Education gender gaps in Pakistan: is the labour market to blame

Aslam, Monazza

Education Gender Gaps in Pakistan: Is the Labor Market to Blame?

MONAZZA ASLAM
University of Oxford

I. Introduction

While Pakistan's large and persistent gender gaps in education are well documented, explaining their existence and obstinacy has proven more difficult. This article tests one—the labor market—explanation for gender gaps in education in Pakistan.[1] On the basis of the investment motive, it contends that if the labor market rewards men's schooling more than women's or if it more generally discriminates between the two genders, parents may have an incentive to invest more in boys' education. In this study, I test whether the rewards to females are less than those to males in Pakistan's labor market, that is, whether the return to educating females is lower than that for men. I also ask more generally whether there is wider gender-differentiated treatment in the labor market, that is, whether much or all of the gender gap in earnings is explained by measurable differences in male and female characteristics.

Private economic returns to schooling attainment are estimated using Mincer's semilogarithmic approach in a regression linking individual earnings with additional years (or levels) of schooling completed (Mincer 1974). As is well known, establishing a causal relationship between education and earnings is problematic. Among the issues to contend with are biases due to omitted

I would like to thank Geeta Kingdon for the ideas, discussions, comments, and suggestions that resulted in this article. This article has also substantially benefited from comments by Harold Alderman, John Knight, Francis Teal, Jim Malcomson, Margaret Stevens, Mary Gregory, and the participants of the Labour Economics Seminar, University of Oxford. I would also like to thank Christopher Colclough and Marcel Fafchamps as well as the participants of the United Kingdom Forum for International Education and Training conference (2005) at Oxford for helpful suggestions. Funding from the Department for International Development's Research Consortium on Education Outcomes and Poverty is gratefully acknowledged. Any errors are mine.

[1] Alternative explanations of gender-differentiated parental treatment are that (*a*) there are pure preferences for sons and (*b*) the returns accruing to parents from daughters' education are lower than those accruing from sons' education (maybe because a daughter's in-laws reap the benefits of her education upon marriage) and economic necessity (or parental selfishness) potentially increases the likelihood that boys are sent to school compared to girls. However, Alderman and King (1998) note that it is difficult to distinguish empirically between these various explanations.

 0013-0079/2009/5704-0003$10.00

variables, measurement error in reported schooling, distinguishing between homogeneous and heterogeneous returns to education, and selection into wage employment. Moreover, while human capital theory hypothesizes a concave education-earnings profile and diminishing returns to human capital production, empirical evidence from various countries has challenged the prevailing view (see Behrman and Wolfe [1984] and Alderman and Sahn [1988] and more recently Ashraf and Ashraf [1993a, 1993b], Kingdon [1998], Kingdon and Unni [2001], Belzil and Hansen [2002], Duraisamy [2002], Nasir [2002], and Söderbom et al. [2004]). This finding raises serious policy concerns and warrants further investigation. Among the policy considerations, key is that primary school completion may not be sufficient for poverty reduction since the poor would need to attain higher education levels to allow them to climb out of poverty. Also, while on the one hand high returns to higher education levels may mitigate the need for subsidizing higher education, on the other hand, there still may be need for subsidizing higher education for the poor who face credit constraints that prevent them from accessing these education levels.

Despite these concerns, Mincerian returns remain popular and have been widely estimated (see Psacharopoulos 1994; Psacharopoulos and Patrinos 2004). Estimates of private returns to schooling, though available for Pakistan, are mostly dated and are often constrained by data (Hamdani 1977; Haque 1977; Guisinger, Henderson, and Scully 1984; Khan and Irfan 1985; Shabbir 1991, 1994; Shabbir and Khan 1991; Ashraf and Ashraf 1993a, 1993b; Nasir 1998, 1999, 2002; Siddiqui and Siddiqui 1998; Asadullah 2005; Riboud, Savchenko, and Tan 2006). There are two consistent findings from past studies in Pakistan: (i) returns to schooling attainment are low as compared to other developing countries and (ii) returns increase with the level of education. The latter finding challenges the dominant view that the earnings function is concave, and I provide evidence to suggest that in fact economic returns to schooling are quite substantial for both men and women.

The estimation of returns to schooling by gender has received less attention in the literature partly because in many countries gender differences are not so large. When estimates are available, the evidence from developing countries is mixed. While some studies find that returns to schooling do not differ significantly by gender (Behrman and Wolfe 1984; Schultz 1993), others discover lower returns to women's schooling (Kingdon 1998) or higher returns (Behrman and Deolalikar 1995; Asadullah 2006). Previous studies in Pakistan mostly compute returns to schooling attainment for males only and, hence, are not able to answer the central question addressed in this study: Does the labor market explain lower female schooling in Pakistan? Two recent exceptions are Nasir (2002) and Riboud et al. (2006). While Nasir implies that the answer to this question is "yes," Riboud et al. find higher returns to women's schooling, suggesting otherwise. These contradictory findings generate a puzzle in the literature. However, as neither of these studies addresses various methodological problems, their estimates could be biased, raising some uncertainty about their findings.

The objective of this article is to estimate returns to schooling attainment by gender in a consistent manner to determine whether childhood and adolescent education investments are affected by how the labor market rewards adult education. Both one-factor and multiple-factor models are used (Blundell, Dearden, and Sianesi 2005). In the former, education is defined as a continuous variable (completed grades of schooling attainment). This is a restrictive specification since it assumes that the return to schooling is the same for different attainment levels. The alternative model (multiple factor) specifies education in level form: each level is allowed to have a different effect on earnings. This is clearly more flexible than a quadratic specification that includes education in years and its square (Blundell et al. 2005). Briefly, four main methods of estimation are utilized: (i) standard ordinary least squares (OLS); (ii) the Heckman two-step procedure, which deals with the sample selectivity issues that arise because earnings are observed only for individuals who participate in the waged labor force and who may therefore form a nonrandom subsample of the population; (iii) two-stage least squares (2SLS) estimates using family background measures (parental education and spouse's education) as instrumental variables for schooling, to deal with endogeneity and measurement error in schooling; and (iv) household fixed-effects estimation to control for unobserved family-specific heterogeneity. For this, estimates are based on spouse pairs, sibling pairs, and parent-child pairs. In all four methods I allow for the possibility that parameters differ between the two genders by estimating separate earnings functions by gender.[2] Latest nationally representative data from the 2002 Pakistan Integrated Household Survey (PIHS) are used for the analysis.

The article is structured as follows. Section II discusses the empirical strategy, and Section III discusses the data. Section IV analyzes the empirical findings, and Section V presents conclusions.

[2] In this study I am unable to deal with measurement error (except when using the instrumental variable method), and although I am able to sign the bias (downward) given the assumption that the error is classical in nature, its magnitude remains unquantified.

II. Empirical Strategy

This study adopts the standard Mincerian approach of estimating earnings functions to compute rates of returns to education by gender. The earnings-schooling relationship can be stated in the form of a semilogarithmic relationship as follows:

$$\ln Y_i = \beta_0 + \beta_1 S_i + \beta_2 \mathbf{X}_{1i} + \beta_3 \mathbf{X}_2 + \varepsilon_i, \tag{1a}$$

$$S_i = \gamma_0 + \gamma_1 \mathbf{X}_{3i} + \gamma_2 \mathbf{X}_4 + \mu_i. \tag{1b}$$

In (1a), $\ln Y_i$ is the log of earnings[3] of individual i, S_i measures years of completed schooling in a one-factor model (or levels of schooling with dummy variables representing various levels of completed schooling in multiple-factor models), $\mathbf{X}_{1i}$ is a vector of observed characteristics of individual i, $\mathbf{X}_2$ is a vector of observed characteristics of the family, and ε_i is the individual-specific error. Equation (1b) models determinants of schooling, where $\mathbf{X}_{3i}$ is a vector of observed characteristics of individual i, $\mathbf{X}_4$ is a vector of household-level covariates, and μ_i is a residual term. The coefficient on schooling, β_1, measures the rate of return to each additional year of schooling (or to a particular level of schooling). This formulation assumes that the rate of return estimate is "homogeneous," that is, identical across all individuals i.

I start by estimating OLS models of earnings functions on male and female wage earners to provide some baseline results. However, OLS estimates of earnings functions potentially suffer from sample selectivity, omitted variables, and measurement error biases. On the first, earnings are observed only for individuals participating in the paid labor force. Moreover, most studies focus on waged workers, whereas many individuals in developing countries, especially men, are self-employed rather than in waged work. Consequently, estimates of returns to education of wage workers are on a potentially nonrandom draw from the population, resulting in sample selection. In most applied work, Heckman's correction for sample selectivity is used. This entails estimating a waged work participation equation, and the predicted probabilities of waged work from this equation are used to derive the selectivity term λ, which is then included in the main earnings function, such as (1a). To identify λ, the participation equation must include exclusion restrictions that are not part of the vector $\mathbf{X}$ in (1a).

The second problem has to do with omitted variable bias. The coefficient on schooling in the earnings function can be interpreted as the causal effect

[3] Wage rates are a better measure of labor productivity since earnings incorporate labor supply decisions and a return to capital. Lack of data on wage rates often prompts the use of earnings.

of education on earnings only if earnings differentials between individuals with varying years of schooling do not reflect differences in unobserved ability (or other unobserved investments in human capital) that happen to be correlated with education. For instance, unobserved inherent ability is clearly a determinant of schooling attainment as well as of earnings and generates endogeneity of schooling in the earnings function yielding inconsistent estimates of returns to schooling.

Finally, when measurement error (ME) is of the classical variety in the schooling variable S_i, it generates a correlation between the error terms in the earnings and schooling functions inducing attenuation bias in the regression coefficient β_1. This problem is compounded in sibling studies since differencing within families reduces the true signal-to-noise ratio in schooling.

Various methods have been used in extant literature to address school endogeneity in an earnings function framework. The instrumental variables (IV) methodology identifies variables (instruments, W_i) that are correlated with schooling (S_i) and uncorrelated with unobserved ability and measurement errors. This method provides a solution to endogeneity with the advantage that it simultaneously addresses ME issues. The key challenge is finding suitable instruments, that is, W_i, that are not part of the vectors $\mathbf{X}$ in equation (1a). Social and natural experiments are useful, and many studies using "institutional variations" in schooling due to such factors as proximity to schools, minimum school-leaving age, and so forth have been used to instrument for schooling. Card (1995a, 1999, 2001) provides a summary of some of the recent studies that use this approach, which include Angrist and Krueger (1991), Butcher and Case (1994), Card (1995b), and Harmon and Walker (1995), among others. The consensus from contemporary research on developed countries is that IV estimates based on natural experiments are as high as and sometimes almost 20% higher than corresponding OLS estimates (Card 2001). The evidence from developing countries is mixed and inconclusive (see Strauss and Thomas [1995] for a review and Maluccio [1998] and Duflo [2001] for returns to education estimates using institutional variation for the Philippines and Indonesia).

However, natural experiment–based IV approaches have exacting data demands, and an alternative is to use nonexperimental IVs for endogenous schooling. As children's schooling outcomes are to a large extent driven by family background (FB), variables such as father's education and mother's education are sometimes used (Trostel, Walker, and Woolley [2002] and Söderbom et al. [2004] are examples of two recent studies). FB variables constitute valid instruments if they affect earnings only indirectly through their effect on schooling, that is, if there is no intergenerational transmission of ability. FB then enters the vector of variables in equation (1b), which directly influence schooling.[4] However, a number of recent studies call into question the assumption of no intergenerational transmission of ability (see Behrman and Rosenzweig 2002, 2005; Plug and Vijverberg 2003; Plug 2004; Black, Devereux, and Salvanes 2005; Behrman et al. 2006; Björklund, Lindahl, and Plug 2006; De Haan and Plug 2006).

Alternatively, a number of studies use FB directly in earnings functions on the grounds that FB proxies omitted ability, school quality, and out-of-school learning environment or reflects nepotistic family connections (see Heckman and Hotz [1989] in Panama, Lam and Schoeni [1993] in Brazil, Krishnan [1996] in Ethiopia, and Kingdon [1998] in India). However, Card (1999, 1825) is critical of the use of FB variables as controls in earnings functions: inclusion of FB in earnings functions may reduce the bias but will still yield an upward-biased estimate of rates of return unless all the unobserved components are completely absorbed in the FB variables (1825–26).

An alternative to the IV technique is to use either repeated observations on the same individual over time or observations from different individuals within the same family to "difference out" the variables generating correlation in the residuals in a "household fixed effects" approach. Arguably, a good part of the unobserved heterogeneity is common to family members. Consequently, differences in unobserved ability and their impact in determining education should be lower *within* rather than *between* families. Earnings functions can be estimated on twin samples, siblings, father-son, or mother-daughter pairs using a household fixed effects or first-differencing approach. By introducing subsamples of households with at least two individuals of a given gender in wage employment (and more stringently households with brothers/sisters, father-son, or mother-daughter pairs in wage employment), the household fixed effects method effectively controls for all household variables that are common across these individuals within a given household. A simultaneous advantage of the household fixed effects procedure is that it controls for selection that is based on additive observables and unobservables (Pitt and Rosenzweig 1990, 978; Behrman and Deolalikar 1995, 106). However, there is some evidence that individual-specific endowments account for a substantial share of earnings variance even after controlling for household-specific endowments (Behrman, Rosenzweig, and Taubman 1994).

Card (1999) provides an excellent summary of findings from twin and sibling studies in developed countries. In almost all instances, household fixed-

[4] Such that if individual ability is an unobservable in the error term in earnings functions (ε_i), family background instruments (Z_i) must not be correlated with the error, i.e., $\mathrm{Corr}(Z_i, \varepsilon_i) = 0$.

effects estimates of the return to education are smaller than naive OLS estimates, suggesting an upward bias in the latter. However, data differencing exacerbates ME problems in sibling studies as part of the true signal is differenced out within families and the return to education is biased toward zero (Griliches 1979). The finding of smaller estimated returns in sibling studies gives credence to the suspicion that these studies suffer potentially severe attenuation bias. However, research in recent years overcomes measurement error problems, and while a majority of studies conclude that fixed-effects estimates corrected for measurement error are still smaller than OLS estimates (Ashenfelter and Rouse 1998; Rouse 1999; Hertz 2003), a study on a sample of twins by Behrman and Rosenzweig (1999) concludes otherwise.

III. Data and Variable Specification

The 2002 PIHS, a nationally representative data set on more than 16,000 households across Pakistan, is used in the analysis. Information, using a household questionnaire, was collected on employment and earnings of all males and females aged 10 and above. I restrict the analysis to adults aged 15–65 reporting waged work employment. Consistent with previous literature, full-time students (currently enrolled in school) are excluded from the sample. This yields a total of 13,519 adult males and females aged 15–65 reporting participation in waged employment.

Table 1 shows the distribution of the labor force by gender in Pakistan. There are striking gender differences in labor force participation rates: whereas 88% of males participate in the labor force, only 26% of females do so. A relatively large proportion of males and females are engaged in self-employ-

TABLE 1
DISTRIBUTION OF THE LABOR FORCE IN PAKISTAN BY GENDER (15–65)

	Male		Female		Total	
Labor Force Status	*N*	%	*N*	%	*N*	%
Unemployed (seeking work) (*a*)	962	3.51	949	3.23	1,911	3.37
Employed ($b = c + d$)	23,095	84.34	6,800	23.14	29,895	52.80
Self-employed (*c*)	11,594	42.34	4,782	16.27	16,376	28.85
Wage employed (*d*)	11,501	42.00	2,018	6.87	13,519	23.81
Total labor force ($e = a + b$)	24,057	87.85	7,749	26.37	31,806	56.02
Out of labor force (*f*)	3,328	12.15	21,638	73.63	24,966	43.98
All persons ($g = e + f$)	27,385	100	29,387	100	56,772	100

Note. Calculated from the 2002 PIHS. The definition of unemployed includes everyone reporting being jobless but seeking work; self-employed includes all defined as employers employing individuals, unpaid family workers, owner-cultivators, sharecroppers, cultivators, and livestock owners; wage employed includes all defined as paid employees; and non–labor force participants (out of labor force) are those reported as jobless and not seeking work.

ment: 42% and 16%, respectively. Gender differences in waged work participation are particularly striking: 42% of males and only 7% of females are engaged in some form of waged employment.

Earnings functions are fitted on the subsample of waged workers, that is, 11,501 males and 2,018 females. Selectivity-corrected earnings functions are fitted on wage work participants, with the reference category (or nonparticipants) including all other individuals (i.e., the unemployed, the self-employed, and nonworkers). The IV estimates are based on subsamples of waged workers who (1) report information on parental education and (2) are married and report spouse's education. Finally, the household fixed effects methodology estimates earnings functions on subsamples of households in which at least one individual of each gender (male/female) is in waged employment (any relation, sibling pairs, or father-son/mother-daughter pairs).

The dependent variable in the participation equation is wage or salaried employment (PAID_EMPLOY) and that in the earnings functions is the natural log of monthly earnings (LN_MONTHLY_Y). The definitions of the variables used in the participation equation and earnings functions are given in table 2. The vector of exclusion restrictions in the work participation equation includes demographic variables (CHILD7, ADULT60, MARRIED, HEAD) and the natural log of unearned household income (LNUNEARNED_Y).

The earnings functions include potential experience and its quadratic (EXP and EXP2). This variable is often computed as age − years of schooling − 5 on the belief that individuals start schooling at the age of 5 and enter the labor market upon completing schooling. The PIHS, however, asks individuals who attended formal schooling the age at which they entered school. Therefore, for individuals with positive years of schooling, EXP is computed as age − completed grades of schooling attainment − age entered school. However, for individuals with zero schooling, it is presumed that they entered the labor market at age 14, and hence potential experience has been calculated as EXP = age − 14.[5]

Table 3 (table 4) presents the means and standard deviations of the variables used in the participation equation (earnings function) by gender separately for

[5] My calculation of EXP tacitly assumes no grade repetition. The available data lack information on grade repetition, containing information only on completed grades of school attainment. I recognize that potential experience calculated in the aforementioned way is overestimated for individuals who repeated grades while at school. Moreover, the cutoff age of 14 used to calculate experience for illiterate individuals is based on the presumption that individuals enter the labor market at 14 and can result in positive years of experience before age 14 for individuals with some schooling but no such experience before age 14 for individuals with no schooling. However, the coefficients of EXP and EXP2 are fairly robust to various specifications (including EXP = age − completed grades of school attainment − 5).

TABLE 2
DEFINITION OF VARIABLES IN WAGE WORK PARTICIPATION/EARNINGS FUNCTIONS

Variable	Description
PAID_EMPLOY	Participation in salaried/waged work during the past month
AGE	Age in years
AGE2	Square of age
HEAD	Head of the household? yes = 1, no = 0
MARRIED	Married? yes = 1, no = 0
LNUNEARNED_Y	Natural log of unearned income (income from boarders/lodgers, zakat, remittances, pensions, gifts and insurance, etc.)
CHILD7	Number of children aged 7 or less in the household
ADULT60	Number of adults aged 60 or above in the household
NO_EDUCATION	Equals 1 if individual reports 0 years of education, 0 otherwise
LESS_PRIMARY	Individual has completed less than 5 years of education (katchi class, 1, 2, 3, or 4 years); equals 1 if has completed less than primary and equals 0 otherwise
PRIMARY	Equals 1 if individual has completed 5 years, 0 otherwise
MIDDLE	Equals 1 if individual has completed 6, 7, or 8 years, 0 otherwise
MATRIC	Equals 1 if individual has completed 9 or 10 years, 0 otherwise
INTER	Equals 1 if individual has completed 11 or 12 years, 0 otherwise
BACHELORS	Equals 1 if individual has completed 13 or 14 years, 0 otherwise
MA_MORE	Equals 1 if individual has completed 15 years of education or more, 0 otherwise
SINDH	Province is Sindh, yes = 1, no = 0
NWFP	Province is North-West Frontier Province, yes = 1, no = 0
BALOCHISTAN	Province is Balochistan, yes = 1, no =0
AJK	Province is Azad Jammu and Kashmir, yes = 1, no = 0
NORTH	Northern areas, yes = 1, no = 0
FATA	Federally Administered Tribal Areas, yes = 1, no = 0
URBAN	Region is urban, yes = 1, no = 0
LAMBDA	Selectivity term, inverse of Mills ratio
LN_MONTHLY_Y	Natural log of monthly earnings (rupees) of individuals in paid employment in the labor market
EXP	Experience (years)
EXP2	Square of experience
EDU_YRS	Number of years of education acquired
EDU_YRS2	Square of years of education
FEDYRS	Father's education (years)
MEDPRIM	Mother's education primary or less equals 1 if mother has positive but less than or equal to primary education, 0 otherwise
MEDPRIMORE	Mother's education more than primary equals 1 if mother has more than primary education, 0 otherwise
SPOUSE_EDU	Spouse's (husband's/wife's) education (years)
READ	Equals 1 if individual can "read in any language with understanding," 0 otherwise
WRITE	Equals 1 if individual can "write in any language with understanding," 0 otherwise
MATHS	Equals 1 if individual can "solve simple (plus/minus) sums," 0 otherwise
PRIVATE	Equals 1 if individual attended private school in the past, 0 otherwise

TABLE 3

DESCRIPTIVE STATISTICS OF VARIABLES USED IN THE PARTICIPATION EQUATION

	Mean Characteristics of Males			Mean Characteristics of Females		
Variable	Participants	Nonparticipants	All	Participants	Nonparticipants	All
PAID_EMPLOY*	1.000	.000	.420	1.000	.000	.069
	(.00)	(.00)	(.49)	(.00)	(.00)	(.25)
AGE	33.378	34.928	34.277	32.452	33.174	33.124
	(12.14)	(15.18)	(14.01)	(12.01)	(13.81)	(13.70)
AGE2	1,261.457	1,450.359	1,371.025	1,197.274	1,291.376	1,284.914
	(906.65)	(1,188.09)	(1,082.88)	(869.53)	(1,045.04)	(1,034.20)
HEAD*	.529	.462	.491	.053	.041	.042
	(.50)	(.50)	(.50)	(.224)	(.20)	(.20)
MARRIED*	.671	.613	.637	.629	.706	.701
	(.47)	(.49)	(.48)	(.48)	(.46)	(.46)
LNUNEARNED_Y	2.402	2.832	2.652	2.900	3.438	3.401
	(4.09)	(4.44)	(4.31)	(4.39)	(4.73)	(4.70)
CHILD7	1.785	1.951	1.881	1.654	2.011	1.987
	(1.79)	(1.98)	(1.91)	(1.84)	(1.95)	(1.95)
ADULT60	.435	.559	.507	.465	.548	.542
	(.65)	(.71)	(.69)	(.66)	(.71)	(.71)
LESS_PRIMARY*	.079	.087	.084	.039	.036	.036
	(.27)	(.28)	(.28)	(.19)	(.19)	(.19)
PRIMARY*	.097	.104	.101	.052	.070	.068
	(.30)	(.31)	(.30)	(.22)	(.25)	(.25)

MIDDLE*	.122	.134	.129	.039	.052	.051
	(.33)	(.34)	(.33)	(.19)	(.22)	(.22)
MATRIC*	.168	.167	.167	.100	.066	.068
	(.37)	(.37)	(.37)	(.30)	(.25)	(.25)
INTER*	.065	.051	.057	.062	.025	.027
	(.25)	(.22)	(.23)	(.24)	(.16)	(.16)
BACHELORS*	.058	.029	.041	.070	.015	.019
	(.23)	(.17)	(.20)	(.25)	(.12)	(.14)
MA_MORE*	.050	.016	.030	.058	.004	.007
	(.22)	(.13)	(.17)	(.23)	(.06)	(.09)
SINDH*	.302	.241	.267	.353	.233	.241
	(.46)	(.43)	(.44)	(.48)	(.42)	(.43)
NWFP*	.137	.171	.157	.075	.193	.185
	(.34)	(.38)	(.36)	(.26)	(.39)	(.39)
BALOCHISTAN*	.165	.164	.140	.085	.128	.125
	(.37)	(.37)	(.35)	(.28)	(.33)	(.33)
AJK*	.032	.028	.029	.027	.042	.041
	(.18)	(.16)	(.17)	(.16)	(.20)	(.20)
NORTH*	.011	.034	.024	.005	.029	.027
	(.10)	(.18)	(.15)	(.07)	(.17)	(.16)
FATA*	.012	.020	.017	.000	.019	.017
	(.11)	(.14)	(.13)	(.00)	(.14)	(.13)
URBAN*	.473	.323	.386	.501	.355	.366
	(.50)	(.47)	(.49)	(.50)	(.48)	(.48)
N	11,501	15,884	27,385	2,018	27,369	29,387

Note. Standard deviations are reported in parentheses.
* These variables are binary 0/1 variables, and their means represent the proportions of ones in the sample.

TABLE 4
DESCRIPTIVE STATISTICS OF VARIABLES USED IN EARNINGS FUNCTIONS (AGED 15–60) IN WAGED WORK

Average Value of Variable	Males	*N*	Females	*N*	*t*-Test (M − F)
LN_MONTHLY_Y	7.783	11,501	6.284	2,018	66.68
	(.01)		(.03)		
EXP	20.492	11,501	20.097	2,018	1.30
	(.12)		(12.79)		
EXP2	577.429	11,501	567.414	2,018	.65
	(5.98)		(14.09)		
EDU_YRS	5.666	11,501	4.326	2,018	10.32
	(.05)		(.13)		
EDU_YRS2	60.295	11,501	51.925	2,018	4.53
	(.70)		(1.82)		
NO_EDUCATION	.361	11,501	.581	2,018	−18.92
	(.00)		(.01)		
LESS_PRIMARY	.079	11,501	.039	2,018	6.42
	(.00)		(.00)		
PRIMARY	.097	11,501	.052	2,018	6.57
	(.00)		(.00)		
MIDDLE	.122	11,501	.039	2,018	11.12
	(.00)		(.00)		
MATRIC	.169	11,501	.100	2,018	7.78
	(.00)		(.01)		
INTER	.065	11,501	.062	2,018	.58
	(.00)		(.01)		
BACHELORS	.058	11,501	.070	2,018	−2.16
	(.00)		(.01)		
MA_MORE	.050	11,501	.058	2,018	−1.58
	(.00)		(.01)		
FEDYRS	2.865	4,155	4.380	493	−7.78
	(4.19)		(4.91)		
MEDPRIM	.057	4,155	.075	493	−1.65
	(.00)		(.01)		
MEDPRIMORE	.040	4,155	.061	493	−2.21
	(.00)		(.01)		
SPOUSE_EDU	2.059	5,638	4.559	943	−16.68
	(.05)		(.17)		

Note. Standard errors are in parentheses beneath the mean values of the variables. Descriptive statistics are computed excluding any individuals in paid employment who are currently enrolled in school. NO_EDUCATION is the reference category for education splines.

wage work participants and nonparticipants. Table 4 shows that in waged employment, males earn very substantially more than females. In logs, male earnings are 24% higher than female earnings. The disparity in earnings is more apparent from table 5, which shows average monthly earnings of waged employees by gender and education level. At all education levels, male earnings are significantly greater than female earnings.

IV. Econometric Results

Earnings functions are estimated using four methods: (1) OLS, (2) the Heckman two-step procedure, (3) 2SLS, and (4) household fixed effects. The results are

TABLE 5
AVERAGE MONTHLY EARNINGS OF WAGE EMPLOYEES, BY EDUCATION LEVEL AND GENDER

Education Level	Male (1)	Female (2)	Gap (M − F) (3 = 1 − 2)	*t*-Test (M − F) (4)	F/M (5 = 2/1)
NO_EDUCATION	2,271.5	581.3	1,690.2	−44.62	.26
	(19.04)	(23.14)			
LESS_PRIMARY	2,258.7	732.1	1,526.6	−9.40	.32
	(46.52)	(118.16)			
PRIMARY	2,539.7	709.0	1,830.7	−10.54	.28
	(52.28)	(91.71)			
MIDDLE	2,599.0	1,054.0	1,545.0	−8.26	.41
	(43.23)	(154.01)			
MATRIC	3,242.5	2,127.5	1,115.0	−7.67	.66
	(45.00)	(127.57)			
INTER	4,109.6	2,512.9	1,597.7	−6.27	.61
	(100.11)	(163.71)			
BACHELORS	5,845.0	3,818.5	2,026.5	−5.39	.65
	(165.40)	(245.64)			
MA_MORE	8,521.9	6,518.3	2,003.6	−3.13	.76
	(282.55)	(301.11)			
All	3,136.3	1,456.8	1,664.7	−25.25	.46
	(26.41)	(48.50)			

Note. Standard deviations are in parentheses beneath the mean values of the variables.

divided into two subsections. Subsection A presents OLS, Heckman, 2SLS, and fixed-effects estimates of earning functions. Subsection B extends the analysis by relaxing the restrictive assumption of linearity in the "years" specification, introducing occupation and industry controls and, finally, decomposing the gender wage gap using Oaxaca's (1973) method. Equations are fitted separately on males and females aged 15–65 in wage employment (except in the household fixed effects models).

A. Earnings Functions

OLS and Sample Selectivity Bias (SSB) Estimates

OLS estimates of returns to schooling attainment are presented in table 6. Columns 1 and 3 report findings for "grades" of education (EDU_YRS). Columns 2 and 4 depict results for education "levels." Focus on columns 1 and 3 first. The key parameter of interest is the point estimate on EDU_YRS—the rate of return to an additional completed grade of schooling. The marginal rate of return to schooling is 7.2% for males and 16.6% for females. The return to schooling attainment for women is more than double that for men in Pakistan. A Wald test confirms that the two coefficients on EDU_YRS in columns 1 and 3 are statistically very significantly different. This baseline

TABLE 6

OLS MINCERIAN EARNINGS FUNCTIONS (MALES AND FEMALES), WITH YEARS OF EDUCATION AND LEVELS OF EDUCATION

	Males (15–65)		Females (15–65)	
Variable	Years (1)	Levels (2)	Years (3)	Levels (4)
CONSTANT	6.223***	6.357***	4.188***	4.307***
	(.03)	(.03)	(.12)	(.13)
EDU_YRS	.072***	…	.166***	…
	(.00)		(.01)	
EXP	.076***	.075***	.073***	.068***
	(.00)	(.00)	(.01)	(.01)
EXP2	−.001***	−.001***	−.001***	−.001***
	(.00)	(.00)	(.00)	(.00)
LESS_PRIMARY	…	.011	…	.334**
		(.03)		(.14)
PRIMARY	…	.136***	…	.342**
		(.02)		(.13)
MIDDLE	…	.271***	…	.958***
		(.02)		(.14)
MATRIC	…	.534***	…	1.505***
		(.02)		(.12)
INTER	…	.762***	…	1.843***
		(.03)		(.12)
BACHELORS	…	1.070***	…	2.294***
		(.03)		(.11)
MA_MORE	…	1.371***	…	2.909***
		(.03)		(.11)
SINDH	.205***	.188***	.281***	.269***
	(.02)	(.02)	(.09)	(.09)
NWFP	−.060**	−.078***	.494***	.479***
	(.03)	(.03)	(.11)	(.11)
BALOCHISTAN	.423***	.386***	.664***	.643***
	(.03)	(.03)	(.13)	(.13)
AJK	.165***	.174***	.711***	.681***
	(.04)	(.04)	(.14)	(.15)
NORTH	.216***	.204***	1.370***	1.392***
	(.05)	(.05)	(.32)	(.32)
FATA	.124**	.096	…	…
	(.06)	(.06)		
URBAN	.200***	.204***	.487***	.503***
	(.02)	(.02)	(.09)	(.09)
R^2	.388	.408	.472	.478
N	11,501	11,501	2,018	2,018
Mean (dependent variable)	7.783	7.783	6.284	6.284

Note. The dependent variable is LN_MONTHLY_Y. Standard errors are in parentheses. Ellipses indicate no observations. NO_EDUCATION is the reference category for education splines and PUNJAB for provinces.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

TABLE 7
RATES OF RETURNS TO ADDITIONAL YEARS OF EDUCATION (MALES AND FEMALES) AT VARIOUS LEVELS OF EDUCATION

	Rates of Return (%)	
Level of Education	Males	Females
PRIMARY	2.7	6.8
MIDDLE	4.5	20.5
MATRIC	13.2	27.4
INTER	11.4	16.9
BACHELORS	15.4	22.6
MA_MORE	15.1	30.7

result implies large and significant gender differences in returns to schooling attainment in Pakistan.[6]

Turn now to columns 2 and 4 for the levels specification. This model relaxes the assumption of linearity of education implicit in columns 1 and 3. Some striking findings emerge. First, the coefficients on education levels are positive and progressively increasing with higher levels of education for both genders, indicating a convex relationship between education and earnings. Second, the coefficients at all education levels are significantly higher for females than for males. The returns to additional years of schooling attainment at various levels have been calculated using the coefficient estimates in table 6 and are reported in table 7.[7] The findings show that returns to female attainment are always higher than returns to male attainment. However, while returns increase for both males and females till INTER, they decline and then increase again at higher education levels for both genders, but more for females than for males.[8] Third, the increase in coefficients with education levels is much sharper for women than for men, suggesting that the earnings profile is more convex for females than for males. Finally, there is a premium in returns from PRIMARY

[6] The sample of 15–65-year-olds includes a very wide age range of individuals. As Pakistan has undergone a significant transformation in schooling for when different cohorts were attending school, I divide the sample of individuals into two cohorts: individuals aged 15–35 and those aged 36–65. The coefficients on the OLS estimates for males in the two cohorts are 0.073 and 0.067 and for females are 0.17 and 0.17, respectively. This suggests that private wage returns to schooling attainment are not significantly different across the two cohorts. Also, one wonders what the selection implications are of excluding full-time students from the sample. I also reestimated the earnings function on a subsample of individuals aged 25–65 with the view that by age 25 a very large proportion of individuals have exited school. The estimated returns do not change for both males and females.

[7] The coefficients in the levels specification in table 6 have to be transformed to arrive at the "returns." The reason is that the number of years of education is different for the various levels of education indicated by the dummy variables, and as measured here, the wage premia for a graduate of a higher level include the premium from a lower level of education.

[8] The dip in returns between MATRIC and INTER is significant for males and females, whereas the increase between INTER and BACHELORS is significant for males but insignificant for females.

to MIDDLE for females (coefficients increase from 0.34 to 0.96), with the increase being substantially smaller for males (0.14 to 0.27).

However, as described before, OLS estimates may be biased because of sample selection and endogenous schooling. Let us turn next to the SSB estimates (table 8), which correct for selection bias by using the Heckman two-step procedure and incorporate LAMBDA into earnings function estimates.[9] The selectivity-corrected earnings functions reported in table 8 include the standard variables: education, experience and its square, and the provincial and regional dummies. Household demographic variables (CHILD7, ADULT60), marital status (MARRIED), and LNUNEARNED_Y (table 6) are used as exclusion restrictions. These variables are believed to determine participation in waged work but do not directly affect labor market earnings. The identifying variables are jointly significant at the 0.1% level (p-value of the F-test is 0.000) as well as individually significant.

The LAMBDA term is large and statistically significantly negative for males (in both years and levels specifications) and significant for females only in the years specification. This would be consistent with more able, entrepreneurial, and motivated persons being less likely to be waged workers. A comparison across columns 1 and 3 and across columns 2 and 4 in tables 6 and 8 reveals the effect of correcting for sample selection. Inclusion of the LAMBDA term reduces the point estimates on years of education from 7.2% to 6.4% for males and 16.6% to 14.2% for females. These differences in coefficients on EDU_YRS across the OLS and SSB specifications are statistically significant for males and females. In the levels specification, the inclusion of the LAMBDA term has no significant attenuating effect on the education coefficients in the female sample (consistent with LAMBDA being insignificant), but in the male sample inclusion of LAMBDA has a significantly attenuating effect on some education-level coefficients (BACHELORS and MA_MORE). Finally, specifying education in levels rather than as EDU_YRS has an attenuating effect on the point estimate of LAMBDA, which falls by a larger absolute value for females than for males. The change in LAMBDA coefficients is significant for males but insignificant for females. Overall, these findings suggest that OLS overestimates the return to education (especially in the years specification).

It is important, however, that the return to female schooling attainment remains significantly greater than that for males even after controlling for selection bias: the marginal return to schooling is 6.4% for males and 14.2%

[9] The probit wage work participation first-stage regressions underlying the Heckman two-step procedure are suppressed because of space limitations. Results are available from the author on request.

TABLE 8

HECKMAN-CORRECTED MINCERIAN EARNINGS FUNCTIONS (MALES AND FEMALES), YEARS OF EDUCATION AND LEVELS OF EDUCATION

	Males (15–65)		Females (15–65)	
Variable	Years (1)	Levels (2)	Years (3)	Levels (4)
CONSTANT	7.295***	6.908***	5.227***	4.310***
	(.08)	(.08)	(.29)	(.46)
EDU_YRS	.064***	…	.142***	…
	(.00)		(.01)	
EXP	.049***	.060***	.062***	.068***
	(.00)	(.00)	(.01)	(.01)
EXP2	−.001***	−.001***	−.001***	−.001***
	(.00)	(.00)	(.00)	(.00)
LESS_PRIMARY	…	.020	…	.334***
		(.02)		(.13)
PRIMARY	…	.149***	…	.343***
		(.02)		(.12)
MIDDLE	…	.283***	…	.958***
		(.02)		(.14)
MATRIC	…	.533***	…	1.504***
		(.02)		(.10)
INTER	…	.732***	…	1.842***
		(.03)		(.15)
BACHELORS	…	.991***	…	2.293***
		(.03)		(.18)
MA_MORE	…	1.246***	…	2.908***
		(.03)		(.28)
SINDH	.085***	.127***	.239***	.269***
	(.02)	(.02)	(.06)	(.06)
NWFP	−.053**	−.070***	.695***	.479***
	(.02)	(.02)	(.11)	(.14)
BALOCHISTAN	.257***	.303***	.778***	.643***
	(.02)	(.02)	(.10)	(.11)
AJK	.062	.117***	.848***	.681***
	(.04)	(.04)	(.17)	(.17)
NORTH	.567***	.395***	1.786***	1.393***
	(.06)	(.06)	(.37)	(.40)
FATA	.160***	.122**	…	…
	(.06)	(.05)		
URBAN	.027***	.111***	.443***	.503***
	(.02)	(.02)	(.06)	(.06)
LAMBDA	−.756***	−.407***	−.472***	−.001
	(.05)	(.05)	(.12)	(.22)
N (uncensored)	11,501	11,501	2,018	2,018
Wald χ^2	3,836.09	5,144.57	1,297.56	1,901.26
p-value (Wald)	.000	.000	.000	.000

Note. The dependent variable is LN_MONTHLY_Y. Standard errors are in parentheses. Ellipses indicate no observations. NO_EDUCATION is the reference category for education splines, and PUNJAB for provinces.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

for females (cols. 1 and 3 in table 8). As with simple OLS, the return to education for women is more than double that for men in Pakistan. This difference is statistically significant. The results in the levels specification reported in columns 2 and 4 are consistent with the OLS findings. Note that the jump in coefficients from PRIMARY to MIDDLE remains even after controlling for selection into waged work.

IV Estimates

However, OLS estimates of returns to schooling may suffer from endogeneity bias. The SSB estimates reported above may also be biased upward because of the classic "ability bias." IV estimates are more robust than OLS and SSB estimates on two counts: (1) they control for endogeneity of EDU_YRS, thereby correcting for any upward "ability biases"; and (2) they are unaffected by measurement error so that the reported findings should be purged of any attenuating effects.

I use parental education as instruments for the subset of individuals reporting father's and mother's education and alternatively spouse's education as an instrument for another subset of married wage workers. Education is instrumented using three variables: FEDYRS (years of education completed by worker's father), MEDPRIM (equals one if mother has completed any year of primary education, zero otherwise), and MEDPRIMORE (equals one if mother has completed more than primary education, zero otherwise).[10] Mothers reporting no education (MEDNONE) are the omitted category. Parental education may be good instruments for own schooling if parents' education positively affects schooling but is not correlated with child ability (which is in the error term of the earnings function), that is, assuming no intergenerational transmission of ability.[11] In the sample of wage workers aged 15–65 from the PIHS (2001–2), between 24% and 26% of the variation in education of males and females, respectively, is explained by father's education.[12] It is important to note, however, that since parental education information is available only

[10] I experimented with a number of instruments for mother's education. The set of instruments that satisfied the overidentification (OID) test and seemed justifiable and theoretically plausible was chosen.

[11] Even assuming no intergenerational transmission of ability, these instruments may still be criticized on the grounds that parental education may have either a direct effect on individual earnings in the labor market (nepotism and family connections) or an indirect effect through its effect on school quality. These arguments make a case for including parental education as control variables in earnings functions rather than using them as instruments for schooling.

[12] I also have data on maternal education for a subset of workers. However, the correlation between mother's education and own education is relatively low (7% and 15% for males and females, respectively).

for those individuals whose parents reside in that household, this clearly non-random process generates sample selection constraints. Moreover, IV estimates are based on selected subsamples of subsets of individuals reporting earnings in waged work.[13]

The IV analysis using parental background is augmented using spouse's education (SPOUSE_EDU = years of education completed by spouse) as another instrument to compare the findings with parental education estimates. This draws on the theory of assortative mating (Weiss 1999): individuals with common social backgrounds, religion, race, and caste are more likely to bond together in marriage. This is accentuated by the high correlation between spouse's education and own education in Pakistan (0.29 for males and 0.51 for females from the 2002 PIHS). Moreover, as social norms in Pakistan are such that almost all individuals above a certain age are most likely to be married, spouse's education as an instrument is less likely to be problematic than parental education.[14]

To determine the empirical validity of the instruments, turn to the first-stage estimates reported in columns 2, 4, 6, and 8 in table 9. For both genders, the first-stage equations reveal that FEDYRS, MEDPRIM, MEDPRIMORE, and SPOUSE_EDU almost always have large, very precisely determined coefficients with the expected signs. The relevance of the instruments can be assessed by examining the significance of the excluded instruments in the first-stage IV regressions (Bound, Jaegar, and Baker 1995). The *p*-values of the *F*-tests in the first-stage regression indicate that the instruments satisfy the relevance condition very well. If the instruments used are not "valid," that is, if $\mathrm{Corr}(Z_i, \varepsilon_i) \neq 0$, the IV estimates will be inconsistent. The only way to

[13] I also corrected the IV estimates for sample selection (into wage work). This was done following the approach outlined in Wooldridge (2002). The idea is to use the same set of instruments in the selection and instrumenting regressions (Z1 and Z2). In the first stage, the probit regression (for waged work) was estimated including all exogenous variables (i.e., including parental education, Z1 and Z2, but excluding EDU_YRS). The earnings regression was estimated on controls Z1, EDU_YRS, and the Mills ratio, using as instruments the controls Z1, Mills ratio, and instrumental variables Z2. The coefficient on EDU_YRS for the male subsample from this regression was highly statistically significant and equaled 0.090. The *p*-value of the OID test was 0.10. For the female subsample, the coefficient on EDU_YRS was 0.164 and was highly significant, and the *p*-value of the OID test was 0.33. These estimates are reported in app. table A1.

[14] Using spouse's education may raise the concern that as using this variable implicitly conditions on marital status, and because marital status may be endogenous, the validity of this variable as an instrument is reduced. However, this is unlikely to be the case in Pakistan because marriage is dictated more by social norms than by individual choice and a vast proportion of adults in Pakistan are married. For instance, almost 91% of individuals aged 30–50 in Pakistan are married. However, spouse's education may not be a good instrument if there is assortative mating on unobserved endowments as reported in Behrman et al. (1994).

TABLE 9

INSTRUMENTAL VARIABLE EARNING FUNCTIONS ESTIMATES (MALES AND FEMALES), USING YEARS OF EDUCATION

	Males (15–65)				Females (15–65)			
Variable	IV (PARENTAL EDUCATION) (1)	First Stage (2)	IV (SPOUSE_EDU) (3)	First Stage (4)	IV (PARENTAL EDUCATION) (5)	First Stage (6)	IV (SPOUSE_EDU) (7)	First Stage (8)
CONSTANT	5.894***	7.516***	6.216***	9.755***		4.623	4.432***	3.647***
	(.08)	(.24)	(.08)	(.37)		(.76)	(.28)	(.75)
EDU_YRS	.099***	…	.106***	…	.169***	…	.176***	…
	(.01)		(.00)		(.03)		(.01)	
EXP	.102***	−.473***	.064***	−.317***	.154***	−.422***	.037*	−.232***
	(.01)	(.03)	(.00)	(.02)	(.03)	(.10)	(.02)	(.06)
EXP2	−.02***	.010***	−.001***	.003***	−.003***	.008***	.000	.003**
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
SINDH	.225***	.381**	.188***	−.011	.230*	−.062	.345***	−.658***
	(.03)	(.16)	(.03)	(.15)	(.13)	(.59)	(.11)	(.25)
NWFP	−.103**	.941***	.006	.375*	−.077	3.823***	.775***	.919*
	(.04)	(.20)	(.03)	(.19)	(.25)	(.96)	(.14)	(.49)
BALOCHISTAN	.486***	−.092	.403***	.174	.728***	1.103	.656***	−.854*
	(.04)	(.19)	(.03)	(.18)	(.26)	(1.14)	(.14)	(.44)
AJK	.029	1.573***	.215***	1.273***	.165	5.366***	1.231***	1.533*
	(.07)	(.36)	(.05)	(.35)	(.25)	(.90)	(.20)	(.92)
NORTH	.162**	.668	.208***	1.399**	…	…	1.623***	−.302
	(.08)	(.63)	(.06)	(.56)			(.26)	(1.43)

FATA	.169*	−1.596***	.142	−.160	…	…	…	…
	(.09)	(.59)	(.09)	(.53)				
URBAN	.096***	.657***	.164***	1.396***	.324*	3.745***	.572***	1.919***
	(.03)	(.13)	(.02)	(.13)	(.19)	(.59)	(.11)	(.25)
FEDYRS	…	.491***	…	…	…	.360***	…	…
		(.02)				(.06)		
MEDPRIM	…	.526			…	2.067***		
		(.29)				(.69)		
MEDPRIMORE	…	1.397***			…	2.486**		
		(.35)				(.99)		
SPOUSE_EDU	…	…	…	.584***	…	…	…	.587***
				(.02)				(.02)
R^2	.282	.323	.303	.383	.458	.510	.478	.578
N	4,155	4,155	5,590	5,590	493	493	903	903
F-test of excluding								
instruments	…	258.62	…	1,305.46	…	23.48	…	610.60
p-value (F-test)	…	.000	…	.000	…	.000	…	.000
OID test	…	4.69	…	…		2.37	…	
p-value (OID test)	…	.100	…	…		.306		

Note. Standard errors are in parentheses and are robust and corrected for clustering at the population sampling unit level. The dependent variable is LN_MONTHLY_Y. Ellipses indicate no observations or not used. PUNJAB is the base category for provinces and MEDNONE (= 1 if mother has no education, 0 otherwise) is the base for mothers' educational dummies; MEDPRIM =1 if mother has primary or less education (but more than zero) and 0 otherwise; MEDPRIMORE =1 if mother has more than primary education, 0 otherwise.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

assess the validity of the instruments is to have a surfeit of instruments and use an OID test (Hansen's *J*-test). This is possible only for the IV sample using parental education since we have both mother's and father's education. In both the male and female samples using parental education instruments, the OID test confirms the validity of instruments used.

Table 9 reports 2SLS estimates. Education is specified as a continuous variable.[15] Columns 1–4 report IV estimates using parental education and SPOUSE_EDU for males, and columns 5–8 report those for females. Columns 2, 4, 6, and 8 report first-stage results. Focus first on columns 1, 3, 6, and 7, the earnings functions estimates. The rate of return to an additional completed grade of school attainment is between 10% and 11% for males and between 17% and 18% for females using either instrument. The main findings are as follows: (1) as before, the rate of return to schooling attainment is always higher for females than for males, and (2) consistent with the findings from numerous other studies, β_{IV} (10%–11% for males and 17%–18% for females) is larger than β_{OLS} (7% and 17% for females).

Household Fixed-Effects Estimates

Let us turn now to household fixed-effects (FE) estimates of returns to education. The results are based on subsamples of at least one male and one female wage worker within a household who are related in any way (e.g., father-daughter, mother-son, brother-sister, or husband-wife) or are siblings (only brother-sister pairs).[16] Table 10 depicts the FE estimates: panel A for all relations and panel B for sibling pairs. As before, education is measured in "grades" and "levels." The set of independent variables remains unchanged with two exceptions. A gender dummy (MALE) is added since the sample includes individuals of both sexes, and interaction terms (EDU_YRS_MALE in grades and LESS_PRIMARY_MALE etc. in levels) are included. These capture the effect of gender on the return to education. Finally, in panels A and B, OLS estimates are reported along with FE estimates for comparative reasons.

Focus first on the grades specification in panels A and B. The returns to male education have been computed as the sum of the coefficients on EDU_YRS and EDU_YRS_MALE. (For example, in panel A, the coefficient on EDU_YRS is 0.14 and that on EDU_YRS_MALE is −0.082. The overall return for males is the sum of 0.14 and −0.082, which equals 0.058, or

[15] With few instruments per subsample, the levels specification cannot be used in 2SLS.

[16] A caveat is in order: most women aged 20 and over are married and no longer reside with their parents. This suggests that the FE sample (especially the subsample of brother-sister pairs) is a select sample from the population.

approximately 6%.) Thus, from FE estimates, the return to female schooling attainment is clearly substantially higher than that to male attainment for all individuals (14% for females compared to 6% for males) and for siblings (15% vs. 11%, respectively).

Note that the FE point estimates on EDU_YRS and EDU_YRS_MALE are lower than the OLS estimates, a finding consistent with previous literature. However, although the FE estimates are lower, they do not collapse and are reasonably close to the OLS estimates. Part of the decline in estimates could be caused by an upward bias in the OLS estimator due to omitted variables. Some part (albeit unmeasured) of the attenuation could be attributable to ME if it is of the classical variety. Although studies such as Hertz (2003) that correct for ME in within-household estimators still find the within-household estimate of the return to education to be smaller than the corresponding OLS estimate, the correction for ME causes a rise from the uncorrected estimate. Data constraints prevent such a correction in the current study, but when a majority of other studies have corrected for random ME, they have mostly found estimates based on FE to be smaller than OLS estimates, which is also what we find, and this gives us confidence in our findings.

Let us turn now to the levels specifications in panels A and B of table 10. First, it is clear that, except for primary education, the FE returns to female attainment are higher than those for males at all levels of completed levels of schooling in panels A and B. This suggests larger labor market incentives for females (than for males) to acquire education. Second, the FE findings confirm that the convexity of the education-earnings profile in previous sections is not an artifact of heterogeneity. Third, although evidence points to $\beta_{\mathrm{FE}} < \beta_{\mathrm{OLS}}$ for a majority of cases in panel A, this is not true for siblings (panel B). Finally, in both the all and siblings samples, the jump in returns from PRIMARY to MIDDLE remains for females; for example, in the all sample, the coefficients for females increase from 0.126 (PRIMARY) to 0.822 (MIDDLE), and those for males are roughly the same for both levels of education (between 0.18 and 0.19).

Summary

Clearly, regardless of the methodology adopted, the results reveal similar findings: the estimated marginal return to additional grades of completed schooling is significantly higher for females than for males. The returns also increase with higher levels of education, pointing to convex education-earnings profiles. I also find that the labor market differentially rewards males and females with relatively low education levels (primary and middle): women with middle schooling are rewarded substantially more than women with primary education.

TABLE 10

HOUSEHOLD FIXED-EFFECTS ESTIMATES OF EARNINGS FUNCTIONS, MALES AND FEMALES (15–65), YEARS AND LEVELS OF EDUCATION

	A. All				B. Siblings (Brother/Sister)			
	Years		Levels		Years		Levels	
Variable	OLS	FE	OLS	FE	OLS	FE	OLS	FE
CONSTANT	4.432***	4.879***	4.587***	5.132***	4.057***	4.143***	4.354***	
	(.08)	(.09)	(.08)	(.08)	(.20)	(.26)	(.21)	
EDU_YRS	.165***	.140***	…	…	.155***	.153***	…	…
	(.01)	(.01)			(.01)	(.02)		
EDU_YRS_MALE	−.093***	−.082***	…	…	−.049***	−.043***	…	…
	(.01)	(.01)			(.02)	(.02)		
MALE	1.659***	1.630***	1.680***	1.628***	1.138***	1.105***	1.173***	1.245***
	(.05)	(.05)	(.06)	(.06)	(.15)	(.15)	(.17)	(.18)
EXP	.073***	.072***	.067***	.069***	.179***	.213***	.167***	.222***
	(.01)	(.01)	(.06)	(.01)	(.03)	(.03)	(.03)	(.03)
EXP2	−.001***	−.001***	−.001***	−.001***	−.004***	−.005***	−.004***	−.01***
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.001)	(.001)
LESS_PRIMARY	…	…	.196	.218	…	…	.443	.802**
			(.15)	(.18)			(.30)	(.35)
PRIMARY	…	…	.126	.050	…	…	.110	.150
			(.12)	(.15)			(.292)	(.37)
MIDDLE	…	…	.822	.707***	…	…	.522**	1.010***
			(.14)	(.17)			(.24)	(.32)
MATRIC	…	…	1.359***	1.204***	…	…	.957***	1.376***
			(.10)	(.13)			(.22)	(.29)
INTER	…	…	1.742***	1.637***	…	…	1.498***	1.777***
			(.12)	(.16)			(.22)	(.30)
BACHELORS	…	…	2.306***	2.060***	…	…	2.078***	2.163***
			(.11)	(.15)			(.22)	(.30)
MA_MORE	…	…	3.012***	2.767***	…	…	2.809***	2.752***
			(.12)	(.19)			(.23)	(.37)

LESS_PRIMARY_MALE	…	…	−.202	−.170	…	…	−.528	−.927**
			(.17)	(.21)			(.39)	(.44)
PRIMARY_MALE	…	…	.051	.139	…	…	.097	−.071
			(.15)	(.18)			(.37)	(.43)
MIDDLE_MALE	…	…	−.669***	−.532**	…	…	−.060	−.401
			(.16)	(.18)			(.31)	(.35)
MATRIC_MALE	…	…	−.888***	−.776***	…	…	−.311	−.571*
			(.12)	(.14)			(.28)	(.31)
INTER_MALE	…	…	−1.051***	−.956***	…	…	−.728**	−.911***
			(.16)	(.18)			(.32)	(.35)
BACHELORS_MALE	…	…	−1.260***	−1.080***	…	…	−.548*	−.223
			(.15)	(.16)			(.32)	(.37)
MA_MORE_MALE	…	…	−1.632***	−1.546***	…	…	−.976***	−.978***
			(.16)	(.16)			(.33)	(.34)
SINDH	.205***	…	.196***	…	.181*	…	.147	…
	(.04)		(.04)		(.10)		(.10)	
NWFP	.171**	…	.150**	…	−.145	…	−.166	…
	(.071)		(.07)		(.17)		(.17)	
BALOCHISTAN	.400***	…	.369***	…	.481**	…	.397*	…
	(.070)		(.07)		(.20)		(.20)	
AJK	.422***	…	.371**	…	.007	…	−.093	…
	(.124)		(.12)		(.21)		(.21)	
NORTH	.472	…	.535*	…	…	…	…	…
	(.324)		(.32)					
FATA	…	…	…	…	…	…	…	…
URBAN	.293***	…	.320***	…	.226**	…	.279***	…
	(.04)		(.04)		(.10)		(.10)	
N	2,423	2,423	2,423	2,423	437	437	437	437
No. groups	…	948	…	948	…	160	…	160
R^2	.551	.522	.567	.541	.522	.501	.560	.528

Note. The dependent variable is LN_MONTHLY_Y. Ellipses indicate no observations or not used. NO_EDUCATION is the reference category for education splines, and PUNJAB for provinces.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

This is not true for males. However, a number of issues warrant further investigation. The following subsections estimate extended earnings functions to investigate the convexity of education-earnings profiles, analyze what factors potentially drive the jump in returns from primary to middle schooling for females and the lack thereof for males, and, finally, use the Oaxaca decomposition method to determine whether there is differential treatment in the labor market other than through differential rewards to education.

B. *Extended Earnings Functions*

Nonlinear Earnings Functions

So far, the continuous specification of the earnings function (grades) has assumed linearity in schooling (captured in EDU_YRS). Theoretical literature suggests that in fact the relationship between completed grades of schooling attainment and earnings may be concave because of diminishing returns to education. While there is empirical support for this concavity (Psacharopoulos and Patrinos 2004), equally convincing evidence dating as far back as the early 1980s suggests otherwise (Behrman and Wolfe 1984; Alderman and Sahn 1988). In recent years, concavity of education-earnings profiles has been further challenged in studies that find returns increasing with higher levels of education, that is, convex profiles (see Kingdon 1998; Card 1999; Söderbom et al. 2004; Belzil and Hansen 2002; Nasir 2002; Schultz 2004; Riboud et al. 2006). For the purposes of this study, first, if the education-earnings profile is indeed concave, imposing linearity as I have done in the continuous specification is too restrictive. Second, documenting the shape of the education-earnings profile is important because of its potential effect on the estimation of the relative rates of return to schooling for women and men.

Using the levels specification, we have already found evidence of convex earnings profiles for males and females with sharper convexities for the latter than for the former. However, as OLS and SSB results in tables 6 and 8 do not control for endogeneity within the levels framework, one wonders whether the finding of convexity is an artifact of endogeneity. Although the FE estimates also reveal sharply convex earnings profiles for males and females, they are based on smaller samples. Hence, this subsection introduces a quadratic term for years of education (EDU_YRS2) within the grades specification to relax the linearity constraint previously imposed.

The endogeneity of EDU_YRS and EDU_YRS2 is tackled using a two-stage control function approach. The main advantage of this approach is that it allows one to control for the endogeneity of a nonlinear variable while simultaneously testing for endogeneity: if the residual term from the first stage is significant, it implies that the unexplained variation in EDU_YRS also

TABLE 11
MARGINAL RETURNS TO EDUCATION (OLS AND CF)

	Males		Females	
Years of Education	OLS (1)	CF (2)	OLS (3)	CF (4)
6	7.1	7.9	12.8	12.9
10	11.9	11.9	18.4	18.5
12	14.3	13.9	21.2	21.3
14	16.7	15.9	24.0	24.1

affects variation in earnings, and the insignificance of the residual term means that one can accept the hypothesis that the schooling variable is not endogenous.[17]

Control function (CF) models can be estimated only on subsamples of individuals reporting parental education, and for comparative purposes, corresponding OLS estimates are reported on the same subsamples. In the first stage of CF estimation, parental education significantly positively determines the completed grades of school attainment of an individual. However, the effect is stronger for females: mother's education (MEDPRIM and MEDPRIMORE) is significantly positive only in the female sample, and the size of the coefficient is double that in the male sample.

It is important to note that the earnings functions estimated incorporating the residuals from the first stage show that even after we control for endogeneity of schooling, the education-earnings profiles for males and females remain convex. Table 11 computes the marginal return to schooling for various years of completed schooling using the OLS and CF estimates and confirms the convexity of the education-earnings profiles. A comparison of OLS and CF estimates reveals that in most cases they are not significantly different from each other. Consequently, in the extensions that follow, I report only OLS and/or SSB estimates.

What Explains the Premium to Women with Middle Education?

Notice in tables 6, 8, and 10 that (1) the returns to low levels of education were high especially for women, and (2) there was a large premium to women for possession of middle-level education and the premium was greater than that for men. One wonders what the labor market realities underlying this result are. One possibility is that women's higher economic benefits from education are realized through better occupational attainment or better industry attachment. In order to test this, I include industry and occupation

[17] CF estimates are not reported because of space constraints. Tables are available from the author on request.

dummies in earnings functions specifications.[18] The results suggest that indeed the effect of education on earnings occurs partly by permitting better occupational attainment. The coefficients on OLS/SSB with occupation/industry controls decline for both males and females. However, this is not the only mechanism since there are large and significant direct returns to certain levels of completed schooling (especially middle) even within occupations and industries. In particular, whereas for men education affects their earnings via occupation/industry at all education levels, for women the effect of occupation and industry association operates only after MATRIC.

Does the Labor Market Discriminate against Women?

In the Pakistani context, the finding that returns to completed grades of school attainment are significantly greater for women than for men potentially raises a puzzle as to why then parents allocate lower education to girls than to boys (Aslam and Kingdon 2008). While it is empirically difficult to disentangle the different reasons for education gender gaps (Alderman and King 1998), one explanation could be that the total labor market return to boys is greater than that to girls. This subsection extends the analysis in this way as a way of probing parental motives further.

I decompose the male-female wage gaps using the technique proposed by Oaxaca (1973). OLS and selectivity-corrected earnings functions (incorporating EDU_YRS2) are estimated to predict earnings. The wage gap is decomposed into two components: (1) that explained by differences in individual characteristics and (2) the residual, unexplained portion, reflecting differences in earnings structure. Using this method, one can decompose the total gender wage gap into the explained and unexplained components. The unexplained component could be seen as the extent of "discrimination" in the labor market. However, if there are important differences in the unobserved or unmeasured characteristics of males and females, then the residual component cannot so validly be termed "discrimination." The Oaxaca decomposition is initially conducted using OLS. As a robustness check, I repeat the exercise using a household FE model on subsamples of at least two waged workers of each gender in a given household.[19]

With OLS, expressed in natural logs, the gross gender wage difference is 1.49. When I standardize by male means, 0.25 of the 1.49 gender wage difference is explained by better male characteristics (such as higher educational attainment) and 1.25 of the gender wage gap remains unexplained. Consequently, almost 84% of the gender wage gap is unexplained. Standardizing

[18] Results are not reported because of space constraints but are available from the author on request.
[19] Results are not reported because of space constraints.

by female means suggests an even greater unexplained proportion (95%). The household FE estimates show that an even larger proportion of the gender wage gap is unexplained. Household FE estimates provide a cleaner test since unobserved or unmeasured characteristic differences between males and females within the family are likely to be much lower than across families, and the large unexplained portion in the FE sample is indicative of high discrimination in the labor market.

However, since male and female hours worked may partly account for the large unexplained component of the gender wage gap, one should ideally perform the decomposition including hours worked in the earnings function. Although I do not have information on hours, I do have data on days worked in the past month. However, mean days worked is very similar for males and females, and including that variable in the OLS and FE regressions causes the unexplained component to remain virtually unchanged (the unexplained portion estimate is 88% in the OLS and 94% in the FE sample).

These estimates of the unexplained portion are large in comparison to international estimates. For example, estimates of the unexplained portion of the wage gap range between 1% and 5% in the United Kingdom (Zabalza and Arrufat 1983), 12% in the United States (Choudhury 1993), and 35%–45% in India (Kingdon and Unni 2001). However, previous estimates in Pakistan range from 63% in 1979 (Ashraf and Ashraf 1993b) to 33% in 1985–86 (Ashraf and Ashraf 1993a) and between 86%–96% and 55%–77% in 1993–94 (Siddiqui and Siddiqui 1998). Clearly, my findings are closest to those of the last study, which is also the latest past study. The findings are suggestive of a pernicious increase in the "unexplained portion" over time. If viewed as discrimination, they would explain not only the low participation of women in Pakistan's labor markets but also the large differentials in intrahousehold education expenditure allocations within households (Aslam and Kingdon 2008). However, these conclusions are subject to an important caveat: the decomposition of the male-female earnings gap is based only on wage earners in a conditional equation (conditional on being a wage earner).[20]

V. Conclusion

This study seeks an answer to the following question: Does the labor market explain lower education of girls than of boys in Pakistan? If the labor market

[20] The large unexplained component in such conditional equations could be partly due to the fact that women's participation is constrained by cultural factors. A decomposition of the male-female earnings gap based on an "unconditional" sample would presumably yield a larger gender gap, with the likelihood that productive characteristics (education, experience, etc.) explain a greater proportion of the gap.

rewards women less than men, scarce resources may be allocated efficiently though inequitably within the household. This question is addressed by estimating returns to schooling attainment for males and females in wage employment in Pakistan using household data from 2002. Four methods are used in an attempt to overcome limitations faced in conventional earnings function analyses: (1) OLS, (2) the Heckman two-step procedure, (2) 2SLS, and (3) household fixed effects. The findings from all four methods consistently point to a sizable gender asymmetry in returns. Females have significantly higher economic incentives to invest in education than males. The estimated return to additional completed years of schooling attainment (EDU_YRS) ranges between 7% and 11% for men and between 13% and 18% for women. By this consideration, the labor market does not explain lower female schooling in Pakistan. If anything, it suggests that there should be a profemale bias in the household decision to educate. However, the Oaxaca decomposition suggests a large element of potential gender discrimination in the Pakistan labor market. While the return to schooling is considerably lower for men than for women, total earnings are dramatically higher for men than for women. While a large part of the male-female earnings differential is not explained by men's and women's differing productive characteristics, one must be cautious in interpreting the residual unexplained earnings differential as labor market discrimination since certain unobserved but relevant characteristics of men and women may not be controlled, such as their quality of schooling (men are more likely than women to have attended private schools), and certain variables may be measured with error, such as the years of labor market experience.

The coexistence of high returns to schooling attainment for women and gender bias against them in household education decisions is a puzzle that demands explanation. One potential explanation is that even if the return to girls' schooling attainment is higher than that to boys' attainment, the part of the return to daughters' attainment accruing to parents may be much lower than that accruing from a son's. The 2002 PIHS shows that only 6% of adult daughters aged over 21 reside in their parental homes, suggesting that a majority are married and living with in-laws/husbands. Any returns from these daughters' schooling attainment would accrue to the in-laws or the husband rather than to the parents. Thus, it would seem that part of the explanation for this puzzle may lie in marriage market considerations. In order to investigate this explanation further, one would need data on transfers received by parents from their male and female offspring. Such data are, to my knowledge, not available. A second explanation for promale education bias, despite such high returns to female education, is that social norms in Pakistan dictate that elderly parents live with their sons. As social security systems in Pakistan are more or less nonexistent, this means that parents invest in sons more as a means of overcoming this flaw in the system. A third explanation for underinvestment in girls' schooling despite higher private returns to education could be that the opportunity cost of schooling for girls may be higher than that for boys. Finally, it may also be that powerful social and cultural (demand-side) factors—such as conservatism of attitudes toward women's education and their labor market work—as well as supply-side constraints limit girls' access to schools and discourage enrollment despite high economic returns to education. The 2002 PIHS shows that school availability continues to be a constraint in rural Pakistan: 34% of the sampled rural communities reported nonavailability of a government girls' primary school (whereas 15% reported no government boys' primary school). Also, many of the aforementioned factors operate simultaneously; for instance, safety concerns often prevent parents from sending daughters to school, and the nonavailability of a single-sex school compounds this constraint further.

Finally, it could be that my estimate of the return to schooling attainment is misleadingly high because it is estimated on the small wage employment sector whereas a relatively large proportion of women in Pakistan are self-employed. Estimating the return to schooling accurately in self-employment is difficult because earnings in self-employment contain a return to physical capital as well and we do not have a good measure of physical capital in order to enable us to isolate the pure return to human capital. Also note that my estimates of private wage returns to schooling attainment for women are likely to suffer even more from the biases associated with sample selection discussed earlier because of the very low proportion of women participating in the wage labor market. However, a recent study by Kingdon and Söderbom (2007) estimates the economic returns to schooling for men and women in wage employment, agricultural self-employment, and nonagricultural self-employment and finds the returns to women's schooling to be larger than those to men's in all occupations.

Hence, this study is unable to explicitly conclude why the gap in years of schooling between males and females in Pakistan is so high despite the private wage returns to schooling attainment favoring women. If we were to focus only on gender differentials in the wage returns to education, we would conclude that women's returns to attainment are higher than men's and, if anything, parents should be allocating more resources to girls than to boys. Since we observe the opposite to be the case, this suggests that parents may have objectives other than maximizing financial returns to family allocation of resources. Given this, one may argue the case for examining social policies in Pakistan that subsidize girls' schooling until the time that family investment in schooling becomes relatively equal for boys and girls. If parents are underemphasizing girls' education because social security systems are lacking, policies may be needed to overcome this institutional failure. Finally, if girls are getting less schooling because the opportunity cost of school time in terms of home production is too high, innovative policies such as night school and monetary incentives may be needed to increase girls' enrollments (Schultz 1995). Note that this study estimates only the private returns to education whereas policy considerations are based also on social returns that include a measure of costs and are, hence, always lower than private returns. The policy implications associated with gender gaps in returns to education might change if gender differentials in education funding reduce the gaps in social returns to education between males and females.

This study also finds sharply convex education-earnings profiles for males and females. These findings are robust to control function estimates. There are several policy implications of convexity of the education-earnings profile. First, the "higher returns at lower education levels" argument has often been used to justify allocating funds to expand primary education. If indeed the returns are greater at higher education levels, the economic efficiency rationale for channeling these funds to primary education may be diluted. However, this is not to say that all rationales for funding primary schooling are eliminated: there is a strong case for primary education in terms of its nonrated returns and also in a rights-based perspective. In any case, the return to primary education includes the benefit that it permits access to further, more lucrative, levels of education. Although the education-earnings profiles in Pakistan are convex, the returns to primary schooling are high compared to those of other developing countries. This may reflect unmet demand within industry sectors that need low-skilled labor, and policy makers may need to promote low-level education as well as adopt policies that encourage these individuals to participate in the labor market (especially women).

Second, and linked to the first, convexity has implications for increasing education inequality. If private returns to schooling increase with higher education, poorer families who educate their children till only, say, primary education will face lower returns whereas richer families who educate children till higher education will reap higher returns. Consequently, the poor are motivated to educate their children less and may also send to school only the more able children, for whom returns are higher. Consequently, education and earnings differentials may widen both across families and within families (Schultz 2004).

Finally, I also find evidence of high wage premia to low education levels, especially for women in Pakistan. These large and significant direct returns to women's schooling attainment at primary and middle levels are not fully explained by occupation and industry attainment and are interpreted to reflect scarcity premia in labor markets.

Appendix

TABLE A1
INSTRUMENTAL VARIABLE EARNINGS FUNCTION ESTIMATES (PARENTAL EDUCATION) CONTROLLING FOR SELECTION INTO WAGED WORK (MALES AND FEMALES)

	Males (15–65)		Females (15–65)	
	IV (1)	First Stage (2)	IV (1)	First Stage (2)
CONSTANT	5.164***	74.762***	3.777**	27.318***
	(.31)	(.92)	(1.48)	(1.09)
EDU_YRS	.090***	…	.164***	…
	(.00)		(.05)	
EXP	.120***	−2.916***	.154***	−.676***
	(.01)	(.04)	(.04)	(.08)
EXP2	−.002***	.075***	−.003***	.012***
	(.00)		(.00)	(.00)
SINDH	.322***	−8.493***	.258*	−1.377***
	(.05)	(.17)	(.15)	(.24)
NWFP	−.098**	.162***	−.074	5.975***
	(.04)	(.13)	(.37)	(.33)
BALOCHISTAN	.589***	−9.611***	.752**	2.876***
	(.06)	(.19)	(.31)	(.56)
AJK	.128*	−5.749***	.136	3.079***
	(.08)	(.21)	(.24)	(.57)
NORTH	−.041	14.909***	…	…
	(.10)	(.43)		
FATA	.051	4.320***	…	…
	(.08)	(.51)		
URBAN	.251***	−11.407***	.360**	.582**
	(.05)	(.19)	(.15)	(.29)
MARRIED	.133***	−1.702***	.274	.309
	(.03)	(.10)	(.20)	(.63)
LNUNEARNED_Y	.002	−.127***	.013	.067***
	(.00)	(.01)	(.01)	(.03)
CHILD7	−.010	1.557***	.018	−.013
	(.01)	(.03)	(.04)	(.08)
ADULT60	−.010	1.590***	−.034	.091
	(.02)	(.06)	(.09)	(.16)
IMR	.587***	−48.355***	.064	−11.125***
	(.20)	(.66)	(.61)	(.44)
FEDYRS	…	.801***	…	.147***
		(.01)		(.03)
MEDPRIM	…	2.147***	…	2.344***
		(.15)		(.32)
MEDPRIMORE	…	3.250***	…	3.164***
		(.16)		(.30)
R^2	.300	.843	.463	.874
N	4,155	4,155	493	493

TABLE A1 (*Continued*)

	Males (15–65)		Females (15–65)	
	IV (1)	First Stage (2)	IV (1)	First Stage (2)
F-test of excluding instruments	…	1,623.22	…	78.34
p-value (*F*-test)	…	.000	…	.000
OID test	…	4.533	…	2.170
p-value (OID test)	…	.104	…	.334

Note. Standard errors are in parentheses and are robust and corrected for clustering at the PSU level. The dependent variable is LN_MONTHLY_Y. Ellipses indicate no observations or not used. PUNJAB is the base category for provinces and MEDNONE (= 1 if mother has no education, 0 otherwise) is the base for mothers' educational dummies; MEDPRIM = 1 if mother has primary or less education (but more than zero) and 0 otherwise; MEDPRIMORE = 1 if mother has more than primary education, 0 otherwise. IMR is the inverse Mills ratio estimated from selection probit separately for males and females.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

References

Alderman, H., and E. M. King. 1998. "Gender Differences in Parental Investment in Education." *Structural Change and Economic Dynamics* 9:453–68.

Alderman, H., and D. E. Sahn. 1988. "The Effects of Human Capital on Wages, and the Determinants of Labor Supply in a Developing Country." *Journal of Development Economics* 29:157–83.

Angrist, J. D., and A. B. Krueger. 1991. "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106, no. 4: 979–1014.

Asadullah, M. N. 2005. "The Effectiveness of Private and Public Schools in Bangladesh and Pakistan." Photocopy, University of Oxford.

———. 2006. "Returns to Education in Bangladesh." *Education Economics* 14, no. 4:453–68.

Ashenfelter, O., and C. Rouse. 1998. "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113, no. 1:253–84.

Ashraf, J., and B. Ashraf. 1993a. "An Analysis of the Male-Female Earning Differential in Pakistan." *Pakistan Development Review* 32, no. 4:895–902.

———. 1993b. "Estimating the Gender Wage Gap in Rawalpindi City." *Journal of Development Studies* 29, no. 2:365–76.

Aslam, M., and G. G. Kingdon. 2008. "Gender and Household Education Expenditure in Pakistan." *Applied Economics* 40:2573–91.

Behrman, J. R., and A. B. Deolalikar. 1995. "Are There Differential Returns to Schooling by Gender? The Case of Indonesian Labour Markets." *Oxford Bulletin of Economics and Statistics* 57, no. 1:97–117.

Behrman, J. R., J. Hoddinott, J. A. Maluccio, E. Soler-Hampejsek, E. L. Behrman, R. Martorell, M. Ramirez, and A. D. Stein. 2006. "What Determines Adult

Skills? Impacts of Pre-school, School-Years and Post-school Experiences in Guatemala." Photocopy, University of Pennsylvania.

Behrman, J. R., and M. R. Rosenzweig. 1999. "'Ability' Biases in Schooling Returns and Twins: A Test and New Estimates." *Economics of Education Review* 18, no. 2: 159–67.

———. 2002. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *American Economic Review* 92 (March): 323–34.

———. 2005. "Does Increasing Women's Schooling Raise the Schooling of the Next Generation? Reply." *American Economic Review* 95 (December): 1745–51.

Behrman, J. R., M. R. Rosenzweig, and P. Taubman. 1994. "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment." *Journal of Political Economy* 102, no. 6:1131–74.

Behrman, J. R., and B. Wolfe. 1984. "The Socio-economic Impact of Schooling in a Developing Country." *Review of Economics and Statistics* 66, no. 2:296–303.

Belzil, C., and J. Hansen. 2002. "Unobserved Ability and the Return to Schooling." *Econometrica* 70, no. 5:2075–91.

Björklund, A., M. Lindahl, and E. Plug. 2006. "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data." *Quarterly Journal of Economics* 121, no. 3:999–1028.

Black, S. E., P. J. Devereux, and K. G. Salvanes. 2005. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." *American Economic Review* 95 (March): 437–49.

Blundell, R., L. Dearden, and B. Sianesi. 2005. "Evaluating the Impact of Education on Earnings in the UK: Models, Methods and Results from the NCDS." *Journal of the Royal Statistical Society*, ser. A, 168, no. 3:473–512.

Bound, J., D. A. Jaegar, and R. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak." *Journal of the American Statistical Association* 90, no. 430:443–50.

Butcher, K., and A. Case. 1994. "The Effect of Sibling Sex Composition on Women's Education and Earnings." *Quarterly Journal of Economics* 109, no. 3:531–63.

Card, D. 1995a. "Earnings, Schooling, and Ability Revisited." In *Research in Labor Economics*, vol. 14, ed. Solomon Polachek, 23–48. Greenwich, CT: JAI.

———. 1995b. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behaviour: Essays in Honour of John Vandercamp*, ed. L. N. Christofides, E. Kenneth Grant, and R. Swidinsky, 201–22. Toronto: University of Toronto Press.

———. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3A, ed. O. Ashenfelter and D. Card, 1801–63. Amsterdam: North-Holland.

———. 2001. "Estimating the Returns to Schooling: Progress in Some Persistent Econometric Problems." *Econometrica* 69, no. 5:1127–60.

Choudhury, S. 1993. "Reassessing the Male-Female Wage Differential: A Fixed-Effects Approach." *Southern Economic Journal* 60, no. 2:327–40.

De Haan, M., and E. Plug. 2006. "Estimates of the Effect of Parents' Schooling on Children's Schooling Using Censored and Uncensored Samples." Unpublished manuscript, Research Center for Education and Labor, School of Economics, University of Amsterdam.

Duflo, E. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91, no. 4:795–813.

Duraisamy, P. 2002. "Changes in Returns to Education in India, 1983–94: By Gender, Age-Cohort and Location." *Economics of Education Review* 21, no. 6:609–22.

Griliches, Z. 1979. "Sibling Models and Data in Economics: Beginnings of a Survey." *Journal of Political Economy* 87, no. 5, pt. 2:S37–S64.

Guisinger, S. E., J. W. Henderson, and G. W. Scully. 1984. "Earnings, Rate of Returns to Education and Earning Distribution in Pakistan." *Economics of Education Review* 3, no. 4:257–67.

Hamdani, K. 1977. "Education and the Income Differential: An Estimation for Rawalpindi City." *Pakistan Development Review* 16, no. 2:144–64.

Haque, Nadeemul. 1977. "An Economic Analysis of Personal Earnings in Rawalpindi City." *Pakistan Development Review* 16, no. 4:353–82.

Harmon, C., and I. Walker. 1995. "Estimates of the Economic Return to Schooling for the United Kingdom." *American Economic Review* 85, no. 5:1278–86.

Heckman, J. J., and V. J. Hotz. 1989. "An Investigation of the Labour Market Earnings of Panamanian Males: Evaluating the Sources of Inequality." *Journal of Human Resources* 21:507–63.

Hertz, T. 2003. "Upward Bias in the Estimated Returns to Education: Evidence from South Africa." *American Economic Review* 93, no. 4:1354–68.

Khan, S. R., and M. Irfan. 1985. "Rate of Returns to Education and Determinants of Earnings in Pakistan." *Pakistan Development Review* 34, no. 3:671–80.

Kingdon, G. G. 1998. "Does the Labour Market Explain Lower Female Schooling in India?" *Journal of Development Studies* 35, no. 1:39–65.

Kingdon, G. G., and M. Söderbom. 2007. "Education, Skills and Labor Market Outcomes: Evidence from Pakistan." Photocopy, University of Oxford.

Kingdon, G. G., and J. Unni. 2001. "Education and Women's Labour Market Outcomes in India." *Education Economics* 9, no. 2:173–95.

Krishnan, P. 1996. "Family Background, Education and Employment in Urban Ethiopia." *Oxford Bulletin of Economics and Statistics* 58, no. 1:167–83.

Lam, D., and R. F. Schoeni. 1993. "Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil." *Journal of Political Economy* 101, no. 4:710–40.

Maluccio, J. 1998. "Endogeneity of Schooling in the Wage Function: Evidence from the Rural Philippines." Working Paper no. 54, Food Consumption and Nutrition Division, International Food Policy Research Institute, Washington, DC.

Mincer, J. 1974. *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.

Nasir, Z. M. 1998. "Determinants of Personal Earnings in Pakistan: Findings from the Labour Force Survey 1993–1994." *Pakistan Development Review* 37:251–74.

———. 1999. "Do Private Schools Make Workers More Productive?" *Pakistan Development Review* 38:937–64.

———. 2002. "Returns to Human Capital in Pakistan: A Gender Disaggregated Analysis." *Pakistan Development Review* 41, no. 1:1–28.

Oaxaca, R. 1973. "Male-Female Differentials in Urban Labour Markets." *International Economic Review* 14, no. 3:693–709.

Pitt, M. M., and M. R. Rosenzweig. 1990. "Estimating the Intrahousehold Incidence of Illness: Child Health and Gender Inequality in the Allocation of Time." *International Economic Review* 31, no. 4:969–89.

Plug, E. 2004. "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees." *American Economic Review* 94 (March): 358–68.

Plug, E., and W. P. M. Vijverberg. 2003. "Schooling, Family Background and Adoption: Is It Nature or Nurture?" *Journal of Political Economy* 111:611–41.

Psacharopoulos, G. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22, no. 9:1325–43.

Psacharopoulos, G., and H. Patrinos. 2004. "Returns to Investment in Education: A Global Update." *Education Economics* 12, no. 2:111–34.

Riboud, M.,Y. Savchenko, and H. Tan. 2006. "The Knowledge Economy and Education and Training in South Asia: A Mapping Exercise of Available Survey Data." Working paper, South Asia Region, World Bank, Washington, DC.

Rouse, C. E. 1999. "Further Estimates of the Economic Return to Schooling from a New Sample of Twins." *Economics of Education Review* 18:149–57.

Schultz, T. P. 1993. "Returns to Women's Education." In *Women's Education in Developing Countries: Barriers, Benefits, and Policies*, ed. Elizabeth M. King and M. Anne Hill, chap. 2. Baltimore: Johns Hopkins University Press.

———, ed. 1995. *Investment in Women's Human Capital*. Chicago: University of Chicago Press.

———. 2004. "Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education." *Journal of African Economies* 13 (suppl. 2, African Economic Research Consortium): ii95–ii148.

Shabbir, T. 1991. "Sheepskin Effects in the Returns to Education in a Developing Country." *Pakistan Development Review* 30, no. 1:1–19.

———. 1994. "Mincerian Earnings Functions for Pakistan." *Pakistan Development Review* 33, no. 1:1–18.

Shabbir, T., and A. H. Khan. 1991. "Mincerian Earnings Functions for Pakistan: A Regional Analysis." *Pakistan Economic and Social Review* 29, no. 2:99–112.

Siddiqui, R., and R. Siddiqui. 1998. "A Decomposition of Male-Female Earnings Differentials." *Pakistan Development Review* 37, no. 4:885–98.

Söderbom, M., F. Teal, G. Kahyarara, and A. Wambugu. 2004. "The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing." Working Paper no. 2003-17, Center for the Study of African Economies, Department of Economics, University of Oxford.

Strauss, J., and D. Thomas. 1995. "Human Resources: Empirical Modelling of Household and Family Decisions." In *Handbook of Development Economics*, vol. 3, ed. J. Behrman and T. N. Srinivasan, 1883–2023. Amsterdam: Elsevier Science.

Trostel, P., I. Walker, and P. Woolley. 2002. "Estimates of the Economic Return to Schooling for 28 Countries." *Labour Economics* 9:1–16.

Weiss, Y. 1999. "The Formation and Dissolution of Families: Why Marry? Who

Marries Whom? And What Happens upon Marriage and Divorce?" In *Handbook of Population and Family Economics*, ed. M. R. Rosenzweig and O. Stark. Amsterdam: North-Holland.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press.

Zabalza, A., and J. Arrufat. 1983. "Wage Differentials between Married Men and Women in Great Britain: The Depreciation Effect of Non-participation." Working Paper no. 382, London School of Economics.