

# **Caste Discrimination and Transaction Costs in the Labor Market: Evidence from Rural North India \***

**Takahiro Ito** †

**June 2008**

## **Abstract**

This paper is an empirical attempt to quantify caste-based discrimination in the labor market using household data taken from rural North India. In the regression analysis, transaction costs associated with entry into the labor market and reservation wages are estimated simultaneously along with market wages. The estimation results provide evidence of the existence of transaction costs in the labor market and discrimination against backward classes with regard to access to regular employment. In line with previous studies, the results suggest that the achievements of India's reservation policy so far have at best been limited. In addition, a comparison between the estimates from the model employed in this paper and conventional (reduced-form) approaches shows that discrimination in labor market entry is likely to be underestimated in the conventional reduced-form approaches.

*JEL classification codes:* D23, J22, J24, J71

*Keywords:* regular employment, casual employment, labor market, India

---

\* I am deeply grateful to Takashi Kurosaki and Daiji Kawaguchi for their advice and encouragement. I have also benefited from many comments and suggestions I received from two anonymous referees of the Journal, Yukinobu Kitamura, Yasuyuki Sawada, Hirokazu Ishise, Masayuki Kudamatsu, Tomohiro Machikita, Masahiro Shoji and participants at various seminars at Hitotsubashi University. An earlier version of this paper received the 2007 Kanematsu Fellowship Award from RIEB of Kobe University. I also thank to three anonymous referees of the Award and participants at the Kanematsu Fellowship seminar for their helpful comments. All mistakes remain my own.

† Graduate School of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601 Japan. Phone: 81-42-580-8000. E-mail: ed044001@g.hit-u.ac.jp.

# **Caste Discrimination and Transaction Costs in the Labor Market: Evidence from Rural North India**

## **1. Introduction**

For rural dwellers in developing countries, off-farm activities are becoming more important in determining their welfare. In rural India, where the labor market is relatively large, off-farm wage employment plays an important role as a source of income (Lanjouw and Shariff, 2004) and as insurance against agricultural risks (Kochar, 1999; Rose, 2001; Ito and Kurosaki, 2006).

While there are a number of studies on the labor market in India, only a few have investigated the role of labor market discrimination.<sup>1</sup> The existence of discrimination may inhibit the expansion of the labor market which could help to reduce the dependence on agriculture. Moreover, the existence of discrimination may distort households' decisions not only with regard to labor allocation but also to human capital investment. In the rural context, both are issues of great importance. Examining labor market discrimination using rural household data therefore can make an important contribution to understanding how rural development can be achieved and rural poverty eliminated.

---

<sup>1</sup> Banerjee and Knight (1985) is a pioneering work on caste-based discrimination in the Indian urban labor market, and recently the journal *Economic and Political Weekly* featured this topic in its October 13th 2007 issue. In addition, Kingdon and Unni (2001) investigated gender-based discrimination in the urban labor market.

This paper is an empirical attempt to quantify caste-based discrimination in the labor market using household data from rural India. The caste system<sup>2</sup> and caste-based discrimination have been officially outlawed. In order to equalize the economic opportunities between castes, Indian government introduced “reservations” (the quotas imposed in government jobs and educational institutions) for lower castes shortly after its independence. The first beneficiaries were members of the scheduled castes (*dalits*, or once known as “untouchables”) and scheduled tribes (SCs/STs), who sit at the bottom of the social hierarchy. Then, the list of eligible groups for reservation was extended, and “other backward classes” (OBCs) were added to the list.<sup>3</sup>

Despite such policy efforts, the economic circumstances of socially backward and upper classes continue to differ substantially (Srinivasan and Kumar, 1999; Thorat, 2002; Borooah, 2005; Kijima, 2006). In the labor market, too, evidence of caste-based discrimination has been found (Banerjee and Knight, 1985; Borooah et al., 2005; Madheswaran and Attewell, 2007; Thorat and Attewell, 2007).

Findings from the previous studies on labor market discrimination are somewhat mixed with regard to discrimination in labor market access (“job

---

<sup>2</sup> Castes (the traditional hereditary classes) in India consist of thousands of endogamous groups called *jatis* (the word literally means “birth”). Members of each *jati* are typically and traditionally engaged in the same occupation.

<sup>3</sup> See Osborne (2001), for the history and evolution of India’s reservation policy.

discrimination”) but suggest that discrimination in wages (“wage discrimination”) does exist and is more likely to be present in unskilled, manual jobs. For instance, Banerjee and Knight (1985), in their examination of wage differentials between SC and non-SC migrant workers in the urban labor market, sought to distinguish two components: discrimination in wages within the same occupation (“wage discrimination”) and discrimination in access to certain occupational position (“job discrimination”). They found that “unexplained” components account for a significant part of observed wage differentials, and that “wage discrimination” dominated “job discrimination.” Likewise, Borooah et al. (2005) examined differences in employment rates between upper and backward castes in the urban labor market and found that “job discrimination” against the backward classes explains only part of the observed differences. On the contrary, Madheswaran and Attewell (2007), which investigated wage differentials between SC and non-SC in the urban labor market, found that the estimated wage gap due to “job discrimination” contributes to the largest part of the observed wage gap.

However, these studies have a common shortcoming that the estimate of “job discrimination” virtually includes the effect of “wage discrimination” because their analyses on labor market participation are based on a reduced-form approach.<sup>4</sup> For

---

<sup>4</sup> This shortcoming is not necessarily due to reduced-form estimation. For example, if a dataset containing information related to firms’ hiring decisions, such as characteristics of job applicants

instance, suppose that members of a low caste are less likely to be engaged in an occupation that pays unreasonably low wages to them. Then, reduced-form estimations cannot identify whether their low participation is due to “job discrimination” or the fact that they do so on a voluntary basis because of “wage discrimination.” The aim of this paper, therefore, is to address this shortcoming by employing a structural model in which the (fixed) transaction costs associated with entry into the labor market and reservation wages are estimated along with market earnings from the observed individuals' labor supply responses.<sup>5</sup>

If individuals knew all particulars about all available jobs, i.e., in the case of perfect information, there would be no need to devote time to looking for a suitable job. In reality, however, this is not the case. The costs of finding employment depend on an individual's characteristics and the features of the region he or she resides in. Moreover, if employers have discriminatory tastes toward certain social classes or utilize social class as supplementary information because of limited information on individual

---

who applied for a job and their results is available, the net effects of “job discrimination” could be estimated. Although such a dataset is not readily obtainable, Thorat and Attewell (2007), as an exception, collected a dataset of (hypothetical) job applicants by sending different types of applications to several job offers placed in newspaper advertisements and examined the existence of “job discrimination.”

<sup>5</sup> For the studies on transaction costs associated with entry into the market, see, for instance, Cogan (1981) on married women's labor supply in the U.S., and Sadoulet et al. (1998) and Key et al. (2000) on corn producers' supply responses to labor and product markets in Mexico. These studies found that transaction costs matter for workers' or households' decisions on market involvement. In addition, Jacoby (1993) found evidence for the existence of transaction costs in the labor market in an analysis of farmers' labor supply in Peru, while Skoufias (1994) obtained a similar result with a sample of Indian rural households.

characteristics, the costs may differ from one class to another. The structural estimation employed in this study enables us to more clearly distinguish between discrimination in wages and in labor market entry. From a policy-making viewpoint, it is quite important to know whether labor market discrimination exists in the form of wage differentiation (“wage discrimination”) or the range of jobs available (“job discrimination”) or both.

The remainder of this paper is organized as follows. Section 2 proposes empirical models to estimate market/reservation wages and transaction costs simultaneously. Section 3 describes the dataset. The sample used in the analysis consists of working-age males in rural areas of Bihar and Uttar Pradesh in North India. Section 4 presents the empirical results. They provide evidence of the existence of transaction costs in the labor market and discrimination against backward classes with regard to access to regular employment, but no evidence of discrimination in wages from regular employment activities is found. Furthermore, a comparison between the estimates from the structural model and the conventional (reduced-form) approaches shows that discrimination in labor market entry is likely to be underestimated when one employs the latter. Section 5 concludes the paper.

## **2. Empirical specification**

Market labor can be broadly classified in terms of employment status into two types: casual and regular employment. In what follows, it is assumed that while there is no cost of entry into the *casual* labor market, transaction costs are involved in finding *regular* employment.<sup>6</sup> Transaction costs in the labor market represent fixed costs associated with labor market entry, such as actual expenditure and/or time spent traveling for finding employment. With transaction costs in monetary terms represented by  $C$ , the reservation-wage model says that the probabilities of being regularly employed and of being casually employed are respectively:

$$\begin{aligned}\Pr(I_r = 1) &= \Pr(W_r - C > W_c, W_r - C > W^R) \\ &= \Pr(\ln W_r - \ln K > \ln W_c, \ln W_r - \ln K > \ln W^R),\end{aligned}$$

$$\begin{aligned}\Pr(I_c = 1) &= \Pr(W_c > W_r - C, W_c > W^R) \\ &= \Pr(\ln W_c > \ln W_r - \ln K, \ln W_c > \ln W^R),\end{aligned}$$

where  $I_k$  is an indicator variable that takes unity if the individual is employed in a type  $k$  job ( $k=r$ : regular,  $c$ : casual),  $W_k$  is the wage earned in the job,  $W^R$  is the reservation wage,<sup>7</sup> and  $K$  is the relative value of the transaction cost defined by

---

<sup>6</sup> In general, specific skills or knowledge are not necessarily required in casual wage labor because this, for the most part, consists of physical labor. In addition, most casual laborers work within their villages, while regular workers often work in neighboring cities (outside the village). For these reasons, transaction costs are likely to be negligible and constant. Thus, the assumption not only simplifies the empirical model but is also plausible in the present context.

<sup>7</sup> As mentioned in the following section, the majority of rural workers in developing countries are engaged in self-employed farming and it is not necessarily assumed in our model that leisure is the sole alternative to wage labor. Thus, the term “reservation wage” is used in the broad sense that it

$$K - 1 = \frac{C}{W_r - C}.$$

The functions of market wages in each type of employment, (relative) transaction costs, and reservation wages are respectively specified as

$$(1) \ln W_k = X\beta_k + e_k, \quad k = r, c.,$$

$$(2) \ln K = Z_K\gamma_K + u_K,$$

$$(3) \ln W^R = Z_R\gamma_R + u_R,$$

where  $X$ ,  $Z_K$ , and  $Z_R$  are matrices of variables affecting market wages, transaction costs, and reservation wages;  $\beta$  and  $\gamma$  are vectors of coefficients to be estimated; and random error terms  $(e_r, e_c, u_K, u_R)$  are independent of  $(X, Z_K, Z_R)$  with zero mean and have a joint normal distribution with covariance matrix  $\Sigma$  given by

$$\Sigma = \begin{pmatrix} \sigma_r^2 & \sigma_{r,c} & \sigma_{r,K} & \sigma_{r,R} \\ & \sigma_c^2 & \sigma_{c,K} & \sigma_{c,R} \\ & & \sigma_K^2 & \sigma_{K,R} \\ & & & \sigma_R^2 \end{pmatrix}.$$

Then, the likelihood function of a market laborer is given by  $\ell^k = \Pr(I_k = 1) f(\ln W_k | I_k = 1)$  ( $k = r, c.$ ). By applying Bayes' rule,  $f(\ln W_k | I_k) = f(I_k | \ln W_k) f(\ln W_k) / f(I_k)$ , and from standard conditional distribution results for joint normal random variables, the likelihood function  $\ell^k$  can be rewritten as

---

determines whether the worker is engaged in wage work or not (including self-employment).



$$\begin{aligned}
\ell^k &= \Pr(I_k = 1 | \ln W_k) f(\ln W_k) \\
&= \Pr(\eta_{1k} > \varepsilon_{1k}, \eta_{2k} > \varepsilon_{2k} | e_k) \frac{1}{\sigma_k} \phi\left(\frac{\ln W_k - X\beta_k}{\sigma_k}\right), \\
&= \Phi_2\left(\frac{\eta_{1k} - \mu_{1k}}{\sigma_{1k}}, \frac{\eta_{2k} - \mu_{2k}}{\sigma_{2k}}, \rho_{12k}\right) \frac{1}{\sigma_k} \phi\left(\frac{\ln W_k - X\beta_k}{\sigma_k}\right), \quad k = r, c.,
\end{aligned}$$

where  $\Phi_2(\cdot)$  is the bivariate standard normal cumulative distribution function,  $\phi(\cdot)$  is the (univariate) standard normal density function,  $\eta_{1r} = -\eta_{1c} = X\beta_r - Z_K\gamma_K - X\beta_c$ ,  $\eta_{2r} = X\beta_r - Z_K\gamma_K - Z_R\gamma_R$ ,  $\eta_{2c} = X\beta_c - Z_R\gamma_R$ ,  $\varepsilon_{1r} = -\varepsilon_{1c} = -e_r + u_K + e_c$ ,  $\varepsilon_{2r} = -e_r + u_K + u_R$ , and  $\varepsilon_{2c} = -e_c + u_R$ . In addition, the means and the elements in the variance-covariance matrix of the conditional distribution of  $\varepsilon$  given  $e_k$  are, respectively,

$$\begin{aligned}
\mu_{\ell,k} &= E(\varepsilon_{\ell,k} | e_k) = \frac{\text{Cov}(\varepsilon_{\ell,k}, e_k) \cdot e_k}{\sigma_k^2}, \\
\sigma_{\ell,k}^2 &= V(\varepsilon_{\ell,k} | e_k) = E(\varepsilon_{\ell,k}^2) - \left(\frac{\text{Cov}(\varepsilon_{\ell,k}, e_k)}{\sigma_k}\right)^2, \quad k = r, c., \quad \ell = 1, 2., \\
\rho_{12k} &= \frac{\text{Cov}(\varepsilon_{1k}, \varepsilon_{2k} | e_k)}{\sigma_{1k} \sigma_{2k}} = \frac{\text{Cov}(\varepsilon_{1k}, \varepsilon_{2k}) - \text{Cov}(\varepsilon_{1k}, e_k) \cdot \text{Cov}(\varepsilon_{2k}, e_k) / \sigma_k^2}{\sigma_{1k} \sigma_{2k}},
\end{aligned}$$

where  $E(\cdot)$ ,  $\text{Cov}(\cdot)$ , and  $V(\cdot)$  represent operators taking the mathematical expectation, the covariance, and the variance of their arguments, respectively. Further,

the likelihood function of a non-participant is

$$\begin{aligned}
\ell^n &= \Pr(I_n = 1) = \Pr(\ln W_r - \ln K < \ln W^R, \ln W_c < \ln W^R) \\
&= \Pr(\eta_{1n} < \varepsilon_{1n}, \eta_{2n} < \varepsilon_{2n}) = \Phi_2\left(-\frac{\eta_{1n}}{\sigma_{1n}}, -\frac{\eta_{2n}}{\sigma_{2n}}, \rho_{12n}\right),
\end{aligned}$$

where  $\eta_{1n} = \eta_{2r}$ ,  $\eta_{2n} = \eta_{2c}$ ,  $\varepsilon_{1n} = \varepsilon_{2r}$ ,  $\varepsilon_{2n} = \varepsilon_{2c}$ ,  $\sigma_{\ell,n}^2 = V(\varepsilon_{\ell,n})$  ( $\ell = 1, 2$ ), and  $\rho_{12n} = \text{Cov}(\varepsilon_{1n}, \varepsilon_{2n}) / (\sigma_{1n} \sigma_{2n})$ . Combining all of these equations, the log likelihood function is then given by

$$(4) \quad LL = \sum \left\{ I_r \cdot \ln \ell^r + I_c \cdot \ln \ell^c + I_n \cdot \ln \ell^n \right\}.$$

Note that  $\ln K$  and  $\ln W^R$  are estimated as latent variables to fit the observed workers' labor supply to the model — in short, to maximize the log-likelihood of the model. Hence, one may think identification problems arise when estimating parameters  $\gamma_K$  ( $\gamma_R$ ) separately from  $\gamma_K$  ( $\gamma_R$ ) and  $\beta_k$  ( $k = r, c$ ). However, due to the assumption that there are no transaction costs involved in finding casual employment, but such costs are present in finding regular employment, identification of  $\gamma_K$  and  $\gamma_R$  are guaranteed in turn by  $\beta_k$ , which can be estimated consistently from the density  $f(\ln W_k)$  in the log likelihood function. Therefore, parameters ( $\beta$  and  $\gamma$ ) can be consistently estimated if the explanatory variables are orthogonal to the error terms, the joint distributional assumption of the error terms is correct, and if the reservation-wage model accurately describe the labor participation decision.<sup>8</sup>

---

<sup>8</sup> Theoretically speaking, identification of  $\beta$  is always guaranteed due to the functional form of the regression model even in the case that all elements in  $X$ ,  $Z_K$  and  $Z_R$  are identical, i.e., there is no exclusion restriction on parameters (identification by non-linearity). However, as mentioned in the following section,  $Z_K$  and  $Z_R$  include at least one variable which is not included in  $X$ , and therefore the identification problem does not arise as long as the assumptions above are satisfied.

### **3. Data**

#### **3.1 Sample and key features**

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 conducted in developing countries. Uttar Pradesh (UP) and Bihar are located in the Ganges Plain of North India and are known for their high poverty incidence. The survey was conducted in 1997/98, covering 1,035 households, 57 villages, and 13 districts in Bihar and 1,215 households, 63 villages, and 12 districts in UP. The sample used in the analysis comprises male household members aged between 15 and 60 and consist of 3,324 individuals.<sup>9</sup> The caste category used in this paper follows the classification of the survey (upper caste, middle caste, agriculture-based backward caste, other backward caste, SC/ST, Muslim upper caste, and Muslim backward caste). Because of the lack of information on sub-caste groups, the caste classification in this paper does not necessarily correspond to the one of the reservation policy. For instance, “other backward castes” in this paper does not exactly correspond to “other backward classes” (OBCs) in the reservation list. OBCs might exactly

---

<sup>9</sup> To focus on labor market participation, students (351), disabled people (35) and members with missing information on work activities and/or other characteristics (227) were excluded from the sample.

correspond to the whole set of “agriculture-based backward castes” and “other backward castes,” or include a subset of the two.

[Table 1]

Information on work activities and wages is available for each household member from January 1997 to December 1997. Market wages are reported by employment status: regular (salaried) or casual employment. Table 1 shows summary statistics for labor market participation and monthly earnings of working-age males by caste. Monthly earnings from regular employment activities in addition to the base salary include other payments such as bonuses, while monthly earnings from casual employment activities are the average earnings in working months.

Obviously, the composition of the employment status and average monthly earnings differ from caste to caste. Individuals belonging to middle and backward castes are less likely to be in regular employment and more likely to be casually employed than those belonging to the upper or Muslim upper castes. Especially, among the scheduled castes, more than half of all male workers are casually employed. In addition, the average earnings of upper caste members are more than thirty percent higher than

those of members of the other caste classes in both wage activities.

On the other hand, there are also large differences in the average schooling years between castes. As can be seen in Table 1, there is a negative correlation between the average years of education and the proportion of casual workers. Given that most casual workers are engaged in agricultural wage work, this may reflect the lack of response of agricultural wages to human capital (Kurosaki and Khan, 2006) and the stigma associated with working as an agricultural laborer in rural India. In the regression analysis, it is examined whether the differences in employment status and wage differentials between castes are mainly attributable to human capital characteristics or other factors, namely caste-based discrimination.

### **3.2 Empirical variables**

The empirical variables used in the regression analysis are summarized in Tables 2 and 3. The first four rows provide information on market earnings (*Log earnings-regular*, *Log earnings-casual*) and employment status (*Regular emp. dummy* and *Casual emp. dummy*). Out of the 2,000 individuals in the sample who did not participate in the labor market, 1,105 (55 percent) were engaged in agricultural self-employed work, 694 (35 percent) in non-agricultural self-employed work, and 201 (10 percent) in domestic work or were not working. Thus, strictly speaking, the

“reservation wage” to be estimated using this dataset mainly reflects the marginal productivity in self-employment activities.<sup>10</sup>

[Table 2]

[Table 3]

Human capital characteristics affecting market earnings, transaction costs and reservation wages include educational attainment, job experience, and age. Actual job experience is not available from the dataset and is therefore measured using the following formula: age minus years of education minus 6 years (*experience*).<sup>11</sup> Age, of course, reflects not only job experience but also human capital accumulation over an individual’s lifetime; however, to concentrate on the effect of human capital accumulation after formal schooling, the variable *experience* is employed in the earnings equation. The effects of education and experience on market earnings are expected to be positive but are likely to depend on the employment status. The impact

---

<sup>10</sup> In this connection, because information on incomes from self-employment activities is not available, the production functions for those activities cannot be estimated directly.

<sup>11</sup> There are 11 cases (less than 0.4 percent of the sample) in which *experience* becomes a negative value (the minimum is -3), and these are replaced by zero.

of education on the cost of entry is expected to be negative, while that of age is expected to be positive. The effect on the reservation wage is expected to be positive both in the case of education and age.

To control for the different roles of household members, other individual-level characteristics are taken into account by employing dummy variables for household heads and firstborn sons (*HH head dummy* and *Firstborn son dummy*) as well as the number of elder brothers (*No. of elder brothers*). Because these variables are likely to have no effects on market earnings and transaction costs, they are included only in the reservation wage function. Being the household head or the firstborn son of the household head may magnify the economic responsibilities of the person in the household, while having many elder brothers is likely to have the opposite effect. Thus, the expected signs are positive for *HH head dummy* and *Firstborn son dummy*, but negative for *No. elder brothers*.

Household level characteristics are family structure, farming assets, and caste membership. Variables for household structure are the number of working age and non-working age members (*No. of working age members* and *No. of non-working age members*). The number of working age (non-working age) members mainly captures the working (dependent) population in the household and the expected impact on

reservation wages is negative (positive). On the other hand, the impact of these variables on market earnings and transaction costs is somewhat ambiguous. If household members' nutritional status is negatively correlated with household size, the effects of these variables on market earnings are expected to be negative and those on transaction costs are positive.

As farming assets, the size of farmland owned by households (*Land size*), the share of farmland that is irrigated (*Irrigation ratio*), and the value of semi-fixed capital in agricultural production (*Agr. capital*) and livestock (*Livestock*) are employed.<sup>12</sup> These variables mainly capture household members' productivity in own-farm activities and the impact of these variables on reservation wages are therefore expected to be positive. Turning to the impact on market earnings and transaction costs, since household members' nutritional status is likely to be correlated with farming assets, they are likely to have a positive effect on market earnings and a negative one on transaction costs.

Caste dummies (with the upper caste as the reference group) capture "unexplained" differences between castes in market earnings, transaction costs and

---

<sup>12</sup> Given that the "reservation wage" mainly captures the marginal productivity in agricultural/non-agricultural self-employed activities, variables affecting the productivity in non-farm enterprise production should be added. Unfortunately, data that would allow us to control for productivity in these activities are not available.



reservation wages. Although the inclusion of caste dummies in the reservation wage function may appear dubious, it seems in fact only natural to assume that workers belonging to a caste that has traditionally suffered from discrimination in the labor market have a negative attitude or low expectations with regard to finding a job. This being the case, the reservation wage of those belonging to an economically disadvantaged class is low in comparison with those belonging to an advantaged class. Table 1 suggests that the impact of belonging to one of the backward classes is negative on market earnings and positive on transaction costs. Therefore, a negative effect is expected on reservation wages.

As regional characteristics, the ratio of the landless within the village (*Ratio of landless*), the average distance to the nearest bank, police station and secondary school (*Distance to facilities*), and a *UP state dummy* are employed. The *Ratio of landless* is a proxy for the number of potential workers, and its impact on market earnings is expected to be negative, although that on transaction costs and reservation wages is ambiguous. The *Distance to facilities* captures proximity to the nearest city and hence the degree of economic development of the village. It is expected that in villages far from an urban area, market earnings and reservation wages are low and the costs of finding regular employment are high. In addition to these variables, the ratio of workers

in regular employment, a proxy for information exchange about jobs among villagers, is used only in the transaction cost function and is likely to lower transaction costs.

#### **4. Estimation Results**

Before embarking on the empirical investigation, some limitations of the analysis should be mentioned. First, several household-level variables to control for family structure and farming assets are excluded from the market earnings and transaction cost equations because of a convergence problem. Therefore, there is a possibility that the estimates suffer from omitted variables bias. For instance, in casual employment activities, workers' nutritional status, which may be captured by these variables, is an important determinant of market wages in the Indian labor market (Deolalikar, 1988; Weinberger, 2003). In addition, it is possible that the estimated effect of human capital on market wages are biased (Heckman and Hotz, 1986; Kingdon, 1998). Unfortunately, the model that includes these variables in all functions failed to achieve a convergence of the likelihood function, but the effects of these variables on market wages and transaction costs are likely to be negligible.

A second limitation is that there is also a possibility of bias due to other omitted variables. If individuals' schooling choices are determined by their ability (or the

possibility of getting a job), this may cause schooling effects to be biased (i.e., there would be “ability bias”) and the exclusion of students from the sample would make the problem worse. To examine this possibility, estimations using limited samples – including only those aged between 18 and 60, and including only those aged between 20 and 60 – were conducted. The coefficients were slightly different from the estimates using the full sample, but there were no systematic differences.<sup>13</sup> This suggests that there is no “ability bias” or, if there is, it is rather small.

Table 4 shows the estimation result of the log-likelihood function (4). As can be seen, the coefficient on most of the explanatory variables takes the expected sign. The null hypothesis of no transaction costs is tested by using a likelihood ratio test. The  $\chi^2$  statistic of 189.75 (p-value is 0.00) indicates rejection of the hypothesis.

#### **4.1 Analysis of the market wage functions**

Estimation results for the market wage functions are presented in the first and second columns of Table 4. The effect of human capital on the market wage differs sharply depending on the employment status. While the return to education is positive in the case of regular employment, it shows an inverted U-shape in the case of casual employment, indicating that marginal returns to education become negative at more

---

<sup>13</sup> These results are available on request.

than eight years of education. The impact of job experience peaks at 44 years for regular employment and at 20 years for casual employment. The fact that human capital does not contribute much toward improving productivity in casual employment activities may reflect the fact that the majority of casual workers are hired for unskilled, manual labor, such as agricultural work.

Looking at caste membership, the coefficients on caste dummies are negative in all cases in the casual labor wage functions and negative in most cases in the regular labor wage function. In casual employment activities, membership of an agriculture-based backward caste or SC/ST significantly decreases wages.<sup>14</sup> This result is consistent with the findings from Table 1. The monthly earnings from casual employment activities for the agriculture-based backward castes are 33.5 percent ( $\approx e^{0.289} - 1$ , 68 percent of the total wage differentials) lower than those for the upper castes, and those for the SCs/STs are 37.4 percent (67 percent) lower.

However, one should guard against any hasty interpretation of this result as evidence of wage discrimination in the casual labor market. In our analysis, workers' occupations are simply classified into two types (casual and regular employment), hence the diversity of occupations within the same status is ignored. An alternative

---

<sup>14</sup> The result also shows that membership of a middle caste has a significant negative effect on wages from regular employment activities. But this result should be treated with caution because the sample contains only eight middle caste workers engaged in this type of job.

interpretation, therefore, is that the socially backward classes are traditionally occupied in low-paying casual employment activities, so that the results possibly reflect occupational segregation rather than wage discrimination. Unfortunately, because of data limitations, it is impossible to investigate this issue further by disaggregating the sample by occupation or industry.

[Table 4]

#### **4.2 Analysis of the transaction cost and reservation wage functions**

Estimation results of the transaction cost and reservation wage functions are presented in the third and fourth columns of Table 4. Schooling years have a negative effect on transaction costs and a positive effect on reservation wages when evaluated at the sample mean (4.84 years), but the effect of *Schooling years* on transaction costs and the effect of *Schooling years squared* in both estimates is insignificant. Age has a significant positive effect on transaction costs but no significant effect on reservation wages.

Turning to caste dummies, individuals belonging to groups characterized as backward face significantly higher transaction costs than those belonging to the upper

castes.<sup>15</sup> The results indicate that membership of one of the four backward classes increases transaction costs by 39 percent (*backward-agr.*) to 79 percent (*SC/ST*), which may suggest discrimination in labor market entry.

However, the results obtained here should be interpreted with care. For instance, it is possible that caste membership captures the effects of caste networks (Munshi and Rosenzweig, 2006). If members of the upper castes are traditionally employed in regular employment activities and hiring through referrals is prevalent, then the negative effect of backward-caste membership may simply reflect the lack of such network referrals. There is another possible explanation as well. As already mentioned, workers belonging to groups that have traditionally experienced discrimination may hold low expectations of gaining a job. In this case, even in the absence of discrimination by employers based on their (or other employees') tastes or beliefs, these workers may not look as hard for a job as members of other castes. The higher transaction costs for members of socially backward castes may result from such self-fulfilling beliefs. However, even in these circumstances, there is no doubt that government policies to combat inequality in employment opportunities have not been successful in the study region. Our results

---

<sup>15</sup> To investigate the robustness of the impact of caste membership on transaction costs, an alternative specification is tested in which caste dummies are excluded from the market earnings and reservation wage equations. This results shows that the coefficients of dummies for the lower castes (except for the *Muslim backward dummy*) are positive and significant in the transaction cost equation. The results are available from the author on request.

show that members of socially backward castes indeed face greater difficulties in finding regular employment than members of the upper castes.

In the reservation wage function, on the other hand, all dummies have negative coefficients. This is consistent with the story mentioned in the previous section. However, this result may simply reflect the low marginal productivities in the self-employment activities that members of the socially backward castes are engaged in because they face discrimination even in these activities.

The effects of the other control variables are as expected. Among regional characteristics, the ratio of salaried workers in the same village significantly decreases the cost of entry. This indicates the importance of social networks in finding regular employment.

#### **4.4 Effects of human capital and caste membership on participation**

Thus far, the discussion has concentrated on the impact of the explanatory on the dependent variables. This subsection focuses on the impact of the explanatory variables on the probability of gaining regular employment. The marginal effects of human capital and caste membership on regular labor market participation, holding all other variables constant at their sample mean, are shown in the first column of Table 5. Furthermore, the effects are decomposed into three parts: those through market earnings

(in the second column), those through transaction costs (in the third column), and those through reservation wages (in the fourth column).

[Table 5]

The table indicates that an additional year of education increases the probability of being in regular employment by approximately 1.0 percentage points. The breakdown of the effect of schooling on the probability of being in regular employment into its components shows that it is largely through the effect on market earnings.

The effect of caste membership through transaction costs lowers the likelihood of members of backward castes to be in regular employment by 7.1 percentage points (70 percent of the total difference in the employment rate) in the case of SCs/STs, by 5.5 percentage points (57 percent) in the case of other backward castes, by 5.1 percentage points (60 percent) in the case of agriculture-based backward castes, and by 4.3 percentage points (57 percent) in the case of agriculture-based backward castes. However, these negative effects through transaction costs are to a great extent attenuated through earnings and reservation wages. This result implies that an analysis of the role of discrimination in labor market entry based on a reduced-form approach is



likely to underestimate the impact of discrimination. In fact, this is confirmed by the probit estimation and the estimation using the Blinder-Oaxaca decomposition procedure (Blinder, 1973; Oaxaca, 1973): the impacts of “job discrimination” estimated using the probit model (in the fifth column) and the Blinder-Oaxaca method (in the sixth column) are considerably smaller than that using the structural model (in the third column).<sup>16</sup>

The comparison of the predicted mean probability and of the percent correctly predicted between our model and the probit model shows that our model predicts the labor participation decision equally as well as the probit model. Thus, when one employs a reduced form model, the possibility of this underestimation should be recognized.

## **5. Conclusion**

This paper examined caste discrimination in the labor market by estimating simultaneously market earnings, reservation wages, and the costs of finding regular employment. The estimation results suggest that socially backward castes do face disadvantages in finding regular employment in the sense that they face higher transaction costs associated with entry into the labor market. On the other hand, there is no evidence for wage discrimination in regular employment activities. Thus, these

---

<sup>16</sup> However, in most cases, there are no statistically-significant differences between the estimates from the structural estimation and reduced-form estimations.

results suggested that caste-based discrimination takes the form of “job discrimination,” which limits the range of available jobs, rather than “wage discrimination,”<sup>17</sup> However, this is at odds with Banerjee and Knight’s (1985) findings, which indicated that “wage discrimination” explained a large part of the “unexplained” component of wage differentials. Although one should be careful about directly comparing their results with ours because the data they used was of 1,115 migrant laborers in Delhi in 1975-76, it should at least be noted that it is possibility that their analysis underestimated the role of “job discrimination.” This paper has shown that the reservation wages of the socially backward castes are lower than those of the upper castes and consequently estimates of “job discrimination” in a reduced-form approach may capture not only discrimination in entry but also such self-selection effects.

Furthermore, it was found that the large differences in educational attainment between castes represent one important cause of inequality in employment status. Thus, as India’s reservation policy has aimed to do, promoting opportunities in education and employment for socially backward castes can be an instrument in eliminating inequality in economic conditions between castes. At the same time, however, the results obtained here suggest that the reservation policy so far has had little effect on rural dwellers in

---

<sup>17</sup> Similarly, studies on racial discrimination in the U. S. also found that discrimination typically takes the form of market segregation rather than price differentiation (see, e.g., Arrow, 1998).

the study region, Uttar Pradesh and Bihar, although it is not clear whether this is the results of some fundamental problems with this policy or reflects the specific circumstances of these two states, which are among the least developed in India. While it is beyond the scope of this paper to provide a detailed assessment of the policy, the fact that backward castes continue to be disadvantaged half a century after it was introduced suggests that its achievements have at best been limited.

Finally, several limitations of this paper should be mentioned. One possible problem is associated with the validity of the assumption introduced in the estimation model that transaction costs in the casual labor market are constant. As mentioned earlier, this assumption seems plausible in the context of rural India, and the comparison of results between our model and the probit model suggest that our model is valid. However, it is impossible to test the sensitivity of the model to the assumption since it is an essential assumption for parameter identification. But the most important limitation of this paper probably is that there is a possibility that estimated caste differences in market wages and participation probabilities capture other omitted differences between castes rather than discrimination. Although this appears to be a common problem in studies on labor market discrimination, the occupational classification in this paper (i.e., casual or regular employment) may exacerbate the problem. Essentially, wage rates and

individual responses may vary not simply by employment status but by sector or type of occupation, but such occupational diversity within the same employment status was ignored in our analysis. Consequently, estimated caste differences may to some extent reflect the effects of occupational segregation. Given that India's caste system is based on hereditary occupations, future analyses should try to employ data disaggregated by occupation or industry.

## Appendix 1: Probit estimation of the probability of being in regular employment

[Table A-1]

## Appendix 2: Estimation of “job discrimination” à la Banerjee and Knight (1985)

The estimation procedure explained below is an application of the Blinder-Oaxaca decomposition method for the trichotomous occupational choice model (multinomial logit model). The probability that an individual  $i$  belonging to the  $g$ th social group ( $g = R$ : reference group,  $C$ : comparison group) will be a regular employee is given by

$$(A-1) \quad P_{i,r}^g = \frac{\exp X_i^g \beta_r^g}{\exp X_i^g \beta_r^g + \exp X_i^g \beta_c^g + \exp X_i^g \beta_n^g} ,$$

where  $X$  is a vector of variables affecting the labor participation decision and  $\beta_k$  is a vector of coefficients corresponding to the  $k$ th type of employment activity ( $r$ : regular employment,  $c$ : casual employment,  $n$ : not employed). Let  $\tilde{P}_{i,r}^C$  be the probability

that an individual belonging to a comparison group ( $C$ ) will be a regular employee if he belongs to a reference group ( $R$ ) and take the average for each probability,

$$P_r^g = \sum_{i=1}^{N^g} P_{i,r}^g / N^g , (g = R, C.) \text{ and } \tilde{P}_r^C = \sum_{i=1}^{N^C} \tilde{P}_{i,r}^C / N^C .$$

The difference between the two groups in the probability of being regularly employed is decomposed as follows:

$$(A-2) \quad P_r^C - P_r^R = (P_r^C - \tilde{P}_r^C) + (\tilde{P}_r^C - P_r^R) .$$

The first component in this equation measures the “unexplained” difference due to differences in coefficients, while the second measures the “explained” difference due to differences in characteristics. Table A-2 shows the estimates for each component in equation (A-2).

[Table A-2]

## References

- Arrow, Kenneth J. 1988. "What Has Economics to Say about Racial Discrimination?" *Journal of Economic Perspectives*, 12: 91-100.
- Banerjee, Biswajit, and Knight, J. B. 1985. "Caste Discrimination in the Indian Labor Market." *Journal of Development Economics*, 17: 277-307.
- Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources*, 8: 436-455.
- Borooh, Vani K. 2005. "Caste, Inequality, and Poverty in India." *Review of Development Economics*, 9: 399-414.
- Borooh, Vani K., Dubey, Amaresh, and Iyer, Sriya. 2007. "The Effectiveness of Jobs Reservation: Caste, Religion and Economic Status in India." *Development and Change*, 38 (3): 423-445.
- Cogan, John F. 1981. "Fixed Costs and Labor Supply." *Econometrica*, 49: 945-963.
- Deolalikar, Anil B. 1988. "Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India." *Review of Economics and Statistics*, 70: 406-413.
- Heckman, James J., and Hotz, V. Joseph. 1986. "An Investigation of the Labor Market Earnings of Panamanian Males Evaluating the Sources of Inequality." *Journal of Human Resources*, 21: 507-542.
- Ito, Takahiro, and Kurosaki, Takashi. 2007. "Weather Risk, Wages in Kind, and the Off-Farm Labor Supply of Agricultural Households in a Developing Country." *Hi-Stat Discussion Paper Series*, No. 226. Institute of Economic Research, Hitotsubashi University, Tokyo.
- Jacoby, Hanan G. 1993. "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra." *Review of Economic Studies*, 60: 903-921.

Key, Nigel; Sadoulet, Elisabeth; and de Janvry, Alain. 2000. "Transactions Costs and Agricultural Household Supply Response." *American Journal of Agricultural Economics*, 82: 245-259.

Kijima, Yoko. 2006. "Caste and Tribe Inequality: Evidence from India, 1983-1999." *Economic Development and Cultural Change*, 54 (2): 369-404.

Kingdon, Geeta Gandhi. 1998. "Does the Labour Market Explain Lower Female Schooling in India?" *Journal of Development Studies*, 35: 39-65.

Kingdon, Geeta Gandhi, and Unni, Jeemol. 2001. "Education and Women's Labour Market Outcomes in India." *Education Economics*, 9: 173-195.

Kochar, Anjini. 1999. "Smoothing Consumption by Smoothing Income: Hours-of-Work Responses to Idiosyncratic Agricultural Shocks in Rural India." *Review of Economic and Statistics*, 81: 50-61.

Kurosaki, Takashi, and Khan, Humayun. 2006. "Human Capital, Productivity, and Stratification in Rural Pakistan." *Review of Development Economics*, 10: 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.

Madheswaran, S., and Attewell, Paul. 2007. "Caste Discrimination in the Indian Urban Labour Market: Evidence from the National Sample Survey." *Economic and Political Weekly*, 42 (41): 4146-4153.

Munshi, Kaivan, and Rosenzweig, Mark. 2006. "Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy." *American Economic Review*, 96: 1225-1252.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, 14: 693-709.

Osborne, Evan. 2001. "Culture, Development, and Government: Reservations in India."



*Economic Development and Cultural Change*, 49 (3): 659-685.

Rose, Elaina. 2001. "Ex Ante and Ex Post Labor Supply Response to Risk in a Low-Income Area." *Journal of Development Economics*, 64: 371-388.

Sadoulet, Elisabeth; de Janvry, Alain; and Benjamin, Catherine. 1998. "Household Behavior with Imperfect Labor Markets." *Industrial Relations*, 37: 85-108.

Skoufias, Emmanuel. 1994. "Using Shadow Wages to Estimate Labor Supply of Agricultural Households." *American Journal of Agricultural Economics*, 76: 215-227.

Srinivasan, K., and Kumar, Sanjay. 1999. "Economic and Caste Criteria in Definition of Backwardness." *Economic and Political Weekly*, 34 (42 & 43): 3052-3057.

Thorat, Sukhadeo. 2002. "Oppression and Denial: Dalit Discrimination in the 1990s." *Economic and Political Weekly*, 37 (6): 572-578.

Thorat, Sukhadeo, and Attewell, Paul. 2007. "The Legacy of Social Exclusion: A Correspondence Study of Job Discrimination in India." *Economic and Political Weekly*, 42 (41): 4141-4145.

Weinberger, Katinka. 2003. "The Impact of Micronutrients on Labor productivity: evidence from Rural India." Contributed paper prepared at the International Association of Agricultural Economists Conference, Durban, South Africa, August 16-22.

**Table 1: Market labor supply of adult males by caste**

| Caste           | NOB  | Employment rate |         |        | Average monthly earnings |         |        | Schooling years |
|-----------------|------|-----------------|---------|--------|--------------------------|---------|--------|-----------------|
|                 |      | All             | Regular | Casual | All                      | Regular | Casual |                 |
| ALL             | 3324 | 0.40            | 0.12    | 0.28   | 1293.0                   | 2564.1  | 768.1  | 4.84            |
| Upper           | 506  | 0.25            | 0.19    | 0.07   | 2804.4                   | 3296.3  | 1373.4 | 9.25            |
| Middle          | 82   | 0.20            | 0.10    | 0.10   | 1319.4                   | 1745.8  | 893.1  | 6.73            |
| Backward-agr.   | 963  | 0.33            | 0.11    | 0.21   | 1403.3                   | 2398.3  | 869.4  | 5.12            |
| Backward-other  | 645  | 0.38            | 0.09    | 0.28   | 1233.0                   | 2359.8  | 863.6  | 4.09            |
| SC/ST           | 818  | 0.59            | 0.09    | 0.51   | 847.4                    | 2267.8  | 600.3  | 2.79            |
| Muslim upper    | 109  | 0.43            | 0.18    | 0.25   | 1513.2                   | 2217.3  | 991.7  | 4.44            |
| Muslim backward | 201  | 0.44            | 0.10    | 0.33   | 1187.3                   | 2326.7  | 830.2  | 2.52            |

Note: (1) Employment status is based on workers' primary job classified based on their annual working days. Note that approximately 20 percent of workers were also engaged in secondary activities.

**Table 2: Definition of variables**

| Name                                 | Definition  | Equation | Expected sign |             |                          |
|--------------------------------------|---|----------|---------------|-------------|--------------------------|
|                                      |   |          | (1)<br>ln W   | (2)<br>ln K | (3)<br>ln W <sup>R</sup> |
| <b>A. Individual level variables</b> |   |          |               |             |                          |
| Log earnings-regular                 | Log of monthly earnings paid to a regular worker  |          |               |             |                          |
| Log earnings-casual                  | Log of average monthly earnings paid to a casual laborer  |          |               |             |                          |
| Regular emp. dummy                   | Dummy variable for those mainly in regular employment   |          |               |             |                          |
| Casual emp. dummy                    | Dummy variable for those mainly in casual employment  |          |               |             |                          |
| Schooling years                      | Standardized years of completed education corresponding to the standard education system in India                           |          | +             | -           | +                        |
| Experience                           | Age minus standardized years of completed education minus 6 years   |          | +             | No          | No                       |
| Age                                  | Age of the person   |          | No            | +           | +                        |
| HH head dummy                        | Dummy variable for the household head   |          | No            | No          | -                        |
| Firstborn son dummy                  | Dummy variable for the firstborn son of the household head  |          | No            | No          | -                        |
| No. of elder brothers                | Number of elder brothers (incl. brothers-in-law)  |          | No            | No          | +                        |
| <b>B. Household level variables</b>  |   |          |               |             |                          |
| No. of working age members           | Number of household members aged between 15 to 60   |          | (-)           | (+)         | +                        |
| No. of non-working age members       | Number of household members other than working age members  |          | (-)           | (+)         | -                        |
| Land size (10acres)                  | Land owned by the household   |          | (+)           | (-)         | +                        |
| Irrigation ratio                     | Ratio of irrigated land to total land size owned by the household   |          | (+)           | (-)         | +                        |
| Agr. capital (Rupees)                | Value of fixed agricultural capital owned by the household  |          | (+)           | (-)         | +                        |
| Livestock (Rupees)                   | Value of livestock owned by the household   |          | (+)           | (-)         | +                        |
| Middle                               | Dummy variable for middle castes  |          | -             | +           | -                        |
| Backward-agr.                        | Dummy variable for agriculture-based backward castes  |          | -             | +           | -                        |
| Backward-other                       | Dummy variable for other backward castes  |          | -             | +           | -                        |
| SC/ST                                | Dummy variable for scheduled castes and scheduled tribes  |          | -             | +           | -                        |
| Muslim upper                         | Dummy variable for Muslim upper castes  |          | -             | +-          | -                        |
| Muslim backward                      | Dummy variable for Muslim backward castes   |          | -             | +           | -                        |
| <b>C. Other</b>                      |   |          |               |             |                          |
| Ratio of landless                    | Ratio of the landless in the village  |          | -             | +-          | +-                       |
| Distance to facilities (Km)          | Average distance to the nearest bank, police station, and secondary school from the village                                 |          | -             | +           | -                        |
| Ratio of regular workers (%)         | Ratio of village-total workers in regular employment (other than the household members) to village-total working age people |          | No            | -           | No                       |
| UP state dummy                       | Dummy variable for Uttar Pradesh  |          | +-            | +-          | +-                       |

**Table 3: Summary statistics of variables**

| <b>Name</b>                          | <b>NOB</b> | <b>Mean</b> | <b>Std. Dev.</b> | <b>Min.</b> | <b>Max.</b> |
|--------------------------------------|------------|-------------|------------------|-------------|-------------|
| <b>A. Individual level variables</b> |            |             |                  |             |             |
| Log earnings-regular                 | 387        | 7.60        | 0.74             | 5.30        | 9.62        |
| (in Rupees)                          | 387        | 2564.1      | 1863.3           | 200         | 15000       |
| Log earnings-casual                  | 937        | 6.36        | 0.73             | 3.65        | 8.72        |
| (in Rupees)                          | 937        | 768.1       | 678.2            | 38.3        | 6111.1      |
| Regular emp. dummy                   | 3324       | 0.12        |                  |             |             |
| Casual emp. dummy                    | 3324       | 0.28        |                  |             |             |
| Schooling years                      | 3324       | 4.84        | 5.20             | 0           | 20          |
| Experience                           | 3324       | 23.49       | 14.20            | 0           | 54          |
| Age                                  | 3324       | 34.32       | 12.57            | 15          | 60          |
| HH head dummy                        | 3324       | 0.51        |                  |             |             |
| Firstborn son dummy                  | 3324       | 0.27        |                  |             |             |
| No. of elder brothers                | 3324       | 0.39        | 1.03             | 0           | 10          |
| <b>B. Household level variables</b>  |            |             |                  |             |             |
| No. of working age members           | 3324       | 4.33        | 2.23             | 1           | 13          |
| No. of non-working age members       | 3324       | 3.22        | 2.43             | 0           | 17          |
| Land size (10acres)                  | 3324       | 0.26        | 0.52             | 0           | 9.3         |
| Irrigation ratio                     | 3324       | 0.65        | 0.43             | 0           | 1           |
| Agr. capital (Rupees)                | 3324       | 7134.48     | 30197.36         | 0           | 373600      |
| Livestock (Rupees)                   | 3324       | 7443.31     | 10692.68         | 0           | 150000      |
| Middle                               | 3324       | 0.02        |                  |             |             |
| Backward-agr.                        | 3324       | 0.29        |                  |             |             |
| Backward-other                       | 3324       | 0.19        |                  |             |             |
| SC/ST                                | 3324       | 0.25        |                  |             |             |
| Muslim upper                         | 3324       | 0.03        |                  |             |             |
| Muslim backward                      | 3324       | 0.06        |                  |             |             |
| <b>C. Other</b>                      |            |             |                  |             |             |
| Ratio of landless                    | 3324       | 0.39        | 0.21             | 0           | 0.99        |
| Distance to facilities (Km)          | 3324       | 5.77        | 3.58             | 0.5         | 20          |
| Ratio of regular workers (%)         | 3324       | 5.45        | 4.27             | 0           | 26.23       |
| UP state dummy                       | 3324       | 0.55        |                  |             |             |

**Table 4: Estimation results for market earnings, transaction costs and reservation wages**

|                                   | Monthly earnings<br>(Regular) |         | Monthly earnings<br>(Casual) |         | Transaction<br>costs |         | Reservation<br>wages |         |
|-----------------------------------|-------------------------------|---------|------------------------------|---------|----------------------|---------|----------------------|---------|
| <b>Individual characteristics</b> |                               |         |                              |         |                      |         |                      |         |
| Schooling years                   | 0.062                         | (3.38)‡ | 0.060                        | (2.63)‡ | -0.044               | (1.45)  | 0.059                | (2.80)‡ |
| Schooling years squared/100       | 0.114                         | (1.04)  | -0.383                       | (1.76)* | 0.349                | (1.48)  | -0.301               | (1.54)  |
| Age (Experience)                  | 0.072                         | (5.52)‡ | 0.014                        | (0.85)  | 0.053                | (5.15)‡ | 0.003                | (0.21)  |
| Age (Experience) squared/100      | -0.081                        | (3.10)‡ | -0.035                       | (1.10)  | -0.031               | (2.44)† | -0.005               | (0.31)  |
| HH head dummy                     | –                             |         | –                            |         | –                    |         | -0.020               | (0.53)  |
| Firstborn son dummy               | –                             |         | –                            |         | –                    |         | -0.051               | (1.16)  |
| No. of elder brothers             | –                             |         | –                            |         | –                    |         | 0.000                | (0.02)  |
| <b>Household characteristics</b>  |                               |         |                              |         |                      |         |                      |         |
| No. of working age members        | –                             |         | –                            |         | –                    |         | 0.008                | (0.70)  |
| No. of non-working age members    | –                             |         | –                            |         | –                    |         | -0.004               | (0.81)  |
| Land size                         | –                             |         | –                            |         | –                    |         | 0.054                | (0.79)  |
| Irrigation ratio                  | –                             |         | –                            |         | –                    |         | 0.147                | (1.66)* |
| Agr. capital                      | –                             |         | –                            |         | –                    |         | 0.032                | (2.13)† |
| Livestock                         | –                             |         | –                            |         | –                    |         | 0.045                | (1.17)  |
| <b>Caste dummies</b>              |                               |         |                              |         |                      |         |                      |         |
| Middle                            | -0.302                        | (2.04)† | -0.812                       | (1.44)  | 0.611                | (1.05)  | -0.705               | (1.29)  |
| Backward-agr.                     | -0.026                        | (0.25)  | -0.289                       | (1.77)* | 0.326                | (1.74)* | -0.403               | (2.54)† |
| Backward-other                    | -0.026                        | (0.22)  | -0.263                       | (1.49)  | 0.357                | (1.69)* | -0.363               | (2.17)† |
| SC/ST                             | 0.000                         | (0.00)  | -0.469                       | (2.17)† | 0.583                | (2.47)† | -0.759               | (3.82)‡ |
| Muslim upper                      | -0.124                        | (0.73)  | -0.124                       | (0.48)  | 0.007                | (0.02)  | -0.272               | (1.19)  |
| Muslim backward                   | 0.143                         | (0.83)  | -0.276                       | (1.38)  | 0.440                | (1.69)* | -0.388               | (2.08)† |
| <b>Regional characteristics</b>   |                               |         |                              |         |                      |         |                      |         |
| Ratio of landless                 | -0.374                        | (2.00)† | -0.119                       | (0.98)  | -0.175               | (0.78)  | -0.214               | (1.63)  |
| Distance to facilities/10         | -0.066                        | (0.53)  | -0.007                       | (0.10)  | 0.007                | (0.05)  | 0.004                | (0.05)  |
| Ratio of regular workers          | –                             |         | –                            |         | -0.020               | (2.08)† | –                    |         |
| UP state dummy                    | 0.044                         | (0.61)  | 0.115                        | (2.05)† | -0.046               | (0.51)  | 0.045                | (0.67)  |
| <b>Intercept</b>                  | 5.588                         | (19.4)‡ | 6.350                        | (14.2)‡ | -0.543               | (1.30)  | 6.562                | (14.4)‡ |
| <b>Standard error</b>             | 0.687                         | (14.5)‡ | 0.720                        | (12.0)‡ | 0.169                | (1.58)  | 0.678                | (13.0)‡ |

Note: (1) Numbers in parentheses are z-values based on clustering robust standard errors using households as clusters.

(2) \* Significant at 10%; † significant at 5%; ‡ significant at 1%.

(3) *Experience* is included in the market earnings functions (the first and second columns), instead of *Age*.

(4) NOB = 3324; log-likelihood = -3887.19.  $H_0$ : zero slope, LR  $\chi^2(62) = 1223.70$ ;  $H_0$ : no transaction costs, LR  $\chi^2(17) = 189.75$ .

**Table 5: Marginal effects on participation in the regular labor market**

|                      | Total<br>(a)+(b)+(c) | Market<br>earnings (a) | Transaction<br>cost (b) | Reservation<br>wages (c) | Probit<br>estimates | Blinder-Oaxaca<br>decomposition |
|----------------------|----------------------|------------------------|-------------------------|--------------------------|---------------------|---------------------------------|
|                      |                      | (Table 4)              |                         |                          | (Table A-1)         | (Table A-2)                     |
| <b>Human capital</b> |                      |                        |                         |                          |                     |                                 |
| Schooling years      | 0.0100<br>(0.0019)‡  | 0.0094<br>(0.0068)     | 0.0044<br>(0.0061)      | -0.0038<br>(0.0024)      | 0.0096<br>(0.0014)‡ | —                               |
| Age                  | 0.0019<br>(0.0005)‡  | 0.0152<br>(0.0082)*    | -0.0135<br>(0.0080)*    | 0.0001<br>(0.0004)       | 0.0026<br>(0.0007)‡ | —                               |
| <b>Caste dummies</b> |                      |                        |                         |                          |                     |                                 |
| Middle               | -0.0485<br>(0.0455)  | -0.0036<br>(0.0223)    | -0.0597<br>(0.0320)*    | 0.0148<br>(0.0130)       | -0.0423<br>(0.0271) | -0.0544<br>(0.1187)             |
| Backward-agr.        | -0.0126<br>(0.0213)  | 0.0174<br>(0.0166)     | -0.0428<br>(0.0206)†    | 0.0128<br>(0.0089)       | -0.0062<br>(0.0162) | -0.0186<br>(0.0096)*            |
| Backward-other       | -0.0375<br>(0.0246)  | 0.0095<br>(0.0152)     | -0.0547<br>(0.0249)†    | 0.0077<br>(0.0074)       | -0.0279<br>(0.0175) | -0.0323<br>(0.0142)†            |
| SC/ST                | -0.0206<br>(0.0254)  | 0.0306<br>(0.0240)     | -0.0709<br>(0.0329)†    | 0.0197<br>(0.0142)       | -0.0138<br>(0.0184) | -0.0217<br>(0.0182)             |
| Muslim upper         | 0.0146<br>(0.0296)   | 0.0010<br>(0.0182)     | 0.0027<br>(0.0122)      | 0.0110<br>(0.0090)       | 0.0430<br>(0.0361)  | 0.0541<br>(0.1221)              |
| Muslim backward      | 0.0042<br>(0.0310)   | 0.0388<br>(0.0262)     | -0.0509<br>(0.0283)*    | 0.0163<br>(0.0101)       | 0.0243<br>(0.0295)  | 0.0136<br>(0.0349)              |

Note: (1) Calculated figures in the table are the average marginal effect of each variable on the probability of being in regular employment. Note that in the case of binary variables, calculated figures are the average change in response probabilities when the variable changes from 0 to 1, and in general the whole effect does not equal the sum of (a), (b) and (c) because of the calculation method. For ease of comparison, however, it is adjusted in such a way that the sum equals the whole effect.

(2) Numbers in parentheses are standard errors calculated using the delta method.

(3) The marginal effects through “Market earnings” are the sum of marginal effects through monthly earnings from regular and from casual employment activities.

(4) The dependent variable for the probit estimation presented in the fifth column is an indicator variable that takes one if a worker is engaged in a regular employment activity and zero otherwise (see Table A-1). While the predicted mean probability with the probit model (Table A-1) is 0.1164, that with our procedure (Table 4) is 0.1162 (the sample mean is 0.1164). In addition, the percentage of observations being correctly predicted with regard to the employment status (using the 50 percent rule) is 0.885 for the probit model and 0.884 for our model. Thus, our model appears to predict the labor participation decision equally as well as the probit model.

(5) In the calculation of the effects of schooling years and age, those through job experience are taken into account using the following relation between them: Experience = age – schooling years – 6. With regards to schooling years, the marginal effects are calculated figures when evaluated at its sample mean (4.84 years).

**Table A-1: Estimation result of participation in the regular labor market  
(Probit model)**

| <b>Dependent variable:</b>        |               |                |
|-----------------------------------|---------------|----------------|
| <b>Regular emp. dummy</b>         | <b>Coef.</b>  | <b>z-value</b> |
| <b>Individual characteristics</b> |               |                |
| Schooling years                   | 0.050         | (2.97)‡        |
| Schooling years squared/100       | 0.056         | (0.50)         |
| Age                               | 0.069         | (3.82)‡        |
| Age squared/100                   | -0.079        | (3.43)‡        |
| HH head dummy                     | -0.103        | (0.80)         |
| Firstborn son dummy               | 0.042         | (0.37)         |
| No. of elder brothers             | -0.020        | (0.44)         |
| <b>Household characteristics</b>  |               |                |
| No. of working-age members        | 0.032         | (1.72)*        |
| No. of non working-age members/10 | -0.015        | (1.03)         |
| Land size                         | -0.089        | (0.93)         |
| Irrigation ratio                  | 0.231         | (2.80)‡        |
| Agr. capital                      | -0.030        | (2.02)†        |
| Livestock                         | -0.117        | (2.92)‡        |
| <b>Caste dummies</b>              |               |                |
| Middle                            | -0.281        | (1.34)         |
| Backward-agr.                     | -0.036        | (0.38)         |
| Backward-other                    | -0.169        | (1.57)         |
| SC/ST                             | -0.081        | (0.75)         |
| Muslim upper                      | 0.221         | (1.31)         |
| Muslim backward                   | 0.131         | (0.87)         |
| <b>Regional characteristics</b>   |               |                |
| Ratio of landless                 | -0.101        | (0.67)         |
| Distance to facilities/10         | -0.025        | (0.25)         |
| Ratio of regular workers          | 0.049         | (7.24)‡        |
| UP state dummy                    | -0.005        | (0.07)         |
| <b>Intercept</b>                  | <b>-3.127</b> | <b>(8.59)‡</b> |

Note: (1) See Table 4, Notes (1) and (2).

(2) NOB=3324; log-likelihood = -1064.56, pseudo  $R^2 = 0.110$ .  $H_0$ : zero slope, LR  $\chi^2(23) = 262.45$ .

**Table A-2: Blinder-Oaxaca decomposition for participation rates in regular employment**

| Comparison group:  | Middle               | Backward-<br>agr.    | Backward-<br>other   | SC/ST                | Muslim<br>upper      | Muslim<br>backward   |
|--|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| Difference in means,<br>$P_r^C - P_r^R$                            | -0.0922              | -0.0755              | -0.0967              | -0.1017              | -0.0062              | -0.0853              |
| Difference due to<br>coefficients,<br>$(P_r^C - \tilde{P}_r^C)$    | -0.0544<br>(0.1187)  | -0.0186<br>(0.0096)† | -0.0323<br>(0.0142)† | -0.0217<br>(0.0182)  | 0.0541<br>(0.1221)   | 0.0136<br>(0.0349)   |
| Difference due to<br>characteristics,<br>$(\tilde{P}_r^C - P_r^R)$ | -0.0378<br>(0.0142)‡ | -0.0569<br>(0.0084)‡ | -0.0644<br>(0.0097)‡ | -0.0800<br>(0.0131)‡ | -0.0603<br>(0.0130)‡ | -0.0988<br>(0.0165)‡ |
| Sample size  | 82                   | 963                  | 645                  | 818                  | 109                  | 201                  |

Note: (1) The reference group is the upper caste with a sample size of 506 individuals.