



HAL
open science

Gender and Household Education Expenditure in Pakistan

Monazza Aslam, Geeta Kingdon

► **To cite this version:**

Monazza Aslam, Geeta Kingdon. Gender and Household Education Expenditure in Pakistan. *Applied Economics*, 2008, 40 (20), pp.2573-2592. 10.1080/00036840600970252 . hal-00582059

HAL Id: hal-00582059

<https://hal.science/hal-00582059v1>

Submitted on 1 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Gender and Household Education Expenditure in Pakistan

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-06-0051.R1
Journal Selection:	Applied Economics
JEL Code:	I21 - Analysis of Education < I2 - Education < I - Health, Education, and Welfare, J71 - Discrimination < J7 - Discrimination < J - Labor and Demographic Economics, J16 - Economics of Gender < J1 - Demographic Economics < J - Labor and Demographic Economics
Keywords:	gender bias, educational expenditure, engel curve, pakistan

powered by ScholarOne
Manuscript Central™

Table 1
Descriptive statistics, by Province

PROVINCE	Proportion of 'girls-only' households (age 0-14)	Mean budget share of education in 'girls-only' households (age 5-24)	Mean budget share of education in 'boys-only' households (age 5-24)	Mean budget share of education in all households (age 5-24)	Average household size (girls)	Average household size (boys)	t-value of difference in (b) and (c)	t-value of difference in (e) and (f)
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
PAKISTAN	14.5	2.0	4.1	4.6	9.4	9.3	-11.8	1.1
PUNJAB	16.7	2.5	4.6	5.5	8.4	8.2	-7.1	1.7
SINDH	13.6	1.9	3.8	3.8	9.8	9.7	-5.4	1.3
NWFP	12.6	1.6	4.5	4.6	10.2	10.2	-6.8	0.6
BALUCHISTAN	13.3	0.6	1.8	2.4	9.9	9.7	-4.6	2.0
AJK	14.0	3.4	6.8	7.9	8.3	8.4	-3.9	-1.0
NORTH	12.1	3.1	6.8	6.9	9.5	9.2	-2.3	2.2
FATA	11.0	0.1	1.6	1.7	10.9	10.8	-3.0	0.6

Note 1: Shaded cells represent significance at 10 % or better

Table 2
Current enrolment rate, by age and gender

PROVINCE	Age 5-9				Age 10-14				Age 15-19				Age 20-24			
	M	F	Gap		M	F	Gap		M	F	Gap		M	F	Gap	
PUNJAB	66	61	5	***	69	58	11	***	36	27	9	***	8	6	2	***
SINDH	52	39	13	***	58	39	19	***	31	18	13	***	9	5	4	***
NWFP	63	45	18	***	79	42	37	***	48	18	30	***	13	5	8	***
BALUCHISTAN	46	31	15	***	65	37	28	***	35	15	20	***	10	3	7	***
AJK	84	72	12	***	91	81	10	***	60	40	20	***	15	7	8	**
NORTHERN AREAS	54	47	7	*	91	70	21	***	74	40	34	***	21	6	15	***
FATA	40	5	35	***	61	2	59	***	19	2	17	***	4	1	3	
PAKISTAN	59	47	12	***	69	49	20	***	38	22	16	***	10	5	5	***

Note 1: *** depicts significance at the 1% level, ** significance at 5 % and * significance at 10%.

Note 2: M denoted 'male' and F denotes 'female'.

Note 3: NWFP = North West Frontier Province, AJK = Azad Jammu and Kashmir and FATA = Federally Administered Tribal Areas.

Table 3
Annual educational expenditure on ALL children and enrolled children only, by age and gender

PROVINCE	Age 5-9			Age 10-14			Age 15-19			Age 20-24		
	M	F	t	M	F	t	M	F	t	M	F	t
ALL (enrolled and non-enrolled)												
PUNJAB	1007	919	2.00	1456	1253	3.61	1499	1045	5.74	611	356	3.40
SINDH	859	762	1.45	1213	1041	2.00	1296	861	4.25	608	318	3.47
NWFP	852	561	5.12	1442	712	10.24	1556	554	10.61	863	226	5.50
BALUCHISTAN	508	280	5.37	813	476	6.25	783	302	6.95	331	74	5.07
AJK	1887	1363	3.33	2590	1840	3.90	2474	1435	4.20	1153	491	2.50
NORTH AREAS	759	559	2.30	1578	1066	4.32	1775	1042	2.98	467	184	1.66
FATA ¹	356	54	5.25	744	10	8.39	577	0	3.92	218	0	1.52
PAKISTAN	874	709	6.34	1338	997	10.27	1389	820	12.79	618	284	7.98
ENROLLED ONLY												
PUNJAB	1535	1503	0.51	2126	2166	-0.51	4208	3878	1.76	7457	6259	1.71
SINDH	1645	1988	-2.51	2083	2306	-3.83	4189	4843	-1.99	7053	6751	1.03
NWFP	1362	1263	0.99	1844	1708	1.19	3285	3039	0.95	6446	4377	2.04
BALUCHISTAN	1109	916	1.91	1262	1284	-0.23	2226	2086	0.59	3212	2874	0.60
AJK	2239	1890	1.90	2843	2292	2.62	4124	3744	0.98	7448	6568	0.71
NORTH AREAS	1421	1215	1.49	1743	1522	1.65	2409	2616	-0.52	2246	3583	-1.04
FATA	919	1228	-0.92	1218	421	1.33	3086	-	-	6040	-	-
PAKISTAN	1495	1513	-0.39	1941	2063	-2.38	3629	3695	-0.57	6260	5646	1.54

Note 1: M denotes 'male' and F denotes 'female'; t depicts the t value. All cells where the gender difference is significant at the 10% level or better are shaded.

Note 2: FATA contains no observations for enrolled girls in the 15-19 and 20-24 age categories.

¹ Despite Table 1 revealing a current enrolment of 2 per cent for females in FATA, the 3 observations on currently enrolled females in the FATA sub-samples reported educational expenditures of 0 in the 10-14 age group.

Table 4A
Difference in Marginal Effects (DME) x 100 of gender variables males5-9 and females5-9,
and p value of the associated test (HH level results)

Province	Sample Size	Probit of ANYEDEXP	Conditional OLS of EDU_SHARE	Combined Probit + Conditional OLS	Unconditional OLS (Conventional Engel Curve)
		(a)	(b)	(c) = f(a,b)	(d)
PAKISTAN	Full	37.35 (0.00)	5.09 (0.46)	2.50 (0.00)	1.77 (0.00)
	Urban	20.84 (0.01)	-0.84 (0.51)	0.91 (0.01)	0.82 (0.50)
	Rural	44.57 (0.00)	1.14 (0.15)	2.85 (0.00)	2.41 (0.00)
Full Sample					
PUNJAB		12.97 (0.19)	1.14 (0.37)	1.78 (0.16)	1.35 (0.19)
SINDH		53.54 (0.00)	-1.32 (0.29)	1.78 (0.08)	1.37 (0.18)
NWFP		60.06 (0.00)	-0.36 (0.77)	2.66 (0.03)	1.53 (0.27)
BALUCHISTAN		61.33 (0.00)	0.73 (0.72)	2.96 (0.01)	1.40 (0.09)
AJK		27.84 (0.04)	3.40 (0.22)	5.51 (0.09)	3.72 (0.24)
NORTH		-18.65 (0.13)	2.92 (0.59)	1.16 (0.81)	4.45 (0.19)
FATA		67.37 (0.17)	7.29 (0.00)	4.51 (0.01)	3.80 (0.01)

Note: The figures in parentheses are p-values of the t-test of the DME and the shaded cells represent significance at 5%. The DME in the conditional OLS equation in Column (b) were transformed as the dependent variable of the conditional OLS equation is the natural log of the budget share of education for the household while the dependent variable in (d) is the budget share of education. Column (b) reports results after transforming the dependent variable of the conditional into absolute terms. The DME have been multiplied by 100. The standard errors of the t-test in column (c) were obtained using bootstrapping in STATA.

Table 4B
Difference in Marginal Effects (DME) x 100 of gender variables males10-14 and females10-14,
and p value of the associated test (HH level results)

Province	Sample Size	Probit of ANYEDEXP	Conditional OLS of EDU_SHARE	Combined Probit + Conditional OLS	Unconditional OLS (Conventional Engel Curve)
		(a)	(b)	(c) = f(a,b)	(d)
PAKISTAN	Full	60.15 (0.00)	3.30 (0.00)	5.80 (0.00)	3.22 (0.00)
	Urban	4.46 (0.61)	2.71 (0.06)	2.63 (0.06)	1.28 (0.42)
	Rural	91.40 (0.00)	3.68 (0.00)	6.70 (0.00)	5.57 (0.00)
Full Sample					
PUNJAB		36.42 (0.00)	1.72 (0.22)	3.90 (0.00)	1.53 (0.28)
SINDH		12.39 (0.00)	2.82 (0.04)	4.63 (0.00)	1.04 (0.56)
NWFP		94.42 (0.00)	6.58 (0.00)	9.85 (0.00)	7.49 (0.00)
BALUCHISTAN		101.90 (0.00)	1.43 (0.63)	5.04 (0.00)	3.61 (0.00)
AJK		25.80 (0.03)	9.28 (0.00)	11.02 (0.00)	10.10 (0.01)
NORTH		-3.43 (0.83)	4.21 (0.39)	3.68 (0.49)	6.01 (0.16)
FATA		193.44 (0.00)	9.22 (0.00)	8.21 (0.00)	8.41 (0.01)

Note: See note in Table 4A.

Table 4C
Difference in Marginal Effects (DME) x 100 of gender variables males15-19 and females15-19,
and p value of the associated test (HH level results)

Province	Sample Size	Probit of ANYEDEXP	Conditional OLS of EDU_SHARE	Combined Probit + Conditional OLS	Unconditional OLS (Conventional Engel Curve)
		(a)	(b)	(c) = f(a,b)	(d)
PAKISTAN	Full	24.70 (0.00)	3.39 (0.00)	3.84 (0.00)	3.13 (0.00)
	Urban	8.85 (0.23)	0.60 (0.72)	1.19 (0.45)	0.20 (0.90)
	Rural	36.40 (0.00)	5.21 (0.00)	5.04 (0.00)	5.57 (0.00)
Full Sample					
PUNJAB		15.31 (0.11)	2.80 (0.14)	3.18 (0.04)	3.03 (0.07)
SINDH		48.50 (0.00)	3.12 (0.04)	4.32 (0.00)	1.94 (0.28)
NWFP		19.31 (0.15)	5.02 (0.00)	4.34 (0.00)	4.76 (0.01)
BALUCHISTAN		15.23 (0.43)	5.60 (0.02)	3.82 (0.05)	1.68 (0.21)
AJK		6.09 (0.44)	5.38 (0.08)	5.67 (0.12)	7.22 (0.06)
NORTH		-45.69 (0.01)	-9.86 (0.18)	-5.40 (0.54)	3.14 (0.48)
FATA		53.87 (0.50)	10.25 (0.08)	5.41 (0.16)	8.05 (0.17)

Note: See note in Table 4A.

Table 5A
Marginal effect of the gender dummy variable MALE and p value of the associated t-test,
age group 5-9 (Individual-level results)

Province	Sample Size	Probit of ANYEDEXP	Conditional OLS of TOTAL_EDU	Combined Probit + Conditional OLS	Unconditional OLS
		(a)	(b)	(c) = f(a,b)	(d)
PAKISTAN	Full	0.142 (0.00)	96.84 (0.00)	195.47 (0.00)	174.20 (0.00)
	Urban	0.067 (0.00)	143.45 (0.00)	216.42 (0.00)	215.20 (0.00)
	Rural	0.169 (0.00)	83.95 (0.00)	168.19 (0.00)	161.00 (0.00)
Full Sample					
PUNJAB		0.059 (0.00)	93.19 (0.00)	130.31 (0.00)	118.6 (0.00)
SINDH		0.171 (0.00)	20.53 (0.52)	160.54 (0.00)	96.0 (0.01)
NWFP		0.192 (0.00)	-24.67 (0.89)	102.09 (0.99)	277.0 (0.00)
BALUCHISTAN		0.170 (0.00)	76.20 (0.17)	159.75 (0.00)	210.0 (0.00)
AJK		0.119 (0.00)	389.64 (0.00)	536.79 (0.00)	436.8 (0.00)
NORTH		0.072 (0.07)	-24.67 (0.89)	102.10 (0.99)	187.4 (0.02)
FATA		0.346 (0.00)	-19.70 (0.91)	-	286.5 (0.00)

Note: The figures in parentheses are p-values of the t-test of the DME of the MALE dummy computed using MALE =1 and MALE =0 and the shaded cells represent significance at 5%. The DME in the conditional OLS equation in Column (b) were transformed as the dependent variable of the conditional OLS equation fitted only on positive expenditure households is the natural log of total expenditure on education for the household while the dependent variable in (d) is the absolute value of total educational expenditure. Column (b) reports results after transforming the dependent variable of the conditional into absolute terms. The standard errors for the t-test in column (c) were obtained by bootstrapping in STATA.

Table 5B
Marginal effect of the gender dummy variable MALE and p value of the associated t-test,
age group 10-14 (Individual-level results).

Province	Sample Size	Probit of ANYEDEXP	Conditional OLS of TOTAL_EDU	Combined Probit + Conditional OLS	Unconditional OLS
		(a)	(b)	(c) = f(a,b)	(d)
PAKISTAN	Full	0.261 (0.00)	174.84 (0.00)	498.88 (0.00)	380.6 (0.00)
	Urban	0.653 (0.00)	263.73 (0.00)	343.65 (0.00)	261.7 (0.00)
	Rural	0.368 (0.00)	136.62 (0.00)	510.91 (0.00)	440.7 (0.00)
Full Sample					
PUNJAB		0.125 (0.00)	49.88 (0.32)	249.13 (0.00)	223.5 (0.00)
SINDH		0.277 (0.00)	199.76 (0.00)	474.34 (0.00)	241.9 (0.00)
NWFP		0.399 (0.00)	428.69 (0.00)	780.06 (0.00)	708.0 (0.00)
BALUCHISTAN		0.375 (0.00)	31.04 (0.74)	429.94 (0.00)	388.1 (0.00)
AJK		0.115 (0.00)	630.47 (0.00)	826.21 (0.00)	763.0 (0.00)
NORTH		0.207 (0.00)	19.40 (0.96)	505.87 (0.06)	536.3 (0.00)
FATA		0.606 (0.00)	-	-	727.6 (0.00)

Note: See note in Table 5A.

Table 5C
Marginal effect of the gender dummy variable MALE and p value of the associated t-test,
age group 15-19 (Individual-level results).

Province	Sample Size	Probit of ANYEDEXP	Conditional OLS of TOTAL_EDU	Combined Probit + Conditional OLS	Unconditional OLS
		(a)	(b)	(c) = f(a,b)	(d)
PAKISTAN	Full	0.192 (0.00)	375.39 (0.00)	613.98 (0.00)	583.3 (0.00)
	Urban	0.087 (0.00)	319.59 (0.00)	398.91 (0.00)	394.7 (0.00)
	Rural	0.239 (0.00)	445.45 (0.00)	671.03 (0.00)	699.4 (0.00)
Full Sample					
PUNJAB		0.093 (0.00)	400.96 (0.00)	390.11 (0.00)	390.7 (0.00)
SINDH		0.168 (0.00)	498.90 (0.00)	518.64 (0.00)	478.0 (0.00)
NWFP		0.316 (0.00)	656.29 (0.00)	931.55 (0.00)	960.8 (0.00)
BALUCHISTAN		0.244 (0.00)	476.30 (0.01)	539.81 (0.00)	547.1 (0.00)
AJK		0.219 (0.00)	202.20 (0.54)	828.24 (0.00)	883.7 (0.00)
NORTH		0.380 (0.00)	-186.22 (0.03)	567.81 (0.55)	536.3 (0.00)
FATA		-	-	-	531.8 (0.02)

Note: See note in Table 5A.

Table 6
Household Fixed Effects: Coefficient of the gender dummy variable MALE and associated t-test by age group (Individual level data).

		Age5-9			Age10-14			Age15-19		
		Probit ANY-EDEXP	Conditional OLS of LNTOTAL_EDU	Unconditional OLS of TOTAL_EDU	Probit ANY-EDEXP	Conditional OLS of LNTOTAL_EDU	Unconditional OLS of TOTAL_EDU	Probit ANY-EDEXP	Conditional OLS of LNTOTAL_EDU	Unconditional OLS of TOTAL_EDU
Pakistan	Full sample	0.135 (17.43)	0.145 (9.01)	200.06 (12.68)	0.237 (26.27)	0.172 (7.49)	462.15 (16.68)	0.164 (17.31)	0.182 (3.95)	551.57 (11.64)
	Urban	0.068 (5.10)	0.138 (6.07)	253.42 (5.93)	0.083 (5.87)	0.128 (4.07)	342.45 (5.80)	0.072 (4.63)	0.193 (3.37)	344.69 (4.07)
	Rural	0.164 (17.30)	0.147 (6.53)	177.47 (13.42)	0.328 (28.95)	0.207 (6.20)	532.85 (19.84)	0.229 (19.67)	0.137 (1.77)	697.89 (12.95)
		Full Sample								
Punjab		0.063 (4.39)	0.104 (3.61)	122.87 (4.69)	0.108 (7.06)	0.106 (3.31)	241.19 (5.02)	0.083 (5.16)	0.073 (1.07)	305.01 (3.45)
Sindh		0.159 (10.14)	0.100 (3.83)	159.80 (5.59)	0.233 (12.21)	0.179 (4.44)	368.96 (6.82)	0.135 (4.40)	0.156 (1.78)	515.23 (6.16)
NWFP		0.179 (9.40)	0.195 (4.20)	200.81 (6.61)	0.368 (5.61)	0.325 (5.59)	771.12 (11.07)	0.272 (12.12)	0.267 (2.25)	863.91 (7.52)
Balochistan		0.124 (6.50)	0.151 (4.08)	167.84 (4.15)	0.350 (7.15)	0.097 (1.32)	472.34 (9.82)	0.262 (7.71)	0.368 (1.11)	342.98 (4.30)
AJK		0.127 (3.04)	0.078 (4.51)	349.46 (3.66)	0.12 (3.37)	0.219 (2.03)	619.25 (3.25)	0.178 (3.11)	0.334 (2.42)	888.27 (2.75)
NORTH		0.112 (2.51)	0.282 (2.82)	299.07 (3.81)	0.272 (4.51)	0.260 (1.71)	665.92 (5.57)	0.221 (3.25)	0.358 (2.28)	559.20 (2.78)

* t-values in parentheses and shaded cells denote significance at the 5 % level or more.

Appendix Table 1
OLS on Budget share, Probit and Conditional OLS, Pakistan

	FULL						URBAN						RURAL						
	EDU_SHARE		ANYEDEXP		LN_EDUSHARE		EDU_SHARE		ANYEDEXP		LN_EDUSHARE		EDU_SHARE		ANYEDEXP		LN_EDUSHARE		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	
Coeff	t-val	ME	t-val	Coeff	t-val	Coeff	t-val	ME	t-val	Coeff	t-val	Coeff	t-val	ME	t-val	Coeff	t-val	Coeff	t-val
CONSTANT	-6.73	-1.39	-	-	-6.66	-8.47	-43.28	-5.06	-	-	-10.60	-8.22	-0.56	-0.07	-	-	-6.97	-5.03	
LNPCE	-0.48	-0.39	0.51	7.31	0.53	2.93	7.25	3.66	0.60	6.16	-0.06	-5.01	-1.55	-0.77	0.65	4.29	-0.65	-1.90	
LNPCE2	0.15	2.00	-0.02	-5.07	-0.02	-1.47	-0.26	-2.19	-0.03	-4.90	0.13	-3.97	0.19	1.47	-0.03	-3.26	-0.03	-1.32	
LNHHSIZE	2.15	18.72	0.40	32.40	0.20	6.10	2.71	12.28	0.30	17.91	0.99	2.81	2.05	16.57	0.44	24.41	0.26	5.88	
M0TO4	-1.26	-1.28	-0.06	-0.65	-0.72	-2.99	-2.73	-1.31	-0.02	-0.15	0.82	-2.55	-0.84	-0.82	-0.10	-0.86	-0.57	-1.88	
M5TO9	7.95	7.56	1.14	12.06	0.69	2.96	10.96	5.19	0.97	7.52	1.46	2.25	6.43	5.67	1.15	9.41	0.60	2.05	
M10TO14	12.42	11.88	1.24	13.18	1.63	6.72	14.24	6.51	0.80	6.61	0.72	4.11	11.11	10.27	1.43	11.74	1.75	5.62	
M15TO19	8.57	7.78	0.46	4.84	1.18	4.91	8.60	3.90	0.36	2.88	-0.06	1.97	8.88	7.57	0.51	4.26	1.64	5.23	
M20TO24	1.10	0.88	-0.18	-1.87	0.04	0.14	1.30	0.56	0.05	0.41	-0.63	-0.15	1.61	1.19	-0.38	-3.05	0.18	0.50	
M25TO60	-1.04	-0.93	-0.27	-2.72	-0.49	-1.87	-2.21	-1.02	-0.14	-1.12	-0.21	-1.66	-0.15	-0.12	-0.34	-2.59	-0.35	-0.96	
M60MORE	-0.75	-0.51	-0.22	-1.76	0.08	0.26	-1.87	-0.63	-0.15	-0.94	-0.71	-0.44	-0.33	-0.21	-0.27	-1.62	0.23	0.53	
F0TO4	-1.04	-1.06	-0.09	-0.94	-0.68	-2.77	-1.85	-0.88	-0.04	-0.32	0.93	-1.94	-0.61	-0.62	-0.12	-0.98	-0.64	-2.02	
F5TO9	6.18	5.87	0.76	8.08	0.60	2.54	10.13	4.69	0.76	5.82	1.11	2.56	4.02	3.64	0.71	5.71	0.37	1.23	
F10TO14	9.20	8.48	0.64	6.73	1.06	4.61	12.96	5.96	0.76	6.11	0.64	3.13	6.51	5.68	0.52	4.17	0.98	3.36	
F15TO19	5.44	4.98	0.21	2.23	0.59	2.47	8.40	3.77	0.27	2.21	0.41	1.77	3.31	2.92	0.15	1.21	0.55	1.77	
F20TO24	2.56	2.44	-0.31	-3.24	0.33	1.32	2.04	0.94	-0.19	-1.61	0.07	1.14	2.49	2.37	-0.34	-2.72	0.21	0.61	
F25TO60	2.07	2.17	0.38	4.29	0.33	1.29	2.29	1.11	0.27	2.34	0.11	0.19	1.72	1.81	0.41	3.48	0.47	1.41	
HEAD_FEMALE	1.28	4.34	0.08	3.46	0.20	4.30	0.94	1.62	0.04	1.37	0.01	1.46	1.60	4.65	0.10	3.37	0.26	4.07	
HEAD_MARITAL	0.06	0.37	0.03	1.90	-0.05	-1.45	0.02	0.05	0.01	0.66	-0.44	-0.12	0.08	0.41	0.04	2.01	-0.07	-1.28	
HEAD_EDU_MISS	-3.43	-11.55	-0.26	-12.40	-0.50	-12.03	-3.52	-8.09	-0.23	-9.64	-0.34	-8.43	-2.99	-7.49	-0.26	-8.08	-0.51	-7.51	
HEAD_PRIMARY	-2.70	-9.05	-0.12	-4.94	-0.36	-9.07	-2.74	-6.42	-0.08	-3.37	-0.14	-6.89	-2.18	-5.41	-0.11	-3.21	-0.33	-5.14	
HEAD_MIDDLE	-1.67	-5.16	-0.07	-2.78	-0.17	-3.99	-1.46	-3.07	-0.06	-2.27	-0.15	-2.89	-1.44	-3.47	-0.06	-1.44	-0.15	-2.19	
HEAD_MATRIC	-1.32	-4.42	-0.05	-2.18	-0.15	-3.85	-1.45	-3.50	-0.02	-1.13	0.10	-3.26	-0.78	-1.89	-0.06	-1.66	-0.08	-1.34	
HEAD_OCCU_MISS	0.32	2.09	0.04	3.23	0.06	1.86	0.50	1.61	0.02	1.23	0.17	1.76	0.37	2.25	0.05	2.61	0.07	1.75	
HEAD_WHITE_COL	1.03	3.97	0.11	6.17	0.07	1.84	1.68	4.45	0.06	3.14	0.13	3.29	0.47	1.37	0.15	5.63	0.00	0.07	
HEAD_SERVICE	0.39	3.24	0.07	6.41	0.05	1.96	0.72	3.35	0.04	2.63	-0.17	3.13	0.23	1.79	0.09	-6.57	0.00	0.12	
URBAN	0.82	5.44	0.04	4.25	0.22	6.95	-	-	-	-	-	-	-	-	-	-	-	-	
SINDH	-1.67	-10.25	-0.15	-12.48	-0.26	-7.50	-1.52	-5.09	-0.08	-4.90	-0.17	-3.99	-1.87	-11.09	-0.19	-7.95	-0.38	-7.23	
NWFP	-0.79	-4.23	0.03	2.22	-0.23	-6.10	-0.53	-1.43	0.03	1.73	-0.20	-3.39	-0.83	-3.92	0.02	0.80	-0.26	-5.20	
BALUCHISTAN	-2.86	-15.57	-0.21	-13.63	-0.65	-9.14	-3.01	-7.82	-0.07	-2.70	-0.49	-6.89	-2.77	-14.57	-0.27	-8.52	-0.76	-7.02	
NORTH	1.33	2.81	0.13	4.86	0.19	2.16	-0.34	-0.38	0.02	0.55	-0.09	-0.70	2.22	4.30	0.18	4.25	0.31	3.06	
FATA	-2.44	-7.73	-0.24	-6.51	-0.59	-4.14	-	-	-	-	-	-	-2.47	-8.27	-0.26	-6.33	-0.63	-4.57	
AJK	1.25	3.26	0.12	5.01	0.23	4.53	1.27	2.18	0.08	3.52	0.17	2.96	1.33	2.70	0.13	4.73	0.25	3.59	
Adjusted R ²	0.29		0.33		0.24		0.27		0.34		0.24		0.26		0.31		0.19		
N	14548		14555		9594		5264		5268		3995		9284		9287		5599		
DEP VAR. MEAN	0.0464		0.6625		-3.1230		0.0673		0.7623		-2.8246		0.0345		0.6054		-3.3357		
P-VALUES: AGE 5-9	0.0001		0.0000		0.4635		0.5058		0.0086		0.5146		0.0000		0.0000		0.1500		
AGE 10-14	0.0000		0.0000		0.0000		0.4200		0.6134		0.0600		0.0000		0.0000		0.0000		
AGE 15-19	0.0000		0.0001		0.0000		0.8971		0.2332		0.7200		0.0000		0.0001		0.0000		

Note: Coefficients have been multiplied by 100 in column (a). The dependent variables are EDU_SHARE (budget share of education), ANYEDEXP equals 1 if household spends anything on education and 0 otherwise and the natural log of EDU_SHARE. Base dummy for Head's education is HEAD_MATRICMORE =1 if head has more than 10 years of education 0 otherwise. The last 4 rows represent the p-values of the F test that the male and female gender-age coefficients in that column are equal. PUNJAB is the excluded province and rural is the base category for the URBAN dummy. Head's education takes form as five dummy variables: (i) dummy capturing missing values for head's education, (ii) with head's education less than or equal to Primary (Grade 'katchi' to 5) (iii) education at least equal to Middle (grades 6, 7 and 8) and (iii) education at least equal to Matric (grades 9 and 10). The base category is head's education more than Matric (including F.A., B.A., Masters etc.). Marital status of the household head takes on various values such as married, widowed, etc. Gender of the head is a dummy variable, equals 1 if head is female. Head's occupation is defined in four occupational dummies: (i) missing (ii) 'white collar workers' include managers, professionals, technicians or clerks, (iii) machine operators and assemblers or belonging to the services or trades industry are grouped into 'HEAD_SERVICE'. Individuals in elementary occupations or skilled agricultural workers are the omitted category.

1
2
3 **Title:** Gender and Household Education Expenditure in Pakistan
4
5
6
7

8 **Authors:** Monazza Aslam*
9
10 Geeta Gandhi Kingdon
11
12
13
14

15 **Affiliation:** Department of Economics, University of Oxford
16
17 Department of Economics, University of Oxford
18
19
20
21

22 Abstract
23
24

25 Pakistan has very large gender gaps in educational outcomes. One explanation could be that
26 girls receive lower educational expenditure allocations than boys within the household, but
27 this has never convincingly been tested. This paper investigates whether the intra-household
28 allocation of educational expenditure in Pakistan favours males over females. It also explores
29 two different explanations for the failure of the extant 'Engel curve' studies to detect gender-
30 differentiated treatment in education even where gender bias is strongly expected. Using
31 individual level data from the latest household survey from Pakistan, we posit two potential
32 channels of gender bias: bias in the decision whether to enrol/keep sons and daughters in
33 school, and bias in the decision of education expenditure conditional on enrolling both sons
34 and daughters in school. In middle and secondary school ages, evidence points to significant
35 pro-male biases in *both* the enrolment decision as well as the decision of how much to spend
36 conditional on enrolment. However, in the primary school age-group, only the former
37 channel of bias applies. Results suggest that the observed strong gender difference in
38 education expenditure is a *within* rather than an *across* household phenomenon.
39
40
41
42
43
44

45 **JEL Classification Number:** I21, J16, J71
46
47
48

49 **Key Words:** Gender bias, educational expenditure, Engel curve, Pakistan.
50
51

52 **Acknowledgments:** This paper has benefited from our discussions with Jean Drèze, Marcel
53 Fafchamps and Måns Söderbom and from comments from seminar participants at the
54 Department of Economics, University of Oxford. Any errors are ours.
55
56
57
58
59
60

* Corresponding author: Monazza Aslam, Wolfson College, Linton Road, Oxford OX2 6UD, United Kingdom.
Telephone Numbers: 00-44-1865-284137, 00-44-7956-970628. Email: monazza.aslam@wolfson.ox.ac.uk

Gender and Household Education Expenditure in Pakistan

1. Introduction

One plausible explanation for girls' very inferior educational outcomes relative to boys in Pakistan would seem to be that girls receive less educational expenditure than boys in the within-household allocation of resources. When it has been tested for other South Asian countries, no consistent evidence of within-household gender differentials in education expenditure has been found. The objective of this study is to test whether the commonly used indirect expenditure (Engel curve) methodology is capable of discerning bias in the within-household allocation of educational expenditures in Pakistan.

The detection of gender bias in intra-household allocation of consumption has relied on two approaches: 1) the direct comparison of expenditure by gender, contingent on availability of individual level data and 2) the indirect Engel curve methodology which utilises household level expenditure data to infer differential treatment, by analysing how changes in household gender composition lead to changes in household consumption or expenditure patterns. Much of the extant literature has, due to lack of individual-level data, relied on the indirect approach. This large literature investigating gender biases in household consumption patterns has raised numerous questions. In particular, the conventional Engel curve approach has failed to detect gender differentiated treatment in household allocations even where *outcomes* bespeak large pro-male differences. Deaton (1997, pp. 239-241) remarks 'It is a puzzle that expenditure patterns so consistently fail to show strong gender effects even when measures of outcomes show differences between boys and girls.' Ahmad and Morduch (2002, p. 17) say "coupled with evidence on [significant gender differences in] mortality and health outcomes, the results on household expenditures pose a challenge in understanding consumer behavior". Case and Deaton (2003) say "it is not clear whether there really is no discrimination or whether, for some reason that is unclear, the method simply does not work".

Several explanations have been advanced for explaining this puzzle. One explanation is by Jensen (2002) who argues that parents' fertility behaviour can lead to girls' educational (and other) outcomes being inferior to boys' without there being any parental discrimination in the within-household allocation of educational (or other) resources¹. Another explanation,

¹ If parents have a preference for having at least one (or some desired number of) boys in the household, they will continue child-bearing till that desired number is reached. This sort of behaviour will lead to girls in the

1
2
3 due to Rose (1999), is that households' inability to smooth consumption in the face of shocks
4 leads to parents sacrificing daughters so that only the wanted girls survive; thus any lack of
5 gender bias in current allocations masks prior gender bias in mortality selection. Yet another
6 explanation by Ahmad and Morduch (2002) suggests two-stage budgeting, namely that
7 parents' choices about aggregate expenditures is separable from their choices about how those
8 expenditures are allocated. In other words, budget share on a commodity might remain
9 unchanged with a change in gender composition of the household but parents might allot
10 different portions of a commodity to sons than daughters. This will not show up in
11 investigations of aggregate expenditures but it will show up in examination of individual
12 outcomes.

13
14
15
16
17
18
19
20
21 Testing these explanations for the failure of the Engel curve method to detect bias
22 requires the availability of individual level data on expenditures. For instance, Jensen's point
23 implies that any observed gender differences in educational expenditure at the individual level
24 could be across-household differences due to endogenously differing household sizes for girls
25 and boys, rather than being due to within-household pro-male parental bias in education
26 expenditure allocations. However, with individual level data on expenditure, a family fixed
27 effects model becomes possible which is a powerful way of purging endogeneity bias and
28 examining whether the gender gaps are a within- or across-household phenomenon. We have
29 individual level data on educational expenditures to permit the estimation of such models.

30
31
32
33
34
35
36
37 The paper has two objectives. Firstly, we test the hypothesis that, in Pakistan, the
38 allocation of household educational resources favours males over females. Secondly, we
39 investigate possible reasons for the failure of extant studies to detect gender bias in contexts
40 where it is expected to exist. Data from the Pakistan Integrated Household Survey (PIHS,
41 2001-2002) are utilised to address both questions.

42
43
44
45
46 Although a large literature documents gender biases in food consumption, only a few
47 studies investigate differential treatments in *educational* expenditure, all these being for India
48 (Subramanian and Deaton, 1990 and 1991; Subramanian, 1995; Lancaster et al., 2003; and
49 Kingdon, 2005). On Pakistan, to our knowledge, no study analyses gender biases in
50 educational allocations².

51
52
53
54
55 As mentioned above, the reliability of the Engel curve approach has been questioned
56 in recent years due to its failure to detect gender-differentiated treatment even where it is

57
58
59 population having more siblings, higher average household size and lower per capita resources than boys.
60 Lower per capita resources due to larger household size imply that girls' outcomes will be worse than boys' even
in the absence of any within-household differential treatment of sons and daughters.

² Studies by Deaton (1997) and Bhalotra and Attfield (1998) focus on food consumption.

1
2
3 strongly expected. Kingdon (2005) proposes two possible reasons for this failure: 1) the Engel
4 curve approach uses the incorrect functional form to model the mechanisms of bias and 2)
5 aggregated household level data mutes the detection of gender biases.
6
7

8
9 On the first issue, the Engel curve technique estimates a single budget share equation
10 encompassing two different mechanisms of bias, assigning equal weight to the two. The two
11 potential mechanisms of bias are: a) in the household's decision of whether to spend anything
12 on a given commodity (the zero-versus-positive expenditure decision, called the 'binary
13 decision' in this paper) and b) in the household's decision of *how much* to spend conditional
14 on spending a positive amount (called the 'conditional expenditure decision' in this paper).
15 Averaging across the two (as is implicit in the Engel curve technique) may dilute biases if
16 gender bias occurs through only one channel rather than both, or if the biases in the two
17 channels are in opposite directions. For example, suppose a pro-male bias exists in
18 households' first decision – i.e. a boy is associated with a larger probability of positive
19 spending on education (i.e. of enrolment). Suppose also that, conditional on enrolment,
20 households spend *more* on daughters' than sons' education either because they belong to a
21 select (e.g. more enlightened) group or because it is genuinely costlier to educate daughters,
22 e.g. more expenditure may need to be incurred for transport and school clothing for girls for
23 safety and modesty concerns. In this case, there will be pro-female expenditure allocation in
24 the second mechanism. Averaging across these two divergent mechanisms may mute gender
25 effects even if there is be pro-male bias in the former mechanism. The researcher would be
26 interested in knowing whether significant bias occurs via either of the two mechanisms
27 separately and whether it is the averaging across the two mechanisms that leads to the
28 conclusion of non-bias. In other words, one would be interested not only in the average
29 unconditional expenditure on girls and boys but also in the *distribution* of the expenditure³.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 To examine this first ('averaging') explanation of the failure of Engel Curve methods,
47 we will estimate Hurdle Models to analyse the two household decisions separately, i.e. the
48 binary and conditional expenditure decisions. This will highlight the two possible
49 mechanisms of bias in intra-household allocations of educational expenditure.
50
51
52
53
54

55
56 ³ The conventional application of the Engel curve technique may fail to pick up bias against girls for another
57 reason as well, namely if the distributional assumptions about the dependant variable and thus the specification
58 of the budget-share equation are wrong. For instance, if the education budget-share for households with positive
59 education spending is distributed log-normally but, because the budget-share equation is fitted on all (zero and
60 non-zero education budget-share) households, the researcher has to use absolute budget-share rather than the log
budget-share as the dependant variable, leading to incorrect standard errors. However, in large samples such as
ours, this is not a particularly important worry.

The second potential explanation for the failure of the Engel curve approach has to do with the nature of the data. Previous studies have, perforce, used aggregated household data to infer discrimination. Typically, expenditure data on food, education and health in household surveys is available for the entire household rather than separately for each individual member. The Engel curve technique attempts to deduce differential treatment from household-level aggregated data. It is possible that using household level data somehow makes it more difficult to detect gender biases in intra-household allocations.

To examine this second (aggregation) explanation, we exploit the fact that we have data on educational expenditure of *each individual* child in a given household. This allows us to test whether data aggregation is responsible for the failure of previous studies to detect gender biases. A few recent studies have attempted to analyse individual-level outcomes to investigate differential treatment by gender in different country environments – Hazarika (2000) for Pakistan, Quisumbing and Maluccio (2000) for Bangladesh, Indonesia, Ethiopia and South Africa, and Kingdon (2005) for India, with only the latter study focusing on educational expenditure allocations and the issues mentioned above.

The paper proceeds as follows. Section 2 describes the models and empirical strategies adopted while section 3 discusses the data and descriptive statistics. The empirical results are discussed in Section 4 and the final section concludes.

2. Model and Empirical Strategy

We begin the analysis with the estimation of a standard Engel curve linking budget shares on educational expenditure with total household expenditure and the demographic composition of the household. We use the Working-Leser specification as follows:

$$w_i = \alpha + \beta \ln (x_i / n_i) + \lambda \ln n_i + \sum \theta_k (n_{ki} / n_i) + \varphi z_i + \mu_i \quad (1)$$

where

w_i is the budget share of education of the i th household. It is = (Exp_educ / Total exp);

x_i is the total expenditure of the household;

n_i is the household size;

$\ln (x_i / n_i)$ is the natural log of total per capita expenditure;

n_{ki} / n_i is the fraction of the household members in the k th age-gender class where $k = 1 \dots K$ refers to the K th age-gender class within household i ;

1
2
3
4 z_i is a vector of other household characteristics such as household head's education,
5 gender and occupation and dummy variables to capture province and region etc.
6 These variables are defined in the note to Appendix Table 1;
7
8
9 μ_i is the error term.

10
11
12 α , β , λ , θ_k and φ are the parameters to be estimated. The Working-Leser specification
13 will be relaxed to allow for non-linearity in log per capita expenditure (LNPCPE). The term n_i
14 allows for an independent scale effect of household size. Since the n_{ki} / n_i fractions add up to
15 unity, one of them has to be omitted from the regression. We allow for 14 age-gender groups:
16 males and females aged 0-4, 5-9, 10-14, 15-19, 20-24, 25-60 and 61 and above (omitting the
17 fraction of women aged 61 and above in the regression analysis)⁴. The age categories 5-9, 10-
18 14 and the 15-19 were chosen to correspond roughly with primary, middle and secondary-
19 school-ages respectively⁵. The remaining age categories represent the infants and young
20 children (0-4), prime-aged adults (25-60) and the elderly (61 and above). The θ_k coefficients
21 capture the effect of household composition on household budgetary allocations. These
22 coefficients tell us what the effect of changing household composition is while holding
23 household size constant, for example by replacing a child aged 5-9 by a child aged 10-14 or
24 by replacing a male with a female in a given age category. The difference across gender can
25 be easily tested using a F-test under the following null hypothesis:
26
27
28
29
30
31
32
33
34
35
36
37
38

$$\theta_{km} = \theta_{kf} \quad (2)$$

39
40
41
42 where m denotes males and f denotes females and k refers to a given age-category. Testing,
43 for example, whether boys aged 10-14 are treated differently from girls aged 10-14, we
44 simply seek whether the coefficient on M10TO14 (proportion of males aged 10 to 14 years in
45 the household) is significantly different from the coefficient on F10TO14 (proportion of
46 females aged 10 to 14 years in the household).
47
48
49
50

51 Existing applications of the Engel curve approach fit OLS equations of the absolute
52 education budget share on the sample of all households (including those with zero education
53
54

55
56 ⁴ These age-gender categories are defined as M0TO4, F0TO4, M5TO9, F5TO9, M10TO14, F10TO14 etc. and
57 are the proportion of males (M) and (F) aged 0-4, 5-19, 10-14 and so in a given household.

58 ⁵ These age-groupings are the same as those used in Subramanian and Deaton (1991) and in Kingdon (2005) for
59 India. While regressions were also estimated for the 20-24 age category (corresponding with higher education
60 ages), we do not report the detailed findings for this age group here (see Aslam and Kingdon, 2005, for these
results). Sample selection issues are stronger for this age category because in this age, a high proportion of girls
are married and do not live in their natal homes.

expenditure). In so doing, they implicitly assume that dependent variable – the budget share of education (EDU_SHARE) - is normally rather than log-normally distributed. The reason for including *all* households in the estimation is that some or much of the bias against girls may occur in the decision of *whether* to enrol a child in school, i.e. in the zero-versus-positive spending decision, $w_i = 0$ vs. $w_i > 0$, rather than only in the decision of *how much* to spend conditional on enrolment.

In much of the existing literature, equation (1) has been estimated using OLS with household budget share of food, education or health regressed on the independent variables. Given the large proportion of households reporting zero education expenditure and the resulting censoring of the dependent variable, OLS is not the appropriate model to apply in the analysis of the education budget share. A simple application of the OLS model to data that is censored yields parameter estimates which are biased downwards (Deaton, 1997)⁶.

Although the Tobit model is a suggested alternative, it is identified only if the assumptions of normality and homoskedasticity are fulfilled (Deaton, 1997). Moreover, it assumes that a single mechanism determines the choice between $w = 0$ versus $w > 0$ and the *amount* of w given $w > 0$. In particular, $\partial P(w > 0 | x) / \partial x_j$ and $\partial E(w > 0 | x, w > 0) / \partial x_j$ are constrained to have the same sign. An alternative to Tobit is the Hurdle model (Wooldridge, 2002, p536-38) which allows the initial decision of $w = 0$ to be separate from the decision of *how much* w is, given positive w .

‘Hurdle Models’ are two-tier models because the ‘hurdle’ or first tier is the decision of whether to choose a positive w or not ($w = 0$ versus $w > 0$) and the second tier the decision of how much to spend conditional on spending a positive amount ($w | w > 0$). A simple Hurdle model can be written as follows:

$$P(w = 0 | x) = 1 - \Phi(x\gamma) \quad (2)$$

$$\log(w | (x, w > 0)) \sim \text{Normal}(x\beta, \sigma^2) \quad (3)$$

where w is the share of family budget spent on education, \mathbf{x} is a vector of explanatory variables, γ and β are parameters to be estimated while σ is the standard deviation of w . Equation, (2) shows the probability that w is positive or zero, while the equation (3) stipulates that conditional on $w > 0$, $w | \mathbf{x}$ follows a lognormal distribution. In our data, the conditional education budget share is indeed lognormally distributed.

⁶ The effect of censored observations (zero consumption expenditure on an item) is a well-discussed issue in the Engle curve literature. For instance, see Beneito (2003) and Yen (2005).

The MLE of γ is the probit estimator using $w = 0$ versus $w > 0$ as the binary response. The MLE of β is just the OLS estimator of which is obtained from the regression of $\log(w)$ on \mathbf{x} using only those observations for which budget share is positive i.e. $w > 0$. The consistent estimator of $\hat{\sigma}$ is just the usual standard error from this latter regression. Because of the assumption that conditional on $w > 0$, $\log(w)$ follows a classical linear model, estimation is fairly straightforward. Using the following properties of a lognormal distribution, it is easy to obtain the conditional expectation of $E(w | \mathbf{x}, w > 0)$ and the unconditional expectation $E(w | \mathbf{x})$:

$$E(w | \mathbf{x}, w > 0) = \exp(x\beta + \sigma^2/2) \quad (4)$$

$$E(w | \mathbf{x}) = \Phi(x\gamma) \exp(x\beta + \sigma^2/2) \quad (5)$$

which can be easily estimated given $\hat{\beta}$, $\hat{\sigma}^2$ and $\hat{\gamma}$. One can obtain the marginal effect of \mathbf{x} on w by transforming the marginal effect of $\log(w)$ and using the exponent. Taking the derivative of the conditional expectation of w with respect to \mathbf{x} , we can obtain the marginal effect of \mathbf{x} on w in the OLS regression of $\log(w)$ conditional on $w > 0$. This is as follows:

$$\partial E(w | \mathbf{x}, w > 0) / \partial \mathbf{x} = \beta \cdot \exp(x\beta + \sigma^2/2) \quad (6)$$

The *combined* marginal effect of \mathbf{x} on w , i.e. taking account of the effect of \mathbf{x} on the probability that $w > 0$ and on the size of $w | w > 0$, can be obtained by taking the derivative of the unconditional expectation of w with respect to \mathbf{x} . We can use the product rule and take the derivative of the unconditional expectation in (7) to obtain the *combined* marginal effect as follows:

$$\begin{aligned} \partial E(w | \mathbf{x}) / \partial \mathbf{x} &= \gamma\varphi(x\gamma) \exp(x\beta + \sigma^2/2) + \Phi(x\gamma) \beta \cdot \exp(x\beta + \sigma^2/2) \\ &= \{\gamma\varphi(x\gamma) + \Phi(x\gamma) \beta\} \cdot \exp(x\beta + \sigma^2/2) \end{aligned} \quad (7)$$

In the analysis that follows, we estimate three equations for each province of Pakistan:

- (i) unconditional OLS equation of the budget share of education (conventional Engel curve) in the household level analysis, and OLS equation of unconditional education expenditure in the individual level analysis;
- (ii) Probit equation of the binary decision whether the budget share of education is positive at the household level analysis, and the probit equation of whether any positive educational expenditure is incurred on the index child in the individual level analysis;
- (iii) conditional OLS of log of budget share of education in the household level analysis, i.e. conditional on positive budget share of education, and OLS of log of conditional education expenditure in the individual level analysis.

Equations (ii) and (iii) together are the Hurdle model estimates. In equation (iii), we attempt to allow for possible sample selectivity bias by estimating a Heckman two-step model (more details later). Each of these three equations are fitted on household and individual level data. The difference between the two lies in the level of aggregation of the data. Household level equations are fitted for households with at least one child aged 5-24 years. At the individual level we estimate the same equations but, instead of the dependent variable in the OLS equations being the budget share of education (as in household level analysis), the dependent variable is education expenditure on the individual child. Also, all the independent variables are the same in household- and individual-level equations except for gender: while household level equations include proportion of household members in 14 age-gender categories, individual level equations simply use age of child and the simple dummy variable MALE for gender.

Lastly, we also estimate all three individual-level equations with family fixed effects. This deals with the potential endogeneity of variables included in all our other equations, i.e. of variables such as household per capita expenditure, household size and household head's occupation. It provides a convincing way of examining whether differential educational expenditures on girls and boys are within- or across-household phenomena in Pakistan.

3. Data and descriptive statistics

We use data from the fourth round of the Pakistan Integrated Household Survey (henceforth PIHS) 2001-2002. The PIHS contains rich information on more than 16,000 households from all regions of Pakistan (GOP, 2002). The analysis is limited to households

1
2
3 with at least one child aged 5-24, which reduces the sample to 14,680 households. Among
4 currently enrolled 5-24 year olds, almost 98 percent reported positive educational
5 expenditures, i.e. enrolment is virtually synonymous with incurring positive education
6
7
8
9 spending. The individual-level analysis is based *at the level of the individual child*, i.e. on
10 57,604 children of school-going age⁷.
11

12 The dependent variable in the conventional Engel curve analysis is the share of
13 educational expenditure in total household expenditure. The PIHS reports individual-level
14 expenditure on each child currently enrolled in school as well as total household level
15 expenditure on various items of consumption including food, leisure, health and education.
16
17
18 The education budget share (EDU_SHARE) variable was created as the fraction of
19 educational expenditure in total household expenditure.
20
21
22

23 In the first instance, we regress the household budget share of education on the log of
24 household per capita expenditure (LNPCE) and its square (LNPCE2), log of household size
25 (LNHHSIZE), the age-gender composition variables, and the z-vector variables including the
26 dummy variables for head's education, marital status and gender, and regional and provincial
27 dummies. This is the pooled sample. To further disaggregate the analysis, we estimate
28 separate regressions for the various provinces and further sub-divide the sample into urban
29 and rural regions to analyse whether gender differential patterns differ across the regions and
30 across provinces, though we report only selected results here⁸.
31
32
33
34
35
36

37 Table 1 shows the sex-ratio in the 0-14 year age group in sample households. There is
38 considerable variation across provinces and regions with Punjab having the highest proportion
39 of girls (49.5 per cent) with the lowest proportions in FATA, followed by Balochistan and
40 AJK. This suggests, a priori, that gender biases in household expenditure allocation are likely
41 to be the highest in these three regions and the least in Punjab. Table 1 also divides
42 households with children aged 0-14 into 'girls only' households and 'boys only' households.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

There is a statistically very significant difference in mean budget share on education in girls-only and boys-only households. Finally, Table 1 computes average household size by gender and province. Average household size is significantly different for boys and girls in Balochistan and Northern Areas. These statistics give some credence to Jensen's (2002) argument that due to parents' fertility behaviour female children will have a larger number of siblings and larger household size than male children, suggesting that girls may get less

⁷ The total educational expenditure (TOTAL_EDU) variable was truncated at Rs. 25,000 to exclude outliers. Only 0.6 per cent of the sample reported expenditures greater than Rs. 25,000.

⁸ See Aslam and Kingdon (2005) for all the disaggregated results.

1
2
3 educational resources not because they are discriminated against *within their own household*
4 but rather because they are more likely than boys to live in larger households.
5
6

7 Table 2 presents current enrolment rates and Table 3 reports the average *unconditional*
8 educational expenditure of all children (enrolled and non-enrolled) and the average
9 *conditional* education expenditure i.e. expenditure on currently enrolled children. These are
10 disaggregated by age-group and gender in each of the provinces and territories of Pakistan.
11
12

13 Table 2 reveals wide disparities in enrolment between males and females across
14 provinces in Pakistan. Table 3 shows very significant differences in average male and female
15 unconditional educational expenditures across the provinces. The Federally Administered
16 Tribal Areas (FATA), Balochistan and North West Frontier Province (NWFP) emerge as the
17 provinces with the largest gender differences. Finally, focussing on conditional expenditure
18 makes clear that once enrolled in school, girls generally do not receive significantly lower
19 educational expenditures than boys. For Pakistan as a whole, in the 10-14 age-group
20 conditional educational expenditure is significantly *higher* on girls (Rs. 2063) than on boys
21 (Rs. 1941). The raw data in Tables 2 and 3 suggests that much of the gender differentiated
22 treatment occurs in terms of parents' decision whether or not to enrol/keep boys and girls in
23 school i.e. in girls' significantly lower probability of positive education expenditure, rather
24 than in lower expenditures conditional on enrolment.
25
26
27
28
29
30
31
32
33
34

35 36 37 **4. Empirical Results**

38 The results of the empirical analysis are divided into two sub-sections. In the first,
39 household level analysis is conducted to explore two main questions: 1) using the
40 conventional Engel curve approach, is there any evidence that the allocation of household
41 educational expenditure favours males over females? And 2) does incorrect functional form
42 explain failure of the conventional Engel curve method in picking up gender bias? This
43 analysis is based on a comparison of the conventional Engel curves with Hurdle Models using
44 household level data. The second sub-section explores whether aggregation of data at the
45 household level can explain failure to detect gender bias where it is expected. To this end, we
46 estimate unconditional OLS and the Hurdle models using individual-level data, which are
47 compared to the results from the household-level analysis.
48
49
50
51
52
53
54
55

56 In the first instance we discuss the main findings on the Pakistan sample as a whole.
57 This is disaggregated by region (urban and rural). The main results in Tables 4, 5 and 6 are
58 also presented by province – Punjab, Sindh, NWFP, Balochistan, AJK, Northern regions
59 (North) and FATA – to allow for area-based differences in expenditure allocations within
60

households⁹. However, to conserve space we do not present the full underlying equations separately for each province in the Appendix Table regressions, but merely report the main results of interest in Tables 4, 5 and 6 from those underlying equations. In the individual-level equations, the standard errors are robust for clustering at the household level.

4.1. Household-level outcomes

4.1.1 Conventional Engel Curve evidence

Appendix Table 1 reports the results for Pakistan as a whole, both for urban and rural areas. Column (a) reports the conventional Engel curve equation, column (b) reports a probit of ANYEDEXP (whether household's budget share of education was positive) and the third column, (c), reports the conditional OLS equation of the log of budget share of education. As the mean of the dependent variable in column (a) at the bottom of Appendix Table 1 shows, on average, households in Pakistan devote 4.6 per cent of the total household budget to education with urban areas spending a larger share (6.7 per cent) as compared to the rural regions (3.5 per cent). This national average masks large differences across provinces and regions. The regional variation is not unexpected given that average incomes and possibly educational preferences vary across provinces.

In column (a), per capita expenditure and its square are significant. The coefficient on household size is highly significant and positive and this was also so across all provinces and regions. This could be evidence of economies of scale but an alternative explanation is that larger households are more likely to have children of school-going age which is why they spend a greater budget share on education¹⁰. Female headed households (HEAD_FEMALE) have significantly higher education budget shares in Pakistan as a whole. As compared to households with more educated heads (in the base category, HEAD_FAMORE), those with heads with primary, middle and matric education have significantly lower education budget

⁹ The provinces were also disaggregated by region (urban and rural). A total of 54 equations have been estimated. There are 4 provinces and 3 territories in Pakistan. We also wish to present results for Pakistan as a whole, thus making 8 geographical units. For 5 of these units, we have broken the unit up into three samples: rural, urban and whole (rural + urban). Thus, in total we have $(5 \times 3) + 3 = 18$ separate samples. For each of these samples 3 different equations have been fitted, implying a total of $18 \times 3 = 54$ equations using household level data. Appendix 1 does not report results by province due to space constraints. Tables 4 and 5 also do not report results by regional categorisation for the different provinces. Disaggregated results are available in Aslam and Kingdon (2005).

¹⁰ The theoretical literature suggests that at any given level of per capita resources, larger households will be better off because they share household public goods, such as housing, consumer durables etc. Larger households should, therefore, be able to allocate larger shares to private goods such as education provided they do not substitute towards the 'cheaper' public goods. In Pakistani households, economies of scale could be especially important given the norm of a 'joint family' system. Deaton and Paxton (1998) did not find evidence of such economies of scale across 7 high and low income countries, though they examined food budget shares.

1
2
3 shares. Relative to households with heads in elementary and agricultural occupations (in the
4 base category), those with heads in white collar and service and trade related jobs are inclined
5 to spend a greater proportion of total household expenditure on education.
6
7

8
9 We now turn to the question of most interest here: what do the conventional Engel
10 curve estimates tell us about gender difference in the allocation of educational expenditure in
11 Pakistan? To address this question, p values of the F tests - for the null hypothesis that the
12 coefficients of the age-gender dummies for males and females are equal - are presented in the
13 last four rows of column (a) of Appendix Table 1. For example, the p value of the F test that
14 the coefficient on M5TO9 equals the coefficient on F5TO9 for Pakistan (full sample) is
15 0.0001, suggesting that education budget share increases by significantly more when an extra
16 boy aged 5-9 is added to the household than when an extra girl of that age is added. This
17 suggests very significant bias against females in education expenditure in the 5-9 age range.
18 There is very significant pro-male bias in the 10-14 and 15-19 age groups as well. Much of
19 this bias manifests itself in rural areas. In equations estimated by province (but not shown for
20 space reasons) bias in the 5-9 age group manifests itself in rural areas of Punjab, Sindh,
21 NWFP and FATA and in urban Balochistan. The reason why there is apparently not much
22 differential treatment among the youngest age group (ages 5-9) in the other areas of Pakistan
23 could be because of incorrect functional form or aggregation issues and we turn to Hurdle
24 Models next to investigate this concern.
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 4.1.2 'Averaging' explanation for the failure of the Engel curve method

40 Appendix Table 1, columns (b) and (c) report Hurdle Model estimates, using
41 household level data. Column (b) presents estimates from the first 'hurdle' - the probability
42 that the household spends *anything* on education (ANYEDEXP), i.e. that it has a positive
43 education budget share. Column (c) presents estimates of the second stage - the natural log of
44 education budget share (LNEDU_SHARE) conditional on positive education budget share.
45
46
47
48

49 As mentioned before, the conditional budget share equation could suffer from sample
50 selectivity bias due to being estimated only for a sub-sample (households with positive
51 education budget share, i.e. with currently enrolled children), which could be non-randomly
52 selected from the population. We attempted to control for selectivity by using the Heckman
53 two-step approach but in the absence of convincing exclusion restrictions, we have not
54
55
56
57
58
59
60

proceeded with this route¹¹. We recognise the possible downward selectivity bias in the coefficients of the gender-age composition variables. However, if selectivity bias affects the male and female demographic variables equally, then we need not worry since our interest is in the *difference* in the coefficients of the male and female demographic variables¹².

In Appendix Table 1, the effect of LNPCE is concave and significant in the probit of ANYEDEXP and also in the conditional OLS equation in the full sample and in urban and rural regions of Pakistan. An increase in household size (LNHHSIZE) also has a positive and very significant coefficient in both the probit and conditional OLS equations. In Pakistan as a whole and in rural Pakistan, female-headed households have both a greater probability of spending a positive amount on education and higher conditional education budget shares.

Since our key objective is an analysis of gender bias, our main interest lies in the effect of the demographic variables on the two outcomes (ANYEDEXP and LNEDU_SHARE) in columns (b) and (c), and on the unconditional budget share outcome (EDU_SHARE) in column (a) in Appendix 1. Table 4 presents the difference in marginal effects (DME) of the demographic variables in the three age categories (ages 5-9, 10-14 and 15-19) calculated from the results in Appendix Table 1. The province values have been calculated similarly but the underlying equations are not reported to conserve space. In keeping with our previous analysis, we disaggregate the results by region.

To see how the DME has been calculated, consider the DME of the demographic variables M5TO9 and F5TO9 for Pakistan as a whole, reported in the probit equation in column 1 in Table 4A. For the full sample, in column (b) of Appendix Table 1, the marginal effect of M5TO9 in the probit equation is 1.1352. The marginal effect of F5TO9 in the same equation is 0.7617, yielding a difference of 0.3735 which is multiplied by 100 to yield a DME of 37.35. The DMEs for the unconditional OLS (conventional Engel curve equations) in column (d) of Table 4A have been calculated similarly using column (a) of Appendix Table 1. The DMEs in columns (b) and (c) of Table 4A have been calculated somewhat differently.

¹¹ Three exclusion restrictions were used in controlling for possible sample selectivity: LAND_OWN (whether household owns any agricultural land), LAND_ACRES (the amount of land owned by the household) and BUSINESS (whether the household is an owner/proprietor of a non-farm business). *A priori*, we might have expected a household owning agricultural land or a business to have a higher demand for child labour, i.e. to affect the school enrolment (or positive education expenditure) decision, but not to affect conditional educational expenditure. However, in no case were the exclusion restrictions jointly significant at the 5 per cent level. The F tests revealed that the p-values of the joint significance of the exclusion restrictions in the probit of current enrolment were: 0.14 (age 5-9), 0.53 (age 10-14) and 0.06 (age 15-19). Only in the 20-24 age-group, the exclusion restrictions were jointly significant (at 4 per cent), but the Lambda term was insignificant ($t = -1.27$).

¹² If girls' unobserved traits are important in parents' decisions about their enrolment/education and boys' traits are not important (or less important) to parents' decisions about their schooling, then any pro-male bias will be over-estimated because the female demographic variables will suffer from greater downward bias in the conditional education budget share equation than will male demographic variables.

Column (b) refers to the DME in the conditional OLS equation of the log of budget share of education (LNEDU_SHARE). Since the dependent variable in this equation was in logs, we transformed the marginal effects of the male and female variables before taking the difference between the two, so as to ensure comparison with column (d), where the dependent variable is EDU_SHARE rather than the log of EDU_SHARE¹³. Column (c) reports the results of the DME of the *combined* marginal effects of the probit and conditional OLS equations. The combined marginal effects were computed using STATA, in the way set out in equation (7) in Section 2. The p value of each DME is reported in the bracket below it. For example, the p value of the DME in column (d) for children aged 5-9 in urban Balochistan is 0.04, suggesting that the DME there is significant at the 4 per cent level. The shaded cells represent DMEs significant at the 5 per cent level or better.

The results of main interest for the Pakistan sample as a whole are in the top panel of Table 4. Several interesting results emerge from an analysis of the DME at the household level. Firstly, looking at Table 4A, the conventional Engel curve results in column (d) shows that in the 5-9 age-group, while conventional Engel curve results suggest a pro-male bias in the full and rural samples only, the results reported in columns (a) and (b) demonstrate differently. In both rural and urban areas, the DME in the probit equation is positive and highly significant: an additional boy in the household has a larger impact on the probability of a non-zero education budget share (ANYEDEXP) as compared to an additional girl, i.e. there is strong pro-male bias in the binary decision of allocating positive educational expenditure. However, the story is somewhat different in column (b): the DME of the gender variables in the conditional budget share equation is negative (albeit insignificant) in the urban sample and, when it is positive, it is insignificant (full and rural sample). Thus, at least in the 5-9 age group in Pakistan, much of the gender bias in household educational expenditure allocation occurs at the stage of the enrolment / drop-out decision for boys and girls, rather than in the conditional decision of how much to spend on enrolled boys and girls. It is clear that averaging the (often) oppositely signed probit and conditional expenditure DMEs – which is implicitly what the Engel curve method does - leads to the conclusion of no bias, and would

¹³ For example for the full sample Punjab, the coefficients on M5TO9 and F5TO9 in the conditional OLS of LNEDU_SHARE was $bm = 0.9426$ and $bf = 0.7840$ respectively. We can obtain the log transformations of these by using the property of the log normal distribution that the conditional expectation of $E(w|x, w > 0)$ equals $\exp(x\beta + \sigma^2/2)$. The $\text{Exp}(\cdot)$ for this sub-sample is 0.1838. Thus the transformed marginal effect for males is $bm * \text{Exp}(\cdot) = (0.9426) * (0.0720) = 0.0679$ and that for females is $bf * \text{Exp}(\cdot) = (0.7840) * (0.0720) = 0.0565$. The difference between the male and female marginal effects is $0.0679 - 0.0565 = 0.0114$. In the table all DME are multiplied by 100 and so the reported DME is 1.14.

miss the fact that there is bias through one of the channels, namely in the enrolment (positive spending) decision.

It is also noteworthy that the Hurdle model in column (c) – which allows the binary and conditional decisions to be modelled separately - has greater power to detect bias than the conventional Engel curve method (column d) which uses a single equation to model bias: there is significant evidence of gender bias in urban areas using the Hurdle model while the Engel curve approach is unable to detect this.

This picture changes quite a lot in the 10-14 age group (Table 4B). In this junior education age group, the Engel curve is good at picking up evidence of pro-male bias in Pakistan, see column (d). This seems to be for two reasons. Firstly the size of the DMEs is greater in the probit equation and secondly, the DMEs in the conditional OLS are almost always positive in this age group and also statistically significant in many cases, i.e. both the binary and conditional expenditure decisions work in the same direction (rather than in opposite directions, as was often the case in the 5-9 age group). Thus, here not only do households favour males in their zero-versus-positive expenditure decision ($w_i > 0$), they also favour them in the *amount* spent conditional on enrolment. The findings are similar in the 15-19 age group (Table 4C).

It is not clear what explains the lack of pro-male bias in conditional education expenditure in the primary age group but its presence in the junior and secondary school age group. Simple tabulations and corresponding t-values for the different educational expenditures indicate that in no age-group is the cost of schooling significantly higher for boys than for girls such that cost-differences explain higher conditional expenditure on boys¹⁴. However, this doesn't entirely rule out lower conditional expenditure on girls due to other supply-side factors. For instance, lack of availability of single-sex schools in rural areas may result in lower conditional education expenditure on girls because of lack of access rather than due to parental discrimination¹⁵. However, there is also evidence (Aslam, 2005) to suggest that male children are more likely to be sent to more expensive private schools in Pakistan and this could be one mechanism by which pro-male biases in conditional expenditure operate.

¹⁴ For instance, tuition fees for males and females aged 5-9, 15-19 and 20-24 are statistically insignificantly different from each other (for the 10-14 age group they are significantly higher for girls). Similarly, the data suggest that expense on transport is significantly greater for girls aged 10-14, 15-19 and 20-24.

¹⁵ Alderman et al. (1996) attribute reduced availability of schools for females in rural Pakistan to lower adult cognitive achievement while Lloyd et al. (2002) suggest that single-sex girls' school availability is a key determinant of parent's decision to enrol girls in school in rural Pakistan.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The results for Pakistan are corroborated when we disaggregate by province in the three age-groups. For example, in the 5-9 age group, the Hurdle model detects pro-male bias in NWFP and Balochistan (and in age group 10-14 in Punjab) which the Engel curve is unable to detect. Punjab, generally known to be a more progressive province, has no significant gender bias in the 5-9 age group (although bias is present in *rural* Punjab even in the 5-9 age group, this is not shown in Table 4A).

The overall results suggest a number of conclusions. Firstly, typically the two ‘discriminatory’ processes highlighted in the probit and conditional OLS equations in the 10-14 and 15-19 age groups reinforce each other. In these age groups in almost all instances, the DMEs in columns (a) are (b) are positive suggesting a pro-male bias in the zero-versus-positive expenditure decision as well as in the conditional expenditure decision. Secondly, this explains why the Engel curve method detects pro-male bias in educational spending at least in the 10-14 and 15-19 age groups. Finally, ‘unpacking’ the two mechanisms of bias sheds some light on the puzzle we started with, namely the inability of the Engel curve method to detect bias even where it is strongly expected. The results above suggest that one of the reasons for the failure of the Engel curve method is its incorrect functional form. If the correct model for the binary decision of whether to make a purchase or not is non-linear and the distribution of conditional expenditure is log normally (rather than normally) distributed, it is incorrect to model these two different decisions within a single OLS budget share equation, and especially so if the effects of the age-gender variables on these two decisions are in divergent directions, as they are in Pakistan in the 5-9 age group.

4.2. ‘Aggregation’ explanation of the Engel curve method’s failure

One of the central limitations of studies investigating gender bias in intra-household allocations has been their reliance, perforce, on aggregated household level data to infer who gets what within the household. Failure of the Engel curve method in detecting differential treatment even where it is expected a priori may be attributable to data aggregation. Using individual level data on education expenditures, we investigate whether this can be a shortcoming of household allocation analyses in urban and rural Pakistan.

In this section we compare the household level Engel curve results with the estimates obtained using individual level data. However, the two sets of results are not directly comparable. Firstly, the dependent variable in the individual level analysis is the educational expenditure on the individual child (TOTAL_EDU) rather than the household’s budget share of education (EDU_SHARE), as in household level analysis. Secondly, instead of using the

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

14 demographic variables ‘proportion of males aged 5 to 9’ (M5TO9), ‘proportion of females aged 15 to 19’ (F15TO19), etc., in individual level analysis we use the simple dummy MALE (equals 1 if child is male, 0 otherwise), to capture the gender of the child. All other independent variables are the ones used in the household level analysis. Of course, the marginal effects of MALE in the individual-level equations will differ from the ‘difference in marginal effect’ of the male and female demographic variables in the household-level equations due to different scaling. However, that does not matter as we are interested primarily in whether the difference in marginal effects (DMEs) of the gender variables in household level equations are statistical significant in those regions/areas where the marginal effect of MALE is significant in the individual level equations.

As before with household level analysis, we estimate three equations at the individual level for each region: (1) probit of ANYEDEXP, (2) OLS of LNTOTAL_EDU (conditional OLS), and (3) unconditional OLS of TOTAL_EDU. These equations are estimated for the three age-groups separately: ages 5-9, 10-14 and 15-19¹⁶. We focus on the MALE gender coefficient and report the marginal effects of this variable in Table 5¹⁷.

In Table 5, columns (a) and (b) refer to the marginal effects of the MALE variable in the probit and the conditional OLS equations, respectively. As before, the marginal effects in (b) have been transformed from logs to absolute values for comparison with (d) where the dependent variable is total educational expenditure. Column (c) reports the combined marginal effects of the probit and conditional OLS equation in the Hurdle model while (d) presents the marginal effect of MALE in the simple OLS of total educational expenditure.

The results based on individual level data confirm the findings from the descriptive analysis. There is a large and significant pro-male bias in the allocation of educational expenditure in all age groups and in both rural and urban regions in Pakistan. Comparing the results in Table 5 with those in Table 4 shows that individual level analysis is far more capable of capturing gender bias than household level analysis. In individual level analysis (Table 5), all cells in the 5-9, 10-14 and 15-19 age groups are statistically significant while in the household level analysis, fewer cells are statistically significant. This suggests that there is no substitute for individual level analysis if one wishes to reliably detect gender bias in the within household allocation of educational resources. Aggregation of data at the household

¹⁶ Although we estimated a total of 288 equations (aged 20-24 was a separate category and the results were disaggregated by region for all provinces), as before we do not report all results. More detailed results are available in Aslam and Kingdon (2005).

¹⁷ The full results of the individual-level regressions are available from the authors.

1
2
3 level, an inherent feature of previous studies on intra-household resource allocation, mutes the
4 true extent of gender bias.
5

6
7 At the household level, much of the evidence for gender bias against girls in the 5-9
8 age-group had manifested itself in the probit equation of ANYEDEXP (Table 4A). There was
9 not much evidence of pro-male bias in the *conditional* OLS estimates or in the conventional
10 Engel curve estimation in this age-group. At the individual level, however, differential
11 treatment against females aged 5-9 is apparent in *both* mechanisms. These findings are largely
12 corroborated by provincial disaggregation.
13
14
15
16
17

18 19 **4.3. Household Fixed Effects: Gender differences within or across households?**

20
21 Jensen (2002) argues that in some developing countries, parents may have “son
22 preferring, differential stopping behaviour”. If parents have a strong preference for male
23 children, they will continue child bearing until one (or their desired number of) male offspring
24 is born. In other words, if early born children are girls, parents will be less likely to stop
25 bearing more children than if the early borns are boys. This type of fertility behaviour will
26 imply that, on average, female children will have a larger number of siblings and larger
27 household size than male children. In larger households, *all* children (male and female) are
28 worse off than in smaller houses, since larger family sizes result in a dilution of household
29 resources across children. Average sibling size in Pakistan is 4.8 for girls and 4.7 for boys,
30 which is significantly different (Table 1). This suggests that girls may get less educational
31 resources not because they are discriminated against *within their own household* but rather
32 because they are more likely than boys to live in larger households. In other words, any
33 observed lower educational expenditures on girls than boys could be an *across-household*
34 phenomenon due to differential household sizes for girls and boys in the population. If
35 household size is endogenously chosen in the way Jensen describes then simply controlling
36 for household size as we have done previously, will not suffice.
37
38
39
40
41
42
43
44
45
46
47
48

49
50 Introducing household fixed effects is a powerful way of controlling for unobserved
51 parental fertility preferences and thus for the endogeneity of household size. Our household
52 fixed effects analysis estimates three equations using individual-level data: 1) a probit
53 equation of ANYEDEXP (whether any positive expenditure was incurred on the child’s
54 education); 2) the equation of the log of educational expenditure (LNTOTAL_EDU)
55 conditional on positive educational expenditure; and 3) the unconditional educational
56 expenditure (TOTAL_EDU) equation. These equations are fitted on the sample of only those
57 households that have at least one child of each gender in the relevant age range (ages 5-9, 10-
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

14, 15-19). Estimates are obtained by age group and province. As before, we discuss the Pakistan results in detail. Of course, controlling for household implies that coefficients only the child variables (age and gender, MALE) are retained.

Table 6 reports results of the household fixed effects estimation. We report the coefficient of the MALE dummy in the three equations with the t-statistic in brackets. We notice the large number of significant values in all decisions – the decision to enrol as well as the conditional and unconditional expenditure decisions – at all age groups, though there are some variations. There is pervasive evidence of significant *within*-household pro-male bias in the allocation of educational resources, and one cannot attribute the results of previous tables simply to differences in household size across the population. The household fixed effects estimates constitute fairly convincing evidence of strong pro-male bias in educational decisions *within* households in Pakistan in all school-going age-groups¹⁸: in the 5-9 age group, a daughter is 13.5 percentage points less likely than a son to have any education expenditure incurred on her education (i.e. to be enrolled in school); this rises to 24 percentage points in the 10-14 age group. Education expenditure allocation differs dramatically for sons and daughters within the household.

5. Summary and Conclusions

In this paper we have examined two questions central to the intra-household allocation literature: (i) does the allocation of household educational resources in Pakistan favour males over females and (ii) what explains the inability of the standard Engel curve approach to detect differential treatment even where discrimination is known to exist? We exploit the latest national sample survey, the Pakistan Integrated Household Survey (PIHS 2001-2002), to address both concerns.

The descriptive statistics reveal large and significant gaps in schooling outcomes for males and females of school-going age in Pakistani households. These gender disparities are more strongly discernible in Balochistan, NWFP and FATA, and in rural areas. Much of the bias in educational expenditures manifests itself in significantly lower probability of girls' enrolment, and hence zero expenditure, rather than in lower expenditures conditional on enrolment.

¹⁸ As a referee of this journal points out, age gaps between siblings may differ for girls and boys within the household since a new born's gender may affect parental decisions about the spacing of the next birth (Angrist, Lavy and Schlosser, 2005). The family fixed effects approach does not address this issue or the possibility of time-varying unobserved household heterogeneity.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
The conventional (Engel curve) approach to discovering differential treatment in intra-household allocations has been questioned in recent years for its inability to detect biases in allocations even when *outcomes* reveal differently. Two possible explanations for this puzzle were tested: 1) that the functional form adopted by the Engel approach is too restrictive and 2) that data aggregation somehow diminishes ability to detect gender biases. To explore the first issue, we estimated Hurdle Models which allow the two potentially ‘discriminatory’ channels of bias – the zero-versus-positive expenditure decision and the conditional spending decision – to be modelled separately, instead of constraining them to be in the same direction. Additionally, the Hurdle Model allows the functional forms of the two decisions to be guided by the underlying distributions of the education budget share. The binary decision is modelled as a probit and the conditional expenditure decision using OLS of the log of budget share since budget share is distributed log-normally. The second explanation – which has to do with aggregation of data at the household level - is tested using unique individual-level data on educational expenditures on each child in the sample.

28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
The results suggest several conclusions. Even using the conventional Engel curve approach, robust evidence of a pro-male bias in educational expenditure is found especially in the 10-14 and 15-19 age-groups. Much of this differential treatment manifests itself in rural areas. The lack of evidence in the 5-9 age-group is puzzling given large gender differentials in enrolment seen in Table 2. The Hurdle Models highlight why this is the case. While there is substantial evidence of strong *pro-male* bias in the binary decision whether to spend anything on education (the probit), there is weak *pro-female* bias in the conditional expenditure decision, i.e. the two potential channels of bias often go in opposite directions. In the older age groups, both channels typically work in the same direction, i.e. reinforce each other. These results hold when using individual level data. Hurdle models are better able to detect gender bias in educational expenditure as compared to the conventional Engel curve approach, especially when using household level data. Controlling for unobserved household preferences by using household fixed effects confirms that the large and significant pro-male biases in educational expenditures in Pakistan are a *within-household* phenomenon.

53
54
55
56
57
58
59
60
Furthermore, a comparison of individual and household level results reveals that aggregating expenditure data across individuals within a household mutes the ability to ‘pick up’ gender effects. The findings suggest that individual level data are far preferable to household level data if one wishes to reliably estimate gender effects.

Whether the substantial gender differences observed in within-household education expenditure allocations constitute pure discrimination remains arguable. Gender differentiated

1
2
3 treatment could in principle be attributed to an investment motive on the part of parents,
4 reflecting differential labour market returns to education for males and females. Evidence on
5 Pakistan (Aslam, 2006) suggests that while returns to education for women are significantly
6 higher than those for men, overall labour market returns are higher for men since the latter
7 have much higher earnings than women¹⁹. Thus, an investment motive seems at least one
8 plausible explanation for gender bias within the household. Even if returns to education were
9 similar for males and females (or even higher for females), the part of the return to a child's
10 education that accrues to the parents is likely to be higher from sons' education than from
11 daughters' since in societies such as Pakistan sons provide old age support for parents while
12 any economic returns to a daughter's education are reaped by her in-laws. This asymmetry in
13 parental incentives to invest in sons' and daughters' education could well explain the
14 observed gender gaps in education expenditure within Pakistani households and has obvious
15 and important public policy implications.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

58 ¹⁹ An Oaxaca decomposition suggests that much of the gender earnings gap is not explained by differences in the
59 observed characteristics between men and women, suggesting a good deal of gender discrimination in the labour
60 market. Studies by Ashraf et al. (1993) and Siddiqui et al. (1998) also find evidence of gender discrimination in
the Pakistan labour market, though they suffer from methodological limitations such as lack of control for
sample selectivity in female work participation and for the endogeneity of education.

References

- Ahmad, A., Morduch, J., 2002. Identifying Sex Bias in the Allocation of Household Resources: Evidence from Linked Household Surveys from Bangladesh. Mimeo, New York University, Department of Economics.
- Alderman, H., Behrman, J.R., Ross, D. and Sabot, R., 1996. Decomposing the Gender Gap in Cognitive Skills in a Poor Rural Economy. *The Journal of Human Resources* 31, 229-254.
- Al-Samarrai, S., Peasgood, T., 1997. Educational Attainments and Household Characteristics in Tanzania. IDS Working Paper 49.
- Angrist, J., V. Lavy and A. Schlosser (2005) "New Evidence on the Causal Link Between the Quantity and Quality of Children", Working Paper 11835, National Bureau of Economic Research. <http://www.nber.org/papers/w11835>, December.
- Appleton, S., Chessa, I., Hoddinott, J., 1999. Are Women the Fairer Sex? Looking for Gender Differences in Gender Bias in Uganda. .Mimeo, University of Oxford, Department of Economics.
- Aslam, M., 2003. The Determinants of Pupil Achievement in Government and Private Schools in Pakistan. *The Pakistan Development Review* 42, 841--876.
- Aslam, M., 2005. The Relative Effectiveness of Government and Private Schools in Pakistan. Mimeo, University of Oxford, Department of Economics.
- Aslam, M. 2006. Returns to Education by Gender in Pakistan, mimeo, Department of Economics, University of Oxford.
- Aslam, M. and Kingdon, G.G. (2005), 'Gender and Household Education Expenditure in Pakistan', Global Poverty Research Group Working Paper Series No. 025.
- Ashraf, J. and Ashraf, B., 1993. Estimating the Gender Wage Gap in Rawalpindi City. *The Pakistan Development Review* 29, 365--376.
- Behrman, J.R., Foster, A.D., Rosenzweig, M.R., Vashishtha, P., 1999. Women's Schooling, Home Teaching, and Economic Growth. *Journal of Political Economy* 107, 682--714.
- Beneito, P. (2003) "A Complete System of Engel Curves in the Spanish Economy", *Applied Economics*, 35(7): 803-16.
- Bhalotra, S., Attfield, C., 1998. Intrahousehold Resource Allocation in Rural Pakistan: A Semiparametric Analysis. *Journal of Applied Econometrics*, 463--480.
- Case, Anne and Angus Deaton. 2003. "Consumption, health, gender and poverty." Working Paper No: 3020, World Bank Poverty Reduction and Economic Management Network.

- 1
2
3 Coale, A.J., 1991. Excess Female Mortality and the Balance of the Sexes in the Population: An
4 estimate of the number of "Missing Females". *Population and Development Review* 17, 517--523.
5
6
7
8 Das Gupta, M., 1987. Selective Discrimination against Female Children in Rural Punjab, India.
9 *Population and Development Review* 13, 77--100.
10
11 Deaton, A., 1989. Looking for boy-girl discrimination in household expenditure data. *World Bank*
12 *Economic Review* 3,1--15.
13
14 Deaton, A., 1997, The Analysis of Household Surveys: A Microeconomic Approach to
15 Development Policy, The Johns Hopkins University Press, Baltimore, Maryland, USA.
16
17 Deaton, A., Paxton, C., 1998. Economies of Scale, Household Size, and the Demand for Food.
18 *Journal of Political Economy* 106, 897--910.
19
20 Dwyer, D., Bruce, J., (eds.) 1988. A Home Divided, Women and Income in the Third World. Stanford
21 University Press, Stanford,CA.
22
23 Gibson, J., Rozelle, S., 2000. Is it Better to be a Boy? A Disaggregated Outlay Equivalent Analysis of
24 Gender Bias in Papua New Guinea. Working Paper No. 23, Department of Agricultural and Resource
25 Economics, University of California Davis.
26
27
28 Glick, P., Sahn, D.E., 2000. Schooling of Girls and Boys in a West African Country: the Effects of
29 Parental Education, Income, and Household Structure. *Economics of Education Review* 19, 63--87.
30
31 GOP (2002) *Pakistan Integrated Household Survey, Enumerator's Manual of Instructions, Round 4:*
32 *2001-2002*, Federal Bureau of Statistics, Government of Pakistan, Islamabad.
33
34
35 Greene, W.H., 2000. *Econometric Analysis*, 4th Edition, New York: Macmillan, Chapter 20.
36
37 Haddad, L., Hoddinot, J., 1995. Does Female Income Share Influence Household Expenditures?
38 Evidence from Cote d'Ivoire. *Oxford Bulletin of Economics and Statistics* 57, 77--96.
39
40 Haddad, L., Reardon, T., 1993. Gender Bias in the Allocation of Resources Within Households in
41 Burkina Faso: A Disaggregated Outlay Equivalent Analysis. *The Journal of Development Studies* 29,
42 260--276.
43
44 Handa, S., 1994. Gender, Headship and Intrahousehold Resource Allocation. *World Development* 22,
45 1535--1547.
46
47 Handa, S., 1996a. Expenditure Behaviour and Children's Welfare: An Analysis of Female Headed
48 Households in Jamaica. *Journal of Development Economics* 50, 165--187.
49
50 Handa, S., 1996b. Maternal Education and Child Attainment in Jamaica: Testing the Bargaining
51 Power Hypothesis. *Oxford Bulletin of Economics and Statistics* 58, 119--137.
52
53 Hazarika, G., 2000. Gender Differences in Children's Nutrition and Access to Health Care in Pakistan.
54 *The Journal of Development Studies* 37, 73--92.
55
56 Hazarika, G., Bedi, A.S., 2003. Schooling Costs and Child Work in Rural Pakistan. *The Journal of*
57 *Development Studies* 39, 29--64.
58
59 Jensen, R., 2002. Equal Treatment, Unequal Outcomes? Generating Sex Inequality Through Fertility
60 Behaviour. Mimeo, Harvard University, John F. Kennedy School of Government.

- 1
2
3
4 Johnson, K., Rogers, B., 1993. Children's Nutritional Status in Female Headed Households in the
5 Dominican Republic. *Social Science and Medicine* 30, 1293--1301.
6
7
8 Kennedy, E., Peters, P., 1992. Household Food Security and Child Nutrition: The Interaction of
9 Income and Gender of Household Head. *World Development* 20,1077--1085.
10
11 Khan, S.R., 1993. South Asia, in King, E., Hill, M., (Eds.), Women's Education in Developing
12 Countries, Washington D.C., Johns Hopkins Press for The World Bank.
13
14 Kingdon, G.G., 1996. The Quality and Efficiency of Private and Public Education: A Case Study of
15 Urban India. *Oxford Bulletin of Economics and Statistics* 58, 57--81.
16
17 Kingdon, G.G., 1998. Does the Labour Market Explain Lower Female Schooling in India? *Journal of*
18 *Development Studies* 35, 39--65.
19
20 Kingdon, G.G., 2005. Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold
21 Allocation of Educational Expenditure. *Economic Development and Cultural Change* 53, 409--451.
22
23
24 Kossoudji, E., Mueller, E., 1983. The Economic and Demographic Status of Female Headed
25 Households in Rural Bostwana. *Economic Development and Cultural Change* 31, 831--859.
26
27 Lancaster, G., Maitra, P., Ray, R. 2003. Endogenous Power, Household Expenditure Patterns and New
28 Tests of Gender Bias: Evidence from India. Mimeo, Monash University and University of Tasmania.
29
30 Lloyd, C. B., Mete., C., Sathar, Z.A, 2002. The Effect of Gender Differences in Primary School
31 Access, Choice and Quality on the Decision to Enroll in Rural Pakistan. Population Council Working
32 Paper presented at the South Asia Meeting of the Econometric Society in Lahore, Pakistan (2002).
33
34 Nicol, C.J., Nakamura, A., 1992. The Effects of Labour Supply on Commodity Demands: Some
35 Canadian Evidence. Mimeo, University of Regina, Department of Economics.
36
37 Parrish, W.L, Willis, R.J., 1993. Daughters, Education, and Family Budgets: Taiwan Experiences. *The*
38 *Journal of Human Resources* 28, 863--898.
39
40
41 Quisumbing, A.R., Maluccio, J.A., 2000. Intrahousehold Allocation and Gender Relations: New
42 Empirical Evidence from Four Developing Countries. FCND Discussion Paper No. 84.
43
44
45 Rose, E., 1999. Consumption Smoothing and Excess Female Mortality in Rural India. *The Review of*
46 *Economics and Statistics* 81, 41--49.
47
48
49 Rose, E., 1996. Gender Bias, Credit Constraints and Time Allocation in Rural India. *Economic*
50 *Journal* 110, 738--758.
51
52 Rosensweig, M. R., Schultz, T.P., 1982. Market Opportunities, Genetic Endowments, and Intrafamily
53 Resource Distribution: Child Survival in Rural India. *The American Economic Review* 72, 803--815.
54
55 Schultz, T. P., 1993. Returns to Women's Education in: King, E., Hill, M., (Eds.), Women's Education
56 in Developing Countries, Washington D.C., Johns Hopkins Press for The World Bank.
57
58 Siddiqui, R., Siddiqui, R., 1998. A Decomposition of Male-Female Earnings Differentials. *The*
59 *Pakistan Development Review* 37, 885--898.
60

1
2
3 Strauss, J., Thomas, D., 1995. Human Resources: Empirical Modelling of Household and Family
4 Decisions in: Srinivasan, T.N., Behrman, J., Handbook of Development Economics, Vol 3, North
5 Holland, Amsterdam.
6

7
8 Subramanian, S., 1995. Gender Discrimination in Intra-Household Allocation in India. Unpublished
9 Mimeo, Cornell University, Department of Economics.
10

11 Subramanian, S., Deaton, A., 1990. Gender Effects in Indian Consumption Patterns. Discussion Paper
12 No. 147, Indira Gandhi Institute of Development Research Bombay, India.
13

14 Subramanian, S., Deaton, A., 1991. Gender Effects in Indian Consumption Patterns. Sarvekshana 14,
15 1--12.
16

17 Thomas, D., 1992. The Distribution of Income and Expenditure within the Household. Yale University
18 Economic Growth Centre Working Paper No. 669.
19

20
21 Udry, C., 1997. Recent Advances in Empirical Microeconomic Research in Poor countries: An
22 Annotated Bibliography. *Journal of Economic Education* 28, 58--75.
23

24 Wooldridge, J. M., 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press,
25 Cambridge, Massachusetts, USA.
26

27 Yen, Steven (2005) "Zero observations and gender differences in cigarette consumption",
28 *Applied Economics*, 37 (16): 1839-49.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60