

## **Personal Earnings Inequality in Pakistan: Findings from the HIES 1993-94**

ZAFAR MUEEN NASIR and RIAZ MAHMOOD

### **I. INTRODUCTION**

The earnings of workers play important role in the well-beings of households' as they account for the largest proportion of total household income. If earnings of workers are distributed unevenly, they contribute significantly to the inequality in the household earnings. It may not be a cause of serious concern if income inequality grows and income of the workers also grows throughout the population and the position of the bottom segment improves. It is however serious when gap between rich and poor increases by worsening the position of poor. To reduce the household income inequality it is therefore important to focus on the distribution of personal earnings and frame a policy.

There are many cause of inequality in personal earnings. As workers income rises at varying rates, it may reflect the decision of household of their investment in human capital and decisions to acquire skills. The factors like education, occupation, gender, regional location, sector of employment, and non-market forces such as discrimination may also play a significant role in the distribution of earnings.

The inequality in the individual earnings is matter of serious concern for both developed and less developed countries. The evidence suggests that greater inequality exist in individual earnings in highly developed country like the United States compared to other OECD countries [Topal (1997)]. A study by Karoly (1992) found a rising trend in personal earnings inequality over the years for U.S. workers. Studies on Pakistani data also confirm the presence of inequality in the personal earnings of workers [Haque (1977) and Hamdani (1977)].

Zafar Mueen Nasir is Senior Research Economist and Riaz Mahmood is Research Economist at the Pakistan Institute of Development Economics, Islamabad.

The present study is designed to analyse the distribution of individual earnings in the Pakistani labour market and find out the possible reasons for the dispersion in the earnings of workers. This study is different from other studies because it analyses inequality in the earned income rather than household income inequality [Azfar (1976); Kemal (1981) and Mahmood (1984)]. The variance of log earnings method is utilised to find out the extent of inequality in the personal earnings. This method is widely used in the literature [Dooley and Gottschalk (1982); Estudillo (1997); Karoly (1992); Moffitt (1990); Mahmood (1984) and Schultz (1982)]. The variance is then decomposed into two parts namely within group and across group variance [Schultz (1982)]. This method allows us to decompose the individual earnings inequality into within and across different groups. The ordinary least square estimation procedure is used to find the determinants of earnings to understand the causes of dispersion in the earnings.

The nationally representative Household Integrated Economic Survey (HIES) data 1993-94 is used for this analysis. The data provides information not only at all Pakistan level but also for many subgroups. It also provides information about many characteristics of the workers, which are important for this study. The paper uses human capital as well as non-human capital factors for explaining the disparity in earnings. The following factors have been investigated to determine the cause of inequality in the personal earnings of individuals.

1. Age distribution of individuals
2. Education and training
3. Employment status
4. Sector of Employment

The study is structured as follows. Section two presents the review of literature, Section three explains the model and its estimation methodology. Section four deals with the data and its characteristics. Section five discusses the results. Major findings are summarised in the concluding section with some policy recommendations.

## II. STUDIES ON EARNINGS DIFFERENTIAL IN PAKISTAN

Haque (1977) addressed the issue of inequality and determinants of personal earnings by using the Rawalpindi city data. His sample consists of only those workers who were employed at the time of the survey. The analysis indicates that the human capital factors explain major part of the earnings differential. Estimating earnings functions with and without these factors draws this conclusion. Using a dummy variable for gender difference, the analysis shows that male earn more than female workers. The study also indicates that the workers in the formal sector earn either same income or slightly less than the workers of the informal sector. Explaining the limitations of his

study, Haque indicates that, because of the sampling design, there is not enough representation of poor living in the city. This may have underestimated the inequality in personal earnings. Moreover, the data utilised by the study is not nationally representative which is a major drawback.

Shabbir (1994) indirectly addressed the issue of inequality in the personal earnings by estimating the Mincerian earnings function from the Population, Labour Force, and Migration survey data 1979. His analysis also indicates that a major portion of the variation in dependent variable is explained by the human capital factors. The study points out that the labour market is not homogenous and workers in different segments of labour market receive different returns. The study suggests some policy measures to reduce inter-regional income disparities. The study is confined to male workers, thus ignoring important segment of the labour force. Although it considers some occupational categories, the role of employer size and industry has not been accounted for.

The present study tries to analyse inequality in personal earnings by using log variance method. This aspect of the labour market has so far been ignored in Pakistan. Moreover, it estimates earnings functions and provides the age-earnings profiles of workers with different education levels. It also tries to find some explanation for the inequality in the earnings of workers within a specific age group with highest inequality compared to other groups. These aspects of the study will make it important contribution in the literature.

### III. MODEL AND METHODOLOGY OF ANALYSIS

The variance method is used to analyse inequality in the personal earnings. The major advantage of this measure is to decompose it into two important components i.e. variance between and within groups. This decomposition is useful in understanding and distinguishing different sources of inequality. This may help in framing policies for the welfare of workers.

The variance of log earnings, our measure of economic inequality, can be written as

$$VLY = \sum_j \sum_i \frac{(\overline{LY} - LY_{ij})^2}{N} \dots \dots \dots \dots \dots \quad (1)$$

Where  $VLY$  represents variance of log monthly earnings of workers,  $LY_{ij}$  is the natural logarithm of the earnings of  $i$ th individual in the  $j$ th group with a positive monthly income, and  $\overline{LY}$  is the mean of the logarithmic incomes of overall sample. We added and subtracted the mean of the logarithmic incomes of the  $j$ th group,  $\overline{Y}_j$ , in Equation 1 and rearranged the terms. This technique allowed us to decompose the variance into its two components namely across groups' variance and within group

variance. This decomposition can be expressed as

$$= \sum_j \left( \frac{n_j}{N} \right) [(\overline{LY} - \overline{LY}_j)^2 + \sum_i \frac{(\overline{LY}_j - Y_{ij})^2}{n_j}] \quad \dots \quad \dots \quad \dots \quad (2)$$

In Equation 2,  $n_j$  is the number of persons in the group  $j$  with a positive income, and  $N = \sum_i n_j$ . The term  $\frac{n_j}{N}$  is the weight or relative frequency of the groups in the population of income recipients. If the weight or the fraction of workers in some group is large, it may increase measured aggregate income inequality of that particular group.

The first component of income inequality in Equation 2 is the square of difference between the group and overall logarithmic mean income. This represents the age-earnings profile of individuals of the sample. The age-earnings profile reflects the investment in human and physical capital of individuals to distribute their earnings opportunities over their life span [Backer (1964) and Mincer (1974)]. One can interpret these profiles as returns to human and non-human capital investments. Educational attainments, vocational and on the job training, labour market experience, and occupation are some of the most important human capital investments. The slope of age-earnings profiles reflects the level of this investment. As one starts making these investments at the start of life cycle, the greater the investment, the steeper is the slope of the profile. For example, the slope of the age-earnings profile of highly educated population will be steeper than low educated people, keeping other things constant. For estimation of these profiles we however need longitudinal data but unfortunately it is not available in most of the underdeveloped world including Pakistan. Therefore we have to rely on cross-sectional data set to obtain the age-earnings profiles.<sup>1</sup> A set of earning functions is included in the study to obtain age-earnings profiles for different educational levels as an explanation for earnings inequality. These functions include dummies for age groups, employment status, training, sex, and other traits of individuals. The general form of the spline earnings equation is

$$\ln w_i = \alpha_i + \beta_i X_i + \delta_i Z_i + u_i \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (3)$$

where  $\ln w_i$  is the natural logarithm of the monthly earnings of the  $i$ th individual;  $\alpha_i$  is the intercept,  $X_i$  is vector of human capital variables including age, education, market experience and training,  $Z_i$  is the vector of other socio-economic characteristics including dummies for employment status, sector of employment, region and gender. The  $\beta$ 's and  $\delta$ 's are the slope coefficients to be estimated; and  $u_i$  is the error term. The

<sup>1</sup>There are significant differences between longitudinal and cross-sectional data sets. Cross-sectional data sets reflect the one time or snap shot of the population whereas in longitudinal data set individuals are tracked over-time. The changes in the economy may be reflected in longitudinal data but not in the cross-section data set.

spline earning functions are estimated by using ordinary least squared technique. The estimates of parameters represent the proportional change in the earnings associated with one year of increase in the variables. The definition and descriptive statistics are provided in Table 1 and 2.

The other component of Equation (2) represents the within group log variance. In other words, it is showing within group inequality in the earnings of individuals. The identification of such group is important from the policy perspective as welfare policies can be designed to lower this inequality. To find out the root cause of such inequality in the earnings of individuals, we have further dis-aggregated the group on the basis of different characteristics of workers. This will shed more light on some of causes of disparity in personal earnings.

Table 1

*Definition of the Variables and Descriptive Statistics*

Variables	Description
LW	Log of monthly earnings
AGE1	One if age of the worker is between 10 and 15 years (zero otherwise)
AGE2	One if age of the worker is between 16 and 25 years (zero otherwise)
AGE3	One if age of the worker is between 26 and 35 years (zero otherwise)
AGE4	One if age of the worker is between 36 and 45 years (zero otherwise)
AGE5	One if age of the worker is between 46 and 60 years (zero otherwise)
AGE6	One if individual belongs to sixth age group (61+) (zero otherwise)
LMAT	One if education is less than 10 years (zero otherwise)
MAT	One if completed 10 years of schooling (zero otherwise)
AMAT	One if completed more than 10 years of schooling (zero otherwise)
TT	One if participated in any training programme (zero otherwise)
SEX	One if worker is Male (zero otherwise)
UR	One if worker lives in Urban areas (zero otherwise)
MS	One if married and lives with spouse (zero otherwise)
EMP	One if individual is employer (zero otherwise)
SEMP	One if individual is self employee (zero otherwise)
EMPL	One if individual is regular employee (zero otherwise)
FORMAL	One if individual works in the formal sector (zero otherwise)
INFORMAL	One if individual works in the informal sector (zero otherwise)

Table 2

*Descriptive Statistics (Means and Sample Fractions;  
Standard Deviations in Parentheses)*

Variable	All Pakistan	Illiterates	Edu<10	Edu=10	Edu>10
Monthly Salary	2155.85 (2317.32)	1589.92 (1637.01)	2054.33 (1861.81)	2738.11 (2600.59)	4363.04 (3685.54)
Age 10–15	0.0269	0.0335	0.0395	0.0035	–
Age 16–25	0.2202	0.1783	0.2768	0.2953	0.1934
Age 26–35	0.2652	0.2377	0.2589	0.3044	0.3744
Age 36–45	0.2275	0.2199	0.2232	0.2260	0.2726
Age 46–60	0.2106	0.2578	0.1663	0.1517	0.1509
Age 61+	0.0500	0.0728	0.0354	0.0191	0.0088
TT	0.0900	0.0140	0.0390	0.0840	0.1210
UR	0.4240	0.2800	0.4630	0.6350	0.7860
SEX	0.9380	0.9260	0.9750	0.9510	0.7860
MS	0.7550	0.7910	0.7100	0.7080	0.7340
EMP	0.0133	0.0080	0.0120	0.0190	0.0360
SEMP	0.3616	0.4350	0.3770	0.2380	0.7340
EMPL	0.6242	0.5560	0.6100	0.7430	0.8480
FORMAL	0.2210	0.0930	0.1690	0.4460	0.6960
INFORMAL	0.7780	0.9060	0.8300	0.5540	0.3030

Source: Household Integrated Economic Survey 1993-94.

#### IV. DATA CHARACTERISTICS

The data used for this paper are drawn from the Household Integrated Economic Survey (HIES) 1993-94. The Federal Bureau of Statistics (FBS) collects this data by direct interviews and the effect of seasonal variations is offset by enumeration on monthly basis evenly distributed over whole year. The data set provides comprehensive information on many characteristics of Pakistani labour force. The data is collected through a series of specific questions from the respondents. The information on earnings, age, education levels, sex, marital status, regional locations, employment status, and sector of employment is particularly helpful.

The HIES 1993-94 data set covers more than 14,000 households and above 100,000 individuals. The sample of this study is however confined to 18,476 individuals who reported work for pay. The individuals included in the study are between 10 and 80 years of age with positive earnings. Some of the unrealistic observations are dropped from the sample.

Of those employed, approximately 94 percent are males and remaining 6 percent are females. A majority of these workers (58 percent) is resident of rural areas. Nine

percent of these workers received either vocational, technical, or on the job training. From both male and female workers, 76 percent are married and live with their spouse. It is also noted that 52 percent of the sample is illiterate. Those who reported less than 10 years of education are 25 percent whereas 11 percent have completed 10 years of education. Approximately 12 percent of the sample have more than 10 years of schooling. Sixty two percent of the sample consists of regular employees and 50 percent of all workers are employed in small non-financial, non-farm establishments that have less than 10 workers on their pay roll.

In general, we find that workers with higher education have higher earnings. Average monthly earnings of illiterates, less than 10 years, 10 years and 10+years of schooling are Rs 1589.92, Rs 2054.33, Rs 2738.11, and 4363.04 respectively. Similarly, those who receive training have earnings higher than those who did not receive any training. These findings are consistent with human capital theory and studies on Pakistani data [Haque (1977); Shabbir and Khan (1991)]. It is further noted that male workers and residents of urban areas earn higher wages compared to their respective counterparts. These findings indicate that earnings are being affected by human as well as non-human capital characteristics of individuals and create inequality in their earnings. The detailed descriptive statistics is provided in Table 2.

## V. EMPIRICAL RESULTS

This section reports the results obtained from the variance and its decomposition, and earning functions. The age spline earnings functions are used to estimate the slope of age-earnings profiles and explain the disparity in personal earnings. These estimates are carried out for overall Pakistan.

Results for decomposition of variance for complete sample are reported in Table 3. Column 1 reports the proportion of workers in different age groups, Column (2) the inequality in earnings between age groups. This is calculated by taking square of the difference of means of the entire sample and the age group. A negative sign shows below average mean earnings of that particular age group. Column (3) shows earnings inequality within age group. The sum of two components of earnings inequality are presented in Column (4) whereas Column (5) represents overall inequality being adjusted by the age group's weight.

Our results indicate that the youngest and the oldest age groups have the greatest contribution to overall inequality in the personal earnings. This inequality is the result of higher inequality in both components i.e. within and across age groups. It is important to point out that the mean earnings of these two groups are lower than the average earnings of the population.

Table 3  
*Components of Variance Decomposition (All Pakistan Workers)*

Age Groups of Earners (in Years)	Proportion of Earners in Age Groups (1)	Between Groups Inequality (2)	Within Group Inequality (3)	Total Cohort Variance (4)	Weighted Variance Share (5)
	$\frac{n_j}{N}$	$(\bar{Y}_j - \bar{Y})^2$	$\sum \frac{(\bar{Y}_j - \bar{Y}_{ij})^2}{n_j}$	(2)+(3)	(1)*(4)
10-15	0.0269	(-) 1.1494	0.8040	1.9534	0.0526
16-25	0.2202	(-) 0.0550	0.4760	0.5310	0.1169
26-35	0.2652	0.0098	0.5750	0.5848	0.1551
36-45	0.2275	0.0481	0.7510	0.7991	0.1818
46-60	0.2106	0.0067	0.9050	0.9117	0.1920
61+	0.0500	(-) 0.0637	1.2170	1.2807	0.0640
Total	1.0004	-	0.7630	-	0.7630

Source: HIES: 1993-94.

### Within Age Group Inequality

As far as within age group earnings inequality is concerned, it is highest in the oldest age group i.e. 61+ years (see Table 3). As expected, the youngest age group 10-15 years also show significantly high incident of inequality in the earnings. There are many reasons for this inequality. These groups are vulnerable to market demand and supply conditions. Rapidly increasing labour force, return migration, rural to urban migration, and substitution of Pakistani workers by different refugee groups are some of the supply side factor. The downsizing of public sector, and slow progress of industrial sector are some of the demand side factors which are shrinking demand and putting downward pressure on the wages of these groups. Moreover, the changing technology is also creating skill mismatch for these groups.<sup>2</sup>

Most of the characteristics of these workers are identical. We have noted that majority of workers in these two groups work for whole year but receive wages much lower than the average wage. Almost 95 percent workers have received no training and mostly work in the informal sector. A higher percentage of these workers live in rural areas and work in the sales or service related occupations. Majority of them are males and mostly illiterate. These characteristics are presented in Table 4. Other reasons include intra industry wage differential and returns to different skills. Some of these reasons need to be investigated as they are out of the scope of this study.

<sup>2</sup>One can postulate these factors but real impact can be determined only by testing them empirically.

Table 4

*Selected Characteristics of Workers in Different Age Groups (Mean Earnings and Number of Workers in the Group)*

Age Group/ Characteristic	10–15 Years		61+ Years	
	Number	Mean Earnings	Number	Mean Earnings
Urban	205	738.25	264	2890.95
Rural	310	706.97	660	1623.13
Male	443	725.52	893	2032.16
Female	72	681.93	31	637.51
Married	3	444.22	786	2053.02
Single	512	721.04	138	1600.03
Training	18	730.10	16	4679.95
No Training	497	424.54	908	1674.34
Farm Sector	119	597.44	480	1686.96
Informal Sector	356	733.47	379	2310.51
Formal Sector	29	1150.31	51	2708.05
Self-employed	69	551.33	637	2038.23
Paid Employee	445	698.97	269	1674.34

Source: HIES: 1993-94.

### Across Age Groups Inequality

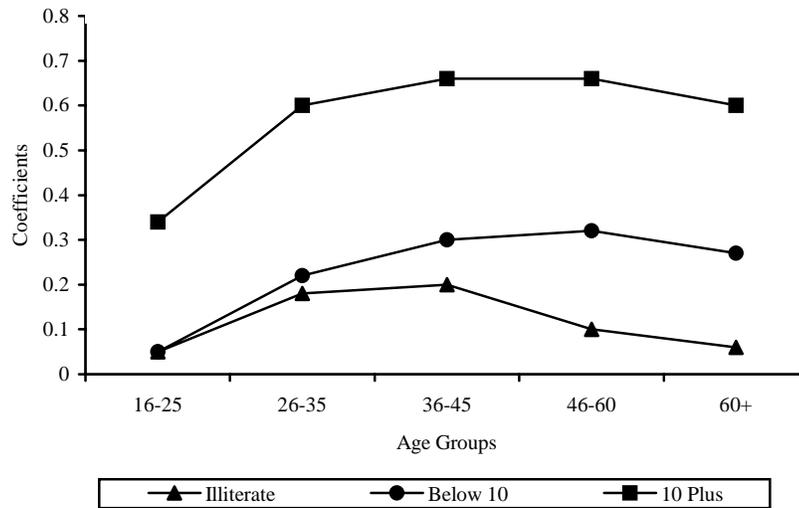
Table 3 shows that the intra age group inequality in the earnings is highest for the first age group i.e. 10–15 years. The oldest age group 61+ shows the second highest inequality in the earnings. The age earnings profiles are estimated by using age spline earnings functions. The results are presented in Table 5 and age-earnings profiles based on these estimates are presented in Figure 1. These earnings profiles reflect the human and physical investments of individuals to spread their earnings opportunities over their lifetime. These functions are estimated for overall Pakistan and three educational categories.

It is observed that earnings remain low till age 25. When educational attainments of workers held constant, the slope of age-earnings profiles consistently increase but the increase is quite significant after age 25. These workers have not only below average education but other skills as well. This suggests that the accumulation of skills continues till age 25 and only after age 25 they receive substantial returns of education and other skills which increases the slope of the age earnings profile. As postulated earlier, the returns differ for different educational achievements. The age-earnings profiles based on the estimates of earning functions presented in Figure 1 shows that the slope of the earnings profile of those who have 10 or more years of schooling is higher than those who have lower educational achievements.

Table 5

*Ordinary Least Squared Estimates of Earnings Equations*  
(Dependent Variable = Log Earnings)

Variables	All Pakistan		Illiterates		Edu<10		Edu>=10	
	Coeff	t	Coeff	t	Coeff	t	Coeff	T
Constant	5.9730*	169.30	5.799*	128.31	6.285*	68.21	6.781*	148.73
Age16–25	-0.0101	-0.358	0.066	1.84	0.052	0.84	0.344*	5.74
Age 26–35	0.1510*	5.74	0.190*	5.77	0.235*	4.01	0.600*	9.81
Age 36–45	0.2600*	9.81	0.237*	7.15	0.340*	5.76	0.687*	7.87
Age 46–60	0.2070*	7.87	0.140*	4.34	0.353*	5.87	0.671*	17.69
Age 61+	0.1521*	6.84	0.092*	3.42	0.281*	4.66	0.580*	6.54
TT	0.0885*	3.22	0.179*	2.65	0.022*	4.03	0.083*	2.58
URBAN	0.4040*	34.61	0.405*	22.99	0.399*	19.01	0.390*	17.61
SEX	0.6770*	30.42	0.810*	29.79	0.473*	7.18	0.192*	5.03
MSP	0.2180*	13.52	0.213*	8.98	0.201*	6.18	0.106*	3.65
SEMP	0.0895*	6.99	0.043*	2.45	0.154*	6.36	0.265*	8.38
EMPLR	0.7280*	15.62	0.544*	6.26	0.778*	8.18	0.922*	14.98
FORMAL	0.1720*	10.91	0.269*	9.62	0.167*	5.62	0.241	9.93
$\bar{R}^2$	0.328		0.198		0.22		0.307	
F-Statistics	601.240		199.20		109.04		163.63	
N	18472		9630		4689		4151	



**Fig. 1. Age Earnings Profiles of Workers.**

## VI. CONCLUSIONS AND POLICY RECOMMENDATIONS

In this study the inequality in personal earnings is analysed using household integrated economic survey data 1993-94. The variance of log monthly earnings and its decomposition is used to see which age groups have highest incident of inequality. The spline earnings functions are estimated to determine the factors playing important role for the inequality in personal earnings.

It is concluded that youngest and oldest age groups have highest contribution to overall inequality in the personal earnings. The comparison across age groups show highest inequality in the youngest age group followed by the oldest age group. It is also observed that earnings are below average for youngest (10–25) and oldest (61+) age groups. The age earnings profiles based on the estimates of earnings functions indicate that earnings are increasing functions of education and one of the main causes of inequality in the personal earnings. Other factors that create disparity in the earnings include sector of employment, regional location, sex, marriage, and other characteristics.

The inequality in the earnings measured within the groups found to be highest in the oldest age group (61+ years) followed by the 46-60 years old age group. Major causes include market supply and demand conditions, skill mismatch, and below average educational achievements of majority of income recipients in these groups. One conclusion can be drawn is that the labour market is structured differently for male and female workers. It is found that education is the most important factor for explaining the disparity in the earnings of Pakistani workers.

We can move into next millennium by improving the plight of Pakistani workers by taking certain steps.

1. As education plays important role for the success in the labour market, we can improve the access to education for all segments of the society.
2. Self-employment schemes, development of small and medium establishments are important steps to provide jobs but benefits should not be concentrated only some specific areas or groups rather, these should be targeted for all the segments of the society and regions.
3. An effective population policy is very important to control the rapidly increasing labour force, which is putting lot of pressure on wages, and adversely affecting income distribution.
4. There is dire need to establish training centre for providing latest skills to workers.

## REFERENCES

- Azfar, J. (1976) The Distribution of Income in Pakistan: 1966-67. *Pakistan Economic and Social Review* 14.

- Becker, Gary S. (1964) *Human Capital*. New York: National Bureau of Economic Research.
- Dooley, Martin, and Peter Gottschalk (1984) Earning Inequality among Males in the U.S. Trends and the Effect of Labour Force Growth. *The Journal of Political Economy* 92:1 59–89.
- Estudillo, J. P. (1997) Income Inequality in the Philippines. *The Developing Economies* I: 1961–91.
- Hamdani, Khalil A. (1977) Education and the Income Differential: An Estimation for Rawalpindi City. *The Pakistan Development Review* 26:2 144–164.
- Haq, Nadeem Ul (1977) Economic Analysis of Personal Earnings in Rawalpindi City. *The Pakistan Development Review* 26:4 687–696.
- Karoly, Lynn A. (1992) Changes in the Distribution of Individual Earnings in the United States: 1967–1986. *The Review of Economics and Statistics* 74:1 107–115.
- Kemal, A. R. (1981) Income Distribution in Pakistan: A Review. Pakistan Institute of Development Economics, Islamabad. (Research Report Series No.123.)
- Mahmood, Zafar (1984) Income Inequality in Pakistan. *The Pakistan Development Review* 23: 365–376.
- Mincer, Jacob (1974) *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.
- Moffitt, Robert A. (1990) *The Distribution of Earnings and Welfare Market*. Burtless. 201–230.
- Schultz, T. P. (1982) Age of Individual, Family Composition and Income Distribution. International Population Conference, Manila. 23–48.
- Shabbir, Tayyeb (1994) Mincerian Earnings Function for Pakistan. *The Pakistan Development Review* 26:4 687–696.
- Shabbir, Tayyeb, and Alia H. Khan (1991) Mincerian Earnings Functions for Pakistan: A Regional Analysis. *Pakistan Economic and Social Review* 24: 99–111.
- Topal, Robert (1997) Factor Proportions and Relative Wages: The Supply Side Determinants of Wage Inequality. *Journal of Economic Perspective* 2: 55–74.