seen from the table, these facilities are not necessarily in the upper caste neighborhoods. Looking at educational facilities, the lower castes have worse access than the upper castes, but only 'Distance to prim. school' shows statistically significant difference (especially between the upper caste and SC/ST). This might be due to the fact that it is the distance to the nearest *public* school in the case of primary schools, while a large share of schools is *private* in the cases of middle and secondary schools (no information about *public* or *private* is available from our dataset). In short, the location of *public* schools could be influenced by a local power structure mainly based on caste. To eliminate the influence of such caste-based political power, distances to middle school (*Distance to mid. school*) and secondary school (*Distance to sec. school*) are used as instruments, and caste dummies are also included as additional regressors.

[Table 3]

Furthermore, it is possible that proximity to middle or secondary schools means proximity to urban areas, since these schools are more likely to locate in areas where there is a demand for higher education. Hence, households residing close to a school may enjoy higher returns to their human capital. Table 3 shows correlation coefficients between village-level wage rates for skilled workers and distance to each facility. As is expected, distances to the nearest schools are negatively correlated with wage rates (the first row), but correlation coefficients calculated using the limited sample of villages having a bank within 5 kilometers, in contrast, are positive (the second row). This indicates that distance to the nearest middle and secondary school could be used as exogenous shifters of educational attainment after controlling for distance to other facilities. Thus, *Distance to police*, *Distance to bank*, *Distance to hospital*, and *Distance to prim. school* are also included as additional explanatory variables. In addition, village-level wage rates for skilled workers (*wage rates*) are also controlled to eliminate any possibility of bias resulted from unobserved household-specific heterogeneity. For the summary statistics of these empirical variables, see Table 4.

[Table 4]

4. Estimation Results

4.1. Estimation of the Average impacts of Education

Table 5 shows estimation results of first- and second-stage OLS estimation. Looking at the first-stage estimation result, the instrumental variables, *Distance to mid. school* and

Distance to sec. school, have negative impacts on schooling years, indicating that households advance their members' education as middle and secondary schools are nearer from their residence. Regarding the validity of these instrumental variables, F statistic of 13.44 indicates the rejection of the null hypothesis that both variables have no effect and the Sargan type test of overidentification restrictions (χ^2 statistic of 1.215) suggests that the both instruments are independent from the error term in the wage equation.

In the second and third columns, estimation results obtained through OLS and IV regressions are reported. The comparison of schooling coefficients shows that the coefficient obtained the IV regression is more than twice the coefficients obtained through the OLS regression: 0.137 for IV and 0.060 for OLS. The results suggest that educational attainment (E_i) is negatively correlated with income-generating ability (a_i), and this can be explainable on the basis of the comparative advantage hypothesis of Willis and Rosen (1979). Consider a simple case, in which people with no formal schooling are hired for unskilled manual labor and those with a formal education are hired for skilled labor, that is, there are only two school levels (S = 0, 1) and two job types (u: unskilled, and s: skilled). Suppose further that the ability term a_i affects earnings differently among different jobs ($a_i^u \neq a_i^s$). In case that people with talent as an unskilled laborer do not necessarily have talent as a skilled laborer or vise verse, people choosing no formal education are more likely to be those with talent as an unskilled laborer. This being the case, the ability term a_i could be negatively correlated with schooling attainment. Another possible explanation of the results — the IV estimate is bigger than OLS estimate — is attenuation bias owing to measurement errors. Since the average schooling years per working-age member in a household may not capture the true value of education in the household, there is a possibility that the result reflects the influence of measurement errors to a large extent. Note that the Hausman test statistic of 2.679, however, cannot rejects the null hypothesis that the difference between the OLS and IV estimate of the average impact of education is not systematic (at the 10% significant level).

[Table 5]

In the last column of Table 5, the estimation results by a control function approach is reported.[13] As can be seen from the table, the control function approach estimates are almost unchanged from the standard IV regression estimates, implying no bias arising from the correlation between b_i and E_i. This is also confirmed by the fact that the coefficient of v_iE_i is statistically insignificant. The results indicate that the rate of returns to schooling (b_i) seems not to differ widely among households, but the heterogeneity in income-generating ability (a_i) appears to be significant. This will be further investigated in the next subsection by adopting quantile regression models.

4.2. Estimation of the Distributional impacts of Education

Figure 2 depicts the distributional impacts of education (in the first panel) and the constant term (in the second panel).[14] In both panels, the IV-QR estimates (solid line) with the 90% confidence interval (shaded area), and the standard IV regression (2SLS) estimate (dashed line) are plotted. In the left panel, the QR estimates (dashed-dotted line)

[13] In the estimation of Equation (4), the estimated residual $\hat{v}_i$ from the first-stage estimation of education E_i is substituted for v_i.

[14] I wrote a STATA program for the estimation of an IV–QR model using the procedure proposed by Chernozhukov and Hansen (2006). As optimization tools are required in the procedure, a combination of the grid search and the simplex method are used. The program is available on request.

are also reported. Coefficient estimates on all variables obtained through the QR and IV-QR model for the 25th, 50th and 75th quantiles are reported in Table 6.

[Figure 2]

Looking at the first panel, the QR and IV-QR estimates of the schooling coefficient appear to vary greatly according to the quantiles of the expenditure distribution, but the IV-QR estimates is more likely to be volatile relative to the QR estimates. The IV-QR estimates of the schooling coefficient range from 0.04 to 0.18, and this result is broadly consistent with the previous studies on the rate of returns to education in India (See, e.g., Kingdon, 1999). Further, the differences in the schooling coefficients between the QR and IV-QR estimates appear to be substantial, implying that the endogeneity of schooling attainment is crucial. The Hausman χ^2 statistic of 39.55 indicates the rejection of the null hypothesis of no endogeneity at the 1% significant level (the first test of Table 7). This is very interesting, since the null of no endogeneity of education cannot be rejected at the 10% level in the analysis of the average impact of education, as discussed in the previous subsection.

The shape of the IV-QR estimates of the schooling coefficient is also interesting: smaller schooling coefficients for those in the lower tail of the expenditure distribution. Although the null hypothesis of no heterogeneity in the impact of education cannot be rejected statistically (the second test of Table 7), the figure implies that poorer households have relatively smaller gains to educational investment. This seems to be at odds with the anticipated pattern, since it is usually expected that the marginal rate of returns to education diminishes with the level of schooling and then poorer households have less educated members. In the context of Indian labor market, however, a number of studies on the Mincerian-type wage regression have found the rate of returns to education is increasing along with the level of schooling (See, e.g. Kingdon, 1998; Duraisamy, 2002; Dutta, 2006). Hence, the result obtained here may reflect the specific circumstance in Indian labor market.

In the last panel of Figure 2, the IV-QR estimates of the constant term, which are expected to partly capture distributional impacts of household-specific ability (a_i), are depicted. The figure shows that the magnitude of household-specific ability increases steadily along with the quantile index and the *F* statistic of 1.66 indicates the rejection of the hypothesis of no heterogeneity in ability (the forth test of Table 7).[15]

[Table 4]

[Table 5]

4.3. Poverty alleviation impacts of education

The previous two subsections focus on the *percentage* increase in per-capita consumption expenditure when the average years of education increases by one year. However, this does not say anything about the *level* impact of education on expenditure, and remember that Figure 1 suggests the existence of heterogeneous impact of education in terms of *level* change in expenditure. This subsection, therefore, investigates the level increase of expenditure with additional year of education.

[15] To investigate the robustness of the result in Figure 2, several specifications are tested in which caste dummies are excluded and/or additional village-level characteristics are included in the equation. Although the magnitude of coefficient estimates are slightly changed, the main result of this paper is

In logarithmic models, like Equation (1), impact in terms of *level* change (not *percentage* change) is calculated by:

$$(7) \qquad \partial Y_i \,/\, \partial E_i = \partial Y_i \,/\, \partial \ln Y_i \times \partial \ln Y_i \,/\, \partial E_i = Y_i \times b_i$$

$$= \exp\{a_i + b_i E_i\} b_i.$$

Thus, the impact of education on consumption expenditure (Y, not $\ln Y$) depend not only on b_i but also on a_i. In short, the constant *percentage* impact of education, which is found in the previous subsections, does not necessary mean the constant *level* impact of education, because heterogeneity in a_i (and, of course, other variables' impacts) also matters in calculating the *level* impact of education. The F statistic of 10.21 rejects the null hypothesis that all coefficients are constant (the sixth test in Table 7).

[Figure 3]

Figure 3 plots simulation results of *level* changes in consumption expenditure when the average schooling years increases by one year, holding other variables constant.[16] The dashed line is the results calculated from the standard IV estimate, and the solid line represents the results calculated from the IV-QR estimates. In the calculation, the distributional impacts based on the standard IV estimate are obtained by calculating the predicted values of Equation (7) for all observations and sorting them. On the other hand, regarding the IV-QR estimates, the distributional impacts are calculated at each quantile point by using

$$(8) \qquad \partial Y \,/\, \partial E = \partial Y \,/\, \partial \ln Y \times \partial \ln Y \,/\, \partial E = Y \times \beta(U) = \exp\{Q_{\ln Y|E}(U)\}\beta(U)$$

---

not changed.

[16] In this simulation, general equilibrium effects — possible changes in wages and/or employment

$$= \exp\{\alpha(U) + \beta(U)E\}\beta(U), \quad U \sim \text{Uniform}(0,1).$$

The figure shows that both the IV regression and IV-QR estimates increase with the quantile index, but the IV-QR estimates have a sharper inclination than the standard IV estimates and gaps between the two appear to be substantial. The IV-QR estimates range from approximately Rs. 30 to Rs. 300, whereas the IV regression estimates range from about Rs. 80 to Rs. 200. Moreover, the gaps at the lower quantiles of the expenditure distribution suggest a possibility that the impact of education on poverty alleviation are overestimated in the IV regression model. A simple calculation based on the IV regression estimates indicates that additional increases in the schooling years of all adults reduce the number of households living on less than 2.16 dollar per day at 1993 PPP (approximately 22 rupees per day in the 1998 price level) from 750 to 498, whereas based on the IV-QR estimates, the number of households living poor declines to 613. The former reduces the poverty ratio by 0.122, while the latter reduces by only 0.066. Thus, it is likely that the standard IV result overestimates the poverty reduction impact of education.

An interpretation of the result — small increases in expenditure with additional schooling at the lower quantiles in the expenditure distribution — is that poorer households are likely to be engaged in agriculture-related unskilled work that is less responsive to human capital enhancements. If this is the case, promoting non-farm employment would be a crucial measure to reduce poverty in this study region, as suggested by Lanjouw and Shariff (2004). Another possibility is the issue of the quality of education. As mentioned earlier, the result also interpretable as evidence that low-income households cannot afford to send their children into high-quality private schools and

---

arising from increases in supply of workers with higher education — are ignored.

consequently their educational returns are low. In this case, improving the quality of public school might be another policy tool to improve living standards of poor households.

5. Conclusion

This paper examined the determinants of living standards by employing several regression models to address heterogeneity in the impact of education and endogeneity of educational attainment. The comparison between the analysis on the *average* impact of education (through the OLS and the IV estimations) and the *quantile* impacts of education (through the QR and IV-QR estimations) provides very interesting results.

In the analysis on the *average* impact of education, the estimation results indicate no evidence of both the heterogeneity and endogeneity. On the other hand, the estimates of the *quantile* impacts of education show that differences in schooling coefficients between the QR and IV-QR are large and the null of no endogeneity of education is statistically rejected at the 1% significance level. Regarding the heterogeneity in educational returns, the estimation results indicate that the impact of education (in terms of *percentage* change in consumption expenditure) is constant over the expenditure distribution. However, the results also provide evidence that the impact of other determinants varies significantly according to the quantiles of the expenditure distribution, and consequently, the IV-QR estimates of the *level* impact of education differs substantially from the standard IV regression estimate.

Thus, there is a sharp contrast between the findings obtained from the standard IV regression and IV-QR, and this implies that focusing not only on the *average* impact but also on the *distributional* impacts of education is quite important to understand the role of education on poverty alleviation. In particular, the simulation exercise based on the estimation results shows that poverty alleviation impacts of education is much smaller when taking into account the heterogeneous impacts of education (and other explanatory variables) than when focusing only on the *average* impact. Thus, for instance, policies trying to achieve “universal primary education” may not have much effect on poverty alleviation without improving the opportunity of wage works responsive to human capital (in most cases, those are non-agricultural wage works) for the rural poor, or the quality of public schools, which a large majority of poor children attend. Although the analysis in this paper cannot figure out what causes the lower educational returns for poorer households, investigating the causes carefully could offer an important and interesting perspective on this topic. This is left for future research.

References

Arias, Omar; Hallock, Kevin F.; and Sosa-Escudero, Walter. 2001. "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twins Data." *Empirical Economics*, 26(1): 7-40.

Banerjee, Abhijit; Iyer, Lakshmi; amd Somanathan, Rohini. 2005 “History, Social Divisions, and Public Goods in Rural India.” *Journal of the European Economic Association*, 3(2-3):639-647.

Behrman, Jere R. and Knowles, James C. 1999. “Household Income and Child Schooling in Vietnam.” *World Bank Economic Review*, 13(2): 211-256.

Card, David. 1999. “The Causal Effect of Education on Earnings.” in Ashenfelter, Orley and Card, David. (eds.) *The Handbook of Labor Economics*, Volume III, Amsterdam, North Holland.

Carneiro, Pedro; Heckman, James J.; and Vytlacil, Edward. 2006. "Estimating Marginal and Average Returns to Education." *Working Paper*, University College London.

Chernozhukov, Victor and Hansen, Christian. 2005. “An IV Model of Quantile Treatment Effects.” *Econometrica*, 73(1): 245–261.

Chernozhukov, Victor and Hansen, Christian. 2006. “Instrumental Quantile Regression Inference for Structural and Treatment Effect Models.” *Journal of Econometrics*, 132(2): 491–525.

Datt, Gaurav and Jolliffe, Dean. 2005. “Poverty in Egypt: Modeling and Policy Simulations.” *Economic Development and Cultural Change*, 53(2): 327–346.

Deaton, Angus and Kozel, Valerie. (eds.) (2005) The Great Indian Poverty Debate. New Delhi: Macmillan India

Duraisamy, P. 2002. “Changes in Returns to Education in India, 1983–94: by Gender, Age Cohort, and Location.” *Economics of Education Review*, 21(6): 609–622.

Dutta, Puja V. 2006. "Returns to Education: New Evidence for India, 1983-1999." *Education Economics*, 14(4): 431-451.

Easterly, William. 2001. *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*, Cambridge, MA: MIT Press.

Garen, John. 1984. "The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable." *Econometrica*, 55(5): 1199-1218.

Government of India (GOI). 2001. *Poverty Estimates for 1999–2000*. New Delhi: The Planning Commission, GOI.

Heckman, James J. and Li, Xuesong. 2004. "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000." *Pacific Economic Review*, 9(3): 155-171.

Heckman, James J. and Vytlacil, Edward. 1998. "Instrumental Variables Methods for the Correlated Random Coefficient Model." *Journal of Human Resources*, 33(4): 974-987.

Ito, Takahiro. 2009. “Caste Discrimination and Transaction Costs in the Labor Market: Evidence from Rural North India.” *Journal of Development Economics*, 88(2): 292–300.

Jolliffe, Dean. 2002. “Whose Education Matters in the Determination of Household Income? Evidence from a Developing Country.” *Economic Development and Cultural Change*, 50(2): 287-312.

Kingdon, Geeta G. 1998. “Does the Labour Market Explain Lower Female Schooling in India?” *Journal of Development Studies*, 35(1): 39–65.

Koenker, Roger. 2005. *Quantile Regression*, New York; Cambridge University Press.

Kurosaki, Takashi and Khan, Humayun. 2006. “Human Capital, Productivity, and Stratification in Rural Pakistan.” *Review of Development Economics*, 10(1): 116–134.

Lanjouw, Peter and Shariff, Abusaleh. 2004. “Rural Non-Farm Employment in India: Access, Income and Poverty Impact.” *Economic and Political Weekly*, 39 (40): 4429-4446.

Pritchett, Lant. 2001. "Where Has All the Education Gone?" *World Bank Economic Review*, 15(3): 367-391.

Rosenzweig, Mark R. 1995. “Why Are There Returns to Schooling?” *American Economic Review*, 85(2): 153-158.

Tilak, Jandhyala B.G. 2007. "Post-elementary education, poverty and development in India." *International Journal of Educational Development*, 27(4): 435-445.

Townsend, Robert M. 1994. “Risk and Insurance in Village India.” *Econometrica*, 62(3): 539–91.

Walker, Thomas S. and Ryan, James G. 1990. *Village and Household Economies in India’s Semi-Arid Tropics*, Baltimore: Johns Hopkins University Press.

Welch, Finis. 1970. “Education in Production.” *Journal of Political Economy*, 78(1): 35-59.

Yang, Dennis, T. 1997 “Education and Off-Farm Work,” *Economic Development and Cultural Change*, 45(3): 613-632.

Willis, Robert J. and Rosen, Sherwin. 1979. "Education and Self-Selection." *Journal of Political Economy*, 87(5): S7-S36.

UNDP (United Nations Development Programme). 2007. *Human Development Report 2007/2008: Fighting Climate Change: Human Solidarity in a Divided World*, New York: Palgrave Macmillan.

Wooldridge, Jeffrey M. 1997. "On Two Stage Least Squares Estimation of the Average Treatment Effect in a Random Coefficient Model." *Economics Letters*, 56(2): 129-133.

Tables

Table 1: Household Characteristics

	Monthly per capita expenditure				
	Quartile				Overall
	Poorest	2^{nd}	3^{rd}	Richest	
Monthly per capita expenditure (Rs.)	409.28	672.81	1024.58	2996.49	1275.25
Land owned (acres)	1.11	1.41	2.06	3.65	2.05
No. of working-age members	3.09	3.30	3.59	3.66	3.41
No. of non-working-age members	3.51	3.10	2.67	2.53	2.95
Schooling years	1.70	2.51	3.10	4.92	3.06
Upper and middle castes	6.4%	9.3%	17.2%	27.5%	15.1%
Backward castes	42.8%	50.7%	46.3%	47.9%	46.9%
Scheduled castes and tribes	40.3%	30.1%	26.7%	14.9%	28.0%
Muslim	10.5%	9.9%	9.7%	9.7%	9.9%

Note: "Working-age" is defined as ages between 15 and 60, and "schooling years" denotes the average number of schooling years among working-age adults.

Table 2: Caste Differences in Access to Facilities

	Upper castes	Non-upper castes		Backward castes		Scheduled castes and tribes	
No. of obs.	263	1799		1097		577	
	mean	mean	difference	mean	difference	mean	Difference
Distance to police	7.570	7.518	−0.052	7.490	−0.080	7.702	0.132
	(4.368)	(4.832)	[0.315]	(4.848)	[0.327]	(4.934)	[0.356]
Distance to bank	5.225	5.082	−0.143	4.923	−0.301	5.624	0.399
	(4.560)	(4.595)	[0.303]	(4.421)	[0.305]	(5.130)	[0.369]
Distance to hospital	20.346	21.441	1.095	21.855	1.509	20.158	−0.188
	(14.406)	(14.489)	[0.956]	(14.305)	[0.983]	(14.904)	[1.097]
Distance to prim. school	0.463	0.579	0.116	0.529	0.066	0.634	0.171
	(0.656)	(0.900)	[0.058]*	(0.829)	[0.055]	(0.948)	[0.065]*
Distance to mid. school	2.719	2.731	0.012	2.692	−0.027	2.651	−0.068
	(2.509)	(2.495)	[0.165]	(2.391)	[0.166]	(2.262)	[0.174]
Distance to sec. school	4.910	5.097	0.187	5.040	0.130	5.275	0.364
	(4.412)	(3.989)	[0.267]	(3.776)	[0.268]	(4.404)	[0.328]

Note: Numbers in parentheses are standard deviations and numbers in brackets are standard errors. * denotes statistically significant at 10% level or better. Distances are in kilometers and missing values are set to the district average.

Table 3: Correlation between Wage Rates and Distance to Facilities

	Police	Bank	Hospital	Prim. school	Mid. school	Sec. school
Whole villages	−0.2953* (117)	−0.3375* (117)	−0.0055 (116)	−0.0204 (106)	−0.2085* (111)	−0.2023* (114)
Villages having a bank within 5 km				0.2778* (65)	0.0507 (71)	0.1138 (71)

Note: * denotes statistically significant at 10% level or better. Numbers in parentheses are the number of observations.

Table 4: Summary Statistics of Empirical Variables

	No. of Obs.	Mean	Std. Dev.	Min.	Max.
Household variables					
Log of expenditure	2062	6.80	0.73	4.60	10.60
Schooling years	2062	3.06	3.35	0	16
Household size	2062	6.36	3.13	1	29
Male ratio	2062	0.52	0.15	0.00	1.00
Dependency ratio	2062	0.43	0.21	0.00	0.83
Age	2062	33.41	7.02	15	60
Age of head	2062	45.95	13.78	17	95
Female head	2062	0.04			
Land owned (Acres)	2062	2.05	4.37	0.00	93.00
Irrigation ratio	2062	0.60	0.45	0.00	1.00
Log of agric. Capital	2062	3.45	3.93	0.00	12.67
Log of livestock	2062	6.44	3.65	0.00	11.92
Backward castes	2062	0.53			
Scheduled castes and tribes	2062	0.28			
Muslim	2062	0.10			
Village-level variables					
Wage rates	2062	64.20	13.94	20	99
Distance to police (Km.)	2062	7.52	4.77	0	20
Distance to bank (Km.)	2062	5.10	4.59	0	20
Distance to hospital (Km.)	2062	21.30	14.48	0	70
Distance to prim. school (Km.)	2062	0.56	0.87	0	5
Instrumental variables					
Distance to mid. school (Km.)	2062	2.73	2.50	0	13
Distance to sec. school (Km.)	2062	5.07	4.04	0	20

Table 5: Estimation Results of First- and Second-Stage Regression

	1st stage estimation		2nd stage estimation					
	OLS		Structure OLS		IV		Control Function Approach	
Dependent variable:	Education		Log of expenditure		Log of expenditure		Log of expenditure	
Schooling years			0.0595	(0.0052)	0.1367	(0.0475)	0.1374	(0.0453)
Household size	0.0569	(0.024)	−0.0285	(0.0056)	−0.0333	(0.0066)	−0.0337	(0.0063)
Male ratio	1.3427	(0.411)	−0.0196	(0.0963)	−0.1224	(0.1189)	−0.1205	(0.1134)
Dependency ratio	−1.1651	(0.309)	−0.6771	(0.0728)	−0.5883	(0.0936)	−0.5858	(0.0895)
Age	−0.0669	(0.010)	0.0040	(0.0023)	0.0092	(0.0040)	0.0092	(0.0038)
Age of head	0.0140	(0.005)	0.0020	(0.0012)	0.0008	(0.0015)	0.0008	(0.0014)
Female head	−0.2802	(0.333)	−0.1518	(0.0783)	−0.1316	(0.0830)	−0.1320	(0.0791)
Land owned	0.1202	(0.016)	0.0211	(0.0037)	0.0125	(0.0066)	0.0125	(0.0062)
Irrigation ratio	1.1087	(0.153)	0.0364	(0.0363)	−0.0510	(0.0655)	−0.0502	(0.0625)
Log of agric. capital	0.0931	(0.021)	0.0061	(0.0049)	−0.0014	(0.0069)	−0.0013	(0.0066)
Log of livestock	−0.0454	(0.020)	0.0131	(0.0048)	0.0169	(0.0055)	0.0168	(0.0053)
Backward castes	−2.8811	(0.164)	−0.0786	(0.0414)	0.1445	(0.1430)	0.1436	(0.1364)
Scheduled castes and tribes	−3.5251	(0.193)	−0.2172	(0.0487)	0.0542	(0.1734)	0.0526	(0.1654)
Muslim	−1.3585	(0.218)	0.0903	(0.0512)	0.2004	(0.0861)	0.1997	(0.0821)
Wage rates	−0.0024	(0.005)	−0.0016	(0.0011)	−0.0015	(0.0012)	−0.0015	(0.0011)
Distance to police	0.0333	(0.016)	−0.0191	(0.0036)	−0.0219	(0.0042)	−0.0219	(0.0040)
Distance to bank	−0.0456	(0.019)	0.0034	(0.0039)	0.0103	(0.0058)	0.0102	(0.0056)
Distance to hospital	−0.0026	(0.004)	−0.0021	(0.0010)	−0.0018	(0.0010)	−0.0018	(0.0010)
Distance to prim. school	−0.0536	(0.071)	−0.0546	(0.0162)	−0.0466	(0.0177)	−0.0471	(0.0169)
Distance to mid. school	−0.0557	(0.028)						
Distance to sec. school	−0.0723	(0.020)						
v (residual from the 1st stage estimation)							−0.0753	(0.0460)
v × Schooling years							−0.0006	(0.0014)
Intercept	6.5601	(0.599)	7.1083	(0.1423)	6.6308	(0.3275)	6.6353	(0.3125)
Adjusted R-square	0.354		0.255		0.181		0.256	

Test for the hypothesis that the coefficients of instrumental variables are zero: $F(2, 2041) = 13.44$ (P-value = 0.000)

Hausman test for the schooling coefficient: $\chi^2(1) = 2.679$ (P-value = 0.102)

Sargan type overidentification test: $\chi^2(1) = 1.303$ (P-value = 0.2537)

Note: Numbers in parentheses are standard errors.

Table 6: Selective Estimation Results of QR and IV–QR models

A. QR estimation	25th percentile		50th percentile		75th percentile	
Schooling years	0.0436	(0.0057)	0.0560	(0.0065)	0.0684	(0.0086)
Household size	−0.0278	(0.0078)	−0.0265	(0.0057)	−0.0164	(0.0064)
Male ratio	−0.0216	(0.1100)	−0.0159	(0.1020)	0.0974	(0.1287)
Dependency ratio	−0.7176	(0.0855)	−0.6798	(0.0736)	−0.6839	(0.1048)
Age	0.0021	(0.0028)	0.0020	(0.0024)	0.0082	(0.0031)
Age of head	0.0020	(0.0014)	0.0022	(0.0013)	0.0002	(0.0016)
Female head	−0.1522	(0.0724)	−0.1628	(0.0755)	−0.2753	(0.0934)
Land owned	0.0221	(0.0040)	0.0314	(0.0092)	0.0377	(0.0117)
Irrigation ratio	−0.0227	(0.0408)	0.0312	(0.0369)	0.0678	(0.0564)
Log of agric. Capital	0.0086	(0.0051)	0.0043	(0.0060)	−0.0056	(0.0082)
Log of livestock	0.0141	(0.0054)	0.0089	(0.0051)	0.0173	(0.0065)
Backward castes	−0.0543	(0.0417)	−0.0061	(0.0467)	−0.0411	(0.0732)
Scheduled castes and tribes	−0.1861	(0.0491)	−0.1037	(0.0544)	−0.1492	(0.0760)
Muslim	0.0699	(0.0581)	0.1141	(0.0526)	0.1781	(0.0693)
Wage rates	−0.0002	(0.0011)	−0.0002	(0.0012)	−0.0038	(0.0015)
Distance to police	−0.0165	(0.0036)	−0.0138	(0.0038)	−0.0201	(0.0053)
Distance to bank	0.0047	(0.0039)	0.0002	(0.0040)	−0.0019	(0.0052)
Distance to hospital	−0.0013	(0.0012)	−0.0015	(0.0010)	−0.0033	(0.0013)
Distance to prim. school	−0.0813	(0.0179)	−0.0565	(0.0182)	−0.0345	(0.0222)
Intercept	6.7122	(0.1570)	6.9001	(0.1556)	7.3327	(0.1974)
B. IV-QR estimation	25th percentile		50th percentile		75th percentile	
Schooling years	0.0394	(0.0445)	0.1573	(0.0629)	0.1567	(0.0394)
Household size	−0.0264	(0.0098)	−0.0263	(0.0113)	−0.0345	(0.0101)
Male ratio	−0.0371	(0.1108)	−0.2052	(0.2137)	−0.0494	(0.1290)
Dependency ratio	−0.7132	(0.1077)	−0.5799	(0.1289)	−0.5160	(0.1170)
Age	0.0016	(0.0043)	0.0087	(0.0040)	0.0071	(0.0030)
Age of head	0.0020	(0.0015)	−0.0003	(0.0019)	0.0013	(0.0018)
Female head	−0.1364	(0.0853)	−0.1851	(0.0871)	−0.2424	(0.0966)
Land owned	0.0218	(0.0048)	0.0082	(0.0160)	0.0156	(0.0119)
Irrigation ratio	−0.0177	(0.0675)	−0.1185	(0.0532)	−0.0079	(0.0641)
Log of agric. Capital	0.0088	(0.0063)	0.0037	(0.0073)	−0.0050	(0.0076)
Log of livestock	0.0144	(0.0058)	0.0153	(0.0078)	0.0144	(0.0064)
Backward castes	−0.0714	(0.1303)	0.2093	(0.1727)	0.1666	(0.1121)
Scheduled castes and tribes	−0.2089	(0.1616)	0.1727	(0.1936)	0.0776	(0.1229)
Muslim	0.0594	(0.0841)	0.2724	(0.1103)	0.2217	(0.0735)
Wage rates	−0.0004	(0.0015)	0.0005	(0.0020)	−0.0036	(0.0017)
Distance to police	−0.0167	(0.0045)	−0.0239	(0.0072)	−0.0242	(0.0070)
Distance to bank	0.0052	(0.0060)	0.0135	(0.0083)	0.0071	(0.0058)
Distance to hospital	−0.0011	(0.0012)	−0.0011	(0.0013)	−0.0025	(0.0014)
Distance to prim. schools	−0.0825	(0.0214)	−0.0728	(0.0220)	−0.0498	(0.0217)
Intercept	6.7557	(0.3720)	6.4287	(0.2283)	7.0968	(0.2339)

Note: Numbers in parentheses are standard errors calculated nonparametrically.

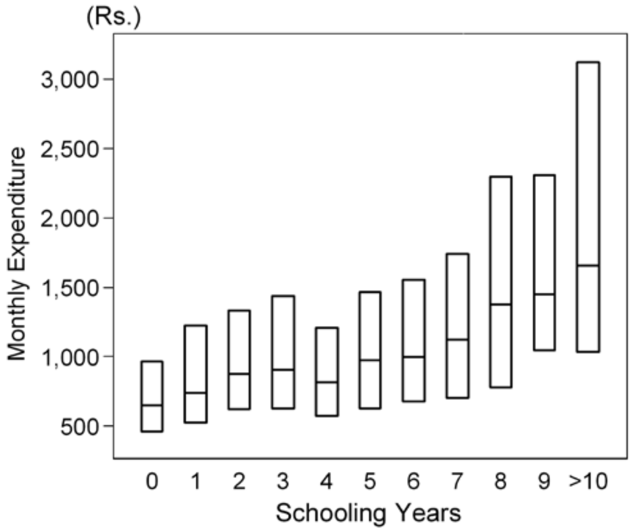
Table 7: Tests for the IV–QR Estimates

Null hypothesis	Statistics	
Tests for the schooling coefficients, $\beta(\tau)$		
1) No endogeneity: $\forall\tau$; $\beta(\tau) = \beta_{QR}(\tau)$	$\chi^2(17) = 39.55$	(P-value = 0.001)
2) No heterogeneity: $\forall\tau \neq \tau'$; $\beta(\tau) = \beta(\tau')$	$F(16, 2042) = 1.16$	(P-value = 0.295)
3) No effect: $\forall\tau$; $\beta(\tau) = 0$	$F(17, 2042) = 2.43$	(P-value = 0.001)
Tests for the intercepts, $\alpha(\tau)$		
4) No heterogeneity: $\forall\tau \neq \tau'$; $\alpha(\tau) = \alpha(\tau')$	$F(16, 2042) = 1.66$	(P-value = 0.048)
5) No effect: $\forall\tau$; $\alpha(\tau) = 0$	$F(17, 2042) = 107.87$	(P-value = 0.000)
Tests for all coefficients, $B(\tau)$		
6) No heterogeneity: $\forall\tau \neq \tau'$; $B(\tau) = B(\tau')$	$F(320, 2042) = 10.21$	(P-value = 0.000)
7) No effect: $\forall\tau$; $B(\tau) = 0$	$F(340, 2042) = 2080.38$	(P-value = 0.000)

Note: Regarding the test for the hypothesis of no endogeneity of "Schooling years," the Hausman test is employed.

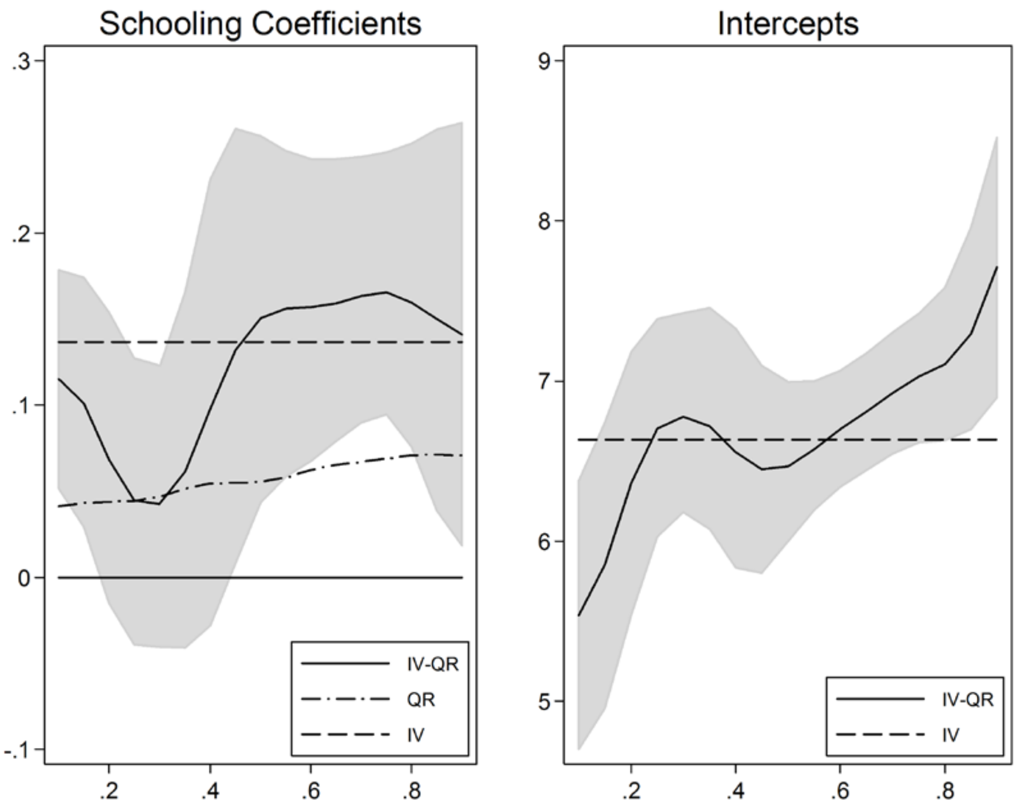
Figures

Figure 1: Boxplots of Monthly Per-Capita Expenditure



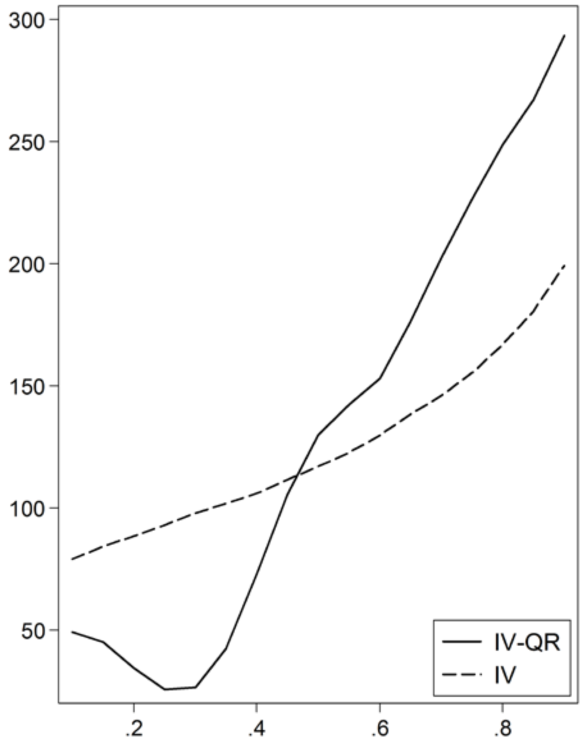
Note: The upper, central, and lower lines of the boxes represent the 75th, 50th, and 25th percentile of the monthly per-capita expenditure, respectively. “Schooling years” on the horizontal axis is the average number of schooling years among working-age adults, and “>10” on the last category represents ten years and above.

Figure 2: IV-QR Estimates of Schooling Coefficients and Intercepts



Note: Coefficient estimates are on the vertical axis, and the quantile index is on the horizontal axis. In both panels, the shaded area is the 90% confidence band estimated non-parametrically. Estimations are implemented at 0.05 unit intervals for $\tau \in [0.1, 0.9]$.

Figure 3: Level Impacts of Education



Note: Estimates of increases in consumption expenditure with additional years of schooling are shown on the vertical axis (in Rupees), whereas the quantile index is on the horizontal axis.