CPWF WORKING PAPER

Improving knowledge for targeting interventions:

Willingness of individuals to participate and calculation of institutional environment indices

Jorge Pubiano and James Garcia
Centro Internacional de Agricultura Tropical,
Cali, Colombia

2009

The CGIAR Challenge Program on Water and Food, Colombo, Sri Lanka.
Impact Assessment Module of the Basin Focal Projects

ISSN 978-92-990053-9-2
You can find the CPWF Working Paper series online at
http://www.waterandfood.org/publications/program-publications/working-papers.html.
Cover design by mcn creative association (Pvt) Ltd.

This paper is a contribution to the synthesis work of the CGIAR Challenge Program on Water and Food's Basin Focal Projects.

It should be cited as:

Rubiano, Jorge, and James Garcia, 2009. Improving Knowledge for Targeting Interventions: Willingness of Individuals to Participate *and Calculation of Institutional Environment Indices.* CPWF Working Paper 05. Colombo, Sri Lanka: The CGIAR Challenge Program on Water and Food. 25pp.

You can find the CPWF Working Paper series online at www.waterandfood.org.

The Challenge Program on Water and Food (CPWF), an initiative of the Consultative Group on International Agricultural Research (CGIAR), contributes to efforts of the international community to ensure global diversions of water to agriculture are maintained at the level of the year 2000. It is a multi-institutional research initiative that aims to increase water productivity for agriculture—that is, to change the way water is managed and used to meet international food security and poverty eradication goals—in order to leave more water for other users and the environment.

The CPWF conducts action-oriented research in nine river basins in Africa, Asia and Latin America. Its research topics focus on crop water productivity, management of fisheries and aquatic ecosystems, promotion of community arrangements for water sharing and integrated river basin management as well as helping to establish and influence institutions and policies for successful implementation of developments in water, food, and environmental management.

The CGIAR Challenge Program for Water and Food's Working Paper series captures the Program's work in progress that is worthy of documentation before a project's completion. Working papers may contain preliminary material and research results - data sets, methodologies, observations, and findings that have been compiled in the course of research, which have not been subject to formal external reviews, but which have been reviewed by at least two experts in the topic. They are published to stimulate discussion and critical comment.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

A method for the calculation of Individual's Willingness to Participate and Institutional Environment Indexes is presented here. A description of the approach, data and procedures is presented using as a case study municipal data of Bolivia. This is used as an example of the way how these two indicators are feasiblecan be estimated to estimate using secondary data. The method can be easily applied to countries with relatively good socio-economic secondary data at municipal or similar administrative unitslevel to produce continental or sub continental pictures of these two key factors. The purpose of developing these two indicators was to provide better information for targeting interventions as pursued by the implementation of Extrapolation Domain Analysis (EDA).

1. ABSTRACT

We present a method for the calculation of individuals' willingness to participate (IWP) and an institutional environment index (IEI). A description of the approach, data and procedures is presented using as a case study of data for municipalities in Bolivia. We use these procedures as examples of how these two indicators can be estimated using secondary data. The method can be easily applied to countries with relatively good socio-economic secondary data at municipal level to produce continental or sub-continental pictures of these two key factors. The purpose of developing these two indicators was to provide better information for targeting interventions as identified by the implementation of Extrapolation Domain Analysis (EDA).

Keywords: Extrapolation domain analysis, socioeconomic data handling, Latin America, institutions

2. INTRODUCTION

2.1 Background and scope

The extrapolation domain analysis (EDA) technique is one of the impact assessment methods being used by CPWF-funded projects. The technique was first used in the assessment of the basin representativity in the Andes region to select the seven Andes systems of basins for CPWF projects (Otero et al. 2006). The method also took advantage of the development of Homologue (Jones et al. 2005) by adding a complementary way to include socio-economic variables in the search for similar sites around the tropical world.

EDA is being used together with the impact pathways and global impact analysis (Bouman et al. 2007) led by CIAT and IFPRI respectively, as a complementary approach to assessing the potential to scale up and transfer the knowledge and successes of CPWF-funded projects to other areas around the tropics. Its implementation is potentially crucial to the successful out-scaling of CPWF projects. Its application has two main objectives. First, to contribute in guiding the CPWF in identifying new areas where change or modification of current water management practices are required to achieve more sustainable and equitable water resource availability. Secondly, it has an important role to play in helping project implementers frame their project scope to achieve maximum positive impact.

A key question in development is that some areas and their communities are more open to innovation than others. National and local government, development NGOs, private corporations, and donors, among many others, have all been interested in targeting where their investments will have the greatest pay-off. Each of these organizations looked for the critical factors that, if modified, could facilitate directed change.

The literature has identified the institutional environment and individuals' willingness to participate as important factors that foster innovation, particularly in rural or marginal areas (Rubiano et al. 2008a). The institutional environment means different things for different people. In this paper we understand it to be that mix of conditions that individuals, organizations and political rules allow interventions to run and achieve successful results. Before implementing a project in a particular site, it is therefore desirable to find situations where local partners with co-financing are strongly involved, farmers are interested and active, and administrative process run at a stimulating pace and in such a way that lessons learned spread out as good news attracting new investors. What can be behind particular conditions like these? What makes individuals willing to participate in projects promoted by external organizations?

This paper presents a systematic attempt to answer these questions. It is a first step, and we invite comments and contributions from those interested in improving the way resources and human effort are allocated in development and research projects.

3. OBJECTIVE

A series of consultations with CPWF project members and others identified the factors that must be taken into account when external actors wish to implement an intervention. In the particular case of the CPWF, we define *intervention* as activities to implement the extrapolation domain of projects' results/findings (Rubiano et al. 2008a, b). The two most important factors that they identified were the quality of the institutional environment and people's willingness to cooperate. Based on this analysis, the overall objective of the study was to advance understanding in the spatial definition of these two important factors because they are key to defining those domains to which projects' results/findings might be extrapolated.

4. IMPROVING KNOWLEDGE FOR TARGETING INTERVENTIONS

4.1 Conceptual framework

A logical procedure for the selection of key measurable variables that influence the success of a proposed intervention requires a framework, or series of concepts, that guide their selection. According to Williamson (1993) and Eggertsson (1996), both cited by Saleth et al. (2004), three main coupled components and their interactions are important determinants of successful performance of interventions: the institutional environment, the economic and political organizations, and the individual decision makers. The institutional environment is a reflection of the transactions made between political and economic organizations to produce particular outcomes. Individuals with particular backgrounds and objectives make up and operate these organizations. If an intervention seeks the right conditions, it has to consider these individuals, and how their actions are reflected in the outcomes of the organizations to which they belong. It also has to consider the surrounding institutional environment and the regulatory framework within which they act.

Attributes that describe each of these elements, fully or partially, could be used to develop an indication of their status and together provide a measure of the institutional environment of a particular place. The assumption is that the institutional environment can be described by measuring the institutions' actual influence compared with the community expectation. For example, a good institutional environment will reflect a positive role of the state at the level of the department, such as the presence of particular organizations that promote development and investment, civil security and efficiency in the management of official entities. Measures of some of these characteristics are available, which, if they are systematically grouped and synthesized, can be expressed in a numerical index. We apply this concept here to both the IWP and the IEI using statistical procedures.

Individuals can be characterized by the condition of their economic well being, their level of education, and their health (Figure 1). Since the objective is to use currently-available information, we used data from the population census to obtain data such as the level of education, the number of students per teacher, education coverage, maximum scholastic level achieved, number of people in official organizations trained to high-level, etc. The reason to use education-related variables is based on, "the impact of scientific and technological progress in improving agricultural production in developing countries is intimately related to the skills and education of the populations in those countries. There are three key groups whose skills and education levels are of fundamental importance: farmers, information providers and researchers" (Pardey et al. 2007). Information about any of these three components is a useful indicator to estimate the potential to improve agricultural development. Health of the current population is also an indicator of the physical quality of their lives, which has potential effects on their capacity and willingness to embark on new or additional tasks.

The assumption for the identification of willingness to participate is that there are particular levels for certain key individual characteristics within the population that influ-

ence the efficiency with which interventions can be successful. These characteristics include individual characteristics such as poverty, health, and education. Wherever these characteristics are at a high level (low level pf poverty), it is very likely that there is no need for interventions, but at the other extreme, where poverty is severe and the quality of education and health is low, it will be difficult for interventions to produce change. Intermediate levels of these characteristics are where interventions are likely to have optimal effects.

It is important to recognize that both institutional support and individuals' willingness to participate are what are termed *agents' variables*. Both determine how effective interventions by external organizations will be at a particular place. They are complementary in the sense that in a good institutional environment people are typically willing to participate. In contrast, in a non-supportive institutional environment people are suspicious and often refuse external interventions. The objective or purpose of the planned interventions is the ultimate factor that defines the target communities for whom it is necessary to determine their individual and institutional status.

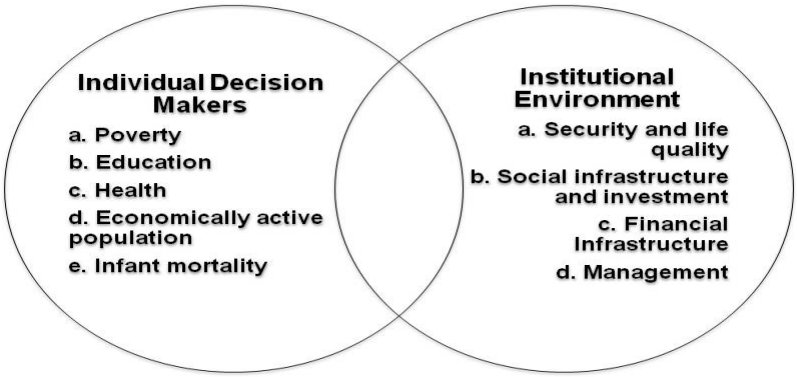

Figure 1.Conceptual framework for the identification of the institutional environment and willingness to participate.

The characteristics of institutions and the organizations relevant to a particular intervention can be assessed in different ways. Restricting ourselves to the water and food sector, indices of efficiency and efficacy of official organizations, transparency, invested resources in particular sectors (such as potable water and sanitation, education, and health), coverage of particular public services, mortality rate, existence of local farmers' organizations, cooperatives, etc. The existence of institutions that support market transactions, influencing growth and promoting the use of people's assets is also a reflection of the level at which individuals belong to organizations supported by legal institutions. Closely associated with these characteristics is the context in which individuals and organizations operate. The status of natural resources, land use, access to markets, the communication technologies in place, openness and flow of information, provision of laws and performance of the judiciary and police to protect property and civil rights, are some of the potential proxy variables from which to infer the context that determine whether an intervention is viable or not. In summary, the status of individuals and their institutional context are the prime factors that determine the institutional environment and people's willingness to participate.

4.2 Methodology

Taking the framework described above, the next task was to seek raw material (data) and select a simple but consistent procedure for the production of the two indices. We used Bolivia, where sub-national data sets were available, as an example of the procedure. We built a single database for the country from national databases we obtained from the internet, local in-country collaborators, and various published reports.

There were a number of problems to resolve. The data are not homogeneous because methods for collection differed depending on the country or region, the dates of measurements are often different, spatial resolution and aggregation are different and in some cases the data are poorly documented. Despite these limitations, the collected data provided a starting point on which to base the evaluation of the IEI and IWP. Disparate sources and variable data are realities, which we addressed here using flexible methodology of proxy variables and statistical scrutiny. This allowed us to integrate the data into a single index that is easy to interpret.

We expect data quality to improve, which will allow simpler approaches than those we used here. Nevertheless we were able to design methods to cope with limitations in the data and to answer the questions we posed. As stated by NRC (2007), "recent advances in the availability of social-spatial data and the development of geographic information systems (GIS) and related techniques to manage and analyze those data give researchers important new ways to study important social, environmental, economic, and health policy issues and are worth further development".

We used the same procedure to generate both indices:

1. Selection of variables and their proxies;
2. Collection of data from a wide range of sources;
3. Organization of the data and their manipulation in a database;
4. Estimation of missing data;
5. Exploration of the data by multiple regression;
6. Cluster and principal components analysis; and
7. Combination of scores to allow spatial representation of the data.

In more detail, and guided by the framework stated above, we selected a preliminary list of variables and their proxies. Our underlying assumption was that the variables we chose for each locality indicated willingness of individuals to participate and availability of institutional support. The unit of analysis we used was the municipality since this is the level at which the most of the information is currently available. Some variables were available at higher administrative levels than the municipality. Where they were relevant we disaggregated them and used them in the analysis.

We searched for relevant data in national databases, ministries, national statistic offices and other reliable sources. A subset of the variables we found is presented in Table 1. The number of dots indicates the number of sources we found for each variable. Given the disparity between sources for each of the variables, we made the analysis for the whole country to avoid the errors caused by mixing potentially incompatible sources. In many cases the data were available in formats that restricted their direct incorporation into spreadsheets or databases, which required manual entry of the data. All the information was carefully verified to remove errors in data entry. Table 2 presents the final set of variables used for the calculation of the indices.

We created an hierarchical database in the Access database program with the higher level being the sum for the country and the lower the municipalities. In some cases departments, provinces, cantons, and sections, which are decreasing order of administrative units from country to municipality, were also included. We identified keys for each spatial unit and matched each with the key in the geographical database of the administrative unit.

We scrutinized each variable to ensure that it provided a clear record of the data units, the collection dates, and the relevance of the data. Depending on the type of variable, we normalized some of them, e.g. as a percentage of the total population when values were given as total of individuals in a class. Once variables were normalized, we estimated missing values using forward stepwise regression, the purpose of which was to use highly-correlated independent but associated variables to estimate the missing data. We also calculated a correlation matrix and used it to identify endogenous or redundant variables (Appendix B). When there were no related variables from which

to infer missing values, we assigned the mean of the higher administrative unit to the municipalities to which they belong. Variables judged to have an unacceptably high number of missing values were excluded from the analysis.

5. RESULTS

5.1 Individuals' willingness to participate index (IWP)

The process described above was made for each of the spheres illustrated in Figure 1. We continue with an explanation of how we managed the data for the sphere that represents IWP. This index is derived from the human condition of individuals within the Bolivian population in terms of their health, education, and poverty status.

Table 1. List of data variables found for Bolivia; data for Colombia, Equador, and Peru are shown for comparison. Black dots are for provincial and red dots are for municipal data sets.

Variables		Country			
		Bolivia	Colombia	Equador	Peru
Poverty	NBI population	● ●		● ● ●	● ●
	NBI households		● ●		● ●
Financial infrastructure	Bank branches	●	●	●	●
	Corruption	● ● ●	●	●	● ●
Management	Municipal quality management	●	● ●	● ●	
	Aqueducts coverage	● ●	● ●	●	●
	Management capacity Index		●		
Life quality	Human development Index	●	● ● ●		●
	Economic active population, employment	●			
Security	Violent deaths	●			
	Displaced population		●		
Social investment	Municipal Investment	●	●		
	Water related investments	●			
	Irrigation management investment	●			
Education infrastructure	Health municipal index				●
	Education	● ● ●	● ●	● ●	
	Lectures in secondary schools	● ●	● ●		●
Training	Literacy	● ●	●	● ●	●
Health	Mortality	●			
	Infant mortality	●	●		
	Nutrition	● ● ● ● ● ●	● ●	● ● ● ●	● ●
Population	Rural population	●	●	●	
	Urban population	●	●	●	
	Total population	●	●	●	●

We used principal components analysis (PCA) to reduce the number of variables and to simplify the multiple dimensions represented by the list of variables. PCA allows the generation of intrinsic variables (components) as the by-products of lineal combinations of the original variables. By using the Kaiser-Meyer-Olkin test, it is feasible to identify the original variables that contribute most to each component. We used a threshold to exclude non-informative variables to each component in an iterative process until the number of variables was reduced to a minimum below which we accounted for no further reduction in the variance.

During this process we were able to gain insight into the key characteristics of each index as reflected in its constituent variables. The contribution of each component in the explanation of the model as calculated is detailed in Appendix C. Note that four main components explain about 90% of the model. Figure 2 is a graphical representation of the contribution of each variable. At the left (negative) side of the figure, infant mortality and the percentage of poor people are the counterparts in component 1 of variables in the right side of the figure that represent education and economic activity. Figure 3 represents the scores of the municipalities in the main two components. Note that only seven municipalities in Bolivia are located in the upper right part of the figure where education and economic activity are the main features. These dots represent the provincial capitals Santa Cruz, La Paz, El Alto, Cochabamba, Oruro, Potosi, and Sucre.

Table 2. List of variables used for the development of the IWP and IEI indices.1

IWP	IEI
Net coverage of secondary education	Education buildings
Net coverage of primary education	Education units
Percentage of poor people	Rooms for teaching
Unsatisfied basic needs	Number of NGOs
Number school years in over 19s	Life expectancy at birth
Total acute malnutrition	Human development index
Total chronic malnutrition	Infant mortality rate
Percentage of total malnutrition	Inequality
Per capita consumption	Mean welfare indicator
Infant mortality rate	Per capita consumption
Net rate migration	Social investment
Population growth rate	Non social investment
Population in 2001	Inadequate housing material
Number of schooling years	Inadequate housing space
Number of enrolled students	Inadequate water and sanitation
Education index	Inadequate energy supply
School attendance rate	Inadequate education
Illiteracy	Inadequate health service
Students enrolled in primary and secondary	Number of financial institutions
Primary fulfillment rate	Census population in 2001
Secondary abandonment rate	
Primary abandonment rate	
Promotion rate	
Losses rate	
Human resources for teaching	
Population economically active	
Male index	
Rural index	
Poverty gap	
Severe poverty	
Widespread poverty	
Estimated population in 2007	

1 Data sources INE - Bolivia. http://www.ine.gov.bo/geoclip/launchmap.php. Censo 2001, Superintendencia de Bancos y Entidades Financieras -SBEF. 2008 http://www.sbef.gov.bo/Inf_Inst.php, Minsiterio de Salud y el Deporte, Sistema Nacional de Inoformación en Salud y vigilancia epidemiológica, Año 2007. http://www.sns.gov.bo/snis/

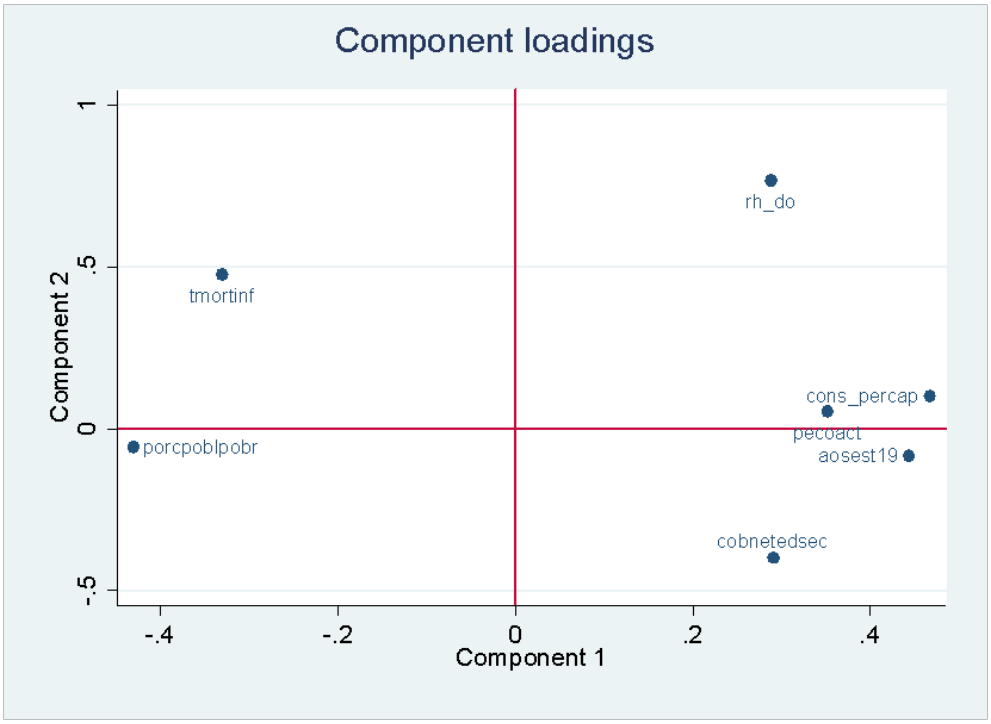

Figure 2. Loadings for two components depicted from a selection of socio-economic variables explaining the IWP index in Bolivia. (tmortinf = infant mortality; porcpoblpobr = percentage of poor people; rh_do = human resources for teaching; cosn_percap = consumption per capita; peoact = population economically active; aosest19 = number of years shooling in the population 19 years and older; cobnetedsec = net coverage of secondary education)

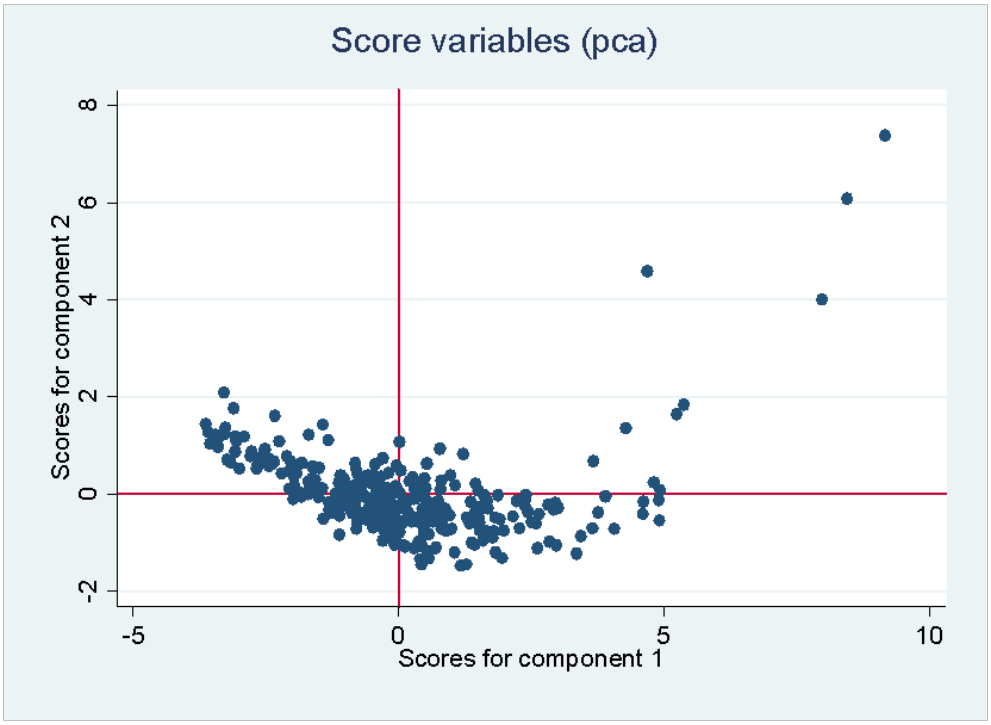

Figure 3. Municipality scores for the two main components of the PCA in the IWP index.

To explain further the meaning of each component, a cluster analysis procedure follows. This method makes possible the identification of individuals (municipalities in this case) that hold similar characteristics of any of the seven finally selected with the PCA. There are several testing rules available to define the appropriate number of clusters (Table 3). These rules, also called indices, are a guide in determining the number of clusters. In a hierarchical cluster analysis, larger values for the first two columns of Table 3 and smaller values for third column indicate more distinct clustering (Milligan and Cooper 1984).

Table 3. Indicators for the selection of the appropriatesize of clusters.

Number in cluster	Duda Je(2)/Je(1)	Hart pseudo T-squared	Calinski/ Harabasz pseudo-F
1	0.3066	703.43	
2	0.6810	143.80	703.43
3	0.5929	88.58	549.68
4	0.5816	126.61	554.02
5	0.3940	58.44	456.48
6	0.4800	112.69	593.86
7	0.4977	89.82	534.05
8	0.7110	14.23	481.90
9	0.2962	4.75	475.32
10	0.3315	18.15	648.19
11	0.4976	70.66	763.36
12	0.4428	46.56	739.47
13	0.4348	31.19	722.28
14	0.0000		727.06
15	0.1138	31.15	846.98

From the three tests shown in Table 3, it seems that a cluster size of 8 is about the optimum for the Bolivia data. Clusters (Figure 4) were then characterized for each of the variables as depicted in Appendix D. The IWP index was then calculated by the weighted sum of the components coefficients for each variable.

We calculated the IWP for Bolivia by the following formulation:

IWP = 0.239*net_coverage_secundary_education +
0.242* number_schooling_years_in_older19 +
0.243*per_capita_consumption +
0.289*human_resources_for_teaching +
0.201*population_economically_active –
0.215*percentage_poor_people –
0.070*infant_mortality_rate

This formulation was then applied to the individual municipalities and represented in Figure 5. Once the information was available in geographical format it was possible to combine this with other biophysical data also found in geographical format as part of EDA.

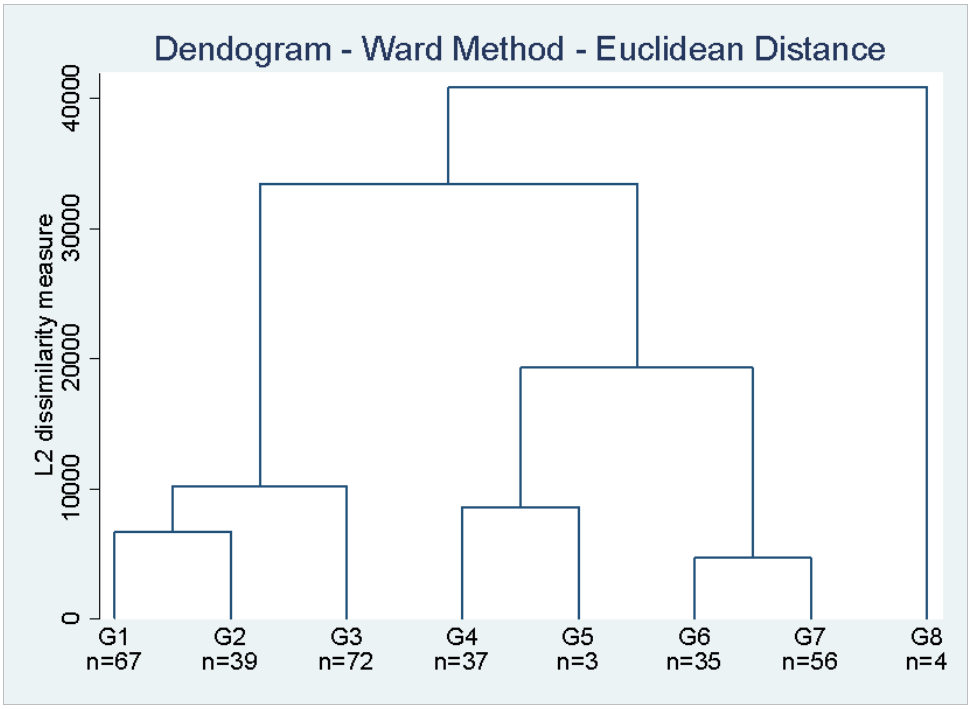

Figure 4. Cluster analysis for 313 municipalities in Bolivia for the identification of the IWP index.

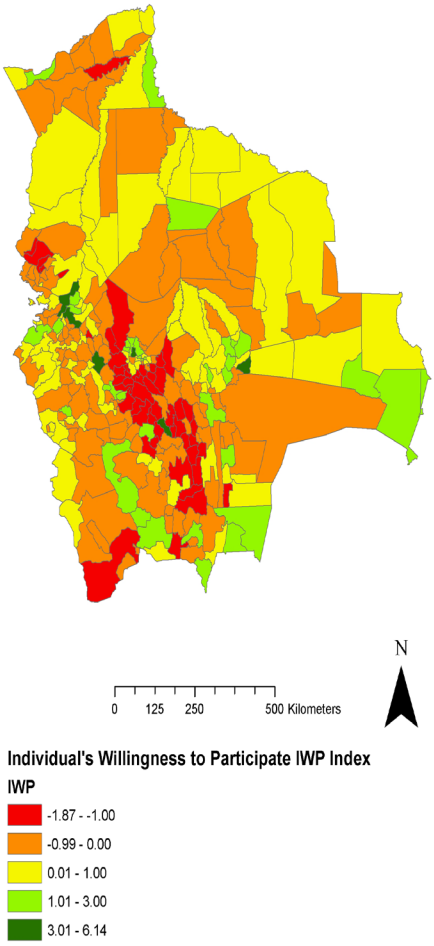

Figure 5. Individual Willingness to Participate Index (IWP) for Bolivia.

5.2 Institutional environment index (IEI)

We sought a wide range of variables to identify the institutional environment in each municipality. For Bolivia, these included the following 20 variables from which those italicized were selected using the same process we used to select variables for IWP: Education buildings and units, rooms for teaching, number and type of NGOs, life expectancy at birth, human development index, infant mortality rate, inequality, mean welfare indicator, per capita consumption, social investment, other investments, housing material, inadequate housing spaces, water and sanitation services, inadequate energy supply, inadequate heath attention, number of financial institutions, and population census.

As in the IWP index, only four components were needed to explain 93% of the variability in the model. Figure 6 shows the two main components and the loadings of each variable in the model. In the lower right, the variables related to economic welfare (human development index, mean welfare indicator, per capita consumption), in the upper left variables related to the basic services such as energy, water and housing material (housing material, water and sanitation services, inadequate energy supply), and in the upper right, variables related to social investment and financial infrastructure (education units, rooms for teaching, social investment, other investments, number of financial institutions).

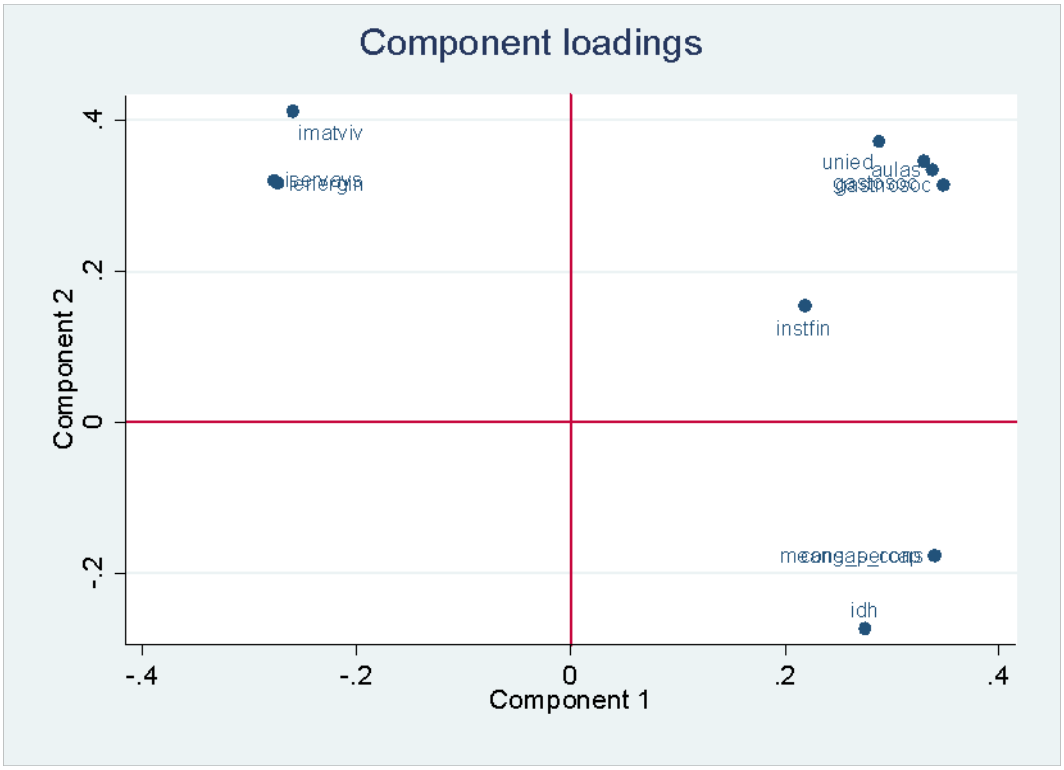

Figure 6. Loadings for two components depicted from a selection of socio economic and institutional variables explaining the IEI in Bolivia.

Figure 7 shows the distribution of municipalities and follows a similar pattern of scores as in the IWP index. This does not mean, however, that the IWP and the IEI indices are the same as we shall show when we plot both groups together. When grouped in clusters and after the application of the tests for the definition of the number of clusters, again we defined eight groups. Figure 8 presents these clusters, which are also characterized in Appendix E. As in the IWP index, the IEI index was calculated by the weighted sum of the components coefficients for each variable. The IEI for Bolivia was calculated by the following formulation:

IEI = (education_units + rooms_for_teaching + HDI + mean_welfare_indicator + per_capita_ consumption + social_investment + other_investments + number_of_financial_institutions)/5

Figure 9 shows the distribution of the IEI index across Bolivia.

5.3 Joint Analysis of the IWP and IEI

The IWP and IEI scores obtained for each municipality in Bolivia were plotted to identify their distribution (Figure 10). Table 4 presents the basic statistics for the combinations that when qualitatively organized produced seven groups, also represented spatially in Figure 11. There were 14 cases in which both indicators scored highly while 139 municipalities scored low in both indices. Only 10 municipalities scored medium in both indicators, while 111 municipalities were located in combinations of low and medium indices. There were some cases in which higher scores in one index had low scores in the other. This is a particular case in which it is possible to find high values in the characteristics of individuals combined with poor institutional conditions. There some instances of the opposite situation, with good institutions, but poor individual characteristics.

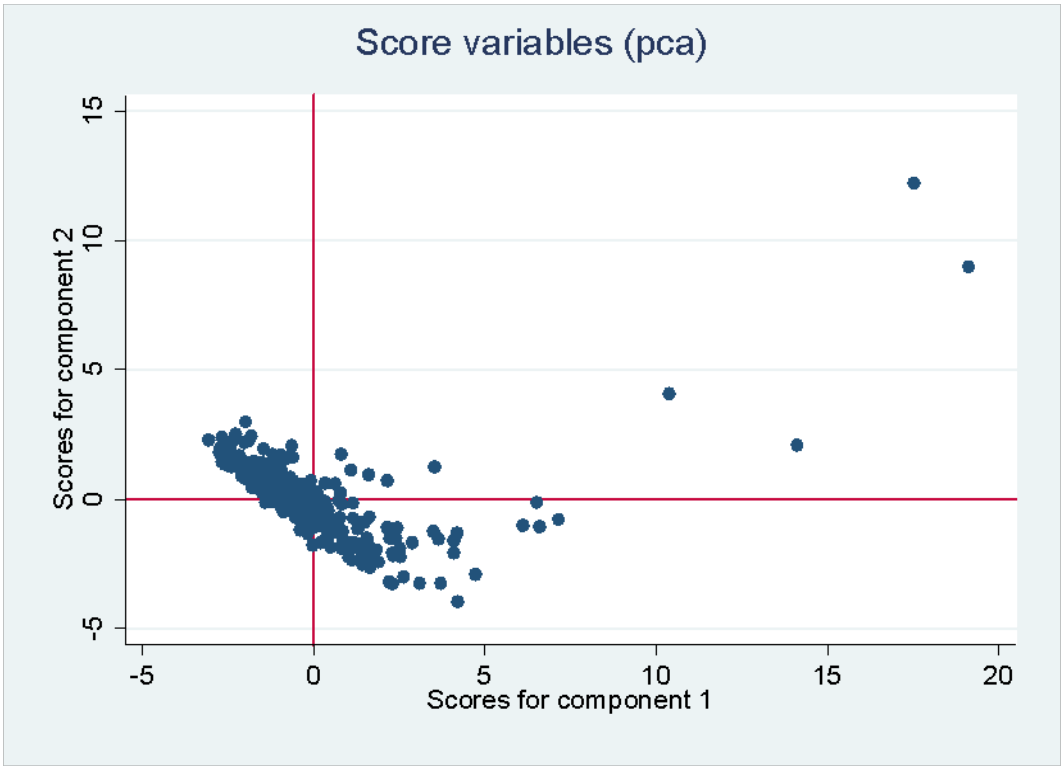

Figure 7. Municipality scores for the two main components of the PCA in the IEI index.

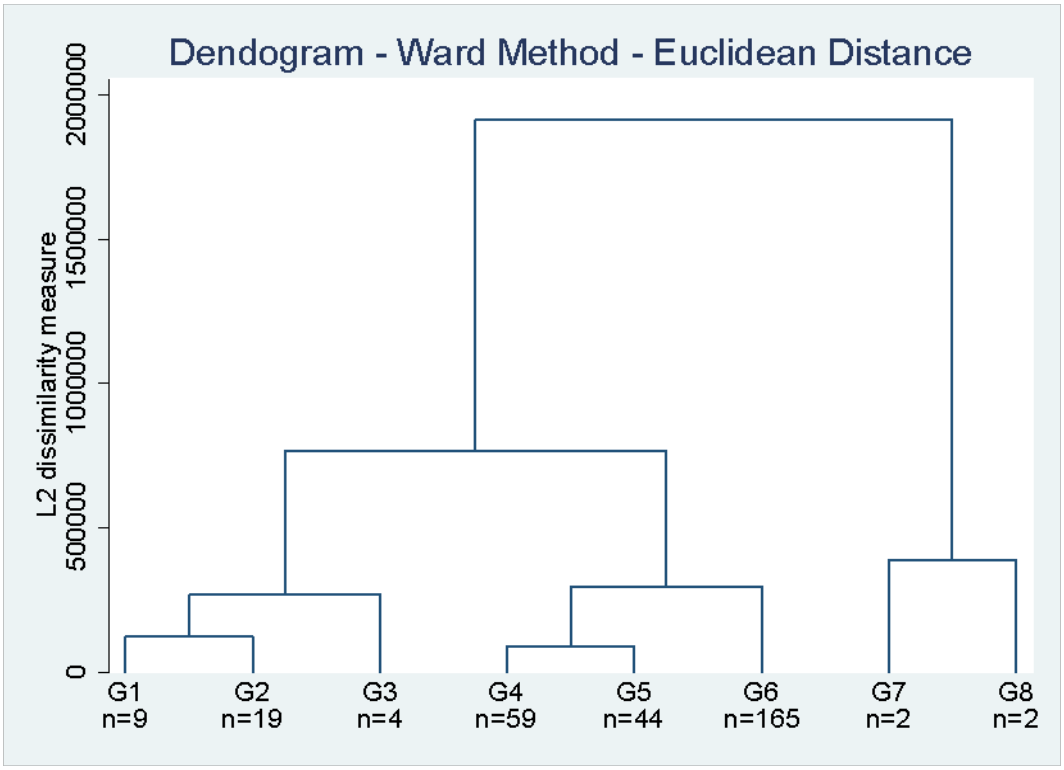

Figure 8. Cluster Analysis for 304 municipalities in Bolivia for the identification of the IEI index.

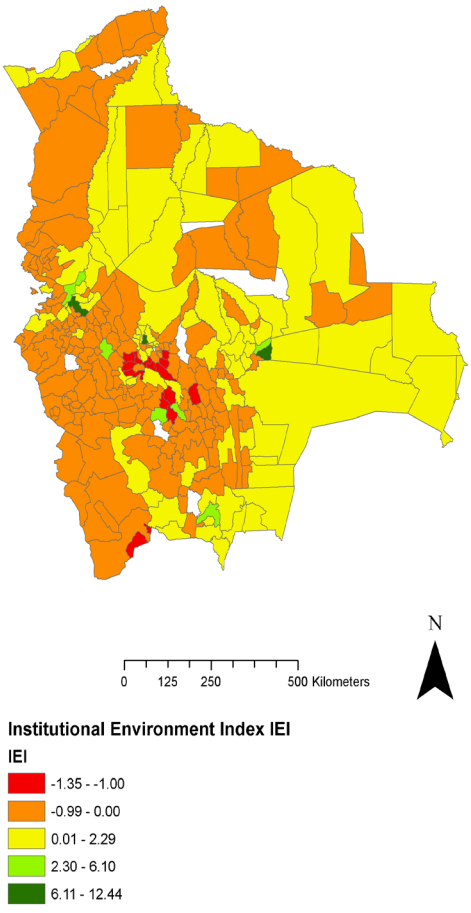

Figure 9. Institutional environment index for Bolivia.

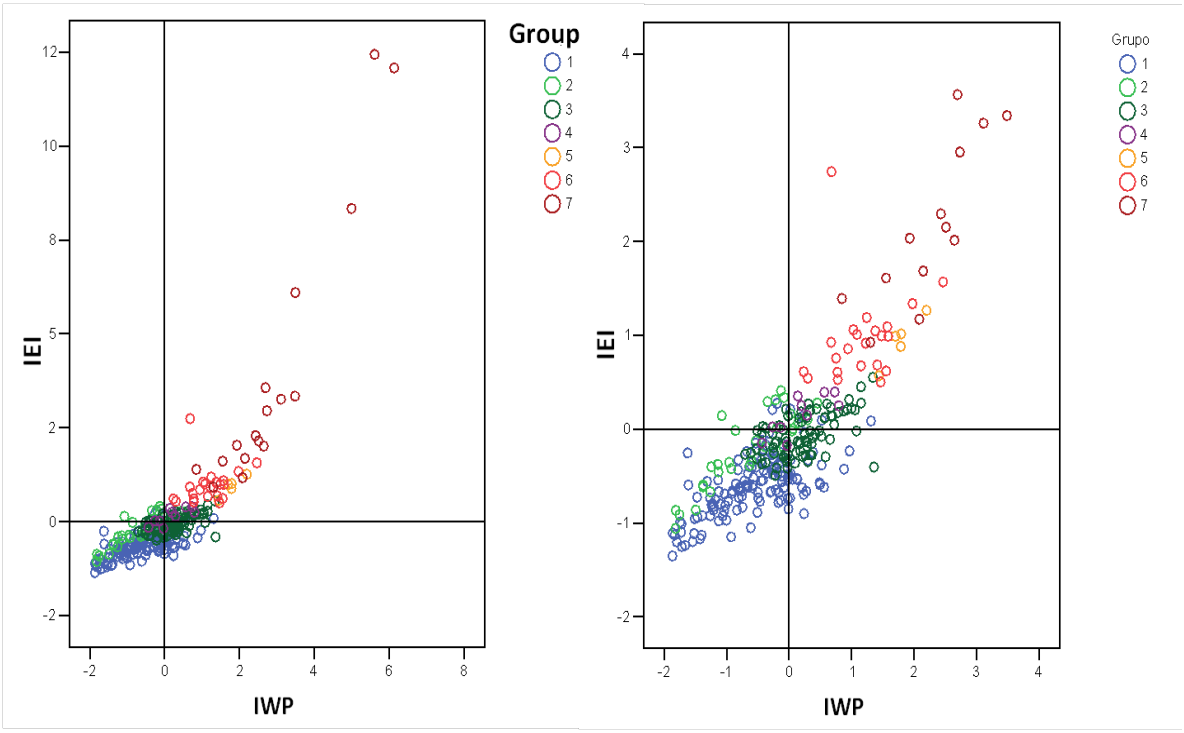

Figure 10. Scores obtained by Bolivian municipalities in the IWP and IEI indices. Coloring groups show the seven combinations found when values are grouped in high, medium and low.

Sensitivity analysis of variables used for the IWP index showed that the variable with the greatest effect on the index is population_economically_active. This variable had an effect up to six-fold greater than the second-most influential variable: the net coverage of education. Within the variables used in the IEI index welfare, and per capita consumption had twice the influence of the next variable, education units.

Table 4. Frequencies of municipalities belonging to each of the IWP and IEI clusters in Bolivia. Colors represents a qualitative classification of clusters in high (green), medium (yellow) and low (blue) values of the IWP and the IEI indices.

IWP Clusters	IEI Clusters								Total
	4	6	2	5	1	3	7	8	
1	6	53		3					62
	9.68	85.48		4.84					100
	10.17	32.12		6.82					20.39
	1.97	17.43		0.99					20.39
2	13	6	5	15					39
	33.33	15.38	12.82	38.46					100
	22.03	3.64	26.32	34.09					12.83
	4.28	1.97	1.64	4.93					12.83
3	14	47		8					69
	20.29	68.12		11.59					100
	23.73	28.48		18.18					22.7
	4.61	15.46		2.63					22.7
6	10	20		5					35
	28.57	57.14		14.29					100
	16.95	12.12		11.36					11.51
	3.29	6.58		1.64					11.51
7	12	38	2	3					55
	21.82	69.09	3.64	5.45					100
	20.34	23.03	10.53	6.82					18.09
	3.95	12.5	0.66	0.99					18.09
4	4	1	12	10	9	1			37
	10.81	2.7	32.43	27.03	24.32	2.7			100
	6.78	0.61	63.16	22.73	100	25			12.17
	1.32	0.33	3.95	3.29	2.96	0.33			12.17
5						3			3
						100			100
						75			0.99
						0.99			0.99
8							2	2	4
							50	50	100
							100	100	1.32
							0.66	0.66	1.32
Total	59	165	19	44	9	4	2	2	304
	19.41	54.28	6.25	14.47	2.96	1.32	0.66	0.66	100
	100	100	100	100	100	100	100	100	100
	19.41	54.28	6.25	14.47	2.96	1.32	0.66	0.66	100

The four values in each cell are (1) frequency; (2) the percentage for the IWP (row) cluster; (3) the percentage for the IEI (column) cluster; and (4) the percentage for the row and column clusters taken together.

6. CONCLUSIONS

We learned many lessons during the process of advancing in the spatial representation of the two objective variables we used here, the IWP and the IEI. We focused mainly in developing the methodology: the approach, the materials used and in the procedures. With regard to the findings in this Bolivian pilot case we can present a few general comments.

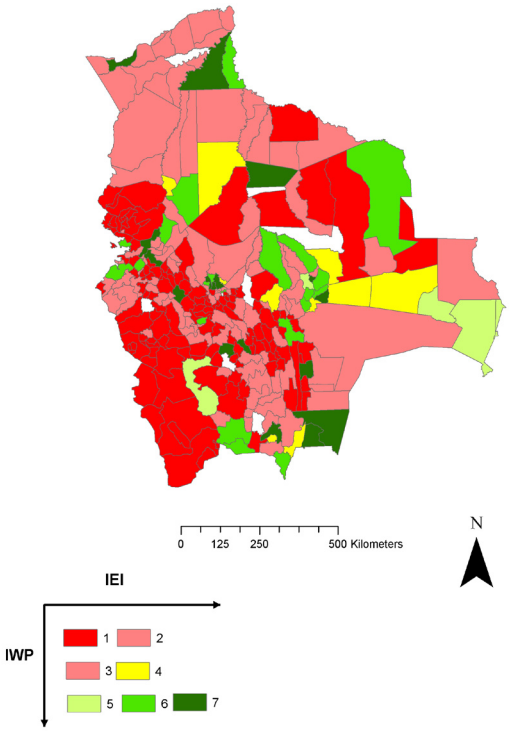

Figure 11. Spatial representation of the seven clustered groups of municipalities accordingly to their scores in the IWP and IEI indices.

There is no doubt that by systematic use of detailed and reliable data is it possible to discover more than any one single variable can express. The method we present here shows that complex characteristics such as the institutional environment and individuals' willingness to participate can be estimated using some of their constituent variables and surrogate variables in a systematic and transparent way. The individual contribution of each variable can also be identified, which allows for targeting potential interventions in specific characteristics that will modify current conditions more efficiently.

The location of the variables along each of the main four components for both indices followed a logical trajectory. We expected to find better institutional environments where the human condition of individuals was also better. The method identified both situations in general, but also particular cases in which there are extreme conditions, e.g. provincial capitals compared with remote municipalities.

We conclude that the data available from statistics offices, Ministries and other social organizations, after careful scrutiny, can be used with confidence. It is advisable to get the same variable from alternative sources to check inconsistencies and solve contradictory information. We believe that a database approach to data management is the best way to prepare the data before undertaking the analytical procedures.

The two statistical techniques that we used here, PCA and cluster analysis, are simple and informative enough for the development of the two indicators. They allow the discrimination of more sensitive variables, simplification of models, and identification of scores for individual municipalities. The procedures are available in many statistical packages and are easy to reproduce.

In terms of the findings on the Bolivian case, it is a country in which living in a provincial capital makes a big difference. These centers concentrate most investment, resources and have the healthiest and more highly-qualified people. Rural areas are in the other extreme with poor conditions in the terms of institutional capacities and individuals' human condition.

7. REFERENCES

Bouman, Bas, Simon Cook, Boru Douthwaite, Claudia Ringler, Jorge Rubiano, and Tingju Zhu 2007. *Impact Potential of the Temperate and Tropical Aerobic Rice (STAR) in Asia.* Internal document prepared by the CPWF Impact Project for the External Review team.
Eggertsson, T. 1996. A note on the economic of institutions. In Lee J. Alston, Thrainn Eggertsson, and Douglas C. North ed. Pp 6-21. Quoted in Saleth and Dinar 2004.
Jones, P. G., W. Diaz, and J. H. Cock. , 2005. *Homologue: A computer System for Identifying Similar Environments throughout the Tropical World. Version Beta a.0.* Cali, Colombia: CIAT, Colombia.
Milligan, G. W., and M. C. Cooper 1984. An examination of procedures for determining the number of clusters in a dataset. *Psychometrika* 50, 159-179.
National Research Council (NRC) 2007. *Putting People on the Map: Protecting Confidentiality with Linked Social-Spatial Data.* Ed. M. P. Gutman, and P. C. Stern. Panel on Confidentiality Issues Arising from the Integration of Remotely Sensed and Self-Identifying Data, Washington D.C.: The National Academies Press.
Otero, M. F., J. Rubiano, V. Soto, and G. Lema 2006. Using similarity analyses to scaling out research. *Water International* 31, 376–386.
Pardey, P., J. James, S. J. Alston, B. S. Wood, E. B. Koo, T E. Binenbaum, T. Hurley, and P. Glewwe 2007. *Science, Technology and Skills.* CGIAR Science Council, International Science and Technology Practice and Policy (InSTePP) Center.
Rubiano, J., S. Cook, and B. Douthwaite 2008a. Adapting to change: How to accelerate impact. *Proceedings of the CGIAR Challenge Program on Water and Food 2nd International Forum on Water and Food, Addis Ababa, Ethiopia, November 10—14, 2008.* Colombo, Sri Lanka: The CGIAR Challenge Program on Water and Food. 183pp.
Rubiano, J., A. Peralta, and N. Johnson 2008b. Scaling-up in watersehd management research projects. In *Proceedings of the CGIAR Challenge Program on Water and Food Second International Forum on water and Food Vol. 1.* , Colombo, Sri Lanka: The CGIAR Challenge Program on Water and Food.
Saleth, R. M., and A. Dinar 2004. *The Institutional Economics of Water. A Cross-Country Analysis of Institutions and Performance.* Washington, DC: The International Bank for Reconstruction and Development/The World Bank.
Williamson, O. E. 1993. Transactions costs economics and organizations theory. *Industrial and Corporate Change*, 2, 107-56. Quoted in Saleth and Dinar 2004.
The World Bank 2002. *Building Institutions for Markets. World Development Report 22825.* Oxford: Oxford University Press.

APPENDIX A

Table A.1. Correlation matrix of variables used in the development of the IWP index.

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	cobnetedsec	1.00																																	
2	codnetedpri	0.41	1.00																																
3	porcpoblpobr	-0.43	-0.16	1.00																															
4	nbi	-0.44	-0.15	0.99	1.00																														
5	añosest19+	0.62	-0.02	-0.70	-0.69	1.00																													
6	iaulas	0.52	0.48	0.06	0.07	0.13	1.00																												
7	irh_do	0.62	0.55	0.02	0.03	0.21	0.86	1.00																											
8	ag_to	0.17	0.07	-0.48	-0.50	0.36	-0.18	-0.17	1.00																										
9	crgl_t	0.16	0.10	-0.21	-0.22	0.23	-0.07	-0.04	0.55	1.00																									
10	pdesnutot	-0.19	0.04	0.10	0.10	-0.26	0.01	0.02	0.10	-0.02	1.00																								
11	cons_percap	0.49	0.12	-0.90	-0.90	0.80	0.00	0.05	0.53	0.27	-0.14	1.00																							
12	tmortinf	-0.39	-0.06	0.62	0.61	-0.65	-0.11	-0.13	-0.23	-0.04	0.11	-0.68	1.00																						
13	tnetmigr	0.02	-0.01	-0.42	-0.37	0.38	-0.20	-0.13	0.16	0.01	-0.10	0.40	-0.35	1.00																					
14	tcrecpobl	-0.16	-0.44	-0.09	-0.08	0.32	-0.37	-0.37	0.08	0.01	-0.33	0.13	-0.19	0.44	1.00																				
15	pob2001	0.16	0.03	-0.41	-0.42	0.33	-0.13	-0.13	0.85	0.64	-0.10	0.49	-0.21	0.14	0.11	1.00																			
16	aos_escol	0.62	-0.01	-0.70	-0.68	0.99	0.14	0.21	0.36	0.23	-0.27	0.80	-0.63	0.38	0.33	0.33	1.00																		
17	alummatr	0.20	0.07	-0.43	-0.44	0.35	-0.12	-0.11	0.87	0.66	-0.11	0.51	-0.23	0.13	0.09	1.00	0.36	1.00																	
18	ind_educ	0.72	0.28	-0.65	-0.64	0.91	0.31	0.40	0.27	0.19	-0.27	0.77	-0.65	0.37	0.13	0.26	0.91	0.30	1.00																
19	tasitesco	0.69	0.22	-0.27	-0.30	0.54	0.40	0.49	0.07	0.13	-0.16	0.37	-0.30	-0.08	-0.03	0.09	0.54	0.11	0.63	1.00															
20	tanalf	-0.55	0.01	0.54	0.51	-0.90	-0.17	-0.25	-0.15	-0.11	0.32	-0.66	0.66	-0.44	-0.33	-0.18	-0.89	-0.19	-0.92	-0.52	1.00														
21	matr_neta_~s	0.63	0.82	-0.42	-0.42	0.27	0.47	0.56	0.24	0.18	0.03	0.41	-0.28	0.02	-0.47	0.19	0.27	0.23	0.55	0.48	-0.22	1.00													
22	ttermiprima	0.88	0.53	-0.40	-0.40	0.53	0.49	0.64	0.13	0.16	-0.12	0.43	-0.35	0.05	-0.19	0.12	0.53	0.15	0.69	0.66	-0.52	0.68	1.00												
23	tabandonosec	0.10	0.10	0.15	0.15	-0.05	0.02	0.10	0.00	-0.01	-0.09	-0.10	0.15	0.01	0.00	0.01	-0.05	0.01	-0.03	-0.03	0.05	0.04	0.13	1.00											
24	tabandonopro	0.25	0.02	0.14	0.13	0.14	0.17	0.24	-0.10	-0.02	-0.17	-0.06	0.13	0.03	0.17	-0.05	0.15	-0.06	0.16	0.18	-0.20	-0.02	0.33	0.53	1.00										
25	tpromocion	-0.04	0.27	0.00	0.00	-0.11	0.12	0.08	-0.01	0.01	-0.10	-0.02	0.07	-0.08	0.05	0.00	-0.06	0.01	-0.12	-0.15	0.14	-0.07	-0.04	-0.16	-0.32	1.00									
26	treproba	0.15	0.28	-0.14	-0.15	-0.02	0.07	0.02	0.27	0.12	0.16	0.10	0.00	-0.17	-0.19	0.15	0.00	0.16	0.02	0.09	0.15	0.30	0.16	-0.01	0.01	-0.07	1.00								
27	rh_do	0.24	0.10	-0.46	-0.47	0.40	-0.09	-0.08	0.87	0.67	-0.11	0.54	-0.24	0.11	0.07	0.97	0.40	0.99	0.34	0.14	-0.21	0.27	0.19	0.01	-0.06	0.01	0.18	1.00							
28	pecoact	0.49	-0.09	-0.61	-0.59	0.78	0.02	0.03	0.26	0.16	-0.30	0.63	-0.41	0.40	0.35	0.27	0.78	0.28	0.69	0.40	-0.72	0.12	0.44	-0.02	0.27	-0.10	-0.06	0.31	1.00						
29	imaculinid	-0.04	0.01	0.03	0.10	0.19	0.02	0.07	-0.18	-0.13	0.00	0.02	-0.20	0.51	0.23	-0.13	0.19	-0.14	0.26	-0.19	-0.38	-0.07	0.03	0.01	0.17	-0.09	-0.23	-0.17	0.19	1.00					
30	irural	-0.11	-0.10	0.17	0.18	-0.10	0.04	0.10	-0.26	-0.15	-0.03	-0.26	0.07	-0.02	-0.01	-0.18	-0.11	-0.20	-0.12	0.02	0.05	-0.16	-0.03	-0.04	0.11	-0.09	-0.12	-0.23	-0.06	0.06	1.00				
31	povgapi	-0.54	-0.12	0.73	0.72	-0.79	-0.12	-0.21	-0.30	-0.16	0.20	-0.88	0.70	-0.42	-0.13	-0.27	-0.81	-0.30	-0.82	-0.42	0.76	-0.37	-0.46	0.10	-0.01	0.04	-0.03	-0.33	-0.65	-0.17	0.22	1.00			
32	povsever	-0.55	-0.11	0.65	0.64	-0.78	-0.17	-0.25	-0.23	-0.13	0.20	-0.82	0.70	-0.40	-0.14	-0.22	-0.79	-0.24	-0.82	-0.46	0.78	-0.34	-0.48	0.11	-0.05	0.06	0.00	-0.27	-0.64	-0.20	0.18	0.99	1.00		
33	povdisp	-0.48	-0.14	0.88	0.87	-0.77	0.00	-0.07	-0.47	-0.23	0.14	-0.97	0.64	-0.43	-0.11	-0.41	-0.78	-0.43	-0.75	-0.33	0.63	-0.41	-0.40	0.09	0.06	0.00	-0.10	-0.47	-0.62	-0.05	0.26	0.92	0.85	1.00	
34	pob2007	0.15	0.03	-0.41	-0.42	0.32	-0.13	-0.13	0.85	0.63	-0.10	0.49	-0.22	0.15	0.12	1.00	0.33	0.99	0.26	0.09	-0.18	0.18	0.12	0.01	-0.05	0.00	0.14	0.97	0.27	-0.12	-0.18	-0.27	-0.22	-0.41	1.00

Table A.2. Key to the variables used in Table A.1.

Variable	Meaning	Variable	Meaning
cobnetedsec	Net coverage of secondary education	ind_educ	Education index
codnetedpri	Net coverage of primary education	tasitesco	Rate of school attendance
porcpoblpobr	Percentage of poor population	tanalf	Iliteracy tate
nbi	Basic insatisfied needs	matr_neta_~s	Net rate of enrolled students
añosest19+	Number of school years age > 19	ttermiprima	Rate of primary school completion
iaulas	School rooms index	tabandonosec	Rate of secondary school abandonment
irh_do	Teachers resource index	tabandonopro	Rate of primary school abandonment
ag_to	Acute malnutrition	tpromocion	Rate of school promotion
crgl_t	Cronic total malnutrition	treproba	Rate of school failure
pdesnutot	Percentage of total malnutrition	rh_do	Human resource index
cons_percap	Per capita compsumtion	pecoact	Economic active population
tmortinf	Infant mortality rate	imaculinid	Male index
tnetmigr	Net rate of migration	irural	Rural index
tcrecpobl	Population growth rate	povgapi	Poverty gap index
pob2001	Population in 2001	povsever	Severe poverty
aos_escol	Number of school years	povdisp	Poverty disparity index
alummatr	Number of enrolled students	pob2007	Population in 2007

APPENDIX B

Table B.1. Principal components correlations and eigenvectors for 313 municipalities in Bolivia in the IWP index.

Principal components/correlation				
Number of observations		313		
Number of components		7		
Trace		7		
Rotation: (unrotated = principal) Rho		1.0000		
Component	Eigenvalue	Difference	Proportion	Cumulative
1	3.916	3.0760	0.559	**0.5594**
2	0.8400	0.0674	0.120	**0.6794**
3	0.7727	0.0703	0.110	**0.7898**
4	0.7023	0.2793	0.100	**0.8902**
5	0.4230	0.1800	0.060	0.9506
6	0.2430	0.1402	0.035	0.9853
7	0.1029	.	0.015	1

Table B.2. Principal components (eigenvectors) for 313 municipalities in Bolivia in the IWP index.

Principal components (eigenvectors)							
Variable	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7
Coverage of secondary education	0.2902	-0.4002	0.5178	0.6745	-0.0892	-0.1468	0.0543
Percentage of poor people	-0.4308	-0.0557	0.2576	0.0468	0.6483	0.2234	0.5221
Number of school years age > 19	0.4434	-0.0832	0.1969	-0.1793	0.1296	0.8095	-0.2314
Per capita consumption	0.4669	0.0999	-0.2292	-0.0491	-0.2579	0.0704	0.8035
Infant mortality rate	-0.3308	0.4763	0.5608	-0.0422	-0.546	0.1997	0.0976
Human resources for teaching	0.2879	0.7682	-0.0433	0.4084	0.3707	-0.0728	-0.1249
Population economically active	0.3513	0.0534	0.5078	-0.5828	0.23	-0.4724	0.0137

Table B.3. Kaiser-Meyer-Olkin test for adequacy for 313 municipalities in Bolivia in the IWP index.

Kaiser-Meyer-Olkin measure of sampling adequacy	
Variable	kmo
Coverage of secondary education	0.8822
Percentage of poor people	0.7925
Number of school yearsage > 19	0.8212
Per capita consumption	0.7391
Infant mortality rate	0.8568
Human Resources for Teaching	0.8186
Population economically active	0.8271
Overall	0.8032

APPENDIX C

Table C.1. Cluster and PCA coefficients for the identification of the IEI index.

Cluster	Coverage secondary education	% poor people	Number of school years age > 19	Per capita consump-tion	Infant mortality rate	Human resources for teaching	Population economically active	pc1	pc2	pc3	pc4	idm	Freq.	Percent	Cum.
1	31.0	91.1	5.1	646	79.4	90	53.0	-0.483	-0.246	0.392	-0.061	-0.26	67	21.4	21.41
2	34.0	86.5	4.6	734	78.7	313	51.3	-0.395	-0.059	0.047	0.447	-0.18	39	12.5	33.87
3	15.0	94.6	3.6	441	97.8	99	49.3	-2.085	0.468	-0.052	0.049	-1.11	72	23.0	56.87
4	43.2	51.1	7.1	1336	58.2	520	55.3	2.790	-0.284	-0.328	-0.084	1.48	37	11.8	68.69
5	57.1	47.3	9.4	1564	66.2	2491	59.2	4.962	1.612	0.700	0.651	3.11	3	1.0	69.65
6	33.9	69.7	5.8	1053	62.6	143	53.8	1.158	-0.467	-0.340	-0.202	0.53	35	11.2	80.83
7	28.0	83.8	5.5	854	72.5	89	53.4	0.161	-0.266	-0.021	-0.299	0.03	56	17.9	98.72
8	48.8	37.4	9.6	2067	53.4	6221	61.0	7.563	5.513	-0.296	2.014	5.06	4	1.3	100

Colours indicate if values are high, medium or low (green, yellow and blue respectively) with respect to the mean of the total of clusters.

APPENDIX D

Table D.1. Cluster and PCA coefficients for the identification of the IWP index.

Cluster	Educa-tion units	Educa-tion rooms	HDI	Mean welfare indica-tor	Per capita consum ption	Social invest-ment	Non social invest-ment	Inad equate housing spaces	Inad-equate water sanit-ation	Inad-equate energy supply	No of financial instit-utions	pc1	pc2	pc3	pc4	IAI	Freq.	%	Cum %
1	70.6	561	0.67	316	1537	24260	21360	40.9	64.8	52.3	5.9	3.474	-1.484	1.101	-0.051	1.70	9	3.0	3.0
2	80.5	384	0.60	219	1060	13660	8870	50.9	68.5	64.1	10.5	1.652	-0.617	0.000	0.263	0.79	19	6.2	9.2
3	188.8	1765	0.68	312	1518	58550	45230	20.4	36.8	23.7	8.5	6.586	-0.752	-0.737	-1.212	3.28	4	1.3	10.5
4	51.7	175	0.54	156	762	3330	2770	67.1	78.3	79.7	1.1	-0.290	0.011	-0.067	-0.078	-0.17	59	19.4	29.9
5	65.1	245	0.54	168	814	6030	4590	64.9	75.3	77.0	1.3	0.082	0.000	-0.139	-0.184	0.02	44	14.5	44.4
6	22.8	78	0.54	145	703	1200	1040	70.6	81.4	84.5	0.4	-0.828	0.001	0.039	0.076	-0.44	165	54.3	98.7
7	288.5	4189	0.69	384	1866	17480	104220	14.3	35.2	9.6	53.0	12.252	3.085	-1.520	0.341	7.22	2	0.7	99.3
8	567.0	6415	0.73	466	2268	374630	185250	52.2	49.8	44.4	74.5	18.313	10.597	-0.174	-0.124	12.26	2	0.7	100

Colours indicate if values are high, medium or low (green, yellow and blue respectively) with respect to the mean of the total of clusters.