UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2016/151

Group inequality and regional development

Evidence from Pakistan

Hadia Majid and Rashid Memon*

December 2016

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: This study explores the patterns and consequences of ethnic and regional inequalities in Pakistan. Using inter-district variation in inequality, ethnic fractionalization, and access to public goods, we extend the literature on public service provisioning by using finer estimates of between-group inequality along multiple dimensions. We also provide the first estimates of horizontal and vertical inequality at the district level in Pakistan. We find that higher group inequality in land, income, and electoral votes is negatively associated with access to public goods. But higher vertical inequality along the same dimensions is positively associated with public goods provisioning.

Keywords: horizontal inequality, publicly provided goods
JEL classification: D63, I24, H42

Acknowledgements: This paper was supported by UNU-WIDER. We are grateful to Maheen Javaid for exceptional research assistance. We are also thankful to Rachel Gisselquist, Carla Canelas, Nishant Chaddha, and to participants at the UNU-WIDER workshop for their comments and feedback.

* Lahore University of Management Sciences, Lahore, Pakistan; corresponding author: hadia.majid@lums.edu.pk.

This study has been prepared within the UNU-WIDER project on 'The politics of group-based inequalities—measurement, implications, and possibilities for change', which is part of a larger research project on 'Disadvantaged groups and social mobility'.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-195-6

Typescript prepared by Lesley Ellen.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Regional and ethnic inequalities in income, health, and education pose significant challenges for developing country governments. These horizontal inequalities are important because they are often accompanied by political tensions, and in extreme cases, civil wars (Kanbur and Venables 2003; Stewart 2008; Deiwiks et al. 2012; Lessman 2015). Recent literature has also pointed to the importance of ethnic inequality for public goods provisioning (Baldwin and Huber 2010; Alesina et al. 2012), arguing that substantial differences between groups make governance and hence the provisioning of public goods difficult. This study replicates this analysis in Pakistan where a number of inequality measures are much more precisely estimated from large national and district level representative household data sets, compared to the much coarser cross-country data used by Alesina et al. (2012).

Pakistan is also an interesting case to study given its politics. A federation of four provinces, each dominated by a different ethnic group, Pakistan is a religiously homogenous but ethnically diverse country (see Appendix A: Maps 1 and 2). The division of British India into Pakistan and India witnessed massive migration streams that rendered Hindus, Christians, and Sikhs a tiny minority in Pakistan.[1] The resultant religious homogenization, coupled with overlapping provincial and linguistic boundaries, provided the space for linguistic ethnicity to emerge as one of the dominant cleavages in society. Despite immense local diversity, regional and resource politics have, over time, led to the consolidation of six[2] major ethnic groups around corresponding major languages–the Baloch, Muhajirs, Pashtun, Punjabi, Seraiki, and Sindhis. In passing, this study also contributes to the literature on ethnic studies in Pakistan by providing the first estimates of ethnic group inequality in Pakistan.

2 Existing work on inequality in Pakistan

Most of the literature on inequality measurement in Pakistan focuses on trends in income and consumption inequality (see, for example, Anwar et al. 2004; Jamal 2006; Shahbaz et al. 2007; Arif and Farooq 2011). Burki et al. (2015), who focus on asset inequality, and Chaudhry and Rahman (2009), who focus on gender inequality in education, are notable exceptions. Apart from some exceptions such as Jamal (2006), we are unware of studies linking economic phenomena, such as public good infrastructure, to changes in inequality in Pakistan.

All studies referred to above (and this one) suffer from several data-related problems. Firstly, the sample sizes are relatively small, increasing the confidence interval of the measures. Secondly households from the upper tail of the income distribution are severely under-represented (Bergan 1967) suggesting that any measure reported here is probably an underestimate of true inequality.

Having said that, the studies focusing on income inequality suggest that income distribution worsened slightly in the 1970s, 1980s, and 1990s as the Gini increased from 0.34 in 1970, to 0.36 in 1987, and to 0.41 in 1998 (Anwar et al. 2004). This worsening has at least partly been attributed to the IMF-supported Structural Adjustment Program of the late 1980s and 1990s (Jamal 2006). However, Burki et al. (2015) argue that it is not possible to compare the Gini from

[1] According to the Pakistan Census of Population 1998, 96 per cent of Pakistan's population is Muslim (GOP 1998: 355).

[2] There were seven large ethnic groups prior to Bangladesh's separation in 1971, with Bengalis being the most populous.

the 1980s to Ginis in the 1990s and subsequent years because of changes in the way income data were collected in subsequent surveys. It is, nevertheless, possible to examine the trends in consumption-based inequality over the 1990–2013 period. These are presented in the Tables 1 and 2. Furthermore, given this concern, this is also the period of concern in this paper.

Table 1: Changes in Gini coefficient inequality in Pakistan

Gini index × 100	1990–91	1992–93	1993–94	1996–97	1998–99	2001–02	2005–06	2007–08	2010–11	2011–12
Urban	32.39	35.97	34.00	33.74	39.18	35.16	33.30	34.77	33.44	34.66
Rural	26.71	28.73	29.34	35.12	26.23	24.79	25.41	27.00	24.84	24.73
Overall	29.79	32.11	32.49	33.89	34.27	30.41	30.55	31.65	29.82	30.96
Middle 50% (deciles 5 to 9)	49.2	47.3	46.9	45.7	47.7	48.9	49.7	48.1	50.1	48.5
Wolfson index	0.112	0.114	0.115	0.107	0.126	0.116	0.120	0.123	0.114	0.118

Source: Authors' adaptation from Burki et al. (2015).

The key insights from Tables 1 and 2 are that: (1) inequality is relatively stable and quite low given daily experiences of inequality—the latter could, possibly, be due to the under-representation of rich households in survey data, but it is the stability of inequality that interests us most and we shall return to it in the next section; (2) urban inequality is higher than rural inequality in each year—this is to be expected, given the diversity of economic activities in urban centres and the diversity in rates of return to these activities; and (3) as can be seen from Table 2, Sindh has the highest levels of inequality.

Table 2: Gini inequality and Wolfson polarization index, by province (Gini/polarization index ×100)

	Punjab		**Sindh**		**KP**		**Balochistan**	
Year	Gini	Wolfson	Gini	Wolfson	Gini	Wolfson	Gini	Wolfson
1990–91	29.70	11.40	31.85	12.34	23.76	8.42	24.86	10.29
1992–93	32.61	11.82	33.60	12.15	27.22	8.09	24.83	9.67
1993–94	33.38	11.78	33.57	12.47	24.83	8.81	27.81	10.73
1996–97	34.78	10.71	33.20	11.91	28.59	8.86	29.01	9.29
1998–99	34.78	12.88	36.61	13.78	28.45	10.44	23.33	8.93
2001–02	30.04	12.14	35.18	12.60	23.33	8.77	22.09	8.86
2005–06	30.39	11.95	33.09	12.91	25.86	10.20	23.54	9.73
2007–08	31.68	12.76	34.31	12.59	26.21	9.67	24.32	9.26
2010–11	30.44	11.98	32.17	11.72	24.80	9.67	19.91	8.11
2011–12	30.96	12.19	34.00	12.23	25.87	9.80	21.90	8.81

Source: Burki et al. (2015).

3 Inequalities: theoretical framework

The discussion above suggests several undercurrents in the research on inequality in Pakistan: firstly, inequality measures are generally limited to income and consumption; secondly, measures are limited to vertical and not horizontal inequality; and thirdly, the stagnation of the measures across time are suggestive of inequality traps (for a detailed discussion of the latter see Burki et al. 2015: chapter 3). In what follows, we suggest that Sen's (1980) capability framework provides a coherent way of not only enlarging the scope/domain of inequality measures, but also suggests mechanisms through which inequality is perpetuated through time, and across generations and groups of individuals.

Sen's answer to the question *inequality of what,* identified the space of capabilities on which to define the metric for measuring progress (Sen 1980). Subsequently, Desai (1995), Alkire and Black (1997), Robeyns (2003), and Nussbaum (2005) have all generated lists of human capabilities in an effort to apply Sen's (1980) framework. Nussbaum (2005), for example, lists: (1) life; (2) bodily health; (3) bodily integrity; (4) senses, imagination, and thought; (5) emotions; (6) practical reason; (7) affiliation; (8) play; and (9) political and material control over one's environment (Nussbaum 2005).

While defining the domain for horizontal inequality, the approach also provides a framework for understanding the persistence in inequality measures in terms of time, across generations, and ethnicities presented above. In passing, note that this persistence is not unique to Pakistan and is prevalent enough for Charles Tilly (1999) to have come up with a name for it—'durable inequality'.[3] Rao (2006) subsequently formalized these ideas in the framework of what he called 'inequality traps', defining them as 'situations where the entire distribution is stable because the various dimensions of inequality (in wealth, power, and social status) interact to protect the rich from downward mobility and the poor from being upwardly mobile' (Rao 2006).

Stewart (2008) is an approachable exposition of how each capability tends to promote other capabilities and how each capability is related to a person's productivity and income. If a low capability such as health leads to low income, which in turn leads to low education and which in turn leads to low health, one obtains a veritable trap. Similarly, access to different stocks of capital may involve similar traps. Different levels of social capital allow access to prestigious schools which not only endow a person with higher levels of human capital but also cultural and social capital which allow the possibility of higher levels of physical (or alternatively economic) capital.

While the discussion above is in the context of vertical inequality, it can be reproduced for horizontal inequalities with some adjustments. There is an enormous literature on wage differentials by race and gender i.e. on the 'capability' of different groups of people to convert human capital into earnings. One possible route for this functioning is the interaction between human capital and social capital: network effects may provide one group with better possibilities than other groups (Munshi 2014). Different groups may also possess different cultural practices and these might reduce the rates of returns to human capital if these practices are different from those of the dominant group (Sowell 1983).

[3] Discussing different mechanisms that allow inequalities to persist, Tilly focuses on the relations of exploitation between dominant and subordinate groups and the maintenance of institutions and practices that support these relationships.

4 Sources of data

We use two datasets—the Population Census, and the Pakistan Social and Living Standards Measurement Survey (PSLM)/Pakistan Integrated Household Survey (PIHS) for our analysis.

4.1 Population census

The census, which was started by the British in 1871, was envisioned as a decennial activity. Post-independence, censuses were carried out in 1951, 1961, 1972, 1981, and 1998. The only census for which *all* district reports are readily available is the 1998 census. The 1998 census reports 52 tables—40 on population data and 12 on housing data for 104 districts in four provinces of the country. Of the 52 tables, 27 are based on complete count information and 27 are based on sample count information (8 per cent single stage stratification).

Several observations merit mention at the outset. Firstly, the information on language, religion, literacy, level of education, and access to water and energy, is based on complete count. Secondly, the census does not make public individual level data so it is impossible to calculate literacy levels by religious or linguistic group.

4.2 PIHS 1996–97/PSLM 1996–97, 2007–08, 2010–11

The PIHS 1996–97 and some rounds of the PSLM are the only data set that contains information on language. The PIHS questionnaire records information on the 'language in which the interview was conducted', and is the only data set that contains information on religion as well language. While other PIHS data rounds conducted in 1995–96, 1998–99, and 2000–01 collected data on language, they do not report it. Efforts to get the data from the Pakistan Bureau of Statistics have been unfruitful.

With regard to the language variable in the PIHS questionnaire, the following observations merit attention: if the interview was conducted in a regional language one can be confident that the respondent identifies with that language ethnically. If, on the other hand, the interview is conducted in Urdu, the national language of Pakistan, the questionnaire does not provide any additional identifiers of ethnicity, and we cannot be certain if the respondent's ethnicity is indeed 'Urdu-speaking'. To resolve this problem one must understand the data collection protocol of the Federal Bureau of Statistics (FBS).

The FBS[4] interviewing team is composed of a man and a woman who interview male and female heads of households respectively. Several teams are sent to cover a primary sampling unit (PSU) and the regional languages spoken in that PSU are kept in mind. It is ensured that, at the very least, the female interviewer speaks the regional language. Since there could be several regional languages spoken in a multi-ethnic PSU, the FBS has several teams with the appropriate language mix. For example, a team interviewing in Sindh is constituted of people who are fluent in Sindhi, Seraiki, Punjabi, and Pashto. Additionally, all team members speak Urdu. The supervisor of each team identifies the households to be surveyed, enquires about the 'language spoken at home', and assigns an interviewer who speaks that language.

[4] The following is based on a telephone conversation with a bureaucrat closely linked to FBS.

5 Dimensions of inequality: trends and patterns

Based on our conceptual framework, we calculate different measures of horizontal inequality for several variables—details on the formulae for each of the HI measures are provided in Appendix B. In particular, we calculate:

- Group Gini (GGini), which compares every group with every other group. Similar to the Gini, which is a measure of overall inequality, the values for the GGini lie between 0 and 1, where the closer the value is to zero the lower the level of inequality.

- Group Coefficient of Variation (GCOV), which measures differences from the mean for each group giving more weight to extremes. Here, a higher value of the GCOV measure implies the greater the difference from the mean, provided that population shares of groups are the same, and therefore the greater the between-group inequality.

- Group Theil Index (GTheil), which again like the GCOV and in contrast to the GGini, compares each group with the mean. And while it is especially sensitive to the lower end of the distribution, unlike the GGini and GCOV, it can take both negative and positive values. Finally, similar to the GGini, the closer the value of the GTheil to zero, the lower the level of horizontal or between-group inequality.

Details on our variables of interest, and how these link to our conceptual framework are provided in Table 3. Essentially, we focus on two main types of capital in our analysis, namely human and physical capital. With regard to the former, we examine differences between groups in terms of educational achievement and nutrition. In terms of physical capital, we emphasize economic capability by considering differences in income level, as well as asset and land ownership.

Table 3: Linking HI measures to the conceptual framework

Type of capital	Corresponding capability	Indicator
Human capital	Education	Highest class completed
	Nutrition	Food consumption
Physical capital	Productivity and income	Income

Source: Authors' illustration.

And while the focus of this section is on examining how ethnic groups at both the national and regional levels vary in their human and physical capital achievements, we also consider the role of social and cultural capital in section 6.

5.1 Educational inequality

Educational inequalities have long been a matter of policy concern in developing countries. Literature inspired by Sen (1980) views education as crucial for a person's functioning and capacity to flourish, if not desirable in its own right. Memon (2013) suggests that human capital differences account for a significant share of the ethnic wage gap. Finally, as discussed in section 1, a large literature finds that inequalities in human capital are correlated with inequalities in income and health.[5]

[5] See Blau and Kahn (2005) and Bedard and Ferrall (2003).

There is much less agreement on how to measure inequality, though. While early attempts to measure inequality focused on years of education, this measure has been criticized because it ignores quality of schooling. A year of schooling in large cities in Pakistan like Karachi or Lahore for example, may not compare with a year of schooling in Baluchistan's Noshki district, one of the country's poorest areas. Andrabi et al (2014), for example, suggest that variations in school quality within a district lead to a difference in children's test scores (achievement is presumably a better measure of education than attainment).

Having noted this important criticism we proceed with measuring inequality of attainment for two reasons. Firstly, since there is no nationally representative data set on test scores, we cannot measure inequality of achievement. Secondly, it seems reasonable to assume that the subaltern group, on average, would not only have a lower level of attainment but also less quality. Similarly, the dominant group, on average, could be expected to have not only a higher level of attainment but also quality. To that extent, inequality of attainment would be a lower bound for the inequality of achievement.

Appendix Table C1 presents the results of the three HI measures implemented on years of education for the 15 to 25 years-of-age cohort. Two main features stand out. Firstly**, inequality is on a downward trend in the 13-year span** for all cleavages, with the decline in inequality being the smallest with regard to the rural–urban divide in the case of the GCOV and GTheil measures that are more sensitive to relative and extreme deprivation. One possible reason for this difference in inequality could be the increasing number of private schools in the country overall but which may be pricing out the poorer people in less developed, rural areas.

Additionally, Appendix Table C2 reports the HI measures for education for different age cohorts. The first noteworthy finding is that across all years and for all HI measures, inequality is increasing in age cohort across ethnicities, i.e. the level of inequality for the youngest age cohort is the lowest across ethnicities while that for the oldest is the highest. Additionally, we see a downward trend in educational inequality across ethnicities over the 13-year span examined here, for all of the age cohorts. Taken together, these findings suggest that over the years, educational facilities may have become more readily available for all ethnicities, manifesting in improved inequality particularly for the younger age groups. This is reflected in the mean achievement levels provided in Appendix Table C3.

5.2 Nutritional inequality

The role of nutrition in productivity and subsequently individual and household welfare has been explored in considerable detail. Strauss and Thomas (1998), for example, highlight the low wage, nutrition, and productivity trap, while Dasgupta and Ray (1986) use the link between consumption/nutrition and labour power to present an argument for efficiency wages. Furthermore, works, such as that by Miguel and Kremer (2004), have highlighted the link between proper nutritional intake and effective education.

When exploring nutritional inequality, most academic work has tended to use daily food diaries which can be converted into calorific intake or have made use of anthropometric measures to proxy for nutritional intake. And while such measures would have been ideal to understand the levels of nutritional inequality within Pakistan, there is no national-level data source available for multiple years that records such data. At the same time, Duflo and Banerjee (2007) garnered some keen insight into the consumption patterns, food, and otherwise, of households by looking at per capita expenditures. We then employ a similar approach and focus primarily on per capita expenditures on food items. However, it is important to note that expenditures data for 2010 are only available in the Household Integrated Expenditures Data (HIES) which is run on a sub-

sample of the PSLM. This sub-sample remains representative at the national, provincial, and urban–rural levels, but not at the district level.

We proceed by dividing food consumption into three categories based on the primary dietary groups for the country—vegetables, lentils/grains, and meat. However, given that the majority of Pakistanis tend to have a non-vegetarian diet it remains unclear to us whether any vegetable intake indicated end-use of the product, i.e. it remained possible that vegetables were simply being used to cook a dish whose primary ingredient were lentils or meat instead. Furthermore, lentil/grain consumption may be driven by regional trends with some types of grains being consumed only in certain regions of the country. In contrast, not only are different types of meat—poultry, fish, mutton, beef—consumed by all regional groups, but, given the high protein content of the food group as well as the relatively high price of the item, we expect inequality in meat consumption to be an excellent proxy for relative (nutrition-based) deprivation of the group. Thus, for per capita food consumption, we focus our analysis on expenditures on meat.

Inequality in overall per capita meat consumption has been declining in the country (Appendix Table D1) quite rapidly. Even when we disaggregate our HI measures by different types of meats, we observe that horizontal inequality has been declining (Appendix Tables D2–D4). Yet, there are also two other points that merit mention: (1) not only did per capita expenditure on poultry/fish see the lowest level of between-group inequality relative to mutton and beef consumption in 1996, but the inequality levels for this consumption product remained the lowest in 2010 as well; and (2) while in 1996 we observe all of the HI measures to be the highest for expenditures on beef relative to mutton and poultry/fish, it is in fact mutton expenditures that are the highest in 2010. In fact, the GTheil for expenditures on mutton is the only instance where we see an upward trend in HI. This is indicative of the increase in mutton prices relative to other types of meat making it especially difficult for more deprived groups to afford the commodity.

5.3 Income inequality

Discussions on inequality in physical capital generally tend to focus on different types of assets including land, as well as other consumer durables. However, we faced severe data limitations when it came to tracking the horizontal inequality in land and other durable assets over the three waves of the PSLM—not only are the data on durable assets only available for the 2007 and 2010 rounds, but detail on land is restricted to agricultural land in all three waves. Thus, in order to get a sense of the productivity capability at the between-group level, we calculated the HI measures for the three waves for income.

Inequality at the between-ethnicity level is fairly constant across the three waves, exhibiting a minor downward trend that becomes pronounced as we move from the 2007 to the 2010 round (Appendix Table E1). Yet, when we turn to the between-province differences we observe a sharp convergence: the GCOV measure for example, falls substantially from 0.24 in 1996 to 0.04 in 2007. In contrast, the urban–rural horizontal inequality is seen to increase sharply from 1996 to 2007. The fall in income inequality across provinces hints at the 'catching up' of the poorer provinces with the richer ones. At the same time, one possible reason for the rising inequality between urban and rural areas could be the self-selection of people into the rural to urban migration stream: over time higher-income people migrate to the cities that provide higher amenities levels in general, leaving the poorer behind.

6 Interactions between dimensions of inequality

As set out earlier, we are particularly interested in how dimensions of group inequality interact with each other. A large literature has generated compelling evidence that public goods provisioning, redistributive policies, and effective governance are all adversely affected by group inequality (Alesina et al. 2001). Theoretically, if, in a diverse society, different groups have different rankings of public goods then the desired good will be more distant to the chosen one (Luttmer 2011). If preferences are in turn shaped by a person's position on the income distribution, group inequality leads to divergent preferences, leading to lower public goods provision and increased political tensions (Deshpande 2000).

A related finding from the literature on urbanization is how the rich including the middle class often want to 'isolate' themselves. This desire might be particularly strong when wealth is correlated with group identity, in which case ethnic inequality may lead to segregation. (Alesina and Zhuravskaya (2011) show that ethnic and linguistic segregation correlate negatively with proxies of effective governance).

If the overlap between political and economic inequality also corresponds to group cleavages then ethnic minorities or less powerful groups could have a limited supply of public goods. In such areas, the rich might compensate for the diminished supply of public good by just directly purchasing the corresponding private group—purchasing water from private providers rather than relying on public supply.

Following Alesina et al. (2012) and Baldwin and Huber (2010), we too explore the link between group inequalities and public services. This, however, forces us to change strategy from the one employed in the preceding sections: group inequality measures at the national or provincial level will not do since this limits the number of observations and robs correlations of any statistical power. For example, correlating nutritional and educational inequality would be tantamount to one number each for three years per measure. To overcome this, we have taken the only data set of the PSLM (the 2010 version), which is representative at the district level. This allows us to calculate inequality measures at the district level. Since there is a large number of districts (105) in the country, this allows for a richer environment for exploring interactions.

In particular, we include between and within ethnic group inequality measures while also controlling for ethnic fractionalization at the district level. While the former two inequality measures capture economic difference between and within groups, the latter serves as an estimate of diversity within each district. By adopting this strategy we are then able to answer whether it is economic differences or just diversity per se that is associated with public goods provision issues.

6.1 Details on variables

We focus primarily on non-excludable public goods for our dependent variable. Specifically, we consider two measures of access to natural gas: a) the average level of gas provisioning in a district; and b) the horizontal inequality of gas provisioning in a district. The choice of public goods is limited by two factors: (1) the availability of the measure at the district level—e.g. while spending on public education is available at the national level, we do not have the corresponding statistics at the district level; and (2) the presence of the measure in both the 2010 PSLM and the 1998 census where the latter allows us to control for path dependency in terms of public good supply for each district.

But more importantly, gas has two distinct characteristics: firstly, gas is associated with provincial politics since two provinces produce the bulk of local natural gas and secondly, natural gas carries a huge annual subsidy of about US$7.4 billion, much more than other sources of energy). Accessing natural gas is therefore quite unique in the sense of attaining rentier status.

Furthermore, gas provisioning requires considerable initial investment on the part of the public sector. This suggests that there may be a role for how strongly a community or group may be able to lobby for the good's provision.

Table 4 provides summary statistics of our public goods' provision levels. The measures are based on household level data which are then aggregated to the district level to represent the proportion of households that have the respective good within the district.

Table 4: Summary statistics—public goods at the district level

	2010	1998
Average gas provisioning	0.249	0.092
	(0.228)	(0.148)
	2010	
Theil of gas provisioning (mean)	0.08	
Theil in gas provisioning	Min 0	Max 0.52

Note: Standard deviations in parentheses.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

While the average provisioning of gas is quite low, there is tremendous variation across districts. Sparsely populated districts have virtually zero average provision levels. Finally, comparing the 2010 levels to 1998 indicates that there has been a marked improvement in coverage.

The Theil measure of gas inequality was constructed as follows. The key feature that we want to capture in this variable is the differential access various ethnic groups have to public goods, using information on whether any given household has gas or not. For each ethnic group in each district, therefore, we calculate the proportion of each ethnic group that has access to gas. So, if 90 per cent of households of one ethnic group have a gas connection, each household of that group gets a value of 0.9. Similarly, if 70 per cent of households of another ethnic group have a gas connection, each household of that group gets a value of 0.7. We then calculate an inequality of these values.

We make use of both vertical and horizontal group-based measures to capture group-level inequality. While the former has been constructed at the district level, i.e. for the entire population, the latter considers the inequality between the different ethnic/linguistic groups within each district.

For measuring vertical and horizontal inequality we make use of the Theil measures for income and land. We choose the Theil since unlike the Gini and the COV this is especially sensitive to extreme instances of inequality.

Our choice of variables capturing economic inequality is driven by the fact that the inequality measures constructed for these variables are expected to drive the mechanisms through which communities negotiate with the state for public goods provision. For example, land inequality, besides capturing wealth, also has a particular social meaning. The British colonial government established different land tenure arrangements with varying levels of inequality and social

equality, mainly as a means to co-opt powerful elite and kinship groups. Even today, social status and power remain tied to land ownership, mediating patron–client networks of access for a wide range of goods and services including both public goods as well as private ones such as formal sector jobs. Similarly, we expect that those ethnicities who are at the higher end of the income distribution will not only also have the necessary networks of access in place to enjoy greater public goods provisioning but may also have very different preferences vis-à-vis the type of public goods that they require relative to poorer groups. Thus, on the one hand, an increase in inequality may mean that poorer groups see spillover benefits from the bargaining that richer groups may be able to do with the state for the provisioning of non-exclusionary public goods. Yet, simultaneously, we may see a decline in provision levels if higher inequality is linked with a divergence in preferences for the public good leading to collective action issues.

Additionally, we make use of the fact that certain districts have higher levels of political competition than others to construct a measure of vote inequality. The Pakistan Election Commission provides data on the voting distribution for each electoral constituency. We construct inequality measures of the votes to capture this voting distribution. If one person gets all the votes, the Gini in that constituency would be 1 and if all candidates get equal votes, the Gini would be 0. Since there are a number of constituencies in each administrative district (the unit of analysis in our regressions) we construct an aggregate measure of inequality as a weighted average of the constituency measures, with population shares serving as weights. Since the number of candidates can vary from one constituency to another, the regression controls for the number of candidates in each constituency. Whether the measure of inequality thus generated is horizontal or vertical is debatable. It could be argued, in a party-based election, where specific parties cater to the interests of specific ethnic groups, that the inequality of votes captures the political aspect of group inequality.

Overall, Table 5 reveals that on average within-group inequality is generally higher than between-group inequality at the district level. Moreover, group-based inequality, both in the horizontal and vertical dimensions, is higher with respect to land. This is unsurprising given the historical background of land inequity and the continued importance of land, especially agricultural, in the socio-economic power context in Pakistan.

Table 5: Summary statistics—group-based inequality measures

	Between Theil	Within Theil
Income	-0.011	0.348
	(0.041)	(0.298)
Land	-0.015	0.716
	(0.071)	(0.535)
Vote inequality (2008 Elections)	1.17 (0.42)	

Note: Standard deviations in parentheses.

Source: Authors' calculation using Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010) and Election Commission of Pakistan (n.d.).

In addition to our inequality measures we capture diversity through a fractionalization index while also controlling for the total number of ethnic groups in each district. The former is a standard Herfindahl Index ($F = 1 - \sum_{i=1}^{N} s_i^2$, where s is the proportion of the group in the population and N is the total number of groups; this then gives the probability that any two people chosen randomly will belong to different ethnic groups). Although we have access to both the 2010 PSLM as well as the 1998 census, we include fractionalization measures based

only on 1998 data thereby providing a control for how past diversity levels in the district may have affected public goods provisioning. We also considered adding several other controls such as the proportion of the population that has completed primary level of education for each district in 2010, the proportion of the district that was urban in both 2010 and 1998, as well as the supply of such public goods as roads and schools. All of these measures are expected to affect each region's ability to effectively bargain with the state for goods provision, while also giving us a sense of the relative strength of the district's labour market and/or level of development of the region which would factor into the state's decision when supplying public services. However, we found that such developmental controls were highly correlated with past levels of gas provisioning within the district.

Table 6 summarizes the additional controls, highlighting that on average there are about two ethnic groups per district (with a minimum of zero and a maximum of seven).

Table 6: Summary statistics—additional controls

Fractionalization (1998)	0.228
	(0.228)
Number of ethnic groups	2.284
	(0.825)

Note: Standard deviations in parentheses.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

6.2 Comparing between- and within-group inequalities

One concern with including both between- and within-group inequalities in the regression analysis relates to the level of correlation between these two measures. We explore this by looking at the overall correlation between the horizontal and vertical measures for each of our variables of interest.

Table 7 reports the correlation levels for the between- and within-group measures for each of our primary variables of interest. As is evident, the horizontal and vertical measures are uncorrelated. Furthermore, upon examining the horizontal inequality for districts with the highest levels of vertical inequality, we do not find that the regions that are the most deprived in terms of the first inequality measure are also the worst performers in terms of the other. In fact, since the number of ethnic groups in our data varies from zero to seven, there are districts that have no ethnic fractionalization and therefore their between-group inequality measures are zero. For these districts then, our variables capturing vertical inequality effectively become the only inequality measures.

Table 7: Correlation matrix

	Theil
Income	0.0146
Land	-0.0026

Source: Authors' calculations using Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

6.3 Regression results: proportion of districts with natural gas

Our results are presented in Tables 8 to 10 below. It merits mention at the outset that all measures of (between-) group inequality (income, land, political) are negatively associated with gas provisioning. The higher the between-group inequality, the lower the proportion of gas in the district. Moreover, while income inequality enters insignificantly in all specifications, land inequality is significant in all specifications and political inequality is significant in all but the richest specification. It is instructive to note from Table 10 columns (2) and (3) that much of the variation explained by the political inequality is explained by levels of (lagged) gas provision in 1998. High proportions of gas provisioning in 1998 predict a higher proportion of gas provisioning in 2010, suggesting a high level of path dependency. One conjecture to explain this finding could be that, for historically low levels of provisioning, the political process could not be harnessed to increase provisioning—marginalized communities have been unable to use the political process to increase access to services. Similarly, relatively well-off communities with historically higher levels of provisioning in 1998 have increased their access to services regardless of political inequality.

The fact that it is land and political inequality that are significant is especially pertinent in the Pakistani context. As mentioned above, land in the Pakistani context has a tremendous relation with existing power structures. The different British land-tenurial systems implied that 'relatively' egalitarian social structures coexisted with 'relatively' egalitarian land distribution.[6] Where thousands of acres of land were allotted to one man, this man provided access to both economic and political resources. Such a land distribution system suggests that the full import of land inequality will be felt at extreme levels of land inequality (hence our use of the Theil, which is more sensitive to the tails of a distribution).

The impact of vertical land inequality is, however, odd. In all but the richest specification, a positive significant coefficient implies that increasing vertical inequality is associated with higher levels of public goods provisioning. A possible conjecture is that higher vertical inequality could lead to the emergence of a strong patron who provides leadership for collective action. This is consistent with a positive but insignificant coefficient on vertical income inequality—it is land rather than income inequality that is amenable for the emergence of a patron.

In addition to between- and within-group inequality, we also find that path dependency matters. Namely, the levels of the goods in 1998 are found to be statistically significant regardless of the specification vis-à-vis the inequality measure. Here it is worth noting that, as per our results, convergence in provision levels, on average, is likely to occur over several decades.

Finally, our lagged measure of ethnic diversity (fractionalization) is significant only once past levels of goods provisioning are included. Given historical provisioning, higher diversity reduces the provisioning of gas. This is consistent with the bulk of literature that suggests that higher diversity reduces the probability of effective collection.

[6] See Bannerjee and Iyer (2005) for a discussion of the different land tenure systems and the social structures that accompanied them.

Table 8: OLS estimates of the effects of income inequality on the level of natural gas provisioning

	(1)	(2)	(3)	(4)	(5)	(6)
Horizontal inc. inequality	-0.371	-0.326	-0.012			
	(0.560)	(0.569)	(0.497)			
No. of ethnic groups	0.024	0.019	0.011	0.028	0.023	0.013
	(0.028)	(0.028)	(0.018)	(0.028)	(0.029)	(0.018)
Fractionalization '98		0.064	-0.199***		0.063	-0.201***
		(0.132)	(0.057)		(0.130)	(0.059)
Gas '98			1.313***			1.309***
			(0.130)			(0.129)
Vertical inc. inequality				0.086	0.082	0.025
				(0.097)	(0.092)	(0.030)
Constant	0.191***	0.188***	0.147***	0.155**	0.153**	0.135***
	(0.059)	(0.059)	(0.034)	(0.069)	(0.068)	(0.036)
Observations	102	102	102	102	102	102
R-squared	0.011	0.015	0.663	0.019	0.022	0.664

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

Table 9: OLS estimates of the effects of land inequality on the level of natural gas provisioning

	(1)	(2)	(3)	(4)	(5)	(6)
Horizontal land inequality	-1.618***	-1.635***	-0.814***			
	(0.307)	(0.301)	(0.268)			
No. of ethnic groups	0.008	0.010	0.008	0.040	0.037	0.012
	(0.021)	(0.024)	(0.016)	(0.029)	(0.030)	(0.018)
Fractionalization '98		-0.029	-0.221***		0.037	-0.200***
		(0.123)	(0.059)		(0.129)	(0.058)
Gas '98			1.174***			1.312***
			(0.126)			(0.130)
Vertical land inequality				0.081*	0.078**	0.003
				(0.043)	(0.039)	(0.022)
Constant	0.207***	0.209***	0.160***	0.100	0.101	0.145***
	(0.049)	(0.049)	(0.031)	(0.076)	(0.075)	(0.042)
Observations	102	102	102	102	102	102
R-squared	0.258	0.258	0.717	0.039	0.040	0.663

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

Table 10: OLS estimates of the effects of vote inequality on the level of natural gas provisioning

VARIABLES	(1) gas_10	(2) gas_10	(3) gas_10
Political inequality	-0.163**	-0.162**	-0.010
	(0.078)	(0.078)	(0.055)
No. of ethnic groups	0.027	0.021	0.011
	(0.028)	(0.028)	(0.018)
Fractionalization '98		0.087	-0.199***
		(0.136)	(0.065)
Gas '98			1.308***
			(0.134)
Constant	0.267***	0.261***	0.153***
	(0.070)	(0.068)	(0.043)
Observations	100	100	100
R-squared	0.058	0.065	0.662

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

6.4 Regression results: inequality in public goods provisioning

In the specifications presented in Tables 11 to 13, the dependent variable is the Theil measure on horizontal inequality in public goods (see above for construction of this measure). Since these regressions are looking at the association between different types of inequality (namely, income/political inequality and inequality in gas provisioning across ethnic groups), we expect that these effects may vary depending on the levels of the good in each district. Thus, we control for the level of the public good within each district by adding the proportion of households with gas (*Gas '10*).

Our regression results appear consistent with the narrative developed in the section above. All measures of the horizontal inequality (land, income, and political) are positively associated with inequality in public goods (while the coefficients on the Theil measures are negative, there are interaction terms in the regression and these have to be netted out). There is, however, a discrepancy in terms of significance. While land inequality is positively associated with public goods inequality, it is insignificantly so. And income inequality (insignificant in the previous specification) is positively and significantly associated with public goods inequality. Political inequality is, however, significantly associated in both specifications.

Measures of vertical inequality are more clearly consistent with results of the first specification. Measures of vertical inequality in all three dimensions—land, income, and political, are negatively associated with inequality in public goods. Moreover, it is land inequality that is significantly associated suggesting, as discussed above, that land is a peculiar repository of influence vis-à-vis the Pakistani state.

Table 11: OLS estimates of the effects of income inequality on the level of natural gas provisioning

	(1)	(2)	(3)	(4)	(5)	(6)
Horizontal inc. inequality	-0.842**	-0.836**	-0.808**			
	(0.373)	(0.374)	(0.396)			
No. of ethnic groups	0.021	0.020	0.020	0.023	0.021	0.021
	(0.018)	(0.020)	(0.020)	(0.018)	(0.020)	(0.020)
Fractionalization '98		0.010	0.017		0.017	0.027
		(0.051)	(0.061)		(0.051)	(0.056)
Gas '98			-0.041			-0.083
			(0.087)			(0.079)
Gas '10	-0.158***	-0.158***	-0.138*	-0.130*	-0.122	-0.119
	(0.046)	(0.046)	(0.080)	(0.077)	(0.086)	(0.088)
Horizontal inc. inequality*	1.719**	1.719**	1.615*			
Gas '10	(0.831)	(0.826)	(0.916)			
Vertical inc. inequality				0.066**	0.070**	0.050
				(0.030)	(0.034)	(0.035)
Vertical inc. inequality*				-0.082	-0.105	0.003
Gas '10				(0.116)	(0.145)	(0.156)
Constant	0.077*	0.077*	0.074*	0.054	0.053	0.054
	(0.041)	(0.042)	(0.043)	(0.045)	(0.045)	(0.045)
Observations	100	100	100	100	100	100
R-squared	0.115	0.115	0.116	0.103	0.104	0.106

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

Table 12: OLS estimates of the effects of land inequality on inequality of natural gas provisioning

	(1)	(2)	(3)	(4)	(5)	(6)
Horizontal land inequality	-0.388**	-0.385**	-0.369*			
	(0.178)	(0.193)	(0.217)			
No. of ethnic groups	0.022	0.021	0.021	0.030	0.030	0.031
	(0.019)	(0.020)	(0.021)	(0.019)	(0.021)	(0.021)
Fractionalization '98		0.004	0.010		-0.007	-0.013
		(0.050)	(0.064)		(0.049)	(0.062)
Gas '98			-0.028			0.031
			(0.086)			(0.111)
Gas '10	-0.168***	-0.167***	-0.154*	-0.032	-0.031	-0.040
	(0.054)	(0.054)	(0.080)	(0.062)	(0.062)	(0.071)
Horizontal land Inequality*	0.499	0.498	0.465			
Gas '10	(0.368)	(0.373)	(0.406)			
Vertical land inequality				0.096**	0.097**	0.099**
				(0.041)	(0.042)	(0.047)
Vertical land inequality*				-0.179**	-0.180**	-0.189**
				(0.074)	(0.075)	(0.092)
Constant	0.082*	0.082*	0.080*	-0.005	-0.005	-0.005
	(0.047)	(0.047)	(0.048)	(0.047)	(0.048)	(0.048)
Observations	100	100	100	100	100	100
R-squared	0.107	0.107	0.107	0.146	0.146	0.146

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

Table 13: OLS estimates of the effects of vote inequality on inequality in natural gas provisioning

	(1)	(2)	(3)
Political inequality	-0.104*	-0.105*	-0.113*
	(0.060)	(0.061)	(0.067)
No. of ethnic groups	0.020	0.017	0.017
	(0.019)	(0.020)	(0.020)
Fractionalization '98		0.038	0.027
		(0.048)	(0.057)
Gas '98			0.063
			(0.122)
Gas '10	-0.294***	-0.299***	-0.349**
	(0.082)	(0.081)	(0.149)
Political inequality*Gas '10	0.271**	0.279**	0.318**
	(0.119)	(0.118)	(0.153)
Constant	0.142***	0.141***	0.150**
	(0.052)	(0.052)	(0.061)
Observations	98	98	98
R-squared	0.128	0.132	0.133

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations using Census of Pakistan 1998 (GOP 1998) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2010–11 (GOP 2010).

7 Conclusion

The empirical analysis provides mixed support for Alesina et al (2012). We find that all measures of horizontal inequality are negatively associated with natural gas provisioning. Higher group inequality in land and in electoral votes is associated with a lower share of the population with access to gas. Conversely, higher group inequality in land and in electoral votes is associated with a higher **inequality** in public goods provisioning. This appears consistent with the idea that higher levels of inequality along one dimension feed into higher levels of inequality in other dimensions. These results also support the hypothesis that higher levels of inequality make it difficult to organize the collective action required to access as subsidized a public service as natural gas in Pakistan.

Our results on vertical inequality are counter intuitive: Higher levels of land inequality are negatively associated with **inequality** in natural gas provisioning. Conversely, higher levels of inequality are associated with higher **levels** of natural gas provisioning. We conjecture that higher levels of vertical land inequality might be associated with the emergence of powerful patrons who mediate access to public goods in exchange for other social transactions.

References

Alesina, A., E. Glaeser, and B. Sacerdote (2001). 'Why Doesn't the US Have a European-style Welfare System?'. NBER Working Paper No. 8524. Cambridge, MA: National Bureau of Economic Research

Alesina, A.F., S. Michalopoulos, and E. Papaioannou (2012). 'Ethnic Inequality'. NBER Working Paper No. 18512. Cambridge, MA: National Bureau of Economic Research.

Alesina, A., and E. Zhuravskaya (2011). Segregation and the Quality of Government in a Cross Section of Countries. *The American Economic Review*, 101(5): 1872–911.

Alkire, S., and R. Black. (1997). 'A Practical Reasoning Theory of Development Ethics: Furthering the Capabilities Approach'. *Journal of International Development*, 9(2): 263–79.

Andrabi, T., J. Das, and A.I. Khwaja (2014). 'Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets'. Unpublished Manuscript. Cambridge, MA: Harvard University, Department of Economics. Available at: https://www.hks.harvard.edu/fs/akhwaja/papers/RC_June2014.pdf (accessed on 1 September 2016).

Anwar, T., S. Qureshi, H. Ali, and M. Ahmad (2004). 'Landlessness and Rural Poverty in Pakistan [with Comments]'. *The Pakistan Development Review*, 43(4): 855–74.

Arif, G.M., and S. Farooq (2011). 'Poverty, Inequality and Unemployment in Pakistan'. Pakistan Institute of Development Economics Background Paper. Karachi: Pakistan Institute of Development Economics.

Baldwin, K., and J.D. Huber (2010). 'Economic versus Cultural Differences: Forms of Ethnic Diversity and Public Goods Provision'. *American Political Science Review*, 104(04): 644–62.

Banerjee, A., and L. Iyer (2005). 'History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India'. *The American Economic Review*, 95(4): 1190–213.

Bedard, K., and C. Ferrall (2003). 'Wage and Test Score Dispersion: Some International Evidence'. *Economics of Education Review*, 22(1): 31–43.

Bergan, A. (1967). 'Personal Income Distribution and Personal Savings in Pakistan: 1963/64'. *The Pakistan Development Review*, 7(2): 160–212.

Blau, F.D., and L.M. Kahn (2005). 'Do Cognitive Test Scores Explain Higher US Wage Inequality?'. *Review of Economics and Statistics*, 87(1): 184–93.

Burki, A.A., R. Memon, and K. Mir (2015). *Multiple Inequalities and Policies to Mitigate Inequality Traps in Pakistan*. Oxfam and LUMS. Available at: http://policypractice.oxfam.org.uk/publications/multiple-inequalities-and-policies-to-mitigateinequality-traps-in-pakistan-346432\

Chaudhry, I.S., and S. Rahman (2009). 'The Impact of Gender Inequality in Education on Rural Poverty in Pakistan: An Empirical Analysis'. *European Journal of Economics, Finance and Administrative Sciences*, 15(1): 174–88.

Dasgupta, P., and D. Ray (1986). 'Inequality as a Determinant of Malnutrition and Unemployment: Theory'. *Economic Journal*, 96(384): 1011–34.

Deiwiks, C., L.-E. Cederman, and K.S. Gleditsch (2012). 'Inequality and Conflict in Federations'. *Journal of Peace Research*, 49(2): 289–304.

Desai, M. (1995). *Poverty, Famine and Economic Development: The Selected Essays of Meghnad Desai, Volume II*. Cheltenham, UK: Edward Elgar Publishing Ltd.

Deshpande, A. (2000). 'Does Caste Still Define Disparity? A Look at Inequality in Kerala, India'. *The American Economic Review,* 90(2): 322–25

Duflo, E., and A. Banerjee (2007). 'The Economic Lives of the Poor'. *Journal of Economic Perspectives*, 21(1): 141–67

Election Commission of Pakistan (n.d.). 'General Elections'. Available at: https://www.ecp.gov.pk/AllResults.aspx (accessed on 14 November 2016).

Federal Bureau of Statistics (1997). 'Pakistan Integrated Household Survey 1996–97'. Available at: http://www.pbs.gov.pk/content/household-integrated-economic-survey-hies-2011-12 (accessed on 1 September 2016).

GOP (Government of Pakistan) (1998). 'Pakistan Census Report'. Islamabad: Population Census Organization, 2000. Available at: http://www.pbs.gov.pk/content/population-census (accessed on 1 September 2016).

GOP (Government of Pakistan) (2007). 'Pakistan Social and Living Standard Measurement Survey (PSLM) 2007'. Federal Bureau of Statistics. Available at: http://www.pbs.gov.pk/content/pakistan-social-and-living-standards-measurement (accessed on 1 September 2016).

GOP (Government of Pakistan) (2010). 'Pakistan Social and Living Standard Measurement Survey (PSLM) 2010-11'. Federal Bureau of Statistics. http://www.pbs.gov.pk/content/pakistan-social-and-living-standards-measurement (accessed on 1 September 2016).

Izady, M. (2000a). 'Languages of Pakistan'. Available at: http://gulf2000.columbia.edu/images/maps/Pakistan_Baluchistan_Linguistic_sm.png (accessed on 1 September 2016).

Izady, M. (2000b). 'Ethnic Groups in Pakistan'. Available at: http://gulf2000.columbia.edu/images/maps/Pakistan_Baluchistan_Ethnic_sm.png (accessed on 1 September 2016).

Jamal, H. (2006). 'Does Inequality Matter for Poverty Reduction? Evidence from Pakistan's Poverty Trends'. *The Pakistan Development Review*, 45(3):439–59.

Kanbur, R., and A.J. Venables (2003). 'Spatial Inequality and Development'. Introduction to an Edited Volume of Papers from the Cornell-LSE-UNU-WIDER Conference on Spatial Inequality and Development.

Lessman, C. (2015). 'Regional Inequality and Internal Conflict'. *German Economic Review*, 17: 157–91.

Luttmer, E.F., and M. Singhal. (2011). 'Culture, Context, and the Taste for Redistribution'. *American Economic Journal: Economic Policy*, 3(1): 157–179.

Memon, R. (2013). 'Gender Prejudice as a Sense of Social Position'. Paper Presented at the British Association of South Asian Studies 6–8 April 2016, Cambridge University, Cambridge, UK.

Miguel, E., and M. Kremer (2004). 'Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities'.*Econometrica*, 72(1): 159–217.

Munshi, J. (2014). *A Method for Constructing Likert Scales*. Available at SSRN 2419366.

Nussbaum, M.C. (2005). 'Women's Bodies: Violence, Security, Capabilities'. *Journal of Human Development*, 6(2): 167–83.

Rao, V. (2006). 'On "Inequality Traps" and Development Policy'. World Bank Other Operational Studies No. 9588. Washington, DC: The World Bank.

Robeyns, I. (2003). 'Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities'. *Feminist Economics*, 9(2–3): 61–92.

Sen, A. (1980). *Equality of What?*. In *McMurrin S Tanner Lectures on Human Values, Volume 1*. Cambridge: Cambridge University Press.

Shahbaz, M., N. Aamir, and M.S. Butt (2007). 'Rural-Urban Income Inequality under Financial Development and Trade Openness in Pakistan: The Econometric Evidence'. *The Pakistan Development Review*, 46(4): 657–72.

Sowell, T. (1983). *The Economics and Politics of Race: An International Perspective*. New York: W. Morrow.

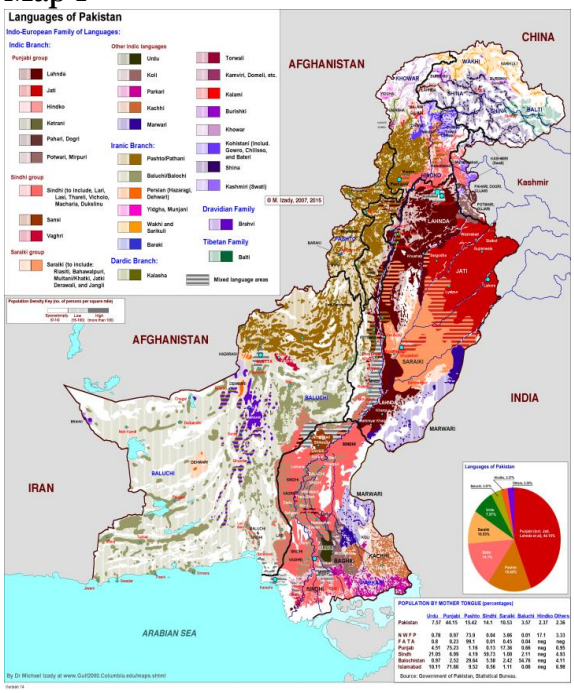
Stewart, F. (ed.) (2008). *Horizontal Inequalities and Conflict: Understanding Group Violence in Multiethnic Societies*. Houndsmills, Basingstoke: Palgrave MacMillan.

Strauss, J., and D. Thomas (1998). 'Health, Nutrition and Economic Development'. *Journal of Economic Literature*, 36(2): 766–817.

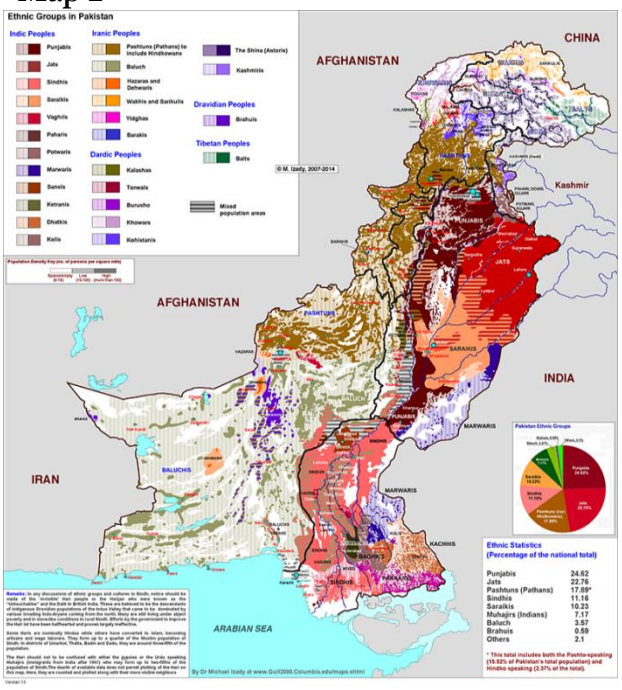
Tilly, C. (1999). *Durable Inequality*. Berkeley, CA: University of California Press.

Appendix A

Map 1



Map 2



Source: Izady (2000a, 2000b) see http://gulf2000.columbia.edu/images/maps/Pakistan_Baluchistan_Linguistic_sm.png and http://gulf2000.columbia.edu/images/maps/Pakistan_Baluchistan_Ethnic_sm.png

Appendix B: HI Fomulae

$$\textbf{GCOV} = \frac{1}{\bar{y}}[\sum_r^R p_r(\overline{y_r} - \bar{y})^2]^{\frac{1}{2}}$$

$$\textbf{GGini} = \frac{1}{2\bar{y}}\sum_r^R \sum_s^S p_r p_s |\overline{y_r} - \bar{y}_s|$$

$$\textbf{GTheil} = \sum_r^R p_r \frac{\bar{y}_r}{\bar{y}} \log(\frac{\bar{y}_r}{\bar{y}})$$

Where $\bar{y}_r = \frac{1}{n_r}\sum_i^{n_r} y_{ir}$ is mean value for group r; R is group r's population size; p_r is group r's population share; y_{ir} is the variable of interest e.g. years of education of the ith member of group r; y_r is the value of y for group; and y is the total of the variable of interest in the sample.

Note that we use population weighted formulae for each of our inequality measures since without population weighting changes in the position of a very small group would have the same effect on the index as that of a large group.

Appendix C: Education inequality

Table C1: Measures of horizontal inequality—years of education (>=15 &<25 years of age)

	Year	Ethnicity	Province	Rural/urban
GCOV	1996	0.2661	0.1031	0.1508
	2007	0.2140	0.1247	0.1524
	2010	0.0546	0.0436	0.1122
GGini	1996	0.1082	0.0850	0.1447
	2007	0.0839	0.0648	0.1291
	2010	0.0212	0.0197	0.0431
GTheil	1996	0.0492	0.0063	0.0428
	2007	0.0532	0.0079	0.0338
	2010	0.0049	-0.0024	-0.0102

Note: Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan, Baluchi and Hazara.

Source: Authors' calculations using Pakistan Integrated Household Survey (PIHS) 1996–97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010–11 (GOP 2010).

Table C2: Measures of horizontal inequality across ethnicities—education years

	Year	15–25 years	25–45 years	Above 45 years
GCOV	1996	0.2661	0.4232	0.5448
	2007	0.2140	0.2955	0.4470
	2010	0.0546	0.0809	0.0980
GGini	1996	0.1082	0.1579	0.1976
	2007	0.0839	0.1095	0.1624
	2010	0.0212	0.0319	0.0380
GTheil	1996	0.0492	0.1051	0.1496
	2007	0.0532	0.0790	0.1243
	2010	0.0049	0.0025	-0.0149

Note: Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan, Baluchi and Hazara.

Source: Authors' calculations using Pakistan Integrated Household Survey (PIHS) 1996–97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010–11 (GOP 2010).

Table C3: Mean and standard deviation for age cohorts

Year	15-25 Years		25-45 Years		Above 45 Years	
	Mean	Std dev	Mean	Std dev	Mean	Std dev
1996	5.19	4.98	3.73	5.20	2.07	4.08
2007	5.87	5.04	5.01	5.84	2.89	4.93
2010	8.56	3.27	9.54	4.41	9.00	4.47

Source: Authors' calculation using Pakistan Integrated Household Survey (PIHS) 1996–97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010–11 (GOP 2010).

Appendix D: Nutritional inequality

Table D1: Measures of horizontal inequality—overall meat consumption

	Year	Ethnicity	Province	Rural/urban
GCOV	1996	0.4675	0.4219	0.1156
	2007	0.3184	0.1035	0.1044
	2010	0.2364	0.1228	0.0730
GGini	1996	0.1681	0.2138	0.1347
	2007	0.1174	0.0542	0.1198
	2010	0.0867	0.0627	0.0819
GTheil	1996	0.0298	0.0848	0.0370
	2007	0.0107	0.0052	0.0289
	2010	-0.0053	0.0073	0.0136

Note: Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan, Baluchi and Hazara. Meat includes chicken, mutton, beef and fish. The measures are based on per capita expenditure on meat consumption

Source Authors' calculation using Pakistan Integrated Household Survey (PIHS) 1996-97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 207) and 2010-11 (GOP 2010).

Table D2: Measures of horizontal inequality—poultry/fish

	Year	Ethnicity	Province	Rural/urban
GCOV	1996	0.3701	0.3285	0.0645
	2007	0.2604	0.0763	0.0664
	2010	0.1482	0.0901	0.0417
GGini	1996	0.1442	0.1656	0.0751
	2007	0.0935	0.0417	0.0874
	2010	0.0592	0.0479	0.0533
GTheil	1996	0.0365	0.0523	0.0116
	2007	0.0358	0.0029	0.0154
	2010	-0.0107	0.0040	0.0057

Note:. Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan, Baluchi and Hazara. The measures are based on per capita expenditure on meat consumption

Source: Authors' calculations using Pakistan Integrated Household Survey (PIHS) 1996–97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010–11 (GOP 2010).

Table D3: Measures of horizontal inequality—mutton

	Year	Ethnicity	Province	Rural/urban
GCOV	1996	0.5582	0.6758	0.1198
	2007	0.2288	0.1545	0.0418
	2010	0.2284	0.2280	0.0275
GGini	1996	0.2008	0.3297	0.1396
	2007	0.0658	0.0792	0.1089
	2010	0.0618	0.1223	0.0767
GTheil	1996	0.0363	0.2139	0.0397
	2007	0.0905	0.0124	0.0249
	2010	0.1083	0.0261	0.0120

Note: Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan, Baluchi and Hazara. The measures are based on per capita expenditure on meat consumption

Source: Authors' calculations using Pakistan Integrated Household Survey (PIHS) 1996–97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010–11 (GOP 2010).

Table D4: Measures of horizontal inequality—beef

	Year	Ethnicity	Province	Rural/urban
GCOV	1996	0.6515	0.4574	0.1457
	2007	0.2249	0.1175	0.0459
	2010	0.1556	0.0571	0.0234
GGini	1996	0.2248	0.2484	0.1697
	2007	0.0863	0.0635	0.0645
	2010	0.0584	0.0287	0.0333
GTheil	1996	0.1283	0.1083	0.0586
	2007	-0.0050	0.0069	0.0084
	2010	-0.0107	0.0016	0.0023

Note: Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan, Baluchi and Hazara. The measures are based on per capita expenditure on meat consumption

Source: Authors' calculations using Pakistan Integrated Household Survey (PIHS) 1996–97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010–11 (GOP 2010).

Appendix E: Income inequality

Table E1: Measures of horizontal inequality—income

	Year	Ethnicity	Province	Rural/urban
GCOV	1996	0.2476	0.2440	0.0245
	2007	0.2460	0.0447	0.2712
	2010	0.2061	0.0519	0.5416
GGini	1996	0.0999	0.1216	0.031
	2007	0.0941	0.0236	0.127
	2010	0.0781	0.0277	0.114
GTheil	1996	0.0380	0.0282	0.001
	2007	0.0154	-0.0006	-0.004
	2010	0.0031	0.0023	-0.013

Note: Individual income (primary+ secondary+ other sources such as pension. Ethnicity is divided into the following categories: Muhajir, East Punjabi, Punjabi, Sindhi, Seraiki, Pathan and Baluchi.

Source: Authors' calculations using Pakistan Integrated Household Survey (PIHS) 1996-97 (Federal Bureau of Statistics 1997) and Pakistan Social and Living Standards Measurement Survey (PSLM) 2007 (GOP 2007) and 2010-11 (GOP 2010).