

Inequality in the Four Provinces of Pakistan

HANS DE KRUIJK*

1. INTRODUCTION

The purpose of this paper is to analyse, compare and explain income inequalities in the four provinces of Pakistan on the basis of decomposition analysis. Overall income inequality is decomposed into various categories of inequalities in such a manner that the relative importance of each category can be quantified. Decomposition analysis searches for the main origins of income inequality not only within or between provinces, but also within and between urban and rural areas, labour income and non-labour income, etc. Identification of origins of inequality is important if policies aiming at reducing inequalities are considered. Clearly, different causes require different policies. A large 'decomposition tree' built for Pakistan recently [2] which provides information on the components of inequality is presented in the accompanying chart. The structure of inequalities in the four provinces is analysed on the basis of this 'decomposition tree'.

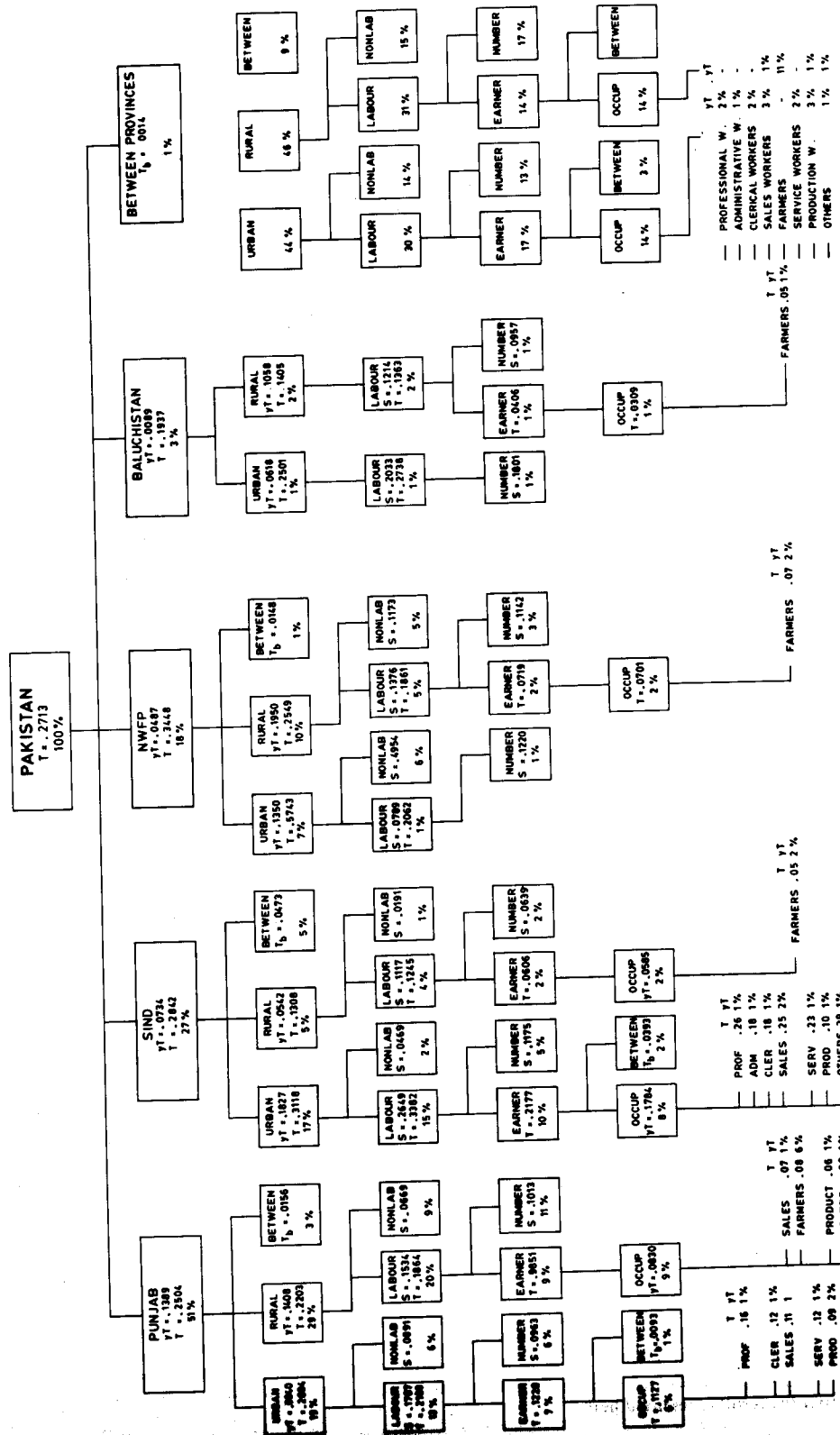
One of the most striking and counter-intuitive findings is that, in relative terms, income inequality between provinces is very minor. Though in absolute terms, average monthly household income in Baluchistan is about Rs 100/- lower than in the Punjab and about Rs 200/- lower than in Sind and the NWFP, this gap explains less than one percent of the overall income inequality in the country. We come back to this issue in Section 2. Overall inequality in Pakistan appears to be almost exclusively due to inequalities within provinces. Inequality is highest within the NWFP, and is followed in a descending order by Sind, the Punjab and, finally, Baluchistan. Reasons for these differences in inequality will be discussed in Section 3.

A detailed *dynamic* analysis of structural changes of distributions within provinces is not feasible yet, because only the Household Income and Expenditure Survey of 1979 gives income distribution data by province. Therefore, studies of the development of income inequality during the 1970s could only be carried out at a higher level of aggregation, i.e. for Pakistan as a whole [3; 4].¹ By comparing distributions relating to different years, such analyses reveal changes in the structure of

*The author is Lecturer in Economics of the Centre for Development Planning, Erasmus University, Rotterdam (the Netherlands). He would like to express his gratitude to Prof. P. A. Cornelisse, Dr S. I. Cohen and Dr S. L. Lodder for their stimulating comments.

¹In an earlier study [4] we show that in both rural and urban areas poverty has declined while at the same time income inequality has increased. We used four inequality measures and also four poverty indicators. All measures and indicators pointed to the same conclusion. That is our earlier papers use aggregate figures for some groups. That is why those results differ slightly from the results presented here.

CHART



income distribution. A number of structural shifts appear to play an important role in explaining increasing inequality in Pakistan over time. Section 3 considers whether the *dynamic* factors thus identified can also explain *static* differences in the level and structure of inequalities in the four provinces of Pakistan in 1979.

The tool of analysis is the Theil coefficient which is a measure of income inequality that is *additively* decomposable. This means that as far as population subgroups are concerned total inequality can be written as the sum of the weighted averages of inequality within subgroups and the inequality between these groups. As far as factor components are concerned, the measure can be split up into the contributions made by the different factors and those made by the participation effect.

In summary, the outline of the paper is as follows. Section 2 discusses the 'decomposition tree'. Section 3 considers whether important factors explaining dynamic changes in inequality also play a role in understanding static differences in inequality in the four provinces. Some concluding remarks are made in Section 4. Appendices contain mathematical formulas and a discussion of the data base.

2. DECOMPOSITION TECHNIQUES

The structure of income inequality can be derived by decomposing income inequality according to occupation, industrial sector, location, production factor, etc. A comparison of different income-distribution structures in different provinces of the same country may provide further insight into causes behind overall income inequality in that country.

Decomposition analysis subdivides the two elements which define income distribution, viz. income and population groups, into various components. Income is disaggregated by source of income, like labour income and non-labour income, while population groups can be split up according to locality, like provinces, or socio-economic groups.

Theil's measure for overall inequality (*T*) is used because it can easily be decomposed. *T* can be written as

$$T = \sum Y_i \ln (Y_i/N_i)$$

in which

y_i = income share of income class *i*, and

N_i = household share of income class *i*.

In a differentiation by provinces, overall inequality can be written as the weighted sum of inequalities within each province plus the inequality between provinces, where weights are the respective income shares of each province.

$$T = \sum (Y_p * T_p) + T_b$$

in which

Y_p = provincial income shares;

T_p = provincial Theil coefficients; and

T_b = inequality between provinces, which can be written as:

$$T_b = \sum Y_p \ln (Y_p / N_p)$$

in which N_p = provincial household shares.

Provincial Theil coefficients are calculated with data from the Household Income and Expenditure Survey (HIES) for 1979 [7]. Provincial income shares are calculated using average household income for urban and rural areas in each province, as reported by HIES, weighted with provincial urban and rural household shares, as reported by the 1981 Population Census. Filling in the figures, we get

inequality within Pakistan	contribution of inequality within Punjab	contribution of inequality within Sind	contribution of inequality within NWFP
0.2713 (100%)	= (.5547 * .2504) (51%)	+ (.2583 * .2842) (27%)	+ (.1412 * .3448) (18%)
	contribution of inequality within Baluchistan	contribution of inequality between provinces	
	(.0459 * 0.1937) (3%)	+ 0.0014 (1%)	

With regard to inequality *within* provinces, it appears that the Theil coefficient of 0.3448 for the NWFP is relatively high and clearly higher than the corresponding coefficient for the Punjab (0.2504). Nevertheless, the contribution of the Punjab to overall inequality is more than fifty percent because of its high income-share. Further decomposition is required to discover reasons for different levels of inequality in the provinces.

Apparently, household income inequality *between* provinces is very small — less than one percent of total inequality. This counter-intuitive result requires

further examination. Table 1 shows that average income in Baluchistan is about Rs 100/- lower than in the Punjab and about Rs 200/- lower than in the NWFP and Sind, which is a substantial difference. However, income differences within provinces are much larger. For example, comparing average monthly household income of the highest income class in urban NWFP (constituting 4 percent of the households) amounting to Rs 13,000/- with the average income of Rs 250/- for the lowest income class (2 percent of the households), it becomes clear that inequalities within provinces are of another order of magnitude. When measuring inequalities between provinces, account is taken only of differences between average incomes, assuming that all households within a certain province earn exactly the same income. Of course, income inequality in a country would be very low if the difference between the highest and the lowest income is Rs 200/-, or 20 percent only.

Table 1
Average Monthly Household Income by Province

	Punjab	Sind	NWFP	Baluchistan
Urban	1,214	1,476	1,607	1,357
Rural	828	794	936	762
Total*	931	1,089	1,032	845

Sources: [7] for Rows 1 and 2; [9].

*Total averages differ from those obtained from HIES [7] data because of the differences in the weights applied. The above-mentioned figures are calculated using average household income for urban and rural areas in each province as reported by HIES, but weighted with provincial urban and rural household shares as reported by the 1981 Population Census.

Provincial Theil coefficients are decomposed further into inequalities within and between urban and rural areas.² In the following formula,

$$T_p = (T_{pu} * T_{pu}) + (Y_{pr} * T_{pr}) + T_{pb}$$

for each province,

Y_{pu} = urban income share;

Y_{pr} = rural income share;

²The order of decomposition slightly influences the results; see [1].

$$T_{pu} = \text{Theil coefficient urban areas;}$$

$$T_{pr} = \text{Theil coefficient rural areas; and}$$

$$T_{pb} = \text{inequality between urban and rural areas.}$$

Filling in data for, for example, Sind and the NWFP (see also the Chart), one can see immediately that the structures of inequality in these provinces are completely different.

$$\begin{array}{rclcl} \text{Sind:} & 0.2842 & = & (0.59 * 0.3118) & + & (0.41 * 0.1308) & + & 0.0473 \\ & (100\%) & & (64\%) & & (19\%) & & (17\%) \end{array}$$

$$\begin{array}{rclcl} \text{NWFP:} & 0.3448 & = & (0.23 * 0.5743) & + & (0.77 * 0.2549) & + & 0.0148 \\ & (100\%) & & (39\%) & & (57\%) & & (4\%) \end{array}$$

Both the level and the relative importance of T_{pb} are higher in Sind than in the NWFP. It is not surprising that the difference between Karachi and the rural areas in Sind appears to be larger than the difference between Peshawar and the rural NWFP (see also Table 1). It is also understandable that inequality in cities is higher than in the countryside. Furthermore, urban inequality has more impact on provincial inequality in Sind than in the NWFP owing to a higher urban income weight. Note, however, that urban income inequality is considerably lower in urban Sind than in urban NWFP. But the reasons why the level of inequality in both urban and rural areas is higher in the NWFP than in Sind are not immediately clear. Therefore, further decomposition is required.

Apart from decomposing T into inequalities within and between groups of households, T is also decomposable according to different sources of income (see Section 2). In this way, the extent to which household income inequality is due to inequalities in labour earnings or to inequalities in property incomes can be located. Formulas for this kind of decomposition are presented in Appendix 1 while the results are shown in the Chart. It can be seen that the above-mentioned high inequality in the NWFP is mainly due to non-labour incomes. We come back to this issue in Section 3.

Another important factor in explaining differences between household incomes is the number of earners per household. Clearly, it makes a great difference whether a household has one earner or two or even more. Therefore, household labour incomes are converted into labour incomes per earner, and labour income inequalities are decomposed into earners' inequalities and into inequalities of the number of earners per household both within urban areas and within rural areas. Again, earners' income inequalities are expressed by a Theil coefficient. (For formulas, see Appendix 1.) Finally, these regional earners' income Theil coefficients are decomposed into inequalities within and between occupational groups.

The Chart shows that the relative contributions of various components are different for each province. Roughly speaking, overall household income inequality in the country is due to the following components:

- inequality between provinces = 1%,
- inequality between urban and rural areas = 9%,
- inequality in non-labour incomes = 29%,
- inequality due to different number of earners per household = 29%,
- inequality between occupational groups = 4%, and
- inequality within occupational groups = 28%.

Factors behind different structures in the four provinces are discussed in the next section.

3. DIFFERENT PROVINCIAL INEQUALITY STRUCTURES

This section presents the main differences in the structure of inequality in the four provinces and examines a number of factors that may be responsible for those differences. Table 2 summarizes the *level* of inequality in urban and rural areas in

Table 2

Theil Coefficients, Income Shares and Household Shares for Urban and Rural Areas, by Province

Province/ Urban, Rural	Theil Coefficient	Income Share (%)	Household Share (%)
Punjab	.25	100	100
Urban	.26	35	27
Rural	.22	65	73
Sind	.28	100	100
Urban	.31	59	43
Rural	.13	41	57
NWFP	.34	100	100
Urban	.57	23	14
Rural	.25	77	86
Baluchistan	.19	100	100
Urban	.25	22	14
Rural	.14	78	86

the four provinces, while the *structure* of inequality is summarized in Table 3. Table 2 shows that in all provinces inequality is higher in urban areas than in rural areas. Among urban areas, inequality is relatively high in Sind and the NWFP. In urban Sind, this is mainly due to high labour income inequality, whereas in NWFP inequality is almost exclusively embodied in non-labour incomes (Table 3). In rural areas, inequality is relatively high in the Punjab and the NWFP (again due to non-labour income inequality). Income inequality is low in Baluchistan, especially when it is realized that half of total inequality is due to a difference in the number of earners per household.

Table 3
Provincial Inequality Structures

(Percentages)

	Punjab	Sind	NWFP	Baluchistan
Total Inequality within Provinces	100	100	100	100
Inequality between Urban and Rural Areas	6	17	4	13
Inequality within Urban Areas				
Non-labour Incomes	13	10	34	6
Labour Incomes				
Number of Earners	10	19	3	17
Earners' Inequality				
Between Occupations	1	6	0	1
Within Occupations	13	29	2	8
Inequality within Rural Areas				
Non-labour Incomes	17	3	26	7
Labour Incomes				
Number of Earners	21	8	19	34
Earners' Inequality				
Between Occupations	3	—	1	3
Within Occupations	16	8	11	11

In analysing the causes behind these interprovincial differences, we make use of an earlier work relating to changes in income inequality in Pakistan over time. It will be considered whether major determinants of increasing inequality in Pakistan during the 1970s are also important in explaining *static interprovincial differences*. The following five structural shifts appear to play a major role in explaining *increasing inequality in Pakistan over time*.

1. Shift from rural areas to urban areas;
2. Within urban areas:
 - increased share of non-labour incomes;
 - expanding modern sector;
3. Within rural areas:
 - increased share of non-agricultural activities;
 - within agriculture: shift from tenancy to ownership.

Possible structural interprovincial differences of this kind are discussed in successive subsections.

Relative Size of Urban versus Rural Sectors within Provinces

Provincial Theil coefficients can be expressed as the weighted sum of inequalities within urban and rural areas (plus inequality between urban and rural areas) weighted by the respective income shares of urban and rural areas. Since inequality is higher in urban areas than in rural areas in all provinces (Table 2), the size of the urban income weight is directly related to the level of inequality within provinces. In other words, if urban Theil coefficients and urban average incomes are, respectively, exactly the same in all provinces (similarly in rural areas), provincial Theil coefficients would differ with the size of their respective urban sectors.

The Punjab and, particularly, Sind have large urban sectors compared with the NWFP and Baluchistan (Table 2). Nevertheless, inequality is highest in the NWFP, notwithstanding the relatively small size of its urban sector. Apparently, the countervailing power of other factors exceeds the impact of the size of the urban sector in the NWFP. The ranking from high inequality to low inequality for the other three provinces, viz. Sind, Punjab and Baluchistan (T is 0.28, 0.25 and 0.19, respectively) is in line with the ranking of the size of their urban sectors (0.59, 0.35 and 0.22, respectively).

Inequality within Urban Areas

Not only the size of the urban sector differs from province to province, but the economic structure of the urban sectors varies as well. Two issues are discussed below, viz. the importance of non-labour incomes and the relative size of the formal sector in urban areas.

Share of Non-labour Incomes

Apart from inequality which can be decomposed according to population groups, the Theil coefficient can also be decomposed according to income sources. In that case, the Theil coefficient is not the sum of inequality within and between

groups, but T is simply split up into two components adding to T . (For formulas, see Appendix 1.) The reason is that one can not speak of a group of households earning income from labour and another group of households receiving non-labour income, because a certain household can get income from both sources. The results of these decomposition exercises are presented in the Chart and summarized in Table 4.

Table 4 shows that the share of non-labour income in total income does not differ very much between urban Punjab, Sind and Baluchistan, but both the share and the level of non-labour income are significantly higher in urban NWFP. The major components of non-labour income are property income, mainly from land and houses, interest and profits. Unfortunately, the *HIES* does not provide detailed information about the size of these components. Since there are no clear reasons to presume that income from land or houses is much higher in the NWFP than in other provinces, it is likely that large trade-margins on a variety of items play an important role in Peshawar. Inequality of non-labour incomes is much higher than inequality of labour earnings. Average monthly non-labour income of the richest four percent of households in urban NWFP is about Rs 9,000! The remaining 96 percent of the households get on average about Rs 210 per month of non-labour income, which is roughly the same amount as similar groups in other provinces receive. Further research is required to discover the peculiarities of this phenomenon.

Table 4 further shows that average labour incomes per earner do not differ much between urban areas. Average earnings are a bit higher in Karachi, but most likely costs of living are also higher in a metropolis. Inequality between earners is lower than inequality between households, because the average number of earners per household increases with household income. The relatively high earners' income inequality in Karachi is discussed below.

Size of the Modern Sector

Four indicators are used to compare the sizes of the formal sectors in urban areas in the four provinces. These are (i) the size of the manufacturing sector, (ii) the number of professional, technical, administrative, executive, managerial and clerical workers in manufacturing, (iii) the size of the banking sector, and (iv) the number of professional etc. workers in the non-manufacturing private sector. All indicators point in the same plausible direction (Table 5): urban Sind (Karachi) has the largest modern sector, followed by urban Punjab. Since both average income and inequality are higher in the modern sector than in the informal sector, the relative size of the modern sector affects the level of earners' inequality (Table 4) in two ways. Firstly, inequality between the two sectors is more pronounced with a relatively large modern sector. (See formulas in Appendix 1.) Secondly, the high inequality within the modern sector gets more weight with a large modern sector.

Table 4
Inequality in Labour and Non-labour Incomes by Province, Urban Areas

Urban Areas	Average Monthly Household Income		Percentage of Non-labour Income		Non-labour Income Inequality		Average Labour Household Income		Labour Income Inequality		Average Labour Income per Earner		Earners' Income Inequality	
	Income	Household	Non-labour Income	Non-labour Income	Non-labour Income Inequality	Household Income	Household Income	Labour Income	Labour Income Inequality	per Earner	Income	Income Inequality	Income	Income Inequality
Punjab	1214	256	22%	0.47	0.22	957	957	575	0.22	575	0.12	575	0.12	
Sind	1476	291	20%	0.31	0.34	1184	1184	675	0.34	675	0.21	675	0.21	
NWFP	1607	564	35%	1.65	0.21	1038	1038	595	0.21	595	0.09	595	0.09	
Baluchistan	1357	312	23%	0.29	0.27	1048	1048	580	0.27	580	0.09	580	0.09	

Table 5

Percentage Distribution of Employed Persons by Sector, Occupational Groups and Provinces, Urban Areas

	Punjab	Sind	NWFP	Baluchistan
Manufacturing	26.67	26.71	16.3	9.79
(prof., techn., admn., cler. workers)	1.41	2.28	1.3	0.36
Financing, Insurance, Real Estate and Business Services	2.37	3.60	1.87	1.30
Prof., Tech., Adm., Cler. Workers in Non-manufacturing Private Sector	5.12	7.53	5.02	5.01

Source: [8].

The ranking from high to low earners' income inequality for urban areas in the four provinces is in line with the ranking of the size of their modern sectors in the sense that urban Sind is definitely Number One and the Punjab is Number Two, while the size of the modern sector as well as earners' inequality are both small in urban NWFP and Baluchistan.

Inequality within Rural Areas

The contribution of rural areas to overall inequality in Pakistan is about 50 percent, while approximately 75 percent of households are living in rural areas [7; 9]. Apparently, inequality is lower in rural than in urban areas. Though Theil coefficients confirm this conclusion, one must be very careful in interpreting this differential in equality. For what is measured is inequality between households *living* in rural areas and not inequality of income *generated* in rural areas. Income transfers between rural and urban areas may cause a bias. Landlords living in urban areas get income from their property located in rural areas. This income is reported as urban income and contributes to urban inequality, whereas it is the result of inequalities in rural areas, like land inequality. Further, urban workers may transfer remittances to family members living in the countryside. The amount of transfers is not known, which hampers the analysing of causes of inequality in the distribution of income in rural areas. Keeping this in mind, we will consider differences in the economic structure of rural areas in the four provinces.

Size of the Agricultural Sector

Table 6 presents for each province the percentage of earners working in agricultural and non-agricultural sectors. Rural Sind has the smallest non-agricultural sector, but it has hardly any impact on interprovincial differences in inequality because differences in earnings between and within occupational groups are very

Table 6

Inequality and Average Income by Sector and Province, Rural Areas

	Punjab	Sind	NWFP	Baluchistan
Percentage of				
Agricultural Workers	63%	82%	55%	64%
Non-agricultural Workers	37%	18%	45%	36%
Average Income				
Agricultural Workers	370	288	433	351
Non-agricultural Workers	353	324	424	379
All Rural Workers	363	294	429	362
Theil Coefficient				
Agricultural Workers	0.08	0.05	0.07	0.05
Non-agricultural Workers	0.09	0.09	0.08	0.03
All Rural Workers	0.08	0.06	0.07	0.04

Source: Calculations based on [7].

small. Income differences between households in rural areas are mostly due to differences in the number of earners per household (see also Table 3).

Different Agrarian Systems

The mode of agricultural production varies greatly with provinces. Feudalism is widespread in Sind; great parts of the NWFP and Baluchistan still have a tribal character; and in the Punjab a mixed system of ownership and tenancy prevails (Table 7). Different agrarian systems lead to different patterns of the level and distribution of agricultural incomes. In a feudal system, the share of non-labour incomes is relatively high (payments from tenants to landlords are classified as non-labour incomes of landlords) and the distribution between landlords and tenants is relatively skewed. Therefore, it is to be expected that average income of agricultural workers is low and the share of non-labour incomes and inequality are relatively high in rural Sind. Table 6 shows that average income of agricultural workers indeed is significantly lower in Sind than in other provinces, but inequality is not higher (see Chart). Probably, the reason is that many landlords are living in cities. In that case, payments from tenants to landlords are income transfers from rural to urban areas, which may be another cause of large inequality between urban and rural areas in Sind. Further research is required to discover the amount of these transfers.

As far as the distribution of labour incomes among agricultural workers living in rural areas is concerned, inequality is low in rural Sind in spite of large numbers of

Table 7

*Distribution of Rural Households According to Activity, Status and Provinces,
Pakistan 1979-80*

Activity/Status	(Percentages)			
	Punjab	Sind	NWFP	Baluchistan
All rural households	100	100	100	100
Activity				
Non-agricultural	37	18	45	36
Agricultural	63	82	55	64
Status				
Landless	18	34	19	36
Not landless	45	48	36	28
Owner	25	19	25	21
Owner-cum-tenant	11	5	5	2
Tenant	10	23	7	5

Sources: [5, for Baluchistan]; [6]; [7]; [8] and [9].

landless agricultural labourers (Table 6). The average size of tenants' cultivated area in Sind is eight acres, with small variation. Their income is low because 50 percent of their produce from eight acres is handed over to landlords. Income of landless labourers may be lower but not much. Inequality between agricultural workers is higher in the NWFP and the Punjab because land inequality between owner cultivators is always higher than land inequality between tenants.

4. SUMMARY AND CONCLUSIONS

A large 'decomposition tree' has been presented, which is a useful framework for analysing income distribution in the four provinces of Pakistan. The following categories of inequalities are distinguished: inequalities between and within urban and rural areas, inequalities of labour and non-labour incomes, inequalities due to differences in the number of earners per household, and inequalities between and within occupational groups. This list is not exhaustive and can be enhanced by inserting more categories. Especially non-labour incomes require further decomposition. By putting all these inequalities into one consistency framework it is possible to localize the inequalities and sense their proportions.

Currently held perceptions on income inequality are based on 'visible' differences like different wage rates in urban and rural areas or different remuneration

rates for different occupations (which happen to explain only a few percentages of total inequality) or inequality between provinces (whose contribution to overall inequality is less than one percent). However, the origin of household income inequality in Pakistan is mainly embodied in two factors, explaining about 30 percent of overall inequality each, viz. inequality of non-labour incomes and differences in the number of earners per household.

One of the main findings of this paper is that not only the level but also the structure of inequality differs substantially among the provinces. About 60 percent of inequality in the NWFP is due to non-labour income inequality. Differences in the numbers of earners per household contribute to 50 percent of total inequality in Baluchistan. Inequality in Sind is mainly due to inequalities within urban areas (relatively large modern sector) and to income differences between urban and rural areas, whereas inequality in Punjab is mainly located in rural areas.

Localizing various categories of inequality is important if policies for reducing inequality are considered, because different types of inequality require different policy instruments.

This static comparative study of the structure of inequality in the four provinces of Pakistan has made use of findings from earlier work on dynamic changes in income distribution in Pakistan at an aggregate level. But a greater advantage of decomposition analysis can be gained by comparing disaggregated charts for different years, so that disaggregated changes in inequalities can be located and analysed. That will be our subject for further research.

Appendix I

MATHEMATICAL EXPRESSIONS

1. Population Subgroups¹

The Theil coefficient (T) can be written as.

$$T = \mu_i \sum_i \frac{n_i}{n\mu} \ln \left(\frac{\mu_i}{\mu} \right) \dots \dots \dots \quad (1)$$

in which

- μ_i = average income of income class i
- n_i = number of households in income class i
- n = total number of households
- μ = average income

Total inequality T can be written as the weighted sum of inequalities within subgroups and inequality between subgroups. If each household belongs to one and only one subgroup (e.g. the province in which the household lives) then inequality T_r for each subgroup r can be calculated by:

$$T_r = \mu_{ir} \sum_i \frac{n_{ir}}{n_r \mu_r} \ln \left(\frac{\mu_{ir}}{\mu_r} \right) \dots \dots \dots \quad (2)$$

in which

- n_r = number of households in subgroup r
- μ_r = average income of subgroup r
- n_{ir} = number of households in income class i of subgroup r
- μ_{ir} = average income in income class i of subgroup r

Inequality between the subgroups T_b is

$$T_b = \mu_r \sum_r \frac{n_r}{n\mu} \ln \left(\frac{\mu_r}{\mu} \right) \dots \dots \dots \quad (3)$$

It follows that total inequality can now be written as

$$T = T_b + \frac{n_r \mu_r}{n\mu} \sum_r T_r \dots \dots \dots \quad (4)$$

¹See [10].

T_b is called the explained part of total inequality. The relative contribution of between-group inequality to total inequality is: $\frac{T_b}{T} \times 100$ percent. The rest of inequality is located within subgroups.

A successive decomposition analysis of unexplained parts requires further breakdowns within subgroups, e.g. by distinguishing more disaggregated subgroups (e.g. urban and rural areas within provinces) or by a breakdown of household incomes into factor components. The latter will be discussed now.

2. Income Factor Components²

In this case *incomes* of households are broken down into incomes from various sources. If two sources are distinguished, viz. labour income and non-labour income, the formula for T becomes:

$$T = [\mu_i^a + (\mu_i - \mu_i^a)] \sum_i \frac{n_i}{n\mu} \ln \left(\frac{\mu_i}{\mu} \right) \dots \dots \quad (5)$$

in which

- μ_i^a = average labour income in income class i , and
- $\mu_i - \mu_i^a$ = average non-labour income in income class i .

The part of total inequality that is attributable to labour income (S_a) is then

$$S_a = \mu_i^a \sum_i \frac{n_i}{n\mu} \ln \left(\frac{\mu_i}{\mu} \right) \dots \dots \dots \quad (6)$$

and the part that is attributable to non-labour income (S_{na}) is

$$S_{na} = T - S_a = (\mu_i - \mu_i^a) \sum_i \frac{n_i}{n\mu} \ln \left(\frac{\mu_i}{\mu} \right) \dots \dots \quad (7)$$

3. Participation within Households

An important determinant of household labour income is the number of earners per household. Of course, it counts whether a household has one earner or two earners (or even more). The contribution of this effect is measured as follows.

As mentioned before, the Theil coefficient for households (T) can be written as follows.

$$T = \sum_i \frac{n_i \mu_i}{n\mu} \ln \frac{\mu_i}{\mu} \dots \dots \dots \quad (8)$$

²See [11].

Similarly, the expression for the Theil coefficient for earners (T_e) is

$$T_e = \sum_i \frac{n_{ie} \mu_{ie}}{n_e \mu_e} \ln \left(\frac{\mu_{ie}}{\mu_e} \right) \dots \dots \dots (9)$$

in which

- n_{ie} = number of earners in income class i
- μ_{ie} = average income of earners in income class i
- n_e = total number of earners
- μ_e = average income of earners

Since $\frac{n_{ie} \mu_{ie}}{n_e \mu_e} = \frac{n_i \mu_i}{n \mu}$ (the income share per income class remains the same)

and: $\mu_{ie} = \frac{\mu_i}{e_i^h}$, and $\mu_e = \frac{\mu}{e^h}$ by definition

in which

- e^h = average number of earners per household
- e_i^h = average number of earners per household in income class i ,

it follows that

$$T_e = T + \sum_i \frac{n_i \mu_i}{n \mu} \ln \left(\frac{e^h}{e_i^h} \right) \dots \dots \dots (10)$$

The contribution of different number of earners per household to inequality of household labour income is therefore

$$-\sum_i \frac{n_i \mu_i}{n \mu} \ln \left(\frac{e^h}{e_i^h} \right)$$

Appendix 2

THE DATA BASE

The only source publishing data on household incomes in Pakistan and covering the entire income range is the *Household Income and Expenditure Survey (HIES)*.

The lowest level of aggregation of the latest published *HIES* 1979 is that for twelve income classes (less than 300 rupees per month, Rs 301–400, Rs 401–500, etc.) in eight regions (urban and rural areas in the four provinces). For each income class in each region, data are presented on the components of household income (labour income and non-labour income), on the number of earners per household, and on the occupation of the earners. Some remarks are due with respect to using this data base for the present analysis.

Firstly, doubts about the *reliability* of data refer mainly to an understatement of incomes accruing to the highest income group, but not to the scope and coverage of the survey which seems excellent. A sample size of almost 20,000 households, which is more than 120,000 persons, for a country with a population of about 80 million persons is very reasonable.

Secondly, this paper makes use of published data in the form of *grouped* data and not of the unpublished individual income data from data tapes. Inequality indicators derived from grouped data underestimate true inequality because inequality within brackets is neglected. The extent of underestimation depends on the number of income brackets. It appears from empirical exercises that when using 10 or more income brackets the underestimation error is no more than a few percentages. Gastwirth (1972) calculated an upper bound (maximum inequality within brackets) and a lower bound (no inequality within brackets) of the Gini coefficient of U.S. household incomes which are classified in 10 income brackets. The difference between the upper bound and the lower bound of the Gini coefficient is 5 percent only. Odink and Imhoff (1984) computed the difference between the Theil coefficient based on grouped data and the Theil coefficient based on individual income data for 5,666 gross incomes of heads of households in 1979 in the Netherlands, ranging from 300 to 250,000 guilders. They found a difference of 0.3 percent, using 30 income brackets. In another exercise, using data on wages of 775 employees of a moderate large Dutch firm in the food industry ranging from 80 to 13,500 guilders, they found an underestimation error of 1.4 percent with 10 income brackets only. Since the *HIES* distinguishes 12 income brackets the Theil coefficient computed from published data should not differ significantly from the true Theil coefficient computed from individual incomes.

Thirdly, the *units of measurement* in this paper are income per household and income per earner. The reason that we measure income inequalities on the basis of total household income and not on the basis of per person household income is

the lack of sufficient information about the extent of economies of scale of households in Pakistan. For total expenditures of an 8-person household are much lower than the eightfold of total expenditures of a one-person household.

Finally, this paper considers inequalities in *income* and not in purchasing power. The latter requires data on price differentials between provinces and between rural and urban areas.

REFERENCES

1. Adelman, Irma, and Amnon Levty. "Decomposing Theil's Index of Income Inequality into, between and within Components". *Review of Income and Wealth*. Series 30, No. 1. March 1985.
2. de Kruijk, Hans. *Income Inequality Decomposition: The Case of Pakistan*. Rotterdam: Erasmus University, Centre for Development Planning. April 1986. (Discussion Papers Series, No. 75)
3. de Kruijk, Hans. *When Poverty Declines and Inequality Increases: The Case of Pakistan during the 1970s*. Rotterdam: Erasmus University, Centre for Development Planning. 1986. (Discussion Papers Series)
4. de Kruijk, Hans, and Myrna van Leeuwen. "Changes in Poverty and Income Inequality during the 1970s". *Pakistan Development Review*. Vol. XXIV, Nos. 3&4. Autumn-Winter 1985.
5. Pakistan. Ministry of Food and Agriculture. Agricultural Census Organization. *Pakistan Census of Agriculture, 1972. Province Report: Baluchistan*. Lahore 1975.
6. Pakistan. Statistics Division. Agricultural Census Organization. *Pakistan Census of Agriculture, 1980: All-Pakistan Report*. Lahore. 1983.
7. Pakistan. Statistics Division. Federal Bureau of Statistics. *Household Income and Expenditure Survey, 1979*. Karachi. 1983.
8. Pakistan. Statistics Division. Federal Bureau of Statistics. *Labour Force Survey, 1978-79*. Karachi. 1982.
9. Pakistan. Statistics Division. Federal Bureau of Statistics. *Statistical Pocket Book of Pakistan, 1982*. Karachi. 1982.
10. Shorrocks, A.F. "The Class of Additively Decomposable Inequality Measures". *Econometrica*. Vol. 48. No. 3. 1980. pp. 613-625.
11. Shorrocks, A.F. "Inequality Decomposition by Factor Components". *Econometrica*. Vol. 50, No. 1. 1982.

Comments on "Inequality in the Four Provinces of Pakistan"

Mr Hans de Kruijk's continuing interest in the phenomena of poverty and inequality in Pakistan is highly commendable, especially as this set of studies seems to be receiving decreasing attention from Pakistani economists. In the PSDE meeting last year, he presented to us the work undertaken by him, jointly with another Erasmus colleague, on changes in poverty and income inequality in Pakistan during the 1970s. The present paper, using the same data and methodology, concentrates on the regional dimensions of economic inequalities in Pakistan. The analysis relates to 1979-1980, the latest year for which the Household Income and Expenditure Survey (HIES) data were available.

Mr de Kruijk, using the Theil Inequality Index, attempts to decompose the total index into its provincial and interprovincial components. He then further decomposes, for each province, the contribution of inequality for urban and rural areas, as well as that between them. Intra-rural and intra-urban inequalities are further decomposed into labour and non-labour income inequalities and the process is carried through, where data are available, to discover the contribution of occupational structure to total inequality.

Some of the results obtained by Mr de Kruijk's analysis seem counter-intuitive, if not counterfactual, the least startling of which is the one pointed out by the author himself, viz. the contribution of interprovincial differences in average incomes being nominal. In the table which reports average monthly household incomes by provinces, the Punjab ranks third in total households, second in rural households and last in urban households. In terms of inequality rankings also, the Punjab is third in total households, second in rural households and third in urban households. Thus, neither in terms of per capita (ignoring the differences in household sizes which are not reported) incomes, nor in terms of inequality indices does the Punjab fare as badly as the popular perception of its being the most prosperous and highly inegalitarian province. Its ranking in terms of level of incomes is distorted partly by the inclusion of Karachi as part of Sind and partly by the more general problem, acknowledged by the author, of urban incomes including a large proportion generated in the rural areas. Karachi is a cosmopolitan city and part of the income of its residents is generated in different parts of the country, including the Punjab. A comparison of the figures in Table 1 with those on gross provincial product on a per capita or per household basis would have served to illustrate this difference. Account

would also need to be taken of income transfers, especially remittances from both within the country and abroad, in measuring both the level and distribution of incomes. I am not certain whether this is adequately done in the data used by the author.

An analysis of interregional differences is a very complex exercise. The author is naturally limited in his task by the quality and content of the data available to him. He does use other available data, such as those from the Agriculture Census and the Labour Force Survey, to account for difference in inequality between the provinces. However, a more rigorous analysis would need a closer interfacing of data on a comparable basis than is available at present. A possible way of enriching the analysis would have been to use district-level data rather than the much more aggregated data on the provinces. It would also be possible then to group together districts according to other criteria than the existing political and administrative boundaries, such as availability of irrigation, degree of urbanization, level of education, etc.

It is unfortunate that Mr de Kruijk has only chosen to analyse regional differences in terms of overall income inequality. In last year's paper, he presented the interesting finding that whereas inequality had been rising in Pakistan in the 1970s, there was a fall in the percentage of the people below the poverty line. It would be interesting to know whether this phenomenon occurred in all the provinces (or districts) or whether it was confined to only a few of them.

Development Planning Division,
ESCAP,
Bangkok (Thailand)

S. M. Naseem