Education Economics

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cede20

Returns to Education: New Evidence for India, 1983-1999

Puja Vasudeva Dutta [a]

[a] National Council of Applied Economic Research, Parisila Bhawan, New Delhi, India

Published online: 19 Jan 2007.

Education Economics
Vol. 14, No. 4, 431–451, December 2006

Returns to Education: New Evidence for India, 1983–1999

PUJA VASUDEVA DUTTA

National Council of Applied Economic Research, Parisila Bhawan, New Delhi, India

ABSTRACT *This paper estimates the returns to education for adult male workers in regular and casual wage employment using Indian national survey data at three points in time spanning almost two decades. Both standard and augmented Mincerian wage equations are estimated using a set of human capital measures and other controls after addressing the issue of potential selection bias. This paper finds that the returns to education are significantly different for the two types of workers—while casual workers face at best flat returns to education, the returns to education for regular workers are positive and U-shaped with respect to education levels. There is also some evidence of a widening wage gap between regular workers with graduate and primary education that could possibly be a consequence of trade liberalization and other reforms pursued during the 1990s.*

KEY WORDS: Rate of return; human capital; India

Introduction

The late 1980s and 1990s were a period of rapid industrial deregulation and trade liberalization in India. This paper exploits three national employment surveys 1983, 1993–94 and 1999–2000 to explore the structure of wages of adult male workers before and after this economic liberalization. The first survey can be interpreted as providing insights into the structure of labour markets prior to liberalization, while the latter two provide the basis for delineating a portrait of these structures after the radical trade liberalization process.

In particular, this paper focuses on the returns to education for two types of adult male workers—those in regular wage employment and those in casual wage employment—in the Indian labour market.[1] Casual empirical analysis reveals that there are considerable differences between the two kinds of wage employment. For instance, Tendulkar (2003, p. 2) refers to 'workers having regular, contractual hired employment' as the 'labour aristocracy because of the privileged service conditions this segment enjoys including high wages', and regular wage employment is often considered to be the most preferred category of work (Das, 2003; Unni, 2001). Regular workers are also covered by labour

Correspondence Address: Puja Vasudeva Dutta, National Council of Applied Economic Research (NCAER), Parisila Bhawan, 11 Indraprastha Estate, New Delhi 110 002, India. Tel: +91 9873197852. Email: pvdutta@gmail.com

0964-5292 print/1469-5782 online/06/040431-21 © 2006 Taylor & Francis
DOI: 10.1080/09645290600854128

Table 1. Distribution of adult male workers by employment status (%)

	Rural				Urban			
Year	Regular workers	Casual workers	Self-employed	Unemployed	Regular workers	Casual workers	Self-employed	Unemployed
1983	10.38	24.59	61.16	3.86	42.26	12.14	39.00	6.60
1993	10.58	24.26	62.20	2.96	40.93	12.90	41.04	5.13
1999	10.80	25.92	59.65	3.64	39.18	13.93	41.24	5.65

Calculations from NSS surveys for adult male workers aged 15–65 years. See Appendix 1 for details.

market regulations that confer some measure of employment security and social security benefits. Only about one-third of the adult male labour force in India is engaged in wage employment —the majority is self-employed (see Table 1). Casual workers tend to have lower physical and human capital relative to regular wage workers, and this is reflected in the low earnings. Although regular workers comprise roughly one-half of all wage workers, their income share is about three-quarters. The raw wage gap between casual and regular workers is substantial and increased during this period—in 1983 an average casual worker in the labour market earned about 35% of the hourly wage earned by an average regular worker, and by 1999 this had fallen to 30%.

There have been several studies on the determinants of wages in India (for a recent review, see Kingdon, 1998). All these studies as well as Kingdon and Unni (2001) use a sample of workers within a city or urban areas of a state at a point in time. In contrast Duraisamy (2002) estimates returns to education using national data from 1983 and 1993–94. These latter two most recent studies also address the issue of selection bias but do so as a dichotomous realization between wage employment and all other categories. This paper extends previous work on the structure of wages in India by analysing national data at three points in time spanning almost two decades and addresses the issue of selection bias in wages of both regular and casual workers considered separately.

This paper is structured as follows. The next section describes the empirical strategy. Following the standard labour economics literature, Mincerian wage regression models are specified that allow for human capital controls. These are then supplemented by augmented Mincerian earnings equations that include other controls such as social exclusion or family background, industry affiliation, seasonality, location and settlement type that have been found in the empirical literature to have an independent effect on the earnings determination process. In addition, the issue of selection bias is addressed using the generalized framework developed by Lee (1983). The third section presents the empirical results for the wage regression models for regular and casual workers. The next section examines the returns to education for both types of workers and the changes in these returns between 1983 and 1999. The final section offers some conclusions and policy implications.

Estimating an Empirical Model

Wage regression models are estimated as standard Mincerian earnings equations allowing for human capital controls. In addition, other controls such as social exclusion or family background, industry affiliation, seasonality, location and settlement type are included in the augmented Mincerian equations. There is considerable empirical evidence that these factors influence the wage determination process (for inter-industry wage differentials see, e.g., Krueger and Summers, 1988; for social exclusion in the Indian labour market, see Banerjee and Knight, 1985; Kingdon and Unni, 2001). Tilak (1994) also points out that, in addition to the above, the failure to control for geographical aggregation biases arising from regional price variations or from combining poor and rich areas in the same sample could potentially lead to inaccurate estimates. The same argument applies to distinguishing between rural and urban settlement types. Before these wage regression models are estimated, the issue of selection bias is also addressed. There is reason to suspect some selection bias as the bulk of the adult male labour force in India is self-employed—about 58–61% of the labour force in rural areas and 39–41% in urban areas. These self-employed individuals cannot be included in our regression models due to lack of data on their earnings. Previous studies (see, e.g., Duraisamy, 2002; Kingdon and Unni, 2001) have found some evidence of sample selection bias using the Heckman procedure by modelling the selection process into wage and non-wage employment. However, casual empirical analysis suggests that there are considerable differences between the two kinds of wage employment—regular and casual. If this selection of individuals into wage employment is systematic, then ignoring the non-random nature of the sample would introduce a selection bias in the wage regression model's estimates. Consequently, the potential problem of selection bias is addressed using the generalized framework popularized by Lee (1983).

Consider the following two-stage model for selection and wage determination (suppressing the *i* subscripts for individuals):

$$w_j = x_j'\beta_j + \mu_j \qquad j = 2,\ 3 \tag{1}$$

$$y_s^* = z_s'\gamma_s - \eta_s \qquad s = 1,\ 2,\ 3 \tag{2}$$

where w is the outcome variable (in this case, log wages) that is observed only for persons engaged in wage employment of two types (denoted by the categorical variable j)—regular wage employment (j = 2) and casual wage employment (j = 3). The latent dependent variable (y_s^*) represents the employment status of the individual: (i) non-wage earners comprising non-participants in the labour market, self-employed and unemployed individuals; (ii) regular wage employment; and (iii) casual wage employment. The vectors x_j and z_s comprise exogenous explanatory variables, s is a categorical variable signifying selection between the above three different alternatives, and μ_j and η_s are random error terms such that $E(\mu_s \mid x_j; z_s) = 0$ and $E(\eta_s \mid x_j; z_s) = 0$. As the selection bias is mediated through observed wages it is sufficient and computationally more convenient to separate employment status into non-wage earners and two different types of wage earners.

If the η_s values are assumed to be independent and identically distributed as Type I extreme value distributions and their differences (i.e., between different employment status) follow a logistic distribution. This gives rise to the Multinomial Logit

(MNL) model, and the probability of an individual being in a selected outcome can be expressed as:

$$P_1 = \frac{1}{1+\sum_{j=2}^{M}\exp(z_j^{'}\gamma_j)}; \quad \text{and} \quad P_j = \frac{\exp(z_j^{'}\gamma_j)}{1+\sum_{j=2}^{M}\exp(z_j^{'}\gamma_j)} \qquad \gamma_1 = 0; j = 2,3 \tag{3}$$

The Theil normalization is applied to the category comprising the non-wage earners. This category's parameters are thus set to zero to resolve an indeterminacy associated with the MNL model. In order to identify the parameters of the wage equations, a set of variables that influence employment status but not wage itself must be included as regressors in the selection equation. Consistent estimates of the parameters (β_j) in the outcome equation can be obtained by replacing the disturbance terms η_j in equation (1) by their conditional expected value obtained from the MNL estimation (equation (3)). This selection bias correction term, $\hat{\lambda}_j$, is similar to the inverse of the 'Mills ratio'. An augmented semi-logarithmic Mincerian specification can then be used to estimate the wage equations:

$$w_j = x_j{'}\beta_j - \beta_j^{*}\hat{\lambda}_j + \upsilon_j \qquad j = 2,\ 3 \tag{4}$$

where $\beta_j^{*} = \rho_j\sigma_{\mu_j}$ the coefficient on the selection bias correction term in the wage equations; ρ_j is the coefficient of correlation between the error terms in the wage equation and the selection equation (the direction of bias is determined by this correlation term); and v_j is the error term for each of the wage equations.

This two-step procedure ensures that the ordinary least squares (OLS) estimates of the coefficients from the wage equations are consistent (Lee, 1983).[2] The sampling distribution for the resultant wage equation estimates is obtained by bootstrapping the wage regression models using 1000 replications. The analysis uses data drawn from large-scale employment surveys undertaken in January–December 1983, July 1993–June 1994 and July 1999–June 2000 (referred to as 1983, 1993 and 1999 in this paper).[3] Appendix 1 describes the data used in this paper. The empirical analysis reported in this paper is restricted to prime-aged adult males.

Empirical Results: Wage Regression Models

We first estimate standard Mincerian wage equations where the natural log of real hourly wages received by an individual is a function of the age (as a proxy for experience) and education of the individual (regression not reported).[4] While the explanatory power of the variables for regular workers is reasonably high in all three years, that for casual workers is very low, suggesting that human capital characteristics do not adequately explain the wage determination process at least for casual workers. In particular, given that a large proportion of casual workers are engaged in agricultural and allied activities, factors arising out of location and rural–urban settlement type are likely to be the main determinants of their wages. As a result, these models are supplemented by augmented specifications that include other controls including marital status, effects of social discrimination or family background operating through caste and religious affiliation, controls for location (settlement type and state of residence) and seasonality effects (proxied by the timing of the interview for the survey), industry affiliation. The summary statistics for these variables are reported in Table A1 in Appendix 1. The results are presented in Table 2 (the estimated coefficients for the six seasonality variables including the interactions with settlement type, 37 industry and 16 state dummies are not reported to conserve space).

The augmented Mincerian wage regression models also include selection bias correction terms constructed from the predicted probabilities from the MNL model as described in the Estimating an Empirical Model section. The

Table 2. Augmented Mincerian wage regression models for regular and casual workers

	Regular wage workers			Casual wage workers		
	1983	1993	1999	1983	1993	1999
Individual characteristics						
Age spline: 15–25 years	0.0139***	0.0106***	0.0130***	0.0106***	0.0136***	0.0137***
	(0.0017)	(0.0024)	(0.0022)	(0.0007)	(0.0009)	(0.0010)
Age spline: 25–35 years	0.0174***	0.0165***	0.0203***	–0.0008	–0.0009	0.0019***
	(0.0008)	(0.0012)	(0.0012)	(0.0005)	(0.0006)	(0.0006)
Age spline: 35–45 years	0.0106***	0.0158***	0.0160***	–0.0015**	–0.0003	–0.0007
	(0.0008)	(0.0011)	(0.0011)	(0.0006)	(0.0007)	(0.0006)
Age spline: 45–55 years	0.0055***	0.0129***	0.0159***	–0.0007	–0.0043***	–0.0033***
	(0.0013)	(0.0014)	(0.0014)	(0.0008)	(0.0009)	(0.0009)
Age spline: 55–65 years	–0.0291***	–0.0304***	–0.0252***	–0.0052***	–0.0053***	–0.0034**
	(0.0031)	(0.0049)	(0.0046)	(0.0013)	(0.0015)	(0.0014)
Married	0.0647***	0.0781***	0.0878***	0.0271***	0.0327***	0.0196***
	(0.0073)	(0.0095)	(0.0102)	(0.0037)	(0.0045)	(0.0043)
Education						
Completed primary school	0.0658***	0.0426***	0.0485***	0.0088	0.0108*	0.0227***
	(0.0072)	(0.0100)	(0.0114)	(0.0057)	(0.0057)	(0.0056)
Completed middle school	0.1361***	0.0933***	0.1090***	–0.0065	–0.0079	0.0165*
	(0.0085)	(0.0120)	(0.0115)	(0.0092)	(0.0092)	(0.0085)
Completed secondary school	0.3486***	0.2640***	0.2945***	–0.0085	–0.0348**	0.0117
	(0.0110)	(0.0160)	(0.0140)	(0.0174)	(0.0154)	(0.0147)
Completed graduate school	0.6192***	0.5385***	0.6025***	–0.0043	–0.0693*	0.0180
	(0.0150)	(0.0229)	(0.0188)	(0.0540)	(0.0381)	(0.0318)
Social exclusion						
Member of scheduled caste or tribe	–0.0472***	–0.0389***	–0.0309***	–0.0033	0.0111*	–0.0060
	(0.0056)	(0.0074)	(0.0072)	(0.0051)	(0.0060)	(0.0059)
Muslim	–0.0163**	–0.0328***	–0.0436***	0.0249***	0.0245***	0.0130**
	(0.0075)	(0.0093)	(0.0092)	(0.0049)	(0.0055)	(0.0053)
Location						
Residence in rural areas	–0.1454***	–0.0468**	–0.0240	–0.0972***	–0.0479***	–0.0329***
	(0.0141)	(0.0207)	(0.0166)	(0.0082)	(0.0079)	(0.0076)
Selection bias correction term	–0.0202	–0.1152***	–0.1149***	0.0610***	0.0682***	0.0391***
	(0.0174)	(0.0270)	(0.0228)	(0.0125)	(0.0129)	(0.0145)

Table 2. (*Continued*)

	Regular wage workers			Casual wage workers		
	1983	1993	1999	1983	1993	1999
Constant	0.4375***	0.6996***	0.8532***	0.4263***	0.4483***	0.5710***
	(0.0579)	(0.0917)	(0.0846)	(0.0279)	(0.0324)	(0.0352)
Number of observations	27 356	26 387	27 295	28 855	26 398	29 805
Adjusted R^2	0.5458	0.4789	0.5298	0.3393	0.3058	0.3382
Standard error of estimate	0.3483	0.4229	0.4477	0.2127	0.2318	0.2364

Dependent variable: natural log of real hourly wages. Standard errors are in parentheses (obtained after bootstrapping with 1000 replications). Significance at ***1%, **5% and *10% level or better. The estimated coefficients on the age splines are not cumulative. Thirty-seven industry dummy variables (food crops is the omitted industry), 16 state dummy variables (West Bengal is the omitted state) and six seasonal dummy variables (including interactions with the settlement type) (being interviewed in the first quarter and the interaction of this with settlement type are the omitted categories).

results for the multinomial model for selection into both types of wage employment—regular and casual—relative to the base category of non-wage earners are not reported here for brevity. Besides the explanatory variables included in the wage equations (other than industry affiliation), as noted earlier, the parameters of the wage equations are identified using variables that capture household structure. These have been traditionally used in the literature, are likely to influence employment status but not wages, and are also the best available instruments in the data. Alternative variables such as land ownership could potentially be endogenous and correlated both with employment status as well as wages, besides not being a good measure for urban areas. The identifying exclusion restrictions used include the household size, the number of persons aged older than 65 years in the household and three dummy variables for whether the household has one child, two children or three or more children aged zero to four years (the omitted category is not having any children aged zero to four years).[5] The majority of the effects estimated are plausible and are significant at the 1% level or better. Individuals who are educated, married with a large number of children and reside in urban areas are more likely to be in regular wage employment. The direction of effect of most of the variables remained stable across all three years with a few exceptions, mostly for the state effects in 1993. A Wald test for the validity of conflating the casual and regular wage employment categories was decisively rejected by the data for all three years.[6] It must be stressed that this estimation is not an attempt at modelling participation in the labour market, but one designed to obtain the necessary tools to control for selection bias in the wage regression models.

We find that the inclusion of the additional controls improves the explanatory power of the model, particularly for casual workers, although the fits appear poorer in the second year. At the same time, the standard error of the estimate increased by about 10 and 2.4 percentage points for regular and casual workers, respectively, between 1983 and 1999, indicating rising returns to unobservable skills during this period, especially for regular workers. The returns to education, the primary focus of this paper, are examined in detail in the next section after briefly discussing the effects of the other control variables.

The majority of the explanatory variables have the expected signs and are significant at the 1% level or better. In particular, the age–earnings profiles for regular and casual workers display a positive relationship between age and real hourly wages and the general shape is in accordance with the prediction of human capital theory and previous empirical research (Murphy and Welch, 1990). The difference in the returns to age for the two types of workers is striking. For regular workers the returns to age peak at the 45–55 age group before declining, whereas for casual workers the returns to age rise very steeply initially for the 15–25 age group and then virtually flatten out. The age–earnings profile has clearly shifted up during these three years for both regular and casual workers—a Wald test of the coefficients on the age splines rejects the null hypothesis of no movement between each pair of years for both types of workers.[7] This steepening of the curve for regular workers and the increasing return to age for casual workers until the age of 35 years combined with the rising standard error of estimates in the wage equations indicate an increase in the returns to unobservable skills that could possibly be related to the liberalization process.

The other controls also have the anticipated effect on wages. Thus, marriage has the expected positive effect for both regular and casual workers. Residing in rural areas significantly reduces the wage received (although this disadvantage declined significantly after 1983). Seasonal effects (interacted with settlement type) are jointly significant while the state and industry dummy variables are almost all significant at the 1% level or better, indicating the presence of inter-regional and inter-industry wage differentials. Social exclusion is captured by mutually exclusive dummy variables for caste and religious affiliation (relative to all other individuals belonging to other religions and castes). It is argued that 'the caste system confines those from lower castes to a limited number of poorly paid, often socially stigmatized occupational niches from which there is little escape …' (Kabeer, 2002, p.3). Ethnicity is also often a source of exclusion—in India this translates into exclusion on the basis of religion and is largely applicable to the Indian Muslims (Das, 2003). Belonging to either of these social groups significantly decreases the wage received by regular workers in all three years, while the opposite is the case for casual workers (see also Banerjee and Knight, 1985). These variables, however, are also likely to capture the effects of omitted variables such as family background (Kingdon, 1998) and/or occupation (Das, 2003).

The selection bias correction terms are significant determinants of wages. Individuals selected into regular wage employment are likely to earn higher wages than a person randomly selected from the population, and this effect has risen substantially in the 1990s from zero in 1983 to a significant 13% in the later years. Conversely, casual workers tend to earn about 6–7% lower wages than an individual selected at random from the population in the first two years. By 1999 this disadvantage had fallen, and casual workers earned about 4% less than a randomly selected individual in 1999. This is plausible as the reference category includes self-employed and unemployed individuals who presumably have the resources to engage in self-owned enterprises or to afford the time taken to obtain regular employment. The wage gap between regular and casual workers computed from adjusted wage data obtained after controlling for various individual characteristics are much lower than the raw wage gap and suggest exactly the opposite trend—casual hourly wages as a proportion of regular wages rose marginally from 55% in 1983 to 57% in 1999. This reflects the decline in the negative selection effect for casual workers despite the rise noted for regular workers.

Returns to Education

This section examines in detail the returns to education for both regular and casual workers and the changes in these returns following the economic liberalization of the 1990s. These rates of return are computed from the standard as well as the augmented Mincerian wage regression models.

In common with other studies the marginal wage effects of education for regular workers are significantly positive and monotonically increasing in education level—a regular worker who has completed primary school earned about 7% higher than one with no education, while a graduate earned as much as 62% higher wages in 1983.[8] For casual workers, acquiring education until primary school raises the wage received by about 1–2%. The estimated marginal wage effects of education above the middle school level are significantly negative in 1993, although this trend is reversed in the last year (significantly so for middle school). The coefficients on the education dummies in the standard Mincerian are higher but rising in education level as in the augmented specification—for instance, a regular worker with primary (graduate) schooling earns about 16% (80%) higher than an uneducated worker. The pattern for casual workers is less clear—while the marginal wage effects of education are rising in education level in the initial year, these are broadly similar across levels in the terminal year (regression not reported).

The private rate of return per year of education at different education levels can be computed using the coefficients from the wage equations. These serve as useful indicators of the productivity of education and the incentive for individuals to invest in their human capital (Psacharopoulos and Patrinos, 2004). If the returns to education are different for different groups participating in the labour market, this will affect the perceived economic benefits of education among these groups (Kingdon, 1998). The National Sample Surveys (NSS) surveys after 1983 do not report the number of years of schooling, only the maximum level of schooling completed that allows the construction of the five education dummy variables used in the selection and wage regression models. Since education policy is a subject under state jurisdiction, the schooling systems (at least until the secondary school) vary somewhat across states. In general, most states follow five years of primary, three years of middle, four years of secondary (including higher secondary) schooling and three years (four if a technical degree) of graduate education (Duraisamy, 2002).[9] Existing literature on the returns to education in India have taken these (or very similar values) to proxy years of foregone earnings. However, Psacharapoulos (see, e.g., Psacharopoulos, 1987) has suggested that it is more practical to take two years rather than five as the period of earnings foregone in the pursuit of primary education as children younger than 10 could neither be expected to work full-time if not in school nor to earn the average industry wage.

The average rate of return to each education level, r_j, can then be estimated as follows:

$$r_j = \frac{\left(\beta_j - \beta_{j-1}\right)}{\left(\mathrm{Y}_j - \mathrm{Y}_{j-1}\right)} \tag{6}$$

where j is primary, middle, secondary or graduate school, β_j is the coefficient in the wage regression models (equation (5)) and Y_j is the years of schooling at education level j. However, the rate of return to primary education is estimated as follows:

$$r_{prim} = \frac{\left(\beta_{prim}\right)}{\left(Y_{prim}\right)}$$

where $Y_{prim} = 2$ years.

Table 3 presents the rates of returns computed from three different specifications of the wage regression model—standard Mincerian (with only human capital variables), augmented Mincerian (with several other individual characteristics but without controlling for selection correction) and the full augmented Mincerian (with controls for individual characteristics and selection bias). The last is our preferred specification given the characteristic feature of the Indian labour market, similar to several other developing countries, where wage workers are only a small fraction of the total labour force and where regular wage employment is considered the most preferred category of work. In this context, it appears preferable to attempt to control for the underlying process by which the set of observations actually observed are generated. This preference is also supported by other Indian studies that also control for selection into wage employment as well as controls for individual characteristics such as location and settlement type, family background, and so on (see, e.g., Duraisamy, 2002; Kingdon and Unni, 2001).

The results suggest that there is an incentive to acquire high levels of education if the individual is in regular wage employment. There is a U-shaped pattern in the returns to education for these workers—the returns to primary education are low relative to secondary and graduate education but higher than those in middle schooling for regular workers in all specifications (except in the standard Mincerian for the first year). The inclusion of individual characteristics in addition to human capital variables in the augmented specification without selection correction yields rates of return that are very similar to those obtained from the standard Mincerian (only the returns to primary education in the first year are statistically different). Introducing controls for selection bias reduces the rates of return to all levels of education significantly, substantially so for primary education.

The estimated returns to education for casual workers are not as well determined as those for regular workers and are substantially (and significantly at least for primary and middle schooling) lower once additional controls, including for selection bias, are included in the wage regression models. For casual workers, while an additional year of primary schooling does contribute to the wage earned, at least during the 1990s, it is much lower than that for regular workers in the preferred third specification. There is no incentive to acquire education higher than primary schooling for a casual worker as there are at best flat returns to education. Indeed, the returns to completing middle school are negative in the augmented specification.[10] This indicates that either there is a low demand for skill or that the acquired skills are not useful in the casual labour market, especially once individual characteristics besides human capital variables are taken into account. It could be argued that the acquisition of certain education levels

Table 3. Average private rate of return to education levels (%)

	Standard Mincerian			Augmented Mincerian, no selection			Augmented Mincerian, full		
	1983	1993	1999	1983	1993	1999	1983	1993	1999
Regular workers									
Completed primary school	8.07***	5.10***	5.60***	5.75***	4.63***	5.30***	3.29***	2.13***	2.43***
	(0.0036)	(0.0047)	(0.0056)	(0.0035)	(0.0046)	(0.0055)	(0.0036)	(0.0050)	(0.0057)
Completed middle school	3.42***	3.14***	3.53***	3.13***	2.92***	3.35***	2.35***	1.69***	2.02***
	(0.0026)	(0.0032)	(0.0036)	(0.0025)	(0.0032)	(0.0036)	(0.0025)	(0.0031)	(0.0034)
Completed secondary school	6.00***	5.37***	6.12***	6.01***	5.58***	6.25***	5.31***	4.27***	4.64***
	(0.0018)	(0.0020)	(0.0021)	(0.0017)	(0.0020)	(0.0021)	(0.0019)	(0.0023)	(0.0022)
Completed graduate school	9.97***	10.94***	12.29***	9.43***	10.82***	12.37***	9.02***	9.15***	10.26***
	(0.0028)	(0.0028)	(0.0028)	(0.0027)	(0.0027)	(0.0028)	(0.0029)	(0.0036)	(0.0032)
Casual workers									
Completed primary school	5.17***	5.09***	5.15***	1.88***	1.90***	1.93***	0.44	0.54*	1.13***
	(0.0024)	(0.0025)	(0.0025)	(0.0021)	(0.0023)	(0.0022)	(0.0028)	(0.0028)	(0.0028)
Completed middle school	0.82***	0.72***	0.72***	0.17	0.17	0.19	–0.51**	–0.63***	–0.2
	(0.0026)	(0.0024)	(0.0022)	(0.0022)	(0.0021)	(0.0019)	(0.0025)	(0.0024)	(0.0021)
Completed secondary school	0.69*	–0.47*	0.16	0.62*	0.16	0.20	–0.05	–0.67**	–0.12
	(0.0036)	(0.0026)	(0.0021)	(0.0032)	(0.0024)	(0.0018)	(0.0034)	(0.0027)	(0.0024)
Completed graduate school	2.18	–2.11*	–0.07	1.13	–0.79	0.92	0.14	–1.15	0.21
	(0.0184)	(0.0115)	(0.0095)	(0.0170)	(0.0112)	(0.0085)	(0.0168)	(0.0113)	(0.0087)

Significance at ***1%, **5% and *10% level or better. Data in parentheses are standard errors computed as follows: $\mathrm{var}(r) = [(1/(Y_j - Y_{j-1}))^2][\mathrm{var}(\beta_j) + \mathrm{var}(\beta_{j-1}) - 2 * \mathrm{cov}.(\beta_j, \beta_{j-1})]$.

might reduce productivity if education causes a worker to find casual work demeaning and to not apply as much effort as less educated workers. However, this also reflects the low proportion of educated individuals in casual wage employment—for instance, less than 1% of casual workers were graduates (about 50 individuals) in 1983. This is supported by the multinomial model, where education at any level has a negative effect on the probability of being in casual wage employment.

Estimates of returns to education can be overstated due to biases arising from omitted variables, especially innate ability, schooling quality, socio-economic background, economic sector and sample selection (Bennell, 1996b; Card, 1999; Heckman and Hotz, 1986; Kingdon, 1998). On the other hand, in cases where the schooling variable is measured with error, the returns are downward biased (Card, 1999). The augmented wage regression models estimated in this paper attempt to deal with this problem to some extent by controlling for selection bias and industry affiliation and socio-economic background using caste and religion affiliation. These models do not include variables capturing innate ability, school quality and family background. Although years of schooling does capture some aspects of schooling—poor quality at low levels of schooling will influence whether or not an individual goes on to the next level—it does so only partially. In the Indian context the quality of schooling varies widely and seems to depend a great deal on the family background, although perhaps not as much for boys (Kingdon, 1998). Kingdon also finds that the returns to schooling (the number of years of education) fall from 10.6% to 8.9% for men once variables capturing family background (such as parental education) are included in the 1995 study of urban Uttar Pradesh. In addition, omitting family background was found to over-estimate the returns to education at the graduate and higher levels as individuals who acquire higher education generally belong to privileged backgrounds so that some part of their return to education arises from their backgrounds (Kingdon, 1998). It is not possible, however, to control for either of these variables with the available data. To the extent that household structure (in the selection model), caste and religion variables capture family background, this bias is less serious. The data also do not provide appropriate instruments for the schooling variable.[11]

The results for regular workers presented in Table 3 are apparently at odds with the conventional pattern of returns to education in other countries. A review of various studies in different countries by Psacharopoulos and Patrinos (2004) suggests that returns are highest for primary education (about 26.6% world-wide and 20% in non-OECD Asian countries) and are decreasing in education level (see also Psacharopoulos, 1994). This empirical regularity has been called into question by several studies, especially in African and some Asian countries; see, for example, Bennell (1995, 1996a, 1996b) for a cross-country review; Siphambe (2000) for Botswana; Glewwe (1991) for Ghana; Sahn and Alderman (1988) for Malaysia; Moll (1996) for South Africa; Gindling *et al.* (1995) for Taiwan; and Hawley (2004) for Thailand. In particular, Kingdon (1998) finds in her review of other empirical work on the returns to education in India (mainly computed from specialized surveys in urban areas of a particular state or city) that the rate of return to education, as in Table 3, tends to rise with education level. In order to place the findings of the current paper in the context of the existing literature, Table 4 summarizes estimates of the rate of return to education for male workers for low-income countries and for India (this paper's estimates for regular workers using both the standard and augmented models are also included to facilitate comparison). It should be noted that these estimates are not comparable as they are drawn from studies that differ with respect to the sample and time coverage, the measurement of wages, the specification of the wage regression models and the treatment of selection, and are based on different assumptions about the number of years at each level of schooling (see footnotes to Table 4 for details). In particular, this paper takes two years to be the period of earnings foregone while in primary education, whereas all other Indian studies take this period to be five years (four years in the case of Kingdon and Unni (2001)).

The estimates in Table 4 suggest that the findings of this paper are broadly consistent with other work on the India labour market (especially if the returns to primary education are computed using five years of earnings foregone as in the other Indian studies) and that the conventional pattern of returns does not necessarily hold for India. The estimates for the private rate of return to education at different levels in the Indian studies indicate rising returns with the level of education at least until secondary schooling, and in most cases until graduate schooling. The exception is Tilak's 1988 study cited in Bennell (1995) where the returns to primary education are higher than those to higher levels of education. However, this sample, unlike the other studies on urban samples, relates to one predominantly rural district in a single state and is also the only study using the cost–benefit rather than the regression-method of estimating returns.

Table 4. The private rate of return to education: a survey

Study	Year	Region	Primary schooling	Middle[a] schooling	Secondary schooling	Graduate
World (all workers)						
Psacharapoulos and Patrinos (2004)	Latest	All countries	26.6		17.0	19.0
		Low income countries	25.8		19.9	26.0
		Non-OECD Asian countries	20.0		15.8	18.2
India (male workers only)						
This study (regular workers)						
Augmented Mincerian[b]	1983	India	3.3	2.4	5.3	9.0
	1993–94	India	2.1	1.7	4.3	9.2
	1999–00	India	2.4	2.0	4.6	10.3
Standard Mincerian[b]	1983	India	8.1	3.4	6.0	10.0
	1993–94	India	5.1	3.1	5.4	10.9
	1999–00	India	5.6	3.5	6.1	12.3
Other studies (male workers)						
Duraisamy (2002)[c]	1983	India	6.1	7.1	13.2	12.2
	1993–94	India	6.2	6.4	12.6	12.2
Author's calculations[d] (based on Duraisamy)	1983	India	7.0	8.1	11.7	15.5
	1993–94	India	7.1	7.4	11.4	15.8
Banerjee and Knight (1985)	1975–76	Delhi (urban)	2.4		6.9	11.4
Bennell (1995) citing Tilak[e]	1978	Andhra Pradesh (rural)	9.9		3.2	7.0

Table 4. *(Continued,*

Study	Year	Region	Primary schooling	Middle[a] schooling	Secondary schooling	Graduate
Kingdon (1998)[f]	1995	Lucknow (urban Uttar Pradesh)	2.6[j]	4.9	17.6	18.2[k]
Kingdon and Unni (2001)[g]	1987–88	Madhya Pradesh (urban)	1.4[j]	6.9	14.2	9.6
Kingdon (1998) citing Unni[h]		Madhya Pradesh (urban)	3.1	9.7	12.0	13.5
Kingdon (1998) citing Unni[h]		Tamil Nadu (urban)	2.9	9.0	17.0	15.6
Kingdon and Unni (2001)[g]	1987–88	Tamil Nadu (urban)	1.1[j]	6.4	12.4	17.1
Santhapparaj (1997)[i]	1989	Madurai (urban Tamil Nadu)	–0.9[j]	0.1[j]	0.2[j]	18.5

Sources: Duraisamy (2002), Kingdon and Unni (2001), Kingdon and Unni (2001) citing Unni 1995 and the author's calculations are computed from the NSS survey data; all others are based on smaller purpose-defined surveys. Bennell (1995), Banerjee and Knight (1985), Duraisamy (2002), Kingdon (1998), Kingdon (1998) citing Unni 1995, Kingdon and Unni (2001), Psacharopoulos and Patrinos (2004), and Santhapparaj (1997).
[a]The middle education level is equivalent to the junior education level reported in Kingdon (1998). [b]These are taken from Table 3. Returns are computed under the assumption of two, three, four and three years of additional years of foregone wages at each education level. [c]Duraisamy's (2002) wage regression models are estimated for all rural and urban wage workers for both years and also include a dummy for technical qualification but do not control for potential selection bias. Estimates for 1993–94 control for selection into wage employment but are almost identical to those reported above. [d]Author's calculations are based on the same wage regression models as Duraisamy. These estimates are somewhat higher than Duraisamy as the dummy variable for technical qualification has not been included due to non-availability of data. [e]Bennell (1995) cites the work of J.B.G. Tilak (1988) *The Economics of Inequality in Education* (New Delhi: Sage). These estimates differ from those reported in Psacharapoulos (1994) as they refer to the adjusted estimates (adjusting for socio-economic background, labour force status, type of employer and economic sector). [f]Estimates from Kingdon's (1998) study is based on a sample of adult males aged 15–59 under the assumption of five, four, three and three years of schooling at each education level and control for selection into wage employment. [g]Kingdon and Unni (2001) do not report the rates of return to these education levels. Those reported in the table have been constructed from the coefficients from the wage equations on the education splines using their mapping of four, four, three and three years of schooling at each of these levels. The wage equations are estimated for adult males aged 15–64 years and control for selection into wage employment. [h]Kingdon (1998) cites the work of J. Unni (1995) Returns to education by gender among wage employees in urban India, *Working paper No. 63* (Ahmedabad: Gujarat Institute of Development Research). [i]Santhapparaj's (1997) estimates includes migrants and natives engaged in wage and self-employment. Returns to education are computed from the wage equation coefficients using the correspondence of five, three, four and three years. [j]These rates are insignificantly different from zero. [k]Taken as the average of all returns to post-secondary levels of education.

Duraisamy (2002) is the only other national study that compares returns to education in India over time. A comparably specified OLS model with education, urban/rural residence and experience on male daily wages for a sample of all workers of all age groups yielded similar estimates of the returns to education and trends in these returns.[12] Part of the differences in Duraisamy's estimates and those from this study could be attributed to the inclusion of industry affiliation controls in this paper as well as using a polychotomous model for selection correction into non-wage earners and regular and casual wage workers.[13] Although Duraisamy does control for selection in 1993–94 it is modelled as a binary outcome between wage and non-wage employment and the estimated returns are almost identical to those estimated without these controls. This suggests that it is important to consider regular and casual workers as competing in separate labour markets and control for selection correspondingly. The primary reason for the differences, however, seems to be the earnings measure—Duraisamy uses real daily wages with no controls for hours worked in the wage regression models, whereas this paper uses real hourly wages.[14]

Estimates of returns to education are often used to inform education policy decisions on the allocation of public investment on different levels of education. The finding of relatively low returns to lower levels of education (especially when individual characteristics other than human capital are also controlled for) do not, however, necessarily imply that educational policy in India should not emphasize primary and middle schooling. First, these estimates relate to adult male workers comprising about 41% of the total labour force and cannot be generalized to the entire population which includes, among others, female wage workers and self-employed. Even within the sample of wage workers, casual workers reap some benefit from primary education but none at all from secondary and graduate schooling. Second, the estimates reported in this paper are private rates of returns that overlook the social benefits of primary and middle schooling, especially for female workers, such as political awareness and health outcomes (Kingdon, 1998). A vitally important indirect benefit of primary education is its role as an input for further education. As a result, investment at this level could influence the rates of return at higher levels. Appleton *et al.* (1996) find that in Cote d'Ivoire and Uganda, although the direct private returns to primary education are low, the value of primary education as an input to post-primary education was quantitatively important. Third, using rate of return calculations to direct investment in education implicitly assumes that there are capacity constraints at each level of schooling and that, given the existing returns to education, the role of investment is to choose which schools (primary, middle, secondary or graduate) to build to meet the excess demand. In countries where poor school quality rather than capacity constraints on education are the main problem, rates of return to higher education levels are likely to be rough guides to policy. In this context the 'deepening' of schooling by increasing quality rather than 'broadening' by increasing quantity is a more appropriate strategy (Behrman and Birdsall, 1993; Glewwe, 1996). There was a rapid increase in schooling infrastructure in India after the 1950s—the gross enrolment ratio for boys in primary school rose from 61% to 104% between 1950–51 and 1999–2000, and for middle school rose from 21% to 67% (Government of India, 2002). There is some evidence that this led to a decline in quality (Duraisamy, 2002). In summary, these estimates of rates of return should be interpreted carefully as private rates of return for a sample of adult male wage workers. Primary and middle education serve as necessary inputs to higher levels of education, and as such it is necessary to understand the reasons for relatively low returns rather than simply directing public investment according to the highest rates of return (Glewwe, 1996).

While the average private rate of return to primary, middle and secondary education for regular workers fell immediately following the trade reforms in 1991 (the *t*-statistics for the change between 1983 and 1993 are –1.88, –1.66 and –3.49 respectively), the returns to graduate education rose during the 1990s (the *t*-statistic for the change between 1993 and 1999 is 2.30). For casual workers, on the other hand, the returns to primary education more than doubled (although still barely 1%) between 1983 and 1999 (*t*-statistic 1.74).[15] As the supply of regular workers educated up to the secondary or graduate school increased throughout this period, the increase in the return to graduate education during the 1990s suggests a corresponding rise in the relative demand for these skilled workers. Several other studies have found evidence of increasing educational returns for the more educated during periods of rapid economic change. For instance, Foster and Rosenzweig (1996) found that during the Green Revolution in India, a period of rapid technical change in agricultural production, increasing educational returns were concentrated among the more educated. Newell and Reilly (1999) also find in their study on transitional economies during the 1990s that the private rates of return to education rose after a period of labour market reforms.

These trends are contrary to the predictions of Heckscher–Ohlin trade theory for an unskilled labour abundant country (although in common with several other studies in developing countries; see Wood, 1997). Possible explanations for this are the Skill-Enhancing-Trade or 'SET' hypothesis (Robbins, 1996) and/or a structure of trade protection that formerly favoured relatively unskilled-labour intensive sectors (Harrison and Hanson, 1999). Alternative explanations for this trend include exogenous skilled labour-intensive technological change (Lawrence and Slaughter, 1993). The rising wage gap between graduate and primary education during the 1990s is reflected in the sharp rise in wage inequality during this period—the Gini coefficient of real hourly wages among regular workers rose from 0.39 to 0.43 between 1983 and 1999.

Conclusion

This paper examines the returns to education for regular and casual workers for three years—1983, 1993 and 1999—that span a period of rapid economic liberalization in India. This paper finds that the benefits of education are two-fold—acquiring education at any level not only raises the probability of obtaining regular wage employment, but also results in higher wages for those in regular wage employment. The main finding of this paper is that the returns to education (and experience) are significantly different for regular and casual workers. Casual workers face at best flat returns to education while the returns for regular workers are positive and U-shaped with respect to education level. Patterns of returns increasing in education level have been observed in several country studies in Africa and Asia (see Bennell, 1995, 1996a) and in national and regional studies within India (Duraisamy, 2002; Kingdon, 1998). There is some evidence of significant changes in the returns to education for regular workers over time, and the widening of the wage gap between graduate and primary education could possibly be a consequence of trade liberalization and other reforms during the 1990s.

The finding of relatively low returns to lower levels of education does not necessarily imply advocating a policy of reallocation of public resources from primary or middle to higher education. The estimates do not take into account the social costs and benefits of each education level. Primary and middle education serve as necessary inputs to higher levels of education, and as such it is necessary to understand the reasons for low returns rather than simply directing public investment according to the highest rates of return. Given that this study does not control for innate ability or quality of schooling due to lack of data, this pattern might imply that more able individuals obtain more schooling or that schooling quality improves with education level. In addition, while these estimates can be generalized for the population of male regular wage workers they cannot be generalized to the entire population of workers. Casual workers do derive some economic returns to primary education but none at all from higher education levels.

From a national perspective education enriches the stock of human capital that serves as a production factor, while from an individual's perspective acquiring education yields economic benefits in the form of higher wages. As a result, education outcomes are inter-linked with economic growth and inequality. There is some evidence that the expansion of primary education resulted in higher growth in India between 1966 and 1996 while the causal relationship between secondary and higher enrolments and growth is weak (Self and Grabowski, 2004). The pattern of returns rising with the education level could exacerbate wage and income inequality. A decomposition of the contribution of the explanatory variables included in the augmented wage regression model estimated here following Fields (2002) revealed that about 40% of wage inequality in all three years was explained by the level of education variables (Dutta, 2005). In addition, the rising returns to graduate education contributed to the rise in wage inequality witnessed during the 1990s. The high private returns to higher education indicate that there is room for the government to shift some of the costs of acquiring higher education to individuals. At present about one-half of the total public expenditure on education goes to elementary education (comprising primary and middle schooling) and government policy lays great stress on achieving elementary education for all (Shariff and Ghosh, 2000)—the Education Bill 2003 sought to make primary education free and compulsory for all. However, the emphasis remains on quantity (as witnessed by the gross enrolment rates that exceed 100%) rather than quality. For instance, the Public Report on Basic Education in India (PROBE Team, 1999) found in the 200 villages surveyed in five states that government-funded schools had low levels of teaching activity coupled with dilapidated infrastructure and poor management (see also Kingdon, 1996). As a result, public investment to raise the returns to primary and middle education even further through investment to improve the quality of primary and middle schooling would be desirable.

Acknowledgements

This research was carried out as part of my D.Phil. thesis at the University of Sussex and was supported by grants from the BFWG Charitable Foundation, the Department of Economics at the University of Sussex, the International Foundation of University Women, Universities UK, the Wingate Foundation and the Royal Economic Society and support from the Poverty Research Unit at Sussex. I am very grateful to my supervisors, Dr Barry Reilly and Professor Alan Winters, for their guidance and suggestions. Comments on earlier drafts from two anonymous referees, participants at the PRUS seminar (2003) and from Patricia Justino, Julie Litchfield, Andy Newell, Yoko Niimi and Howard White were also much appreciated.

Notes

1. In this paper we focus on the wage determination process for adult male workers; we extend this analysis to female workers in our other work.
2. Although there has been some criticism of this approach in the literature (see Bourguignon *et al.*, 2004), the Lee correction was chosen over alternative parametric and semi-parametric approaches

(see Vella, 1998) because of its simplicity, computational convenience and transparent interpretation of the selection effect. The Bourguignon *et al.* alternative allows for the estimation of one outcome equation only, whereas in this paper there are two outcome equations—wage regression models for regular and casual workers. It should be noted that the parameter estimates of the wage equations using power series approximations for the selection term following the semi-parametric approach advocated by Newey (1999) were very similar to those obtained using the Lee correction.
3. The employment survey for 1987–88 could not be used as over 76% of observations on rural wages for persons participating in wage employment are missing (see also Duraisamy, 2002).
4. Results are available on request.
5. As the choice of identifying variables is necessarily *ad hoc* the MNL model was estimated for different specifications of identifying variables. The parameter estimates in the wage equations are not sensitive to the choice of the identifying variables and the coefficient on the correction term itself was not materially different across specifications. On balance, these instruments were also not found to strongly influence wages in most specifications in most years.
6. The χ^2_{40} statistics are 19 087.91, 21 281.50 and 22 068.91 for the three years, respectively. At the same time, the results of the Small–Hsiao test for the independence of irrelevance alternatives assumption are inconclusive, implying correlations in the extreme value errors. The Hausman–McFadden test could not be conducted as the matrix of the differences in the variance–covariance matrices of the restricted and full models was not positive definite.
7. For the standard Mincerian specification, the χ^2_5 statistics are 244.66 between 1983 and 1993, 152.28 between 1993 and 1999, and 794.73 between 1983 and 1999 for regular workers, and 47.30, 22.78 and 119.85 for the three years for casual workers. The corresponding figures for the augmented specification are 95.78, 24.54 and 193.21 for regular workers and 22.36, 20.52 and 66.42 for casual workers.
8. The omitted category for the education dummy variables is those who are illiterate or have less than two years of formal education.
9. The 1983 survey has additional information on the number of years of schooling and somewhat confirms this correspondence—on average, individuals that had completed primary, middle, secondary and graduate education had done so in 5.18, 8.17, 10.79 and 14.16 years, respectively.
10. The standard specification yields returns to primary education for casual workers that are comparable with those for regular workers in the 1990s as well as a positive, but very small, return to middle schooling (and to secondary schooling in 1983). Note, however, that the human capital variables as captured by age and education explained only about 5% of the variation in log real hourly wages for casual workers in the standard Mincerian specification (regression not reported).
11. Card (1999) concludes in his survey of the empirical literature that the OLS estimates of returns to education are not widely different from instrumental variables (IV) estimates obtained after correcting for ability bias and measurement error.
12. The education effects are marginally higher than those reported by Duraisamy as the current paper is unable to include a technical education dummy in the analysis—all values for this variable equalled one in the 1983 survey.
13. The estimates from the standard Mincerian model used in this paper are closer to Duraisamy's estimates than the augmented model (see Table 4).
14. Although using the constructed real hourly wage variable (see Appendix 1) introduces measurement error in the dependent variable, this would be captured in the error term in the wage regression models. On the other hand, using the reported real weekly wage as the dependent variable without controlling for the hours worked or including this as a control might lead to biased parameter estimates due to omitted variable bias or because of correlation of measurement error with the error term.
15. A broadly similar trend for returns to regular workers is apparent from the standard Mincerian model—with the returns to primary and secondary schooling declining between 1983 and 1993 while the returns to graduate education rose during both subperiods. For casual workers, on the other hand, there is no significant change in the returns to primary education over time, although the returns to secondary and graduate education decline between 1983 and 1993. A Wald test of shifts in the education coefficients from the wage equations between 1983 and either of the later two years (although not for the 1990s) is decisively rejected for regular workers. For casual workers, on the other hand, the null hypothesis of no change during any pair of years cannot be rejected.

References

Appleton, S. *et al.* (1996) Primary education as an input into post-primary education: a neglected benefit, *Oxford Bulletin of Economics and Statistics*, 58(1), pp. 211–219.

Arbache, J. S. *et al.* (2004) Trade liberalisation and wages in developing countries, *Economic Journal,* 114(493), pp. F73–F96.

Banerjee, B. and Knight, J. B. (1985) Caste discrimination in the Indian labour market, *Journal of Development Economics,* 17(3), pp. 277–307.

Behrman, J. and Birdsall, N. (1993) The quality of schooling: quantity alone is misleading, *American Economic Review,* 73(5), pp. 926–946.

Bennell, P. (1995) Rates of return to education in Asia: a review of the evidence, Institute of Development Studies Working Paper 24, Brighton: Institute of Development Studies.

Bennell, P. (1996a) Rates of return to education: does the conventional pattern prevail in sub-Saharan Africa?, *World Development,* 24(1), pp. 183–199.

Bennell, P. (1996b) Using and abusing rates of return: a critique of the World Bank's 1995 education sector review, *International Journal of Educational Development,* 16(3), pp. 235–248.

Bollinger, C. R. and Chandra, A. (2005) Iatrogenic specification error: a cautionary tale of cleaning data, *Journal of Labor Economics,* 23(2), pp. 235–257.

Bourguignon, F. *et al.* (2004) Selection bias correction based on the multinomial logit model: Monte-Carlo comparisons, DELTA Working Paper No. 2004–20, Paris: DELTA.

Card, D. (1999) The causal effect of education on earnings, in: O. Ashenfelter and D. Card (Eds) *Handbook of Labor Economics,* Vol. 3, Chapter 30, pp. 1801–1863 (Amsterdam Elsevier Science BV).

Das, M. B. (2003) Ethnicity and social exclusion in job outcomes in india: summary of research findings, unpublished paper, World Bank Institute.

Duraisamy, P. (2002) Changes in the returns to education in India, 1983–94: by gender, age-cohort and location, *Economics of Education Review,* 21(6), pp. 609–622.

Dutta, P. V. (2005) Accounting for wage inequality in India, *Indian Journal of Labour Economics,* 48(2), pp. 273–295.

Fields, G. S. (forthcoming) Accounting for income inequality and its change, *Research in Labor Economics.*

Foster, A. D. and Rosenzweig, M. R. (1996) Technical change and human-capital returns and investments: evidence from the green revolution, *American Economic Review,* 86(4), pp. 931–953.

Gindling, T. H. *et al.* (1995) Changing returns to education in Taiwan: 1978–91, *World Development,* 23(2), pp. 343–356.

Glewwe, P. (1991) Schooling, skills, and the returns to government investments in education: an exploration using data from Ghana, *Living Standards Measurement Study Working Paper No. 76* (World Bank).

Glewwe, P. (1996) The relevance of standard estimates of rates of return to schooling for education policy: a critical assessment, *Journal of Development Economics,* 51(2), pp. 267–290.

Government of India (2002) *Selected Education Statistics, 2001–2002* (New Delhi: Ministry of Human Resource Development, Government of India).

Harrison, A. and Hanson, G. (1999) Who gains from trade reform? Some remaining puzzles, *Journal of Development Economics,* 59(1), pp. 125–154.

Hawley, J. D. (2004) Changing returns to education in times of prosperity and crisis, Thailand 1985–1998, *Economics of Education Review,* 23(3), pp. 273–286.

Heckman, J. and Hotz, V. (1986) An investigation of the labour market earnings of panamanian males: evaluating the sources of inequality, *Journal of Human Resources,* 21(4), pp. 507–542.

Kabeer, N. (2002) Safety nets and opportunity ladders: addressing vulnerability and enhancing productivity in South Asia, Overseas Development Institute Working Paper 159, London: Overseas Development Institute.

Kingdon, G. G. (1996) The quality and efficiency of private and public education: a case-study of urban India, *Oxford Bulletin of Economics and Statistics,* 58(1), pp. 55–80.

Kingdon, G. G. (1998) Does the labour market explain lower female schooling in India?, *Journal of Development Studies,* 35(1), pp. 39–65.

Kingdon, G. G. and Unni, J. (2001) Education and women's labour market outcomes in India, *Education Economics,* 9(2), pp. 173–195.

Krueger, A. B. and Summers, L. H. (1988) Efficiency wages and the inter-industry wage structure, *Econometrica,* 56(2), pp. 259–293.

Labour Bureau (various years) *Indian Labour Statistics* (Chandigarh: Labour Bureau Government of India).

Lawrence, R. Z. and Slaughter, M. J. (1993) International trade and American wages in the 1980s: giant sucking sound or small hiccup?, *Brookings Papers on Economic Activity: Microeconomics,* 2, pp. 161–211.

Lee, L.-F. (1983) Generalised econometric models with selectivity, *Econometrica,* 51(2), pp. 507–512.

Moll, P. (1996) The collapse of primary schooling returns in South Africa, 1960–90, *Oxford Bulletin of Economics and Statistics,* 58(1), pp. 185–210.

Murphy, K. M. and Welch, F. (1990) Empirical age–earnings profile, *Journal of Labour Economics*, 8(2), pp. 202–229.

Newell, A. and Reilly, B. (1999) Rates of return to educational qualifications in the transitional economies, *Education Economics*, 7(1), pp. 67–84.

Newey, W. K. (1999) Two step series estimation of sample selection models, MIT Department of Economics Working Paper 99–04, Cambridge, MA: MIT.

PROBE Team (1999) *Public Report on Basic Education in India* (New Delhi: Oxford University Press).

Psacharopoulos, G. (1987) Earnings and education in Brazil: evidence from the 1980 census, World Bank Policy Research Working Paper 2881, Washington, DC: World Bank.

Psacharopoulos, G. (1994) Returns to investment in education: a global update, *World Development*, 22(9), pp. 1325–1343.

Psacharopoulos, G. and Patrinos, H. A. (2004) Returns to investment in education: a further update, *Education Economics*, 12(2), pp. 111–134.

Robbins, D. (1996) HOS hits facts: facts win. Evidence on trade and wages in the developing world, Harvard Institute for International Development Discussion Paper 557, Cambridge, MA: Harvard University.

Sahn, D. and Alderman, H. (1988) The effects of human capital on wages and the determinants of labour supply in a developing country, *Journal of Development Economics*, 29(2), pp. 157–183.

Santhapparaj, A. S. (1997) Earnings and household income of the migrants and natives: the role of human capital investment and social factors, *Manpower Journal*, 32(4), pp. 27–46.

Self, S. and Grabowski, R. (2004) Does education at all levels cause growth? India, a case study, *Economics of Education Review*, 23(1), pp. 57–65.

Shariff, A. and Ghosh, P. K. (2000) Indian education scene and the public gap, *Economic and Political Weekly*, 15 April, pp. 1396–1406.

Siphambe, H. P. (2000) Rates of return to education in Botswana, *Economics of Education Review*, 19(3), pp. 291–300.

Tendulkar, S. D. (2003). Organised labour market in India: pre and post reform. Paper presented at the Conference on Anti Poverty and Social Policy, Alwar, India, 2–4 January.

Tilak, J. B. G. (1994) *Education for Development in Asia* (New Delhi: Sage Publications).

Unni, J. (2001) Gender and informality in labour market in South Asia, *Economic and Political Weekly*, 36(26), pp. 2360–2377.

Vella, F. (1998) Estimating models with sample selection bias: a survey, *Journal of Human Resources*, 33(1), pp. 127–169.

Wood, A. (1997) Openness and wage inequality in developing countries: the Latin American challenge to East Asian conventional wisdom, *World Bank Economic Review*, 11(1), pp. 33–57.

Appendix 1. Data Appendix

The three large-scale employment surveys for January–December 1983, July 1993–June 1994 and July 1999–June 2000 (referred to as 1983, 1993 and 1999 in the paper) provide comprehensive national coverage and provide a wealth of information on numerous socio-economic issues at the household and individual level.

Wages

Nominal weekly wages include payment in cash and kind. Some observations (about 1–2% in the three years) had to be dropped from the sample as there were missing observations on wages, hours worked and industry affiliation. It is assumed that the excluded observations are random as the mean observable characteristics of the workers excluded do not differ significantly from those retained in the sample, although this does not take possible differences in unobservables into account. The wage distribution was then trimmed by 0.1% at the top and bottom tails. This is necessarily an *ad hoc* measure: some researchers prefer to 'winsorize' the wage distribution using specific values (Krueger and Summers, 1988), while others prefer to trim the distribution at the tails (Arbache *et al.*, 2004) as adopted here. It is acknowledged that there are potential problems associated with both procedures that could exacerbate coefficient bias (Bollinger and Chandra, 2005). These nominal wages were deflated to 1983 prices using official state-level monthly consumer price indices (base year 1960–61) for agricultural labourers (CPIAL) for rural wages and industrial workers (CPIIW) for urban wages (Labour Bureau, various years). As the hours worked are not reported, the intensity of work for each day of the week preceding the survey is used to construct the weekly hours worked. This variable takes one of three values—no work, part-time if worked between one and four hours during the day, or full-time if worked more than four hours during the day. The number of hours worked each day is coded zero if no work, four hours if part-time work and eight hours in full-time work reported for that day. This is then aggregated for all seven days to obtain a measure of hours worked in the week, subject to a maximum of 48 hours a week. The real hourly wage is constructed by dividing the real weekly wage by the number of hours worked. While a reported hourly wage variable would have been ideal these are the best data available to us, and we believe that the possible measurement error introduced by using this constructed variable is less serious than that that would ensue if we ignored the variation in hours worked altogether.

Variables Influencing Wages

Individuals were divided into three mutually exclusive categories using current weekly status: (i) non-wage earners (i.e., non-participants in the labour market, self-employed and unemployed individuals); (ii) regular wage employment; and (iii) casual wage employment. The standard quadratic form for age is not used as this did not fit the data well and, following Murphy and Welch (1990), age splines at 10-year intervals were included instead as a proxy for labour force experience. Marital status is a dummy variable coded one if currently married and zero if never married, widowed, divorced or separated. There is information on the highest level of schooling completed (but not on the number of years of schooling) so dummy variables corresponding to the following education variables were constructed: primary school, middle school, secondary school and graduate and above. The reference category is individuals who are illiterate or have less than two years of formal or informal schooling. Dummy variables for caste and religious affiliation were constructed from household data; the omitted category is all other households. Seasonality effects are captured by dummy variables for the quarter in which the households were interviewed. These quarterly dummies were also interacted with the dummy variable for the rural sector. The variable for industry affiliation was constructed based on the individual's current weekly industrial classification. In order to ensure adequate observations in each industry the three-digit National Industrial Classification codes are aggregated into 38 industries.

Table A1. Summary statistics: wage regression models

	Regular workers			Casual workers		
	1983	1993	1999	1983	1993	1999
Natural log of real hourly wages (Rs.)	1.2679	1.5173	1.6919	0.6936	0.8513	0.9632
	(0.5176)	(0.5867)	(0.6539)	(0.2620)	(0.2786)	(0.2910)
Age	34.8790	36.7696	37.1110	32.4766	33.0999	33.0796
	(11.0472)	(10.8995)	(11.2464)	(12.2919)	(12.1414)	(11.9523)
Age spline: 15-25 years	24.1235	24.3587	24.3413	23.3198	23.5189	23.5630
	(2.1711)	(1.8793)	(1.8730)	(2.9473)	(2.8005)	(2.7283)
Age spline: 25-35 years	6.2395	6.9669	6.9569	5.1112	5.3953	5.4257
	(4.2634)	(4.0333)	(4.0819)	(4.5413)	(4.4880)	(4.4887)
Age spline: 35-45 years	3.2483	3.9034	4.0739	2.6734	2.7623	2.7481
	(4.2377)	(4.3920)	(4.4414)	(4.0875)	(4.1097)	(4.0790)
Age spline: 45-55 years	1.1109	1.4049	1.5646	1.1031	1.1495	1.0790
	(2.7388)	(2.9901)	(3.1579)	(2.8418)	(2.8970)	(2.8054)
Age spline: 55-65 years	0.1570	0.1357	0.1742	0.2690	0.2739	0.2639
	(0.9775)	(0.8399)	(0.9289)	(1.3331)	(1.3273)	(1.3383)
Married	0.7760	0.8030	0.7889	0.7186	0.7334	0.7218
Literate ‡	0.0987	0.0823	0.0715	0.1514	0.1670	0.1557
Completed primary school	0.1472	0.1027	0.0969	0.1437	0.1447	0.1465
Completed middle school	0.1745	0.1612	0.1694	0.0744	0.1127	0.1542
Completed secondary school	0.2560	0.3106	0.3300	0.0207	0.0458	0.0740
Completed graduate school	0.1403	0.2342	0.2408	0.0016	0.0036	0.0054
Member of scheduled caste or tribe	0.1770	0.1555	0.1793	0.4092	0.4253	0.4310
Muslim	0.0937	0.0888	0.1024	0.1143	0.1088	0.1221
All others ‡	0.7293	0.7556	0.7183	0.4765	0.4659	0.4470
	(0.4443)	(0.4297)	(0.4498)	(0.4995)	(0.4988)	(0.4972)
Residence in rural areas	0.3098	0.2850	0.2849	0.7958	0.7640	0.7500
Selection bias correction term	1.1441	1.1121	1.1626	1.1901	1.1453	1.1460
	(0.5268)	(0.5122)	(0.4830)	(0.3908)	(0.4301)	(0.4113)
Total number of observations	27,356	26,387	27,295	28,855	26,398	29,805

Notes: Figures in parentheses are standard deviations. In the case of dummy variables the mean refers to the percentage of observations falling in each category (standard deviations not reported). 6 seasonal, 16 state and 37 industry dummy variables are also included in the wage regression models but not reported to conserve space. ‡ Omitted dummy variables.