Income inequality in rural Pakistan-sources and decompositions

Shams, Khadija

January 2012

Income inequality in rural Pakistan – sources and decompositions*

Khadija Shams†

June 7, 2012

Abstract

This paper analyses income inequalities in rural Pakistan. Using household survey data, we decompose income inequality according to its different sources. We calculate Gini coefficients and Theil indices both within and across provinces and districts. A partial correlation analysis extends our descriptive investigation to reveal the different impacts of the various income sources on overall income inequalities. The unique focus on rural areas and the more disaggregated (district level) approach permits more nuanced policy implications. We find that inequality between districts is higher than within districts. Non-farm and transfer income have the strongest impact on income inequality across districts both in economic and econometric terms. While the former source of income is inequality increasing, transfer income tends to reduce inequalities. Our results highlight in particular the need for factor mobility to facilitate transfer income.

JEL Classification: O12; O15; Q15.

Keywords: Income inequality, Gini coefficients, partial correlation, rural Pakistan.

*I am grateful to Luis Angeles, Alexander Kadow and seminar participants at the University of Glasgow for helpful comments. A considerable part of this paper was published in the *FWU Journal of Social Sciences, 2010, Vol. 4, No. 1, pp. 18-38*; and I also wish to acknowledge the corresponding Editor as well as anonymous referees.

†University of Glasgow, Department of Economics, Adam Smith Building, Glasgow, G12 8RT, United Kingdom; Tel: +44 (0)141-330 2592; Fax: +44 (0)141-330 4940; k.shams.1@research.gla.ac.uk

1 Introduction

All over the developing world policymakers are interested in devising new strategies for rebalancing skewed income distributions and reducing poverty. The choice of such strategies crucially hinges on an improved understanding of the sources of income inequality. Why do certain types of incomes go to particular groups of people? And what roles do variables such as land-ownership, migration and education play in improving the income distribution and lifting people out of poverty?

Using primary household survey data (see Shams (2012) for details), we identify different types of income to disentangle the impact of each source on overall income inequality in Pakistan's four provinces. Given that almost two thirds of Pakistan's population live in rural areas we confine ourselves to those households (Government of Pakistan, Statistics Division, 1998). Our dataset is wide-ranging, providing rather detailed information on different districts across the country. Other authors such as Adams and Alderman (1992) and Glewwe (1986) use panel data which covers up to three years, but are less detailed on the household level. We resort to a cross-sectional, yet well-designed, framework.

In comparison to similar studies conducted for countries in the region, our paper differs in two aspects. Regarding methodology, we decompose overall income inequality both by region and income sources. Furthermore, we exploit our rather detailed dataset fully, considering additional sources of income such as rental and livestock income. Our cross-sectional analysis suggests that while livestock income does not affect overall income inequality in rural Pakistan whatsoever, other types of income such as transfer, agricultural, nonfarm and rental contribute equally. A similar study has been conducted by Glewwe (1986) who finds that non-labour income (defined as the income derived from crop production by the landowners) is largely responsible for overall income inequality in rural areas.

This paper is structured as follows. Section 2 briefly discusses the related literature. In Sections 3 and 4, we first identify the different income sources and decompose them in a next step accordingly. Section 5 provides a decomposition of inequality by regions. We conduct a correlation analysis between the different income sources in Section 6. Section 7 concludes and outlines potential policy implications.

2 Related literature

Adams (1994) uses three year panel data to analyse the impact of non-farm income on income inequality in rural Pakistan. He describes the importance of rural non-farm income for the poor by decomposing total rural income into the following five sources: non-farm, agricultural, livestock, rental and transfer income. The decomposition shows that non-farm income represents an inequality-decreasing source of income. The study then decomposes the sources of non-farm income. This analysis reveals that while non-farm unskilled labour income has an equalising effect on the income distribution, non-farm government income has a disequalising effect.

Ahmad (2000) uses micro data based on the *Household Integrated Economic Survey (HIES)* from 1992-93 to calculate the distribution of income in rural and urban areas of Pakistan. He finds that Gini coefficients show a more favourable distribution in rural areas compared to urban areas. In a related study, using the same HIES from 1992-93, Ahmad (2002) examines income inequality among various occupations in Pakistan such as (i) legislators/officials, (ii) teaching/health and other professionals, (iii) skilled construction, manufacturing, craft and related trades workers, and (iv) unskilled agricultural, fishing and related workers. Ahmad (2002) finds that within the given occupational groups in the four provinces of Pakistan, the highest level of inequality is observed among skilled workers (Gini coefficient of 0.299), followed by inequalities amongst the group of legislators/officials (Gini coefficient of 0.273). In contrast to that, the Gini coefficient among unskilled workers is 0.180 and is the lowest within the professional class (Gini coefficient of only 0.136). This finding may have been due to the fact that many of them were government employees and the wage structure was more equal in the government sector.

Anwar (2005) provides a series of Gini coefficients based on a consistent methodology using grouped household income data over 17 HIES conducted by the Pakistan Federal Bureau of Statistics (FBS) during 1963 to 2002.[1] The calculated Gini coefficients are generally higher in the urban than in the rural areas. Anwar (2005) conjectures that this was because of the urban labour force being more diversified in terms of skills and education and therefore the wage incomes being more unevenly distributed than in rural areas. Moreover, income from self-employment was more

[1]The household survey years are 1963-64, 1966-67, 1968-69, 1969-70, 1970-71, 1971-72, 1979, 1984-85, 1985-86, 1986-87, 1987-88, 1990-91, 1992-93, 1993-94, 1996-97, 1998-99 and 2001-02.

deviated in urban areas than in rural areas as urban self-employment ranged from wealthy businessmen to poor workers, whereas the bulk of the rural self-employed were a rather homogeneous group being mostly employed in informal sector enterprises.

Nugent and Walther (1982) use panel data in ungrouped (i.e. disaggregated) form to examine the sources of income inequality in India. The study is based on only three sources (agricultural, nonagricultural and transfer). The paper examines the dramatic decline in income inequality observed in rural India between 1968-69 and 1970-71. The authors describe the effects of changes in weather, technology and other factors on the distribution of incomes, which indirectly influences the labour market. Nugent and Walther (1982) find that periods of bad weather crowd out workers from the low-income agricultural group and increase inequality. Good weather, on the other hand, opens up more job opportunities for the workers from the low-income group and hence reduces imbalances in the income distribution.

Glewwe (1986)'s decomposition analysis is based on only two income sources (labour and non-labour) in Sri Lanka. It is assumed that all profits and non-monetary income are non-labor income. Specifically, non-monetary income is mainly agricultural produce consumed in the household. Additionally, incomes of those who are self-employed are counted as profits. The results suggest that labor income inequality accounts for slightly more than half of overall inequality in urban and estate sectors but not in the rural sector, where total inequality attributes largely to non-labor income and income from the sale of agricultural products (accounted for as profits). The conclusions drawn relate non-labor income inequality to an unequal distribution of land and capital in rural areas. Labor income inequality is explained by education and high wages that are paid to government employees in urban and estate sectors.

Adams and Alderman (1992) use a decomposition analysis to estimate sources of agricultural income inequality in rural Pakistan. They find that imbalances in land ownership are not the main drivers behind inequality in agricultural income. Instead, income from returns to labor and crop profits contribute the most. According to their analysis, policy makers concerned about inequality in rural Pakistan should pay attention to find ways to reduce the disparities between abilities, for instance by teaching agriculturalists managerial and technical skills.

Anwar et al. (2004) suggest that unequal landownership is one of the most im-

portant causes of poverty and inequality in rural Pakistan as land is the principal asset in an agrarian economy. Numbers of landless households are substantially high in Pakistan. About 67% of households do not own any land. About 18% of the households own less than 5 acres of land and about 10% of the households own 5 to 12.5 acres of land, which merely provides a subsistence level of living. A very small proportion of households holds large farms. Strikingly, barely 1% of the households owns more than 35 acres of land suggesting a highly skewed landownership pattern (Anwar et al., 2004).

Overall, the various studies appear to highlight the importance of a meaningful land reform in order to tackle the poverty and inequality issues in rural Pakistan. Policies that stress labour income and advocate education and training as the main means for addressing inequality show only half of the picture. After all, how is inequality supposed to be improved through education, if access to quality education is dependent on asset and ownership inequality? Can a child coming from a landless household access a top school whose graduates enjoy high income and other privileges? In fact, this kind of purely education-focused policy may even trigger inequality in the first place. Therefore, any resulting policy implications appear only meaningful when related to land reforms. The experience of Taiwan and South Korea clearly shows that land reforms even at the early stage of development contribute to shared economic growth which does not add to inequality.

3 Income sources

3.1 Overall income shares

The rather detailed dataset allows for a differentiation between various income types. Moreover, we can identify the impact of each of the different sources of income on income inequality at different levels of aggregation. As laid out in the questionnaire, individual households' total income may stem from up to five sources, y_i:

- **Transfer income (y_t)** includes internal and international remittances, government pensions and *zakat* (payments to the poor). Transfer income includes income earned from migration, both within and outside of Pakistan. Income

earned from the first type of migration is treated as internal remittances; income from the latter as external remittances.

- **Agricultural income (y_a)** includes net income (cash as well as in kind) from all crop production plus wage earnings from agricultural labour.
- **Nonfarm income (y_{nf})** includes wage earnings from nonfarm labour including self employment, government and private-sector employment.
- **Rental income (y_r)** includes rents received from ownership of assets including land, machinery (tractors, threshers), buildings, and water.
- **Livestock income (y_l)** includes net returns from traded livestock (cattle, poultry) plus imputed values of home-consumed livestock.

On a purely descriptive level, we can calculate each income source's share in total monthly household income data, denoted by S_i. According to Table 1, transfer income is the most important source of income, accounting for nearly one third of mean monthly household income. This seems to suggest that much of transfer income may come from migration to urban areas. Migration hence seems to have an important effect. In addition, the households may also tend to rely on help both from the government but especially from within their own families, neighbours and other peers. Since we are exclusively looking at rural backgrounds, agricultural income obviously is the major source of labour income. Nonfarm, livestock and rental income are of roughly equally (low) importance.

y_i	S_i
y_{nf}	18%
y_a	23%
y_t	30%
y_l	15%
y_r	14%

Table 1: *Monthly income shares. $N = 600$ households. Source: Survey 2008.*

3.2 Distribution of sample households by income quintiles

The distribution of the sample households by income quintiles is shown in Table 2. Two points are worthwhile noticing: First, the shares of transfer, agricultural and livestock income are relatively higher in lower income quintiles as compared to the higher income quintiles and may be considered as major sources of income for the households belonging to lower income quintiles. In contrast, the shares of rental and nonfarm income tend to be higher in higher income quintiles. Second, the share of livestock income is lowest in the monthly income of the households reported in the top income quintile, whereas the share of rental income is lowest in the bottom quintile. Overall, we may say that the distribution is actually very similar for the bottom 4 quintiles and the differences really come with the richest quintile.

Sample Distribution (%)	**Income Groups (in PKR)**	S_{nf}	S_t	S_a	S_r	S_l
Bottom quintile	1596-2526	16.4	31.9	23.9	9.0	18.8
2nd quintile	2550-3116	16.4	30.6	23.2	10.2	19.6
3rd quintile	3136-3359	13.2	33.1	23.6	10.3	19.8
4rth quintile	3365-4298	16.6	30.8	22.3	13.5	16.8
Top quintile	4371-6938	22.8	27.1	21.7	21.2	7.2

Table 2: *Distribution of sample households by income quintiles.*
Source: Survey 2008.
Note: All income shares are expressed in percentages.

3.3 Income shares by district

Table 3 breaks down total monthly household income y into the different shares S_i^j by the j districts surveyed. Compared with Table 1 which provides a more aggregate perspective, we see that the same ranking of shares applies on the district level as well.

According to Table 3, the share of transfer income is the highest in all the districts. Agricultural income and nonfarm income are other significant sources of income for Pakistan's rural population. Rental and livestock income contribute the least.

District	S_t	S_a	S_{nf}	S_r	S_l
Attock	26.94	22.00	26.11	18.85	6.10
Sahiwal	27.99	22.08	23.31	11.56	15.06
Layyah	28.50	22.65	20.33	10.43	18.09
Rahimyarkhan	29.56	23.30	18.33	10.49	18.32
Thatta	33.69	25.22	14.23	07.46	19.40
Badin	32.76	22.89	16.82	10.48	17.05
MirpurKhas	27.89	21.14	17.02	25.94	08.01
Malakand	29.38	21.84	12.66	16.10	20.02
Dir	34.01	24.05	12.06	08.84	21.04
Kalat	33.77	24.08	11.99	09.04	21.12

Table 3: *Share of different income sources; by district.*
Source: Survey 2008.
Note: All income shares are expressed in percentages.

4 Inequality decomposition by income sources

4.1 Decomposing overall income inequality

While calculating the shares of the different income types provides insights in the allocation of income, it does not make any statement about distributional aspects. The most common measure of income inequality is arguably the Gini coefficient G, ranging from 0 to 1, where a higher number indicates more inequality. In the extreme case of $G = 1$, the entire income goes to a single economic unit. The Gini coefficient allows for quick and easy comparison across countries, or in our case, provinces and districts. According to the World Bank (2005), Pakistan's overall Gini coefficient amounts to 0.33, with more recent data suggesting a downward trend.[2] In comparison to Pakistan's geographic neighbours Bangladesh and India, we find that the entire subcontinent is characterised by a similar degree of income inequality (World Bank, 2005). Yet, numbers in developed economies such as the UK ($G = 0.36$) and the US ($G = 0.41$) are higher and, in contrast to Pakistan, seem to stagnate.[3]

[2] In its latest Human Development Report, the United Nations Development Programme (UNDP) provides a Gini coefficient of $G = 0.31$ for Pakistan (UNDP, 2007).

[3] The more recent study by UNDP (2007) reports identical coefficients for relevant benchmark countries such as the UK and the US. Figures for Bangladesh ($G = 0.33$) and India ($G = 0.37$), in comparison to Pakistan, seem to be on the rise which may be attributed to higher and more persistent economic growth lately.

Based on our sample data, we obtain the following Gini coefficient for rural Pakistan (rPk), denoted by G_{rPk}:

$$G_{rPk} = \sum_{i=1}^{5} w_i R_i G_i = 0.189. \tag{1}$$

Comparing G_{rPk} with Pakistan's overall Gini coefficient, G_{Pk}= 0.306 (CIA, 2008); we may say that rural inequality is lower than overall inequality, which is expected as we don't have the inequality coming from rural-urban differences. But cities (or the urban areas) themselves may have low inequality. Ahmad (2000) finds that income inequality in rural areas tends to be lower compared with urban areas. Furthermore, it appears that there is a positive relationship between the level of skilled labour and income inequality (Ahmad, 2002). The more equitable distribution of income on the countryside may hence be explained by the lower level of skilled labour there.

We need to decompose the Gini coefficient G to measure how much each particular income source contributes to overall income inequality in rural Pakistan (and also at the district and provincial level, accordingly). For that matter we break-down each household's total monthly income in the same n sources as before:

$$y = \sum_{i=1}^{n=5} y_i,$$

where y is the total monthly household income and y_i represents the income type i.

Decomposing the Gini coefficient rests on three elements: First, the weight of income source i, w_i, defined as

$$w_i = \frac{\mu_i}{\mu},$$

where μ_i represents the mean monthly income of source i and μ is the mean of total monthly household income. Second, the correlation ratio between income source and total income R_i defined by the following ratio of covariances

$$R_i = \frac{cov(y_i, r)}{cov(y_i, r_i)},$$

where r is the ranking of total income such that higher income receives a higher rank and r_i expresses the corresponding ranking of income sources. Both r and r_i follow a

dense ranking from 1, 2, and so on.[4] The third part is the Gini coefficient associated with y_i given by

$$G_i = \frac{2}{\mu_i} cov(y_i, r_i).$$

The relative concentration coefficient is given by

$$g_i = R_i \frac{G_i}{G_{rPk}}.$$

This decomposition procedure further allows us to determine the individual contribution of each income source to overall income inequality. The contribution of income source i to overall income inequality in rural areas, c_i, is easily calculated by

$$c_i = w_i R_i \frac{G_i}{G_{rPk}} = w_i g_i, \tag{2}$$

where obviously all individual shares sum up to 1.

y_i	w_i	R_i	G_i	c_i
y_a	0.23	0.99	0.17	0.21
y_l	0.15	0.13	0.16	0.02
y_{nf}	0.18	0.88	0.31	0.26
y_r	0.14	0.97	0.37	0.27
y_t	0.30	0.95	0.17	0.26

Table 4: *Contribution to overall income inequality by income source. Source: Survey 2008.*

Table 4 shows that agricultural, nonfarm, rental and transfer income each account for about one quarter of total income inequality. Livestock income accounts for hardly any inequality at all. Looking at the correlation ratios (R_i) for agricultural, rental and transfer income, we can conclude that all these sources have a strong positive relationship with the total income rank, followed by nonfarm income. In contrast, the correlation ratio between livestock income and total income is small and positive. This is the result of a low covariance between livestock income and corresponding total income rank.

[4] In dense rankings, items that compare equally receive the same ranking number, and the next item(s) receive the immediately following ranking number. Equivalently, each item's ranking number is 1 plus the number of items ranked above it that are distinct with respect to the ranking order.

According to the individual Gini coefficients G_i, rental and nonfarm incomes are the most unequally distributed income sources. Agricultural and transfer income are relatively equally distributed. Inequality in livestock income is relatively small. A similar result is obtained by Ahmad (2002).

4.2 Decomposition of overall income inequality by province

Disaggregating the data further, we calculate in the following the Gini coefficients for the $k = 4$ provinces using the procedure outlined above. We see that the level of income inequality in Sind represents closest the overall situation in rural Pakistan. Yet, for the other provinces there is a considerable degree of dispersion. While the plain Gini index suggest hardly any income inequality in the relatively poor provinces of NWFP and Baluchistan, Punjab's coefficient is quite high considering that we only look at the rural sector.

Given our results in Table 5, we observe that in Punjab and Sind nonfarm, agricultural, transfer and rental each contribute for about one fourth of overall income inequality. Livestock income hardly contributes to the overall income inequality in both provinces. This suggests an important policy conclusion: policy makers who are concerned with income inequality in Punjab and Sind provinces would be well-advised to pay more attention to livestock as it has the smallest share of total income and hardly contributes to overall income inequality in the two provinces. In NWFP and Baluchistan, however, the situation is quite the opposite. For instance, in NWFP the rental income has the largest share (that is around 50%) in overall income inequality; followed by livestock, agricultural and nonfarm income. Transfer income, on the other hand, represents an inequality decreasing source of income. The inequality in rental, livestock and agricultural income may be attributed to unequal landownership in NWFP. In Baluchistan, rental income contributes the least to the overall income inequality, while livestock and transfer income make up the largest share (that is around 40% each) and nonfarm income accounts for 20% of the overall income inequality in the province. Surprisingly, agricultural income appears as inequality decreasing source of income. This may suggest that bringing more land under cultivation may improve the distribution of income in Baluchistan.

	Punjab	**Sind**	**NWFP**	**Baluchistan**
Gini index	0.24	0.18	0.04	0.02
Income source share (w_i)				
Nonfarm	0.232	0.164	0.123	0.120
Agricultural	0.223	0.226	0.229	0.241
Transfer	0.279	0.308	0.316	0.338
Livestock	0.123	0.136	0.205	0.211
Rental	0.143	0.166	0.127	0.090
Concentration index (g_i)				
Nonfarm	0.304	0.199	0.043	0.029
Agricultural	0.230	0.136	0.025	-0.005
Transfer	0.217	0.145	-0.002	0.022
Livestock	0.016	0.011	0.046	0.035
Rental	0.368	0.394	0.182	0.011
Relative contribution (c_i)				
Nonfarm	0.298	0.187	0.124	0.190
Agricultural	0.217	0.175	0.134	-0.070
Transfer	0.255	0.255	-0.017	0.409
Livestock	0.008	0.009	0.221	0.414
Rental	0.222	0.374	0.538	0.056

Table 5: *Inter-provincial rural household income disparities. Source: Survey 2008.*

4.3 Decomposition of overall income inequality by district

We cannot dismiss a priori the possibility of considerable disparities at district/tehsil levels compared to a pure province perspective. Although evidence for spatial differences at a disaggregated level is rather scanty, this may be an issue for the data at hand. This impression is endorsed by Pasha and Hasan (1982) who observe that statements about inter-provincial levels of development tend to hide major intra-provincial disparities.[5] We therefore calculate Gini coefficients for each district to clarify the contribution of each income source to overall income inequality within districts. This also allows us to draw important conclusions about the provinces. Indeed, according to our results in Table 5, income inequality in Punjab is much higher as compared to NWFP, but still both provinces experience similar levels of inequal-

[5]The mean monthly income fluctuates quite strongly across the 10 districts (standard deviation of 1238.14) which could conceivably affect any decomposition effort that is based on monthly income data.

ity in some of their districts such as Attock and Layyah in Punjab and Malakand in NWFP. Although the aforementioned districts share the same overall inequality, the income sources contribute differently to overall inequality within each district (see Table 6).

Table 6 shows that Sahiwal and Thatta have relatively higher inequality in terms of overall monthly household income. In Sahiwal, nonfarm income is the most unequally distributed source of income and contributes the most to overall income inequality in rural Pakistan. This may be due to a rather low literacy rate (only 40%) of the population which is the main reason for the highest income inequality in the nonfarm sector. Rental income contributes the least to overall income inequality. Similarly, the table indicates that in Thatta agriculture is the most unequal source of income and contributes with 38.4% to overall income inequality, which is highest among the other income sources. According to our sample survey, 70% of the population in Thatta are small farmers (having land between one or two acres) and their livelihood continues to revolve around farm activities like agriculture and livestock. Some of these farmers are making best use of their land, while others are still using old and primitive techniques for production (as revealed by personal communication with the respondents), which may be one of the main reasons for higher income inequality in the farm sector.

The field study also provides a socio-economic profile of each district. According to the survey, the average monthly household income of Rahimyar Khan district is the lowest among all the districts analysed. Livestock income is the most unequal source of income in Rahimyar Khan as well as in Kalat district and contributes most strongly to an increasing overall income inequality in both districts (Table 6).

	Attock	Sahiwal	Layyah	Rahimyar-Khan	Badin	Thatta	Mirpur-Khas	Dir (lower)	Mala-Kand	Kalat
Overall Gini coefficient of monthly household income (G)	0.049	0.081	0.048	0.053	0.016	0.075	0.033	0.024	0.043	0.020
Source income weight(w_i)										
Nonfarm	0.261	0.233	0.203	0.183	0.142	0.168	0.170	0.127	0.121	0.120
Agricultural	0.220	0.221	0.227	0.233	0.252	0.229	0.211	0.218	0.241	0.241
Transfer	0.269	0.279	0.285	0.296	0.337	0.328	0.279	0.294	0.340	0.338
Livestock	0.061	0.151	0.181	0.183	0.194	0.170	0.080	0.200	0.210	0.211
Rental	0.188	0.116	0.104	0.105	0.080	0.105	0.260	0.161	0.088	0.090
Relative concentration coefficients of income sources(g_i)										
Nonfarm	0.605	1.916	0.415	1.131	2.129	0.851	0.055	-1.623	1.246	1.595
Agricultural	1.262	1.222	1.246	0.773	-1.738	1.524	-0.051	1.586	1.040	-0.294
Transfer	0.987	0.704	0.979	0.727	3.807	0.655	1.982	1.604	-0.328	1.214
Livestock	0.349	0.423	0.660	1.546	0.530	1.266	1.999	1.371	1.553	1.964
Rental	1.469	0.179	2.242	1.113	-3.034	0.340	1.091	-0.254	2.517	0.633
Source income contribution to overall income inequality(c_i)										
Nonfarm	0.158	0.447	0.084	0.207	0.358	0.121	0.009	-0.196	0.158	0.191
Agricultural	0.278	0.270	0.283	0.180	-0.398	0.384	-0.011	0.381	0.227	-0.071
Transfer	0.265	0.197	0.280	0.215	1.249	0.221	0.553	0.545	-0.096	0.410
Livestock	0.021	0.064	0.120	0.283	0.090	0.246	0.160	0.288	0.310	0.414
Rental	0.278	0.022	0.233	0.117	-0.319	0.026	0.284	-0.023	0.405	0.057

Table 6: *Decomposition of overall income inequality by district using the Gini coefficient.*
The more unequally distributed the income source, the higher the coefficient of concentration. The concentration coefficient shows how much a given income source "pushes" up the overall income inequality. The relative concentration coefficient is given by $g_i = R_i \frac{G_i}{G_{rPk}}$. *The "negative" sign indicates that the source income decreases as the total income increases and thus "pushes" down the overall income inequality . The contribution of each income source (c_i) to overall income inequality is determined by* $w_i g_i$, *where* $w_i = \frac{\mu_i}{\mu}$ *and* $g_i = R_i \frac{G_i}{G_{rPk}}$.
Source: Survey 2008.

Similarly, Table 6 shows that in Attock and Layyah agricultural income has maximum contribution to overall income inequality due to the uneven distribution of land in favour of the rich, while livestock income contributes the least to income inequality in the two districts. Table 6 indicates that transfer income is the most unequally distributed income source in Badin, Mirpurkhas and Dir and is mainly responsible for pushing up overall income inequality. Furthermore inequality in transfer income is the result of internal and external remittances (according to personal communication with the respondents). Evidently, the uneven land distribution in rural Pakistan forces the poor to seek the bulk of their livelihood by migrating within Pakistan or even by emigrating away from the country.

The study reveals that 60% of the households in Malakand are landless which may explain why rental income is the most unequally distributed income source. Rental income contributes with 40.5% to overall income inequality in the district which is the highest contribution as compared to the other income sources (see Table 6).

5 Inequality decomposition by regions

5.1 Decomposition of income inequality within and between provinces

Although the Gini index provides a transparent and easy way of comparing income inequality across regions, it is flawed in that it is not perfectly decomposable in some situations (Shorrocks, 1982). One should therefore be careful in interpreting this index in empirical studies. Other inequality measures seem more suitable for inequality decompositions. Shorrocks (1980) derives an entire class of measures which are additively decomposable under relatively weak restrictions on the form of the index. The subclass of mean independent measures turns out to be a single parameter family which involves the square of the coefficient of variation and two entropy formulae proposed by Theil.

In the context of additive decomposability, the generalised entropy (GE) class of inequality indices is a good alternative to the Gini index. Unlike the Gini coefficient, the members of this class are perfectly decomposable without a residual term. The

most common index within the GE class is the Theil Index (I) as introduced in Equation 3:

$$I = \frac{1}{N} \sum^{N} \frac{y}{\bar{y}} ln \left(\frac{y}{\bar{y}} \right) = 0.055, \tag{3}$$

where, as above, N is the sample size, y is the individual household's monthly income (summable over N) and $\bar{y}$ the corresponding mean. Given the design of our survey, the sample is readily partitioned to differentiate between inequality within (I_w) and between (I_b) regions, where it obviously holds that $I = I_b + I_w$.

Using this methodology, we can first decompose the Theil Index on a more aggregate province level as follows:

$$\begin{aligned} I &= \underbrace{\sum_{k=1}^{K=4} \frac{N_k}{N} \left(\frac{\bar{y}_k}{\bar{y}} \right) ln \left(\frac{\bar{y}_k}{\bar{y}} \right)}_{Between} + \underbrace{\sum_{k=1}^{K=4} \left(\frac{N_k}{N} \frac{\bar{y}_k}{\bar{y}} \right) I_k}_{Within} \\ &= 0.0016 + 0.0534 = 0.055. \end{aligned} \tag{4}$$

The results of Equation 4 suggest that inequality on this level of aggregation is almost exclusively driven *within* rather than between the provinces.

5.2 Decomposition of income inequality within and between districts

We can decompose the Theil Index for the district level in a similar fashion. Decomposing income inequality within and between districts helps to further disentangle total inequality in the given sample.

On the more disaggregated level we have $J = 10$ districts. The Theil Index decomposition hence takes the following form:

$$\begin{aligned} I &= \underbrace{\sum_{j=1}^{J=10} \frac{N_j}{N} \left(\frac{\bar{y}_j}{\bar{y}} \right) ln \left(\frac{\bar{y}_j}{\bar{y}} \right)}_{Between} + \underbrace{\sum_{j=1}^{J=10} \left(\frac{N_j}{N} \frac{\bar{y}_j}{\bar{y}} \right) I_j}_{Within} \\ &= 0.0535 + 0.0013 = 0.055, \end{aligned} \tag{5}$$

where N_j is the number of households in district j, $\bar{y}_j$ is mean monthly household income in district j and I_j is the resulting Theil Index in district j.

	I_b	I_w	I
Overall	–	–	0.055
Provinces	0.002	0.053	0.055
Districts	0.054	0.001	0.055

Table 7: *Regional comparison of Theil indices*

Results are summarised in Table 7. We find that between-district inequality is greater than within-district inequality. This may seem at odds with the results for the province level, where exactly the opposite effect occurs. Yet, as Table 7 illustrates, if the entire degree of inequality is determined within the provinces on a more macro level, this necessarily implies that the "whole action" occurs between the districts which are *nested within* the provinces. Therefore, while overall measures are obviously the same (up to some rounding margin) regardless of the level of aggregation, it appears much more meaningful to look at the district level rather than comparing between provinces as commonly done. We may also suggest that on the whole, all the rural Pakistan gives the same picture irrespective of which province is chosen as its sample representation.

To substantiate this point further, we can easily determine how much total inequality is explained by between-district inequality by looking at the following ratio R_b:

$$\begin{aligned} R_b &= \frac{I_b}{I} \\ &= \frac{0.0535}{0.0550} = 0.97. \end{aligned} \tag{6}$$

According to Equation 6, 97% of total inequality is explained by inequality between the districts. The remaining 3% is explained by inequality within the districts. We therefore focus in the remainder on inequality arising at the district level.

6 Partial correlation analysis

In this Section, we consider partial correlation coefficients to measure the statistical relationship between income inequality (G_j) and the shares of different income sources ($S_{i,j}$) across the districts, where all variables are expressed in terms of *percent-*

ages and G_j represents the Gini coefficient as measure of overall income inequality in the j^{th} district and the share $S_{i,j}$ is defined as the ratio of average income of type i in district j over total average income in district j:

$$\begin{aligned} S_{a,j} &= \frac{\bar{y}_{a,j}}{\bar{y}_j} \times 100 \\ S_{t,j} &= \frac{\bar{y}_{t,j}}{\bar{y}_j} \times 100 \\ S_{nf,j} &= \frac{\bar{y}_{nf,j}}{\bar{y}_j} \times 100 \\ S_{r,j} &= \frac{\bar{y}_{r,j}}{\bar{y}_j} \times 100 \\ S_{l,j} &= \frac{\bar{y}_{l,j}}{\bar{y}_j} \times 100 \end{aligned}$$

Similar to a regression analysis, partial correlation seeks to measure a relationship between dependent and independent variable, whilst eliminating potential effects of a third variable. The partial correlation coefficients measure in this case the degree of statistical association between income source and district-wide Gini coefficient, where the latter one is considered as the dependent variable.

Partial Correlation of G_j with		
Variable	Corr.	Sig.
$S_{a,j}$	0.6494	0.114
$S_{t,j}$	-0.7354*	0.060
$S_{nf,j}$	0.7354*	0.060
$S_{r,j}$	0.6291	0.130
$S_{l,j}$	0.6517	0.113

Table 8: *Correlation Analysis*

Results are shown in Table 8. Columns two and three report the partial correlation coefficient and the corresponding significance level, respectively. We find a strong positive and statistically significant correlation of 0.74 between the share of nonfarm income and income inequality across the districts. Thus, nonfarm income appears as inequality increasing source of income. In contrast to this, we detect a similarly strong (and also statistically significant) but negative correlation (-0.74)

between the share of transfer income and income inequality. Unlike nonfarm income, transfer income is a source of income that is capable of reducing income inequalities in rural Pakistan.

7 Conclusion and Policy implications

This paper analysed the impact of various sources of income on income inequality in Pakistan. Unlike other studies, we exclusively focus on the rural sector. Our unique survey dataset allows for detailed comparisons and decompositions at different levels of aggregation. The descriptive analysis suggests that it is much more meaningful to analyse income inequalities in rural Pakistan at the district rather than province level. We find that transfer income is the most important source of total monthly household income, accounting for almost 30%. Shares of rental and livestock income are the lowest in the sample. Shares of transfer, agricultural and livestock income are relatively higher in lower income quintiles as compared to higher income quintiles. Shares of rental and nonfarm income tend to be higher in higher income quintiles.

We consider both Gini coefficients and Theil indices to measure income inequality and the corresponding contributions of the various sources of income. Agricultural, nonfarm, rental and transfer income each contribute to income inequality in the sample area to a similar degree, whereas livestock income hardly affects the results. Decomposing overall income inequality within and between districts implies that inequality between the districts is greater than within the districts. The Theil index decomposition suggests that almost the entire degree of inequality (97%) can be explained by inequality between the districts.

The correlation analysis confirms the conjectured importance of transfer income in reducing income inequalities. While the derived coefficient for transfer income is negative, nonfarm income is found to be an inequality-increasing source of income. Considering both effects together appears to neutralise each other's role in driving income inequalities.

With regards to policy implications, given that agricultural income and rental income each contribute one fourth to overall income inequality in rural Pakistan, the policies should aim first to correct both of these sources of inequality. It is well known

that addressing these inequality sources requires addressing asset inequality. Land is the main source of agricultural and rental income, and Pakistan is characterised by high land ownership inequality as identified by Anwar et al. (2004) and one of the main causes of poverty and inequality in rural Pakistan. Therefore, the policy makers should seriously consider land reforms in rural Pakistan in order to enhance pro-poor growth.

Secondly, as an inequality-decreasing source of income, policy makers should also put more emphasis on transfer income in fighting inequalities across the districts. They should try to foster its flows both within Pakistan and from abroad but particularly try to increase factor mobility within the country's urban sector.

In order to reduce spatial imbalances, policy makers should try to provide stronger technical education for the rural (and often poor) unskilled labour force so as to increase mobility and to improve job opportunities within Pakistan and abroad. To raise the overall income level, policy makers should take steps to help poorer households send migrants abroad. One supporting measure that may be considered in that regard is the establishment of "rural migration centers" to process visas, work contracts and loan arrangements for prospective external migrants.

References

Adams, R. H. (1994). Non-farm income and inequality in rural Pakistan: a decomposition analysis, *Journal of Development Studies* **31**: 110–133.

Adams, R. H. and Alderman, H. (1992). Sources of income inequality in rural Pakistan: a decomposition analysis, *Oxford Bulletin of Economics and Statistics* **54**: 591–608.

Ahmad, M. (2000). Estimation of distribution of income in Pakistan, using micro data, *The Pakistan Development Review* **39**(4): 807–824.

Ahmad, M. (2002). Income inequality among various occupations/professions in Pakistan – estimates based on household income per capita, *The Lahore Journal of Economics* **7**(1).

Anwar, T. (2005). Long-term changes in income distribution in Pakistan: Evidence based on consistent series of estimates, *Discussion Paper Series No.3, Centre for Research on Poverty and Income Distribution (CRPRID)* .

Anwar, T., Qureshi, S. K. and Ali, H. (2004). Landlessness and rural poverty in Pakistan, *The Pakistan Development Review* **43**(4): 855–874.

CIA (2008). The World Factbook.

Glewwe, P. (1986). The distribution of income in Sri Lanka in 1969-70 and 1980-81: a decomposition analysis, *Journal of Development Economics* **24**: 255–274.

Government of Pakistan, Statistics Division (1998). Population Census.

Nugent, J. and Walther, R. (1982). Short-run changes in rural income inequality: A decomposition analysis, *Journal of Development Studies* **18**: 239–269.

Pasha, N. and Hasan, T. (1982). Development ranking of districts in Pakistan, *Pakistan Journal of Applied Economics* **1**(2).

Shams, K. (2012). *Income inequalities and well-being in rural Pakistan*, PhD thesis, University of Glasgow.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures, *Econometrica* **48**: 613–625.

Shorrocks, A. F. (1982). Inequality decomposition by factor components, *Econometrica* **50**: 193–211.

UNDP (2007). Human Development Report.

World Bank (2005). World development indicators.