Inter-American Development Bank
Banco Interamericano de Desarrollo
Latin American Research Network
Red de Centros de Investigación
Research Network Working paper #R-404

Does Geography Explain Differences in Economic Growth in Peru?

by

Javier Escobal
Máximo Torero

July 2000

Escobal D'Angelo, Javier.
Does geography explain differences in economic growth in Peru? / by Javier Escobal, Máximo Torero.

p. cm. (Research Network Working papers ; R-404)
Includes bibliographical references.

1. Peru--Economic conditions--1980- 2. Economic geography. 3. Economic development--Effect of geography on. 4. Cost and standard of living--Peru--Effect of geography on. I. Torero Cullen, Máximo. II. Inter-American Development Bank. Research Dept. III. Title. IV. Series.

330.985 E746--dc21

82000
Inter-American Development Bank
1300 New York Avenue, N.W.
Washington, D.C. 20577

Abstract

In Peru, a country with an astonishing variety of different ecological areas, including 84 different climate zones and landscapes, with rainforests, high mountain ranges and dry deserts, the geographical context may not be all that matters, but it could be very significant in explaining regional variations in income and welfare. The major question this paper tries to answer is: what role do geographic variables, both natural and manmade, play in explaining per capita expenditure differentials across regions within Peru? How have these influences changed over time, through what channels have they been transmitted, and has access to private and public assets compensated for the effects of an adverse geography?

We have shown that what seem to be sizable geographic differences in living standards in Peru can be almost fully explained when one takes into account the spatial concentration of households with readily observable non-geographic characteristics, in particular public and private assets. In other words, the same observationally equivalent household has a similar expenditure level in one place as another with different geographic characteristics such as altitude or temperature. This does not mean, however that geography is not important but that its influence on expenditure level and growth differential comes about through a spatially uneven provision of public infrastructure. Furthermore, when we measured the expected gain (or loss) in consumption from living in one geographic region (i.e., coast) as opposed to living in another (i.e., highlands), we found that most of the difference in log per-capita expenditure between the highland and the coast can be accounted for by the differences in infrastructure endowments and private assets. This could be an indication that the availability of infrastructure could be limited by the geography and therefore the more adverse geographic regions are the ones with less access to public infrastructure.

It is important to note that there appear to be non-geographic, spatially correlated, omitted variables that need to be taken into account in our expenditure growth model. Therefore policy programs that use regional targeting do have a rationale even if geographic variables do not explain the bulk of the difference in regional growth, once we have taken into account differentials in access to private and public assets.

***JEL** classification: D91, R11, Q12*

(*) Both authors are principal researchers at the Grupo de Análisis para el Desarrollo (GRADE). Send correspondence to jescobal@grade.org.pe or mtorero@grade.org.pe .

This research was supported by the IDB Research Network. We are greatly indebted to John Gallup, Eduardo Lora, Alejandro Gaviria and colleagues of the Geography and Development project for numerous helpful comments on the different drafts of this work. Their comments and criticism improved it substantially. We are also grateful to Jorge Aguero for excellent research assistance on this project. All remaining errors are our own.

Table of Contents

List of Tables

List of Maps

List of Graphs

List of Figures

1. Introduction

In *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*, Landes (1998) argues that Europe's temperate climate encouraged hard work and capitalist development, while the heat of the tropics brought reliance on slaves (Eichengreen, 1998). Engerman and Sokoloff (1998), trying to explain why the United States and Canada have been so much more successful over time than other New World economies, suggest that the roots of these disparities in the extent of inequality lay in differences in the initial factor endowments of the respective colonies. Why do we see areas with persistently low living standards, even in growing economies? Will the legacy of these differences persist?

One view is that differences arise from persistent spatial concentrations of individuals with personal attributes inhibiting growth in their living standards. This view does not ascribe a causal role to geography per se; in other words, identical individuals will, by this view, have the same growth prospects regardless of where they live. Alternatively, one might argue that geography has a causal role in determining how household welfare evolves over time. By this view, geographic externalities arising from natural geographic characteristics, local public assets, or local endowments of private assets, entail that living in a well-endowed area means that a poor household can eventually escape poverty. Yet an otherwise identical household living in a poor area experiences stagnation or decline. If this is so, then it is important for policy to understand what geographic factors do matter to growth prospects at the micro level (Jalan and Ravallion, 1998; Engerman and Sokoloff, 1998).

Peru has an astonishing variety of ecological areas. Only a few countries offer so many climate zones and landscapes, with rainforests, high mountain ranges and dry deserts. Peru contains a total of 84 of the world's 104 known living ecological regions and 28 different climates. This geographic diversity, its link to development, and the important differences in the welfare of the different regions makes Peru a good case study in attempting to ascertain what role geographic variables, both natural and manmade, play in explaining per capita expenditure differentials across regions within Peru.

As shown in Table 1, when comparing the income per capita and consumption per capita differences between the diverse regions of the country, it is clear that Peru has one of the highest degrees of inequality between regions in Latin America. According to Fallon (1998) and our own estimates based on the Peruvian Living Standard Measurement Study (LSMS) of 1997, Peru has a larger dispersion of per capita income by region than Colombia, Brazil Chile or Mexico. Only Argentina is reported as having larger regional income disparities. Furthermore, this dispersion is also very large within the different geographical regions of Peru.

This paper attempts to understand whether geographic externalities arising from natural geographic characteristics have a causal role in determining how household welfare evolves. The paper is divided into six major sections. The second section gives a detailed description of Peru's geography and specifically the main areas in which geography might play a fundamental role in economic development. It also makes a first attempt to analyze whether there is a correlation between geographic variables and earning levels. Additionally, it analyzes whether the differences observed across the different regions in Peru are also correlated to the changes in geography and therefore to geographic externalities. In the third section we try to formally answer whether geography is a determinant of the evolution of welfare across households over time. We develop a model of consumption and consumption growth at the household and province level, respectively. This model not only takes in the local effect of geographic variables but also includes *spatial* *econometric techniques* to ascertain the presence of persistent spatial concentrations forced by geography. In addition, we also analyze whether the presence of positive geographic externalities arising from local public assets or local endowments of private assets implies that the effect of natural geographic characteristics can be overcome and therefore a poor household can eventually escape poverty. To be able to analyze the partial effects of each of these types of assets (geographic, private and public assets) we also develop a methodology to break down the partial effects of each of these variables.

Section Four details the main databases constructed for this paper and the methodological issues regarding the databases. We use the national censuses for 1972, 1981 and 1993, the LSMS surveys for 1991, 1994, 1996, 1997, information from the district infrastructure census, geographical data-sets, and information from the third National Agrarian Census of 1994. In Section Five the results are presented and in Section Six we detail the major conclusions of the study.

Table 1
Regional Income Per Capita Dispersion in Latin American Countries
(*Selected Years*)

	YEAR	DISPERSION
Colombia	1989	0.358
Brazil	1994	0.424
Chile	1994	0.470
Mexico	1993	0.502
Peru	1997	0.561
Argentina	1995	0.736

(1): Unweighted coefficient of Variation

Source: World Bank (1999) and LSMS 1997

2. Basic Characteristics of Peruvian Geography

Leading historians and economists have long recognized geography as having a crucial role in economic development, even though geography has been neglected in most recent empirical studies of comparative growth across countries and of comparative growth within countries.[1]

Specifically, in the case of Peru the enormous diversity of its geography makes it an extremely interesting case study for analyzing the importance of these variables to economic growth within the country.[2] Peru is located in the Tropical Zone of the globe, but because of variations in relief (particularly elevation, as shown in Map 1) and such factors as rain shadows, bodies of water (i.e., marine currents such as “El Niño” and the Humboldt Current) and wind patterns, the country comprises a multitude of microclimates. Although many geographic factors interact, it can be said that throughout most of Peru the orography and the morphologic structure of the Andes has

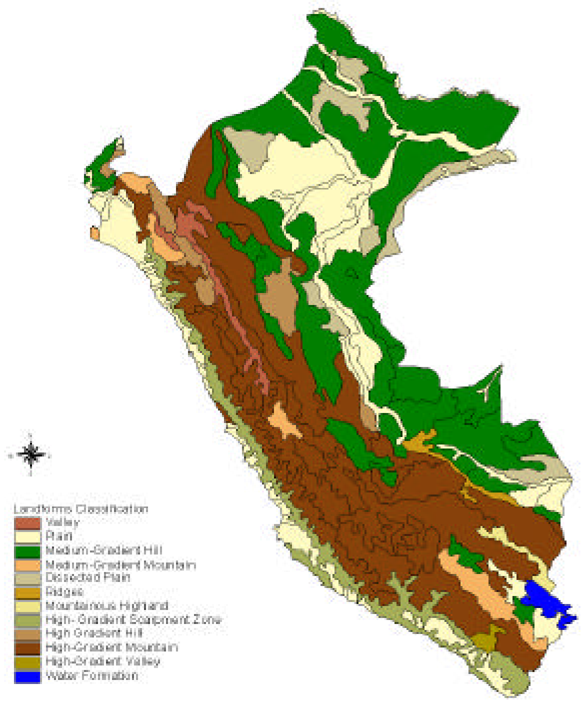
[1] There are few studies estimating the economic importance of geography within a region or a country, though Bloom and Sachs (1998), for example, make a great contribution for the case of Africa and Engerman and Sokoloff (1998) for Canada and U.S.

[2] There are several papers (for example Hall and Jones, 1996, 1997, 1998; Gallup and Sachs, 1998; Moreno and Trehan, 1997; and Davies and Weinstein, 1996) that have tried to answer the question of the importance of geography in explaining the levels of economic activity across countries

conditioned the local climate, the type and use of the land, and the agricultural activities of the country.

The entire coastal area of Peru (around 11% of its territory but with 49% of the total population[3]) is one of the driest regions on the surface of the Earth. Cold waters off the coast and the proximity of the high Andes, as well as wind patterns out of the South Pacific high-pressure system, contribute to the virtual lack of rainfall in this region (see Map 3). However, this cold humid desert results in pleasant living conditions for those not bothered by the lack of rainfall.

Map 1
Major Landforms in Peru



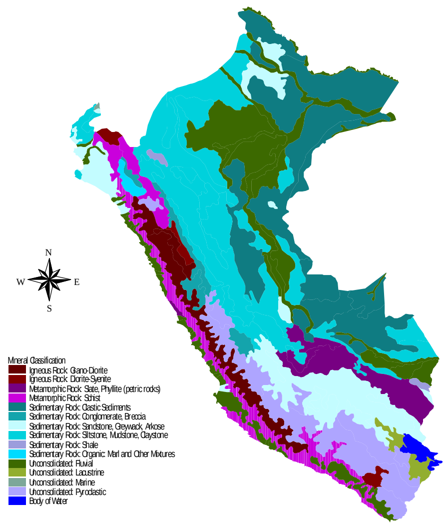
Many separate ranges, surrounding several areas of high plateau, make up the Andes in Peru, which account for 31% of Peruvian territory. Passes through these mountains are usually high and difficult, especially in the southern Andes, which can be considered a barrier to trade and transportation. Climatic conditions also make vast areas of the Peruvian Andes relatively inhospitable (see Maps 3 and 4).

A large part of Peruvian territory (about 58%) lies in the Amazon Basin. Most of this area is covered by dense forest that has slowed the development of the region. In some of these areas annual floods raise the water level more than 15 meters (50 feet) and inundate thousands of square miles of land. These floods deposit alluvial silts that renew the soils of the flooded areas (see Map 3).

---

[3] On the contrary, the *Selva* (jungle or Amazon region) represents 58% of the territory but holds only 7% of the population.

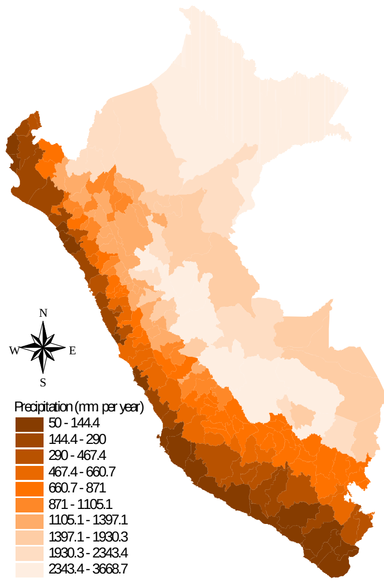
Map 2
Underlying Surface Composition in Peru



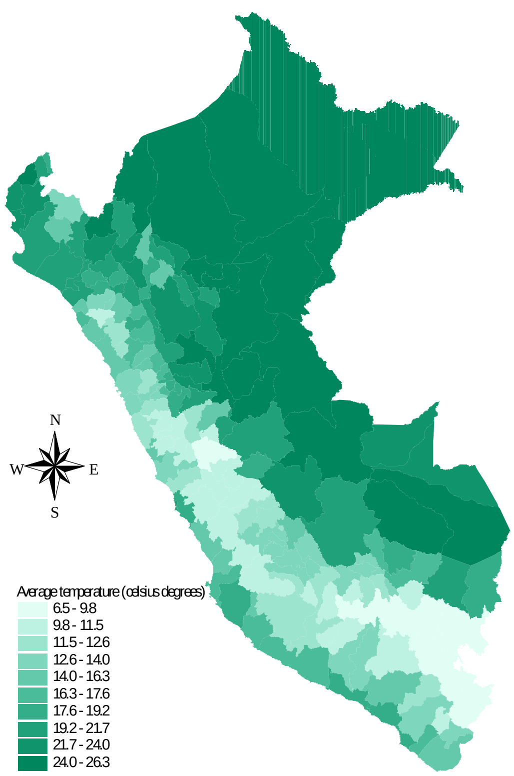
The distribution patterns of vegetation and soils in Peru are closely related to the distribution patterns of landforms and climate. That is, tropical-forest types of vegetation and soils are found mainly in the Amazon Basin, while desert types are found mainly along the coast of Peru. Soils in most tropical forests are poorly developed and low in fertility except in areas subject to annual flooding.

Peru is also well known for its mineral reserves. It has the world's second largest proven reserves of silver, third largest of tin, fourth of lead, seventh of copper and eighth of gold. As can be seen in Map 2 above, a large proportion of Peru's mineral surface composition is sedimentary rock where petroleum deposits are usually found and igneous and metamorphic rock where gold, silver, and copper deposits are found.

Map 3
Precipitation



Map 4 Temperature



Peru has a long tradition of geographic analysis and study of geography's links with development. Initially, following the Spanish tradition, the country was classified into three distinct zones: the *costa* (coast or plains), the *sierra* (basically the Andean mountain range) and the *selva* (the jungle or Amazon). However, many authors[4] have shown that this classification scheme is not sufficient to encompass Peru's actual geographic diversity.

As can be seen in Figure 1, depending on the transversal cut, Peru's geographic heterogeneity is quite high and landscapes can differ widely. Based on these findings, Pulgar Vidal (1946) divided Peruvian territory into eight distinct "natural regions" (see Table 2). The geographical pattern of these zones is depicted in Map 5.

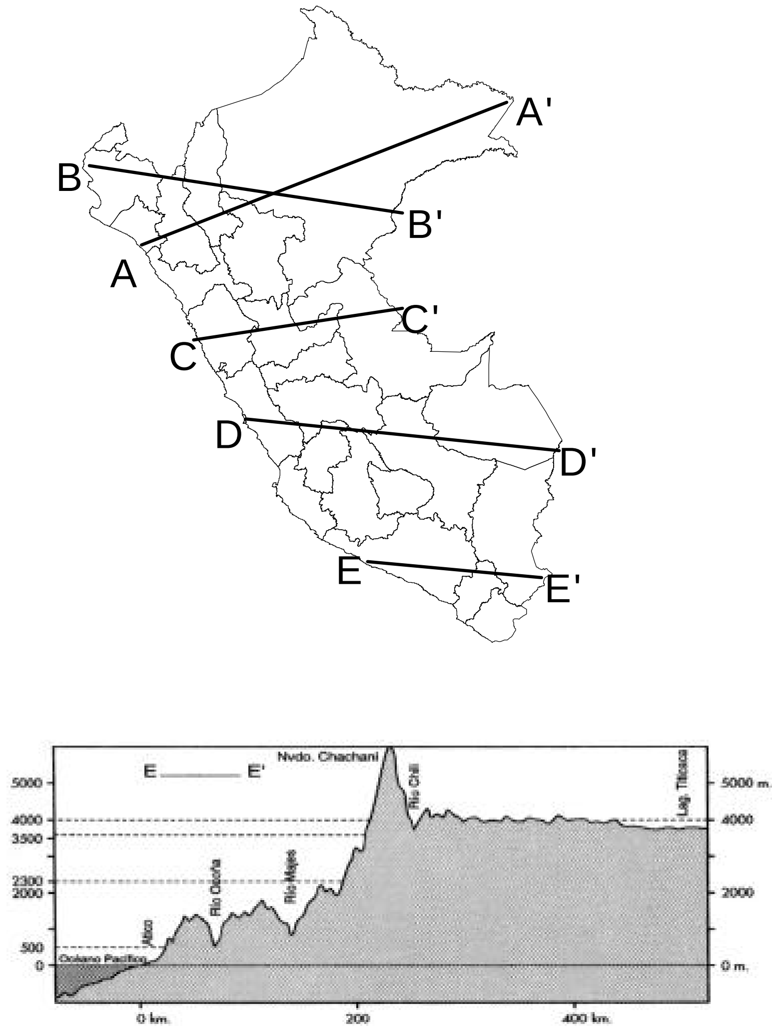
Table 2. Peru's Eight Natural Regions

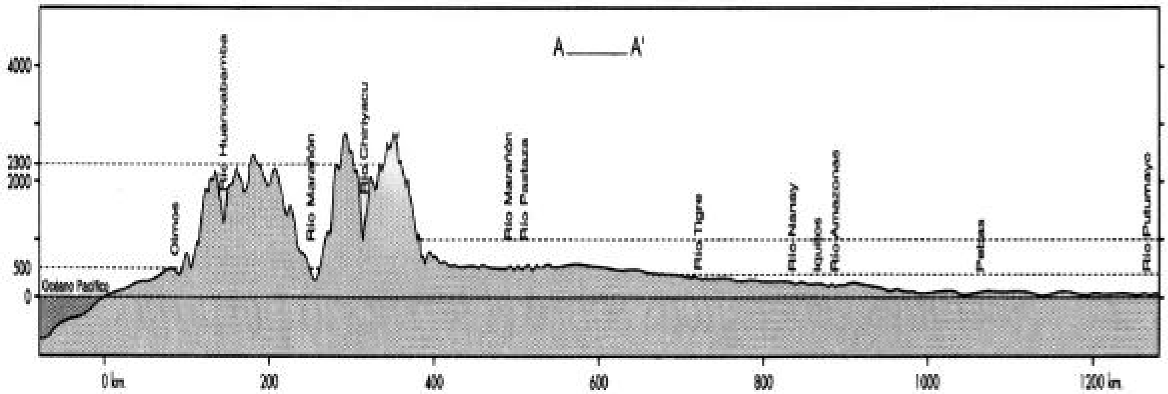
Region	Description
Costa/Chala (coast/plain)	Below 500 meters over sea level (m.o.s.l.) on the western side of the Andes. Mainly desert.
Yunga (warm zone)	On both sides of the Andes, between 500 and 2,300 m.o.s.l. on the western side and 1,000 and 2,300 m.o.s.l. on the eastern side. Typically formed by valleys.
Quechua (temperate zone)	On both sides of the Andes, between 2,300 and 3,500 m.o.s.l. Typically formed by knolls and medium to steep hillsides.
Suni or Jalca (cold lands)	On both sides of the Andes, between 3,500 and 4,000 m.o.s.l. Typically consisting of steep terrain.
Puna (high altitude plateau)	On both sides of the Andes, between 4,000 and 4,800 m.o.s.l., immediately below the snow line.
Janca or Cordillera	At the top of the Andean range, between 4,800 and 6,768 m.o.s.l. Not a continuous area. Usually no permanent settlements are found in this area. (Only 1 district capital of the 1,879 districts in Peru is at an altitude higher than 4,800 m.o.s.l.)
Selva Alta (high altitude jungle)	On the eastern side of the Andes, between 400 and 1,000 m.o.s.l. Consisting of mountainous forest with valleys.
Selva Baja (low altitude jungle)	On the eastern side of the Andes, below 400 m.o.s.l.

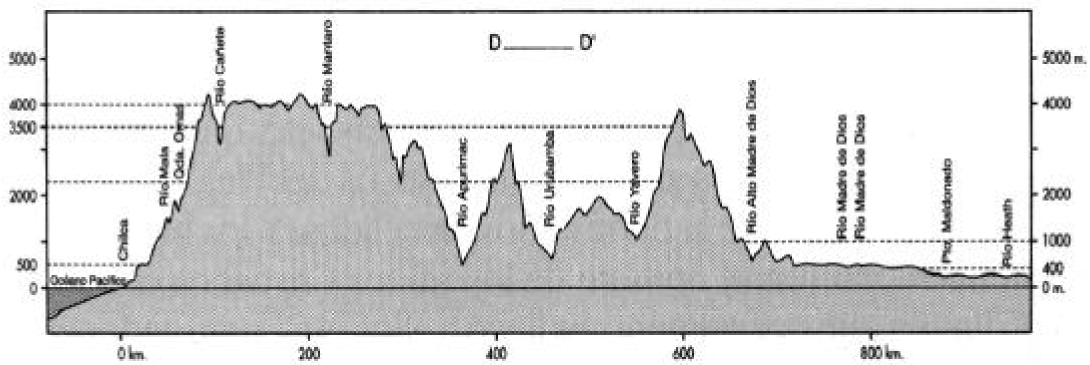
Despite the fact that there have been many efforts to link Peruvian geographical diversity to key issues such as settlement location or construction of administrative or political regions, very little has been done to analyze the links between this geographic diversity and development, economic growth or poverty. The only exception is the construction of "poverty maps" made by the Government to help target social programs. One of the most recent efforts in this regard is the construction of poverty indices at the provincial and district level by the Fondo Nacional de Compensación Social (FONCODES), the public agency in charge of poverty alleviation programs. Although these maps are "geographic" in nature, no effort has been made to link them to geographic variables, such as trying to find out whether there is any kind of poverty trap due to the negative externalities of certain "geographic endowments."

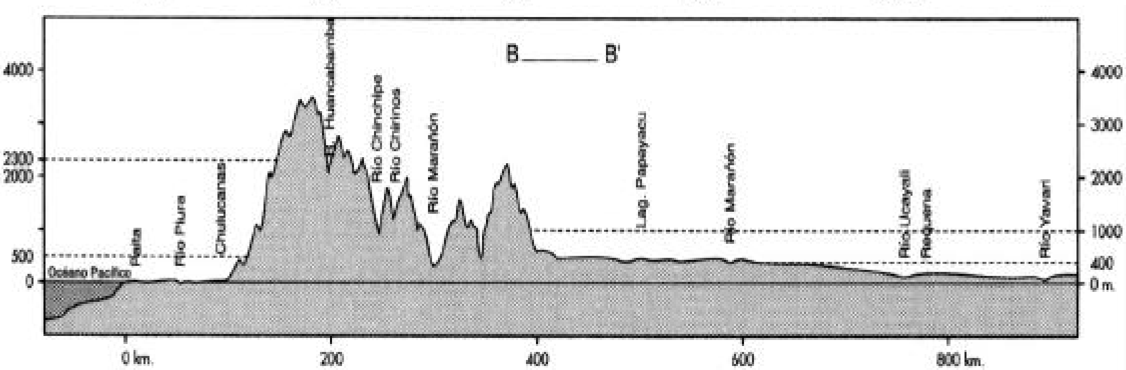
[4] A literature review on this topic can be found in Pulgar Vidal (1946) and in Peñaherrera (1986).

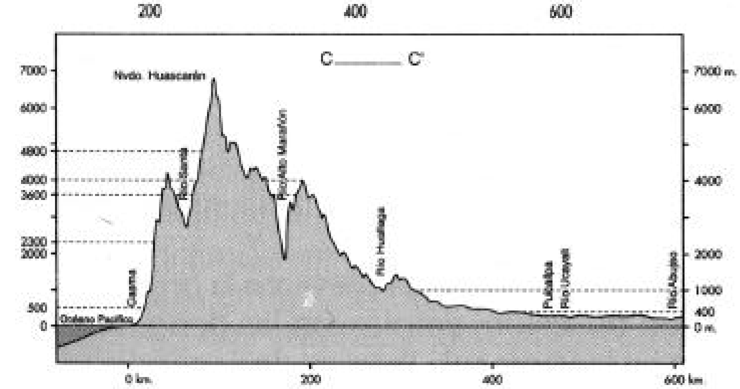
Figure 1. Five Landscape Profiles of Peru



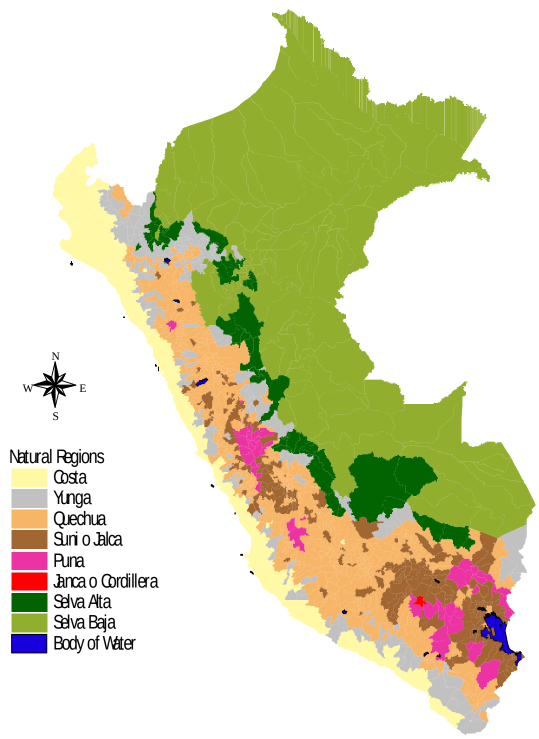

A ____ A'
4000
2300
2000
500
0
Océano Pacífico
Olmos
Río Huancabamba
Río Marañón
Río Chiriyacu
Río Marañón
Río Pastaza
Río Tigre
Río Nanay
Iquitos
Río Amazonas
Pebas
Río Putumayo
0 km.
200
400
600
800
1000
1200 km.


D ____ D'
5000
4000
3500
2000
500
0
5000 m.
1000
400
0 m.
Océano Pacífico
Chilca
Río Mala
Río Cañete
Río Mantaro
Río Apurimac
Río Urubamba
Río Yavero
Río Alto Madre de Dios
Río Madre de Dios
Río Madre de Dios
Río Heath
0 km.
200
400
600
800 km.


B ____ B'
4000
2300
2000
500
0
3000
1000
400
0 m.
Océano Pacífico
Paita
Río Piura
Chulucanas
Huancabamba
Río Chinchipe
Río Chirinos
Río Marañón
Lag. Papayacu
Río Marañón
Río Ucayali
Requena
Río Yavarí
0 km.
200
400
600
800 km.


C ____ C'
7000
6000
4800
4000
3600
2300
2000
500
0
7000 m.
1000
400
0 m.
Nvdo. Huascarán
Océano Pacífico
Casma
Río Santa
Río Alto Marañón
Río Huallaga
Pucallpa
Río Ucayali
0 km.
200
400
600 km.

Map 5
Eight Natural Regions of Peru



The next question to ask, then, is whether there is geographic concentration of poverty in Peru. Map 6, the Poverty Map for Peru, graphically answers this question by showing the poverty indices at the provincial level in Peru based on a "poverty index" constructed by FONCODES.[5] As shown on the map, there are huge welfare disparities across the country, and there is a heavy concentration of very poor people along the most geographically adverse regions, as in the *sierra* and *selva*.

Table 3 also shows how there is a negative relation between the main geographic variables (altitude, rainfall, and temperature) and household economic welfare. The higher the altitude, the larger the number of poor households in the specific region (districts). As expected, temperature shows a non-linear relationship, such that poverty increases in areas with very low levels of temperature and with extremely high levels of temperature. The precipitation variable, however, does not display a clear relationship.

On the other hand, these welfare disparities can also be attributed, at least in part, to a significant dispersion of asset ownership or access. As can be seen in the following table, most of the access to public assets and services is at least two or three times as high in urban areas as compared to rural areas. In the case of access to sanitation connection, differences are even greater (see Table 4).[6]

Even though access to public goods and services has increased dramatically in rural areas during the last four years, new access continues to be biased in favor of urban areas. Two thirds of the new electricity, sanitation and health services are placed in urban areas. Only in education does the pattern of new public goods placed in rural areas surpass that of urban areas.

---

[5] This index was constructed at the district level by weighting socioeconomic indicators reflecting: extreme poverty (infant mortality, children with chronic malnutrition), indicators of education (illiteracy rate, school attendance rate), labor market indicators (proportion of working children, percentage of illiterate adults), housing indicators (percentage of households living in overcrowded housing, percentage of houses with precarious roofing), and basic services indicators (access provided by public networks to water, sanitation and electricity).

[6] Poverty maps provide a detailed description of the spatial distribution of poverty within the country and are a crucial tool for research in trying to explain the relationship between poverty or inequality and indicators of development. On the other hand, it is important to mention that they must be interpreted within their limitations given that their quality is limited by the sparseness of the desegregated data. Some improvements on these methodologies can be found in Hentschel *et al*. (1998).

Map 6
Poverty Indices at the Provincial Level in Peru

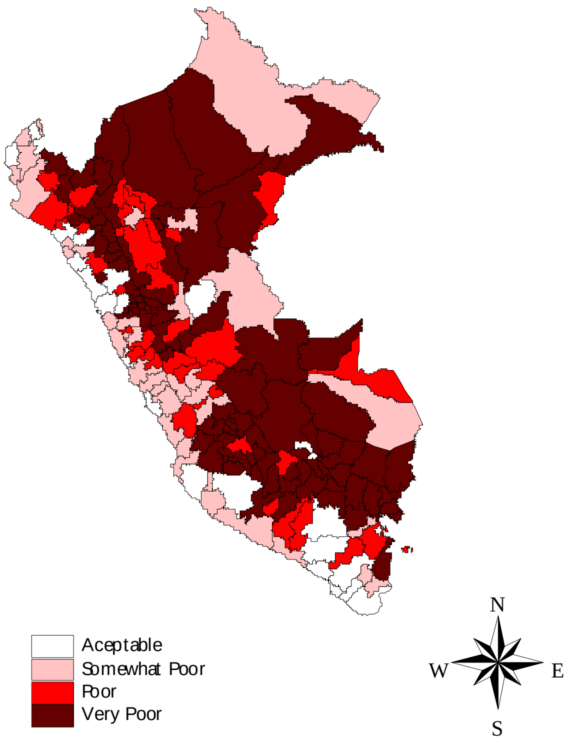


Table 3
Geography and Economic Welfare
(Percentage of poor households)

	1985	1994	1997
Altitude (m.o.s.l)			
0-500	41.4	37.5	46.1
500-1000	43.5	38.2	48.6
1000-2300	51.9	37.0	53.8
2300-3500	57.7	43.7	59.7
3500-	52.1	62.5	63.3
Precipitation (mm per year)			
0-100	35.3	33.2	40.7
100-200	54.0	33.4	42.8
200-400	46.0	65.3	58.7
400-600	59.4	69.8	61.9
600-1000	51.5	49.2	63.1
1000-1400	67.0	42.8	59.4
1400-2000	63.4	43.4	58.4
2000-2800	60.3	70.4	55.8
2800-	42.7	34.4	54.7
Temperature (celsius degrees)			
0-5	52.7	67.6	65.4
5-10	49.1	44.2	57.8
10-15	40.6	34.4	43.1
15-20	55.1	43.0	53.1
20-	61.7	46.8	55.9

Source: Authors' calculation based on 1985-86, 1994 and 1997 LSMS.
Poverty line is obtained from Escobal, *et.al.*(1998).

Table 4
Regional Differences in Access to Services and Assets: Peru 1997

	URBAN	RURAL	RATIO
Family Size	6.1	6.3	1.0
Years of Education (head)	8.6	4.5	1.9
Years of Education (adults)	8.1	5.0	1.6
Drop-Out Rates, Secondary School	12%	15%	0.8
Access to Electricity (%)	97%	30%	3.2
Access to Water, public network (%)	89%	43%	2.1
Access to Sanitation Connection (%)	84%	12%	7.3
Access to Credit (%)	37%	23%	1.6
Memo: Poverty rate	40%	65%	

Source: LSMS 1997

Table 5
Distribution of New Access to Basic and Social Services
Peru: 1994 – 1997

	URBAN	RURAL	RATIO
Water, Public Network	57%	43%	1.3
Electricity	72%	28%	2.6
Sanitation Connection	78%	22%	3.5
Outpatient Health	74%	26%	2.8
School Enrollment	33%	67%	0.5

Source: LSMS 1994 and 1997

Given the above evidence, the major question this study will try to answer what causal role do geographic variables, both natural and manmade, play in explaining per capita expenditure differentials across regions within Peru. How have these influences changed over time, how important will they be in the future, and through what channels have those influences been transmitted? It is additionally important to ask whether access to private and public assets plays a crucial role in reducing the negative effects of an adverse geography. The next section describes how we plan to formally answer these questions.

3. Analytical Framework to Test the Effects of Geography

The main question this paper tries to answer is whether geography has any effect on living standards after controlling for observable non-geographic characteristics of the households and whether access to public and private assets compensates for the effects of an adverse geography. To address this question, we have divided the analysis into three stages.

The first stage analyzes the evidence of regional income differences and to what extent these differences have been hampered (or facilitated) by local or neighboring, and natural or manmade, geographic endowments. We analyze the evolution of geographic patterns and the importance of clustering in some areas by using spatial econometric techniques, such as the Moran I statistic (see

Annex 1). We measure for the presence, over time, of spatial concentration of per capita expenditure and geographical, private and public assets and test for their significance.

In the second stage, to formally answer whether geography has a causal role in determining how household welfare evolves over time, we developed an estimable micro model of consumption levels and growth.

To model changes in consumption over time we use three census databases at the provincial level (see Annex 2 for details on how consumption is estimated for the census databases). This analysis also allows us to see what geographic factors matter to growth prospects at the micro level (Jalan and Ravallion, 1998, and Engerman and Sokoloff, 1998).

Our explanatory variables include a set of individual characteristics such as human assets (x), a set of private assets (z), a set of public assets at the district level (r) and a set of variables comprising specific geographic characteristics such as climate, soil characteristics and altitude (g). Specifically the change in consumption equation is:

$$\Delta c_p = \alpha + \beta x_{p,0} + \phi z_{p,0} + \gamma r_{p,O} + \varphi g_p + \varepsilon_p \qquad (1)$$

in which the subscript *p* refers to province-level averages of the respective variables, and the subscript zero refers to information of the initial period. We include each of the groups of regressors incrementally, and lastly we estimate the full model. We run a set of models including, one by one, each of the groups of explanatory variables: geography (g), neighboring public assets (r), private assets (z), and individual characteristics (x). We then identify the direct externality effects of the presence of each of them. Additionally, according to the hypothesis of the presence of spatial concentration we analyze the importance of neighboring province effects by measuring the significance of spatial autocorrelation[7] in each of our specifications and test how it decreases as we include additional groups of regressors (see Annex 1 for the spatial autocorrelation tests used).

We model spatial dependence as a nuisance (a nuisance since it only pertains to the errors). Formally, this dependence is expressed by means of a spatial process for the error terms, either of an autoregressive or a moving average form (see Anselin, 1988 and 1990, and Anselin, Varga, and Acs, 1996). Such an autoregressive process can be expressed as:

$$\begin{aligned} \Delta c_p &= \alpha + \beta x_{p,0} + \phi z_{p,0} + \gamma r_{p,O} + \varphi g_p + \varepsilon_p \\ \varepsilon_p &= \lambda W \varepsilon_p + \xi \end{aligned} \qquad (2)$$

with $W\varepsilon$[8] as a spatially lagged error term, λ as the autoregressive coefficient and ξ as a well-behaved (i.e., homoskedastic-uncorrelated) error term.

---

[7] Spatial autocorrelation, or more generally, spatial dependence, is the situation where the dependent variable or error term at each location is correlated with observations on the dependent variable or values for the error term at other locations.

[8] For *N* districts observed, $\boldsymbol{W_i}$ is the *i*th row of an (N*N) matrix $\boldsymbol{W}$ that assigns neighboring districts to each district . The $\boldsymbol{W}$ used can be characterized by $W=\{w_{ij}\}$ such that $w_{ij}=1$ if *i* and *j* are neighboring districts, $w_{ij}=0$ otherwise, and $w_{ii}=0$ for all *i*. The rows of *W* are then normalized such that each observation's neighboring districts have the same

As a consequence of the spatial dependence, the error term no longer has the usual diagonal variance matrix but instead takes the following form (Anselin, 1988 and 1990):

$$E[\varepsilon\varepsilon'] = \Omega = \sigma^2[(I - \lambda W)'(I - \lambda W)]^{-1} \tag{3}$$

Therefore, Ordinary Least Squares (OLS) estimates are no longer efficient, but they are still unbiased. Furthermore, given that the lambda coefficient is unknown, the regression coefficients cannot be estimated using Generalized Least Squares (GLS), and therefore in our last specification we estimate the lambda coefficient jointly with the regression coefficients using full maximum likelihood estimation techniques.[9]

In order to identify the effects of geography on households we also use the LSMS household surveys and perform an estimation of the levels of consumption and an estimation of the growth of consumption using two household panels, one for 1991-1994 and another for 1994-1997. The specification used is very similar to the one in equation (1). We include again as regressors a set of individual characteristics such as human assets (x), a set of private assets (z), a set of public assets at the district level (r) and a set of variables comprising specific geographic characteristics such as climate, soil characteristics and altitude (g). Specifically the equation we estimate is:

$$c_i = \alpha + \beta x_i + \phi z_i + \gamma r_d + \varphi g_d + \varepsilon_i \tag{4}$$

in which the subscript *i* refers to a household and the subscript *d* refers to district-level information.[10] Additionally, to analyze the effects of geography on the income distribution of the households we perform quantile regressions.

We also develop a micro model for consumption growth allowing for constraints on factor mobility and externalities, whereby geographic factors, in the specific region or neighborhood regions, can influence the productivity of a household's own capital. For this purpose, we follow Islam (1995) and estimate the following model:

$$\begin{aligned} &\Delta c_{it} = \gamma c_{it-1} + \beta_1 x_{it} + \beta_2 z_{it} + \beta_3 g_{it} + \beta_3 r_{it} + \varepsilon_i \\ &\text{where:} \\ &c_{it-1} = \ln c(t_1) \\ &\Delta c_{it} = \ln c(t_2) - \ln c(t_1) \\ &\gamma = (1 - e^{-\lambda\tau}) \end{aligned} \tag{5}$$

---

amount of influence, that is $\sum_j w_{ij} = 1$, for all *i*. In addition it will be assumed that each neighboring district of a given district carries equal weight, $w_{ij} = w_{ik}$ for non-zero elements (neighbors) *k* and *j* for firm *i*. If more information were available about the amount of influence each district yields, this could be incorporated into the *W* matrix (regarding the different possible structures see Anselin, 1988).[8]

[9] For a more extensive technical discussion of the relative merits of the various estimators suggested in the literature, see Anselin (1988,1990).

[10] In contrast to our previous specification we cannot correct for the presence of spatial autocorrelation because we don't know the exact location of the households and therefore we can't construct the spatial matrix (W).

This methodology will allow us to test over time the effect of geographic variables as well as the convergence rate. As mentioned by Jalan and Ravallion (1998), “one should not be surprised to find geographic differences in living standards in this setting. For one thing, restrictions on labor mobility can perpetuate spatial concentrations of households with poor endowments. But geography can also have a deeper causal role in the dynamics of poverty in this setting. If geographic externalities alter returns to private investments, and borrowing constraints limit capital mobility, then poor areas can self-perpetuate. Even with diminishing returns to private capital, poor areas will see low growth rates, and possibly contractions.”[11]

Lastly, the third stage follows Ravallion and Wodon (1997) and tries to use the results of the previous specifications and break down the geographic effects into their component elements. For this purpose, we compute the expected gain (or loss) in consumption from living in one geographic region (e.g., coast) against living in another geographic region (i.e., mountains) specifying how much of the gain is explained by geographical variables, location (urban or rural areas), infrastructure and private assets:

$$(\bar{X}_M - \bar{X}_C)\bar{\beta} \qquad (6)$$

where $\bar{X}_{M,C}$ are the sample means for mountain and coast regions, for example, and $\bar{\beta}$ is the parameter of the respective variables under analysis (i.e., geographical, location, infrastructure and private assets). This breakdown represents the differential impact on a household's standard of all non-excluded variables in the two regions.

4. The Data

To be able to answer the major questions outlined in the previous section we have developed four different databases: census, household surveys (LSMS), and a panel database from the LSMS surveys, all of which were linked to a geographical database (see Data Sources below).

We have used the Population and Household Census of 1972, 1981 and 1993 to construct a set of variables that allows us to analyze the kind of changes that have emerged in the geographical pattern of Peru’s most important socioeconomic variables during the last three decades. Additionally, using the methodology of Jesko *et al*. (1998), we estimate a household-level expenditure equation using the information from the 1985-86 and 1994 LSMS surveys (see Annex 2 for details on the estimation) which allowed us to model the determinants of per capita expenditure growth at the provincial level. This, in turn, allows us to determine what role geographic variables, both natural and manmade, play in explaining per capita expenditure differentials across regions in Peru.

We also used the cross-sectional LSMS household surveys, given that they had vast information on household characteristics, income and expenditures, as well as on household access to private and public services. This cross-sectional micro data is therefore used in our second methodological strategy to test for geographic effects on living standards at a point in time. For example, see Borjas (1995) on the effects of neighborhood on schooling and wages in the U.S. and

[11] See Jalan and Ravallion (1998) for formal tests of poverty traps.

Ravallion and Wodon (1997) on the effects of geography on the level of poverty in Bangladesh, as well as on the importance of public and private assets in explaining regional poverty variations.

Lastly, in order to apply Jalan and Ravallion's methodology we built up a panel between 1991, 1994 and 1997 using the LSMS surveys. The advantage of having standard panel data with time invariant fixed effects on households, allowing for latent household heterogeneity, will protect against spurious geographic effects that arise solely because geographic variables proxy for omitted non-geographic, but spatially autocorrelated, household characteristics.

5. Empirical Results

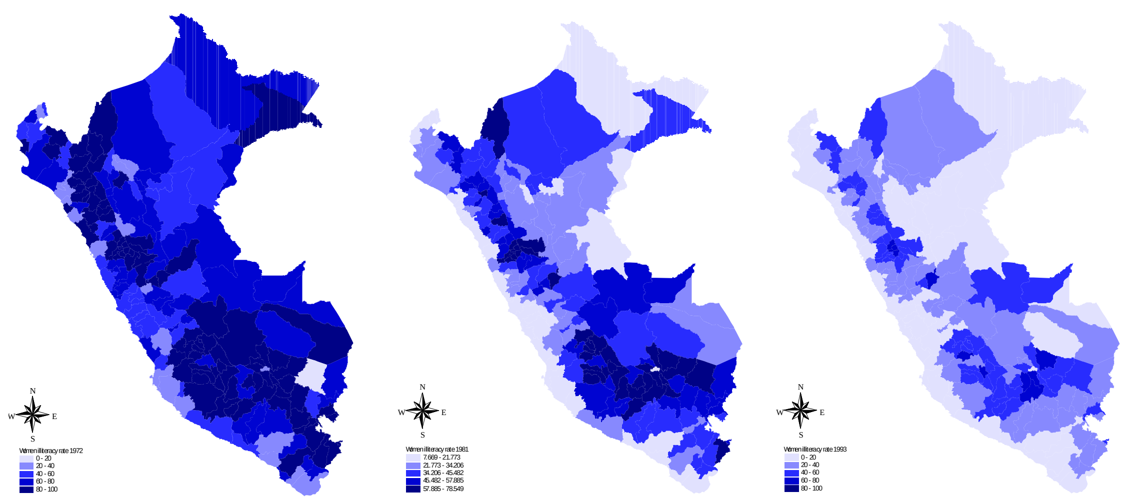
5.1. Peru's Geography and Regional Differences in Expenditure

In this section we analyze the kind of changes that have emerged in the geographical pattern of Peru's most important socioeconomic variables during the last three decades. In addition we analyze changes in expenditure estimates, at the province level, between three Census dates (1972, 1981 and 1993).

We analyze 24 variables at the provincial level for a panel of three Census years (1972, 1981 and 1993), as well as 160 additional variables at the provincial level and 88 additional variables at the district level for variables that were available only for 1993 and beyond. Annex 2 describes these variables as well as the databases that generate them. We have included in this section some of the maps generated with these variables.

It is interesting to note that there are several types of evolution in the geographic patterns. There are cases such as the one depicted in Map 7 that show a dramatic reduction of illiteracy rates among women but, at the same time, the high rates are clustered in some areas (like the southern sierra and other high altitude zones). This kind of pattern can also be found in other key socioeconomic variables, such as total illiteracy rate or household size.

Map 7. Illiteracy Rate of Women

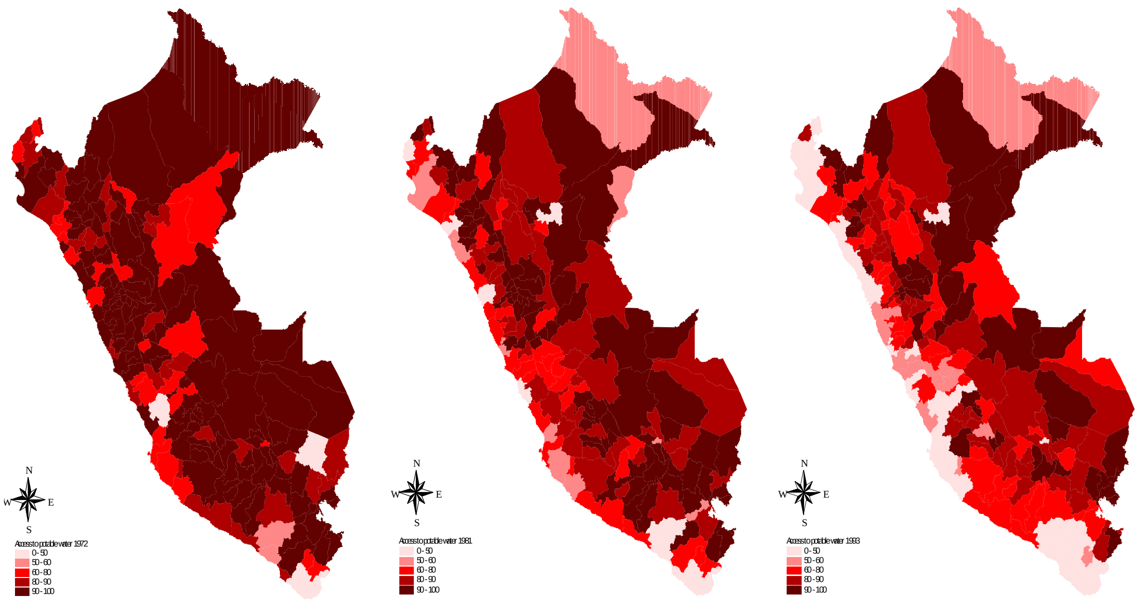


1972 1981 1993

There are other variables, such as percentage of households without access to potable water, percentage of households without access to sanitation services or percentage of households without access to electricity, that display during the 1972-1981 period a significant reduction in the coastal areas and afterward some clustering of high values, especially in the southern sierra and high altitude jungle regions, and no distinguishable pattern in the rest of the country (access to potable water is depicted in Map 8)

Map 8

Households Without Access to Potable Water by Years



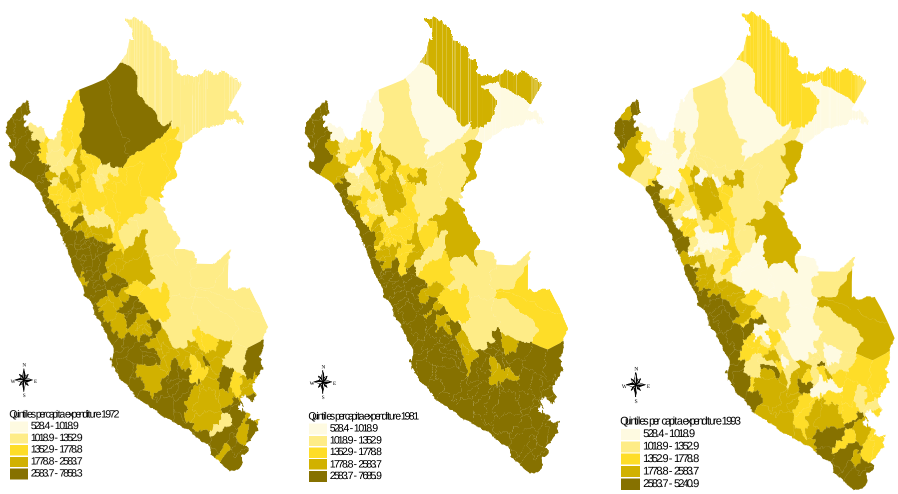
1972 1981 1993

In order to more comprehensively analyze the changes that occurred in these geographic patterns we have constructed a per capita expenditure variable at the provincial level. Following a procedure similar to that of Hentschel *et al.* (1998), we used household data to construct expenditure functions using the Peruvian LSMS surveys of 1985 and 1994. We used the 1985 expenditure function to construct provincial level expenditure estimates, using data taken from the 1972 and 1981 Census as explanatory variables. We used the 1994 expenditure function to construct the provincial level expenditure estimates based on data taken from the 1993 Census. The exact procedure and data involved in these calculations can be found in Annex 2.

The geographical evolution of Peru's per capita expenditures between 1972 and 1993 can be viewed in Map 9. Here it is evident that higher per capita expenditure is to be found along low altitude coastal regions. This pattern, which is already clear using 1972 data, is even more apparent as time passes. It is interesting to note that the Gini coefficients are extremely low (0.118 in 1972, 0.088 in 1981 and 0.187 in 1993). It must be noted, however, that inter-regional expenditure variance is very low, at least when compared to within-region variance, making these Ginis perfectly consistent with a national Gini coefficient of 0.42 and 0.38 in 1985 and 1994, respectively.

Map 9

Distribution of Per-Capita Expenditure by Year

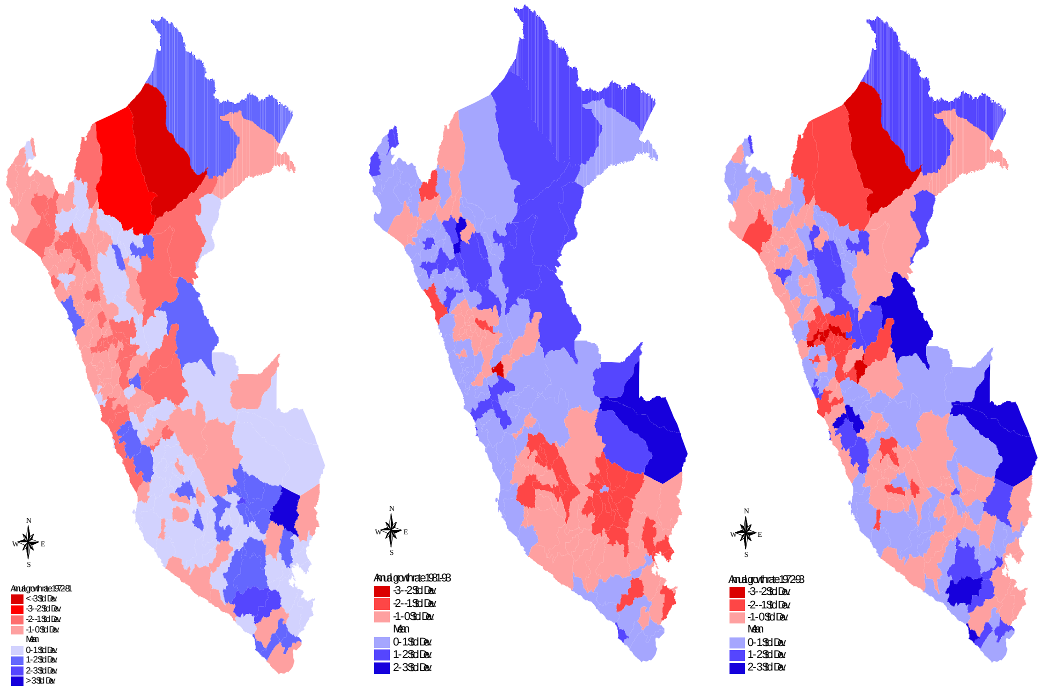


1972 1981 1993

Map 10 shows the pattern of distribution of inter-annual per capita expenditure growth rates between Census years. Here it can be noted that the provinces whose per capita expenditures have grown faster tend to be clustered, as do those provinces showing little or even negative growth. Provinces showing high growth tend to be clustered in the higher altitude jungle. Table 6 confirms the graphical analysis, showing high and statistically significant Moran Index and Geary Index values for all three Census years. In addition, high Moran and Geary Index values can also be found for per capita expenditure growth.

Map 10

Change in Per Capita Expenditures (%)



1972-81 1981-93 1972-93

Table 6

Spatial Autocorrelation of Province-Level Expenditure Variables

Variables	Moran Index	Prob. a/	Geary Index	Prob. a/
Per capita expenditure				
1972	0.4131	0.00	0.6078	0.00
1981	0.5709	0.00	0.3993	0.00
1993	0.4888	0.00	0.4565	0.00
Change in per capita expenditure				
1972-81	0.3708	0.00	0.6186	0.00
1981-93	0.4990	0.00	0.4616	0.00
1972-93	0.2427	0.00	0.7308	0.00

a/ Probablity to reject null hypothesis (absence of spatial autocorrelation)

Source: Author's calculation based on province estimates.

Table 7 shows some of the most significant spatially autocorrelated variables in our data set. Using the Moran and Geary Indexes, we find that aside from some obviously spatially correlated variables such as annual precipitation or altitude of the province or district capital, critical socioeconomic variables such as household size, percentage of households headed by women or total and female illiteracy rates tend to be heavily clustered, showing high values in high altitude zones and low values in coastal areas. A similar situation can be found in other variables such as percentage of houses with inadequate flooring or overcrowded housing, malnutrition rates, and school dropout rates and schooling years. One summary welfare variable, per capita expenditure for 1993, displays a high and statistically significant Moran Index value and Geary Index. It is also interesting to note that the variable of soil depth, constructed to show agricultural land potential, also has a highly spatial autocorrelated pattern.

Aside from some obvious variables, such as those related to urban areas (urban density or number of towns per province, for example) there are very few variables that do not show a clear geographical pattern. Only three variables deserve some mention: Change in household size between 1972 and 1981; the growth of the illiteracy rate between 1981 and 1993; and the growth in per capita expenditures between 1972 and 1981, which do not show any geographical pattern measured by the Moran spatial autocorrelation index or the Geary Index. (see Annex 3)

Table 7

Highly Spatial Autocorrelated Variables

Variables	Moran Index	Z-Value	Geary Index	Z-Value
South latitude	0.9302	20.21 *	0.057	-18.76 *
North longitude	0.8870	19.27 *	0.093	-18.04 *
Precipitation	0.7573	16.47 *	0.259	-14.73 *
Household size 1993	0.7495	16.30 *	0.241	-15.10 *
Temperature (average)	0.7486	16.29 *	0.256	-14.79 *
Temperature (min.)	0.7469	16.25 *	0.255	-14.83 *
Temperature (max.)	0.7422	16.15 *	0.265	-14.62 *
Altitude of the district capital (meters over sea level)	0.6693	14.57 *	0.322	-13.47 *
% household head that are female 1993	0.6560	14.28 *	0.325	-13.43 *
Inadequate floor	0.6518	14.19 *	0.339	-13.16 *
Soil depth	0.6422	13.99 *	0.328	-13.37 *
Total illiteracy rate 1981	0.6352	13.83 *	0.356	-12.82 *
Overcrowded houses 1993	0.6286	13.69 *	0.339	-13.15 *
Household size 1981	0.6130	13.35 *	0.377	-12.39 *
Per capita expenditure in 1981	0.6084	13.26 *	0.399	-11.95 *
Perimeter of the province	0.6032	13.14 *	0.390	-12.12 *

Note: p<0.01=*, p<0.5=~, where p is the probablity to reject null hypothesis (absence of spatial autocorrelation)

: Authors' calculation based on National Census of Populations 1972, 1981 and 1993.

5.2. Testing the Causal Role of Geography on the Evolution of Welfare: Provincial Level Data

As we have seen in Section 3, it is possible to derive a connection between the asset endowment of an individual household and its expenditure level. Following the same reasoning we can derive a connection between the level of private and public assets that can be found at some level of spatial aggregation (here the provincial level) and the per capita expenditure level that can be found in that area.

Table 8 shows the econometric results of what could be called the determinants of per capita expenditure growth at the provincial level. To reduce any possible endogeneity bias in explaining 1972-1993 per capita expenditure growth rates, we have chosen initial asset endowments as independent right hand-side variables. To this basic data set we have added several key geographical variables to check whether they can provide some explanation of causes of expenditure growth. Table 8 shows the Moran spatial autocorrelation index for the four different specifications that were evaluated: (1) only private assets; (2) private assets plus geography variables; (3) the previous variables plus public assets; and (4) all the variables plus changes in access to key public assets.

We have used the log difference of per capita expenditures as a dependent variable. The reason for this choice (as opposed to using percentage changes) is related to functional form issues. If there is any misspecification in the per capita expenditure equations (which have been estimated as semi-log functions) the log difference of per capita expenditures will clean the bias, provided that these variables have similar effects over the years.

As can be seen in Table 8, when geographic variables are included as the only explanatory variables, altitude and longitude prove to be highly significant in explaining expenditure growth. In particular, it can be shown that the higher altitude provinces tend to have slower expenditure growth rates. When we add the variable of basic needs, which encompasses the absence of critical public infrastructure (sanitation, water, telephone and electricity) we can see that altitude remains significant but its negative impact diminishes considerably. This effect can be viewed as demonstrating the importance of public infrastructure to lower negative geographic externalities. It is important to note that when we add private assets (some of which are obviously correlated with public assets) the importance of geography almost vanishes. This effect can be seen in Map 11, where we have graphed the pattern of geographic residuals of each model. This initial finding will be followed up more rigorously in the next section.

It is interesting to note that, despite the fact that this expenditure growth function has included all relevant geographic variables at hand, the residuals continue to show spatial autocorrelation. As can be seen in Table 9, although the Moran Index diminishes as we include explanatory variables it remains significant. This fact suggests the idea that there may be non-geographic non-observables that may be affecting the provincial expenditure pattern. This is consistent with Ravallion and Wodon (1997) when they show that sizable geographic differences in living standards can persist even if we take into account the spatial concentration of households with readily observable non-geographic characteristics conducive to poverty.

The last column in Table 8 shows the estimated parameter values corrected for spatial autocorrelation.[12] The results confirm that when public and private assets, as well as household characteristics, are included in the regression, the impact of geographic variables is dampened.

Finally, in Table 10 we can find the spatial breakdown of the regression model according to the Anselin (1995a) technique (see Annex 1). Here residuals are clustered in four groups: large residuals values clustered around large value areas, small residuals values clustered around small value areas, large residuals values located around small value areas, and small residuals values located around large value areas. The results confirm that geography and public asset access variables tend to lower spatial autocorrelation, and geography variables are the ones that (at the marginal level) most account for per capita growth patterns.

5.3. Testing the Causal Role of Geography in the Evolution of Welfare: Household Data

To be able to identify specific effects of geography on households we use the LSMS household surveys and estimate the levels of consumption and growth of consumption using two household panels, one for 1991-1994 and a second for 1994-1997. The specification used is detailed in equations (4) and (5). As mentioned previously, we include as regressors a set of individual characteristics as human assets (x), a set of private assets (z), a set of public assets at the district level (r) and a set of variables taking in specific geographic characteristics such as climate, soil characteristics and altitude (g).

[12] The likelihood-ratio test for spatial error dependence for the equation in the last column in Table 7 has a value of 3.67 with 1 degree of freedom, which confirms that the estimation has been properly corrected for spatial autocorrelation. For alternative methods of correcting for spatial autocorrelation see Annex 5.

Table 8

Determinants of Per Capita Expenditure Growth Rate: 1972-93

(OLS estimations with robust standard errors, at province level)

Variables at initial period	Models (1)	(2)	(3)	(4)	(5)
Intercept	4.8269 *	4.6892 *	4.3913 *	-0.0277	-0.3270
	(1.631)	(1.563)	(1.585)	(1.385)	(1.706)
Altitude	-1.1081 *	-0.7872 ~	-0.5096	0.2616	0.4580
	(0.385)	(0.377)	(0.447)	(0.385)	(0.389)
Latitude	-0.0226	-0.0308	-0.0288	-0.0231	-0.0170
	(0.017)	(0.017)	(0.017)	(0.019)	(0.019)
Longitude	-0.0561 *	-0.057 *	-0.0543 *	-0.0182	-0.0171
	(0.018)	(0.017)	(0.018)	(0.015)	(0.015)
Soil slope	-0.0012	0.0016	0.0021	0.0033	0.0035
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Soil depth	-0.003	-0.0017	-0.0018	0.002	0.0023
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Igneous rock	-0.2143	-0.2944 ~	-0.3102 *	-0.3197 *	-0.2757 *
	(0.126)	(0.123)	(0.123)	(0.100)	(0.106)
Metamorphic rock	0.0732	0.0536	0.0863	-0.1318	-0.1362
	(0.149)	(0.145)	(0.146)	(0.122)	(0.122)
Temperature	-0.0191	-0.0045	-0.0043	-0.0114	-0.0082
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)
Basic needs		-0.0561 *	-0.0393 ~	-0.0222	-0.0225
		(0.013)	(0.020)	(0.017)	(0.016)
High*basic needs			-0.1110	0.0045	-0.0149
			(0.097)	(0.090)	(0.080)
School attendance rate				0.0143 *	0.0144 *
				(0.003)	(0.003)
Household headed by women (%)				-0.0109 ~	-0.0134 ~
				(0.005)	(0.005)
Working children (%)				0.0533 *	0.0462 ~
				(0.020)	(0.018)
Household size				0.0783	0.1057
				(0.133)	(0.128)
Household size growth a/				-0.2624	-0.2208
				(0.140)	(0.136)
Number of migrants				0.0171	0.0101
				(0.029)	(0.029)
Spatial autocorrelation					0.2305 ~
					(0.102)
Number of observations	190	190	190	190	190
Adjusted R-squared	0.122	0.195	0.197	0.486	0.526

a/ Instrumental variables are shown in Annex 2.

Note: Standard deviation in parenthesis and p<0.01=*, p<0.5=~

Model 1: Geography

Model 2: Geography + infrastructure.

Model 3: Geography + infrastructure.+Geo*infra.

Model 4: Geography + infrastructure.+Geo*infra+private assets

Model 5: Geography + infrastructure.+Geo*infra+private assets, modelling first-order spatial error autocorrelation.

Source: Authors' calculation based on 1972 and 1993 Population and House Censuses.

Table 9

Spatial Autocorrelation of Growth Regression Residuals, by Model

Type of Association	Regression Model Residuals 1	2	3	4
Moran Index	0.1091	0.1005	0.0973	0.0816
Z-value	3.1226	2.9658	2.9357	2.7877
Probability	0.0018	0.0030	0.0033	0.0053

Model 1: Geography

Model 2: Geography + infrastructure.

Model 3: Geography + infrastructure.+Geo*infra.

Model 4: Geography + infrastructure.+Geo*infra+private assets

Authors' calculation based on table 7.

Table 10

Spatial Association of Growth Regression Residuals, by Model

(Number of provinces)

Type of Association	Regression Model Residuals 1	2	3	4
Positive association	111	102	102	100
	(72.1)	(65.3)	(63.2)	(63.2)
Large values in large value areas	49	48	52	52
	(40.5)	(34.2)	(34.2)	(33.7)
Small values in small value areas	62	54	50	48
	(31.6)	(31.1)	(28.9)	(29.5)
Negative association	79	86	88	90
	(27.9)	(34.7)	(36.8)	(36.8)
Large values in small value areas	38	43	44	45
	(11.6)	(17.9)	(17.4)	(17.4)
Small values in large value areas	41	43	44	45
	(16.3)	(16.8)	(19.5)	(19.5)
Total	190	188	190	190
	(100.0)	(100.0)	(100.0)	(100.0)

Model 1: Geography

Model 2: Geography + infrastructure.

Model 3: Geography + infrastructure.+Geo*infra.

Model 4: Geography + infrastructure.+Geo*infra+private assets

Note: column percentages are shown parenthesis

Source: Authors' calculation based on Table 7.

Map 11

Spatial Distribution of Regression Residuals by Model

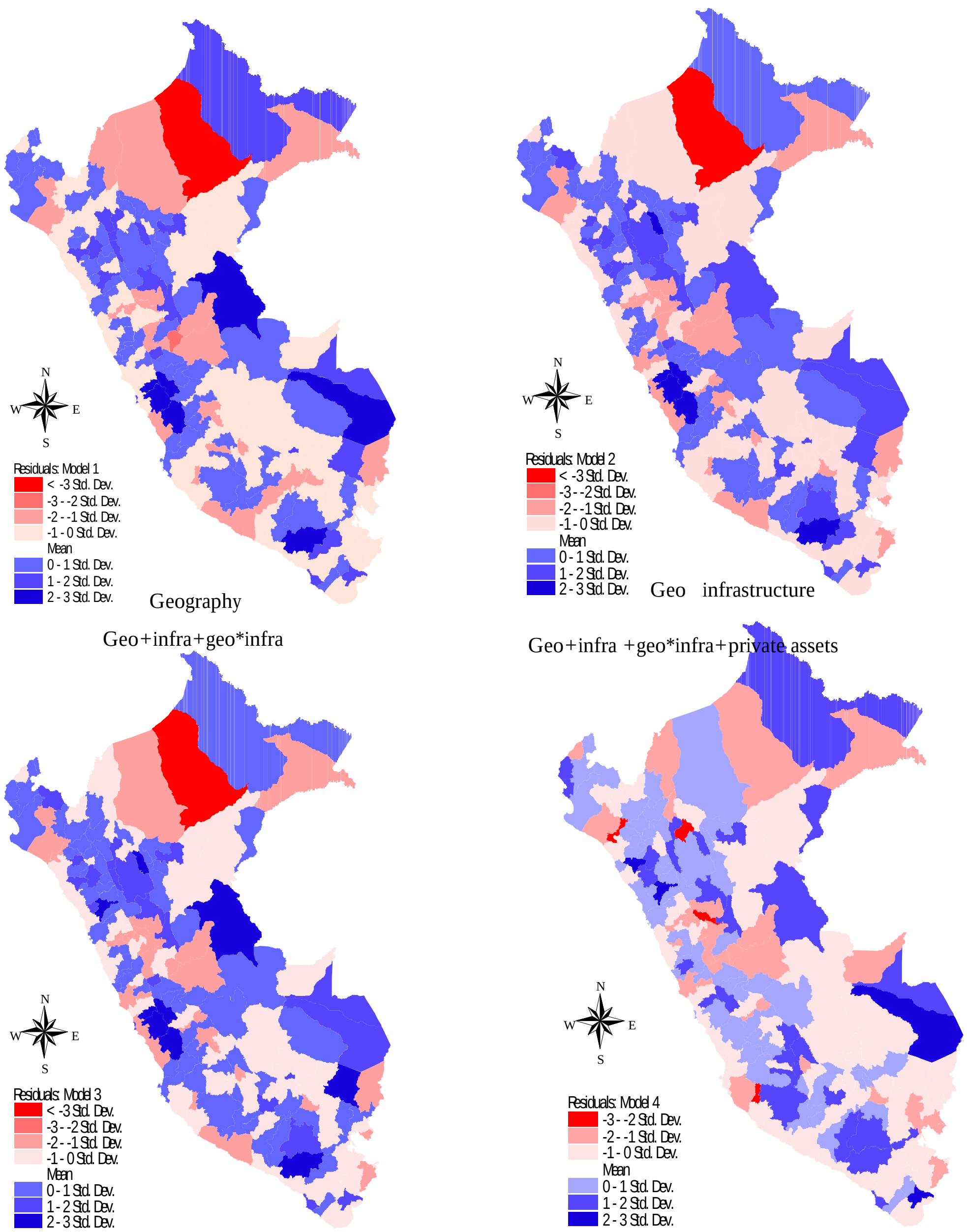


Table 11 shows the results of the determinants of current consumption expressed in logs, and, as in Section 5.2, we use four different specifications. The first specification includes only geographic variables (Model 1), the second includes geography plus location variables (urbanization and distance to capital), the third adds public assets to the previous variables, and Model 4 includes variables that measure the possession of private assets.

When geographic variables are included as the only explanatory variables, the negative and non-linear effect of temperature appears to be significant in explaining the level of consumption of the households. Therefore, as previously shown in Table 3, poverty increases for households located in regions with low temperatures and in regions with extremely high temperatures. On the other hand, as we add variables for presence of infrastructure and control for the private assets of the households this variable loses significance (see column 4). A similar pattern is found with the presence of sedimentary rock, which could imply relatively poor soil. In the first model these variables have a negative and significant effect as expected, but as we include public and private assets its negative effect is reduced and the variable loses significance.

Furthermore, when adding the variable of basic needs, which as previously mentioned, encompasses the absence of critical public infrastructure (sanitation, water, telephone and electricity) as well as overcrowded housing, we can see that the negative effects of temperature (temperature square) and of sedimentary rock diminish considerably.

Altitude, on the other hand, despite having a negative sign, is not significant, as was shown in the province-level model for consumption growth. Nevertheless, when we correlate altitude with urbanization the coefficient is significant and positive, showing the marginal positive effect that urbanization has on high altitude regions.

The variable that measures the potential presence of mineral resources underlying the surface (igneous rock) moves from negative and insignificant to positive and significant after we control for the presence of public and private goods. This could be an indicator that as more private and public resources are present it becomes easier for the households in the region to be able to profit from this type of natural resource that requires high levels of investment and infrastructure to be exploited. Similarly, soil depth becomes positive and significant when public infrastructure variables are included; this again could be an indication that the presence of public infrastructure facilitates the exploitation of the land in regions endowed with a significant depth of soil.

Finally, and as expected, the most important variables measuring private assets, such as education, labor experience, migration experience and household size, come to have the expected signs and be significant.

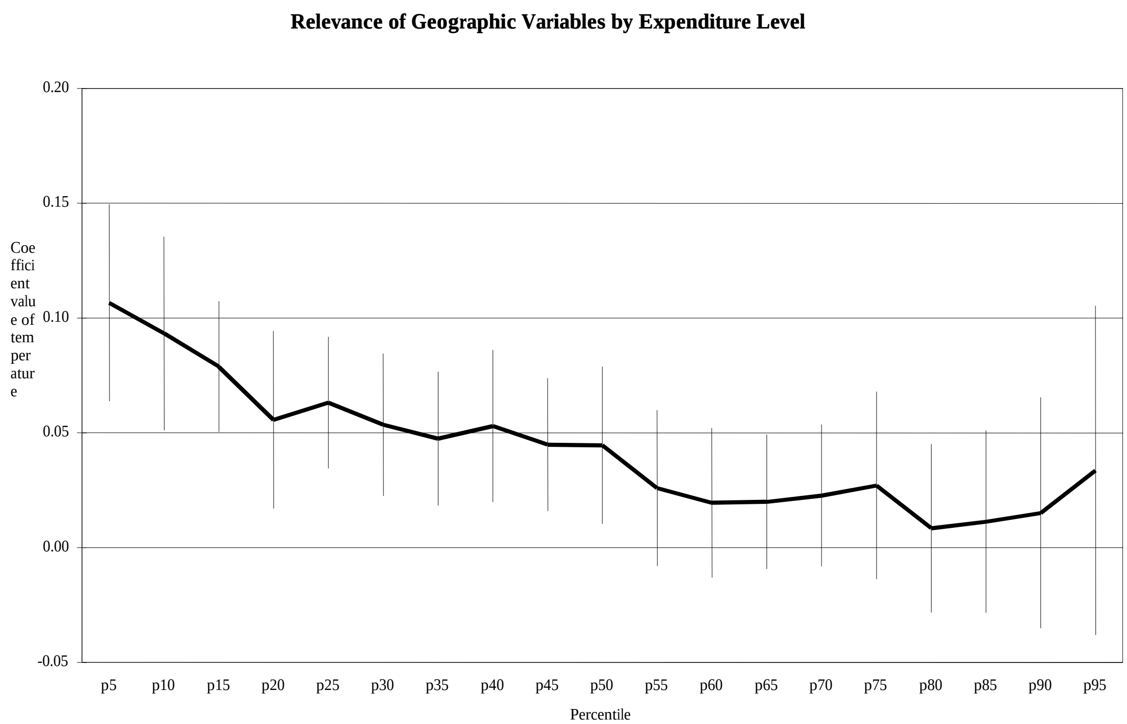
In attempting to assess whether the impact of our explanatory variables was different between poor and rich households, in Table 12 we present the results of an econometric exercise in which we run quantile regressions.[13] By calculating regressions for different quantiles, it is possible to explore the shape of the conditional distribution. This is of great interest for the present study

[13] Quantile regressions are also used to analyze the presence of heteroskedasticity Quantile regressions other than median can be defined by minimizing:

because it will allow us to determine whether richer households are less affected by adverse geographic characteristics.

In Table 12, we present the results of our full consumption level specification for the 10th, 20th, 60th, 80th and 90th percentiles. Although there are not great differences in the magnitude of the coefficients, there are some important findings. First, for the poorest percentiles, when the main geographic variables (temperature, land depth and altitude) are compared with urbanization, they play a more important role in explaining the lower levels of consumption of the lowest percentiles (10th) compared to the effect they have on the 80th and 90th percentiles. For example, the temperature square variable is negative and significant for the 10th percentile, while it is not significant for the 90th percentile. This result is depicted in Graph 1, which clearly shows how the confidence interval (represented by the vertical lines) increases significantly as we move from the poorest to the richest percentiles.

Graph 1



$$\phi_q = -(1-q)\sum_{y \le x'\beta}(y_i - x_i'\beta) + q\sum_{y > x'\beta}(y_i - x_i'\beta)$$

$$= \sum_{i=1}^{n}[q - 1(y_i \le x_i'\beta)](y_i - x_i'\beta)$$

where $0<q<1$ is the quantile of interest, and the value of the function 1(z) signals the truth (1) or otherwise (0) of the statement z. For further details see Deaton (1997)

Table 11

Determinants of Per-capita Expenditure at Household Level: 1994

(OLS estimation with robust errors including geographic variables)

	Models			
Variables	(1)	(2)	(3)	(4)
Intercept	6.2476 *	5.3807 *	6.1735 *	6.1749 *
	(0.1874)	(0.2166)	(0.2190)	(0.1802)
Altitude	-0.2417	0.2718	-0.2204	-0.1226
	(0.1323)	(0.2915)	(0.2915)	(0.2292)
Temperature	0.0733 *	0.1058 *	0.0676 *	0.0378 *
	(0.0184)	(0.0185)	(0.0172)	(0.0142)
Temperature squared	-0.0018 *	-0.0024 *	-0.0014 *	-0.0006
	(0.0005)	(0.0005)	(0.0005)	(0.0004)
Igneous rock	-0.1033	0.1066	0.0414	0.1129 ~
	(0.0711)	(0.0728)	(0.0692)	(0.0524)
Sedimentary rock	-0.1892 *	-0.1322 *	-0.0937 ~	-0.0142
	(0.0409)	(0.0415)	(0.0390)	(0.0309)
Land depth	0.0001	0.0018 ~	0.0030 *	0.0012
	(0.0008)	(0.0008)	(0.0008)	(0.0006)
Urbanization		0.3920 *	-0.0623	-0.1205
		(0.0900)	(0.1018)	(0.0800)
Distance to province capital		-0.0003	-0.0005	-0.0006
		(0.0006)	(0.0006)	(0.0005)
Urbanization*altitude		0.6970 ~	1.0291 *	0.6072 ~
		(0.3513)	(0.3470)	(0.2745)
Per capita schools in town			0.3598 *	0.1613
			(0.1141)	(0.0951)
Per capita medical centers in town			0.2752	0.3368
			(0.2983)	(0.2428)
Basic needs			-0.2183 *	-0.0704 *
			(0.0104)	(0.0099)
Household size				-0.1158 *
				(0.0042)
Schooling years (household head)				0.0417 *
				(0.0029)
Schooling years (other members)				0.0429 *
				(0.0033)
Potential labor experience				0.0057 *
				(0.0007)
Household head gender				-0.0132
				(0.0255)
Number of migrants				0.0158 ~
				(0.0073)
Spell of illness (household head)				0.0005
				(0.0079)
Savings				0.0310 *
				(0.0068)
Value of durable goods				0.0033
				(0.0022)
Observations	3623	3623	3623	3623
Pseudo R-squared	0.037	0.071	0.176	0.492

Note: Standard deviation in parenthesis and p<0.01=*, p<0.5=~

Model 1: Geography.

Model 2: Geography + localization.

Model 3: Geography + localization + infrastructure.

Model 4: Geography + localization + infrastructure + private assets.

Source: Authors' calculation based on LSMS 1994.

Table 12

Quantile Regressions of (Log) Per Capita Expenditure: 1994

(At household level)

Variables	Percentile: 10	20	60	80	90
Intercept	4.8091 *	5.3829 *	6.6526 *	7.0426 *	6.9805 *
	(0.2790)	(0.2569)	(0.2146)	(0.2401)	(0.3279)
Altitude	-0.0248	-0.0819	-0.1628	-0.3209	0.1202
	(0.3922)	(0.3453)	(0.2602)	(0.2896)	(0.3738)
Temperature	0.0933 *	0.0557 *	0.0195	0.0084	0.0151
	(0.0215)	(0.0197)	(0.0166)	(0.0187)	(0.0256)
Temperature squared	-0.002 *	-0.0009	-0.0001	0.0001	-0.0002
	(0.0006)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
Igneous rock	0.2338 *	0.1043	0.0772	0.0908	0.1196
	(0.0865)	(0.0789)	(0.0614)	(0.0677)	(0.0916)
Sedimentary	0.0052	-0.0165	-0.0266	0.0184	0.0453
	(0.0507)	(0.0465)	(0.0360)	(0.0406)	(0.0542)
Land depth	0.0032 *	0.0023 ~	0.0011	0.0007	0.001
	(0.0011)	(0.0009)	(0.0007)	(0.0008)	(0.0012)
Urbanization	-0.0872	-0.1099	-0.2073 ~	-0.202 ~	-0.0259
	(0.1414)	(0.1280)	(0.0932)	(0.0998)	(0.1295)
Distance to province capital	0.0009	0.0001	-0.0006	-0.0005	-0.0007
	(0.0008)	(0.0007)	(0.0005)	(0.0006)	(0.0008)
Urbanization*altitude	1.0585 ~	0.9463 ~	0.6216 ~	0.4445	0.1177
	(0.4821)	(0.4284)	(0.3112)	(0.3409)	(0.4445)
Per capita schools in town	0.2197	0.2551	0.0254	0.0261	0.2235
	(0.1691)	(0.1478)	(0.1108)	(0.1240)	(0.1682)
Per capita medical centers in town	0.6409	0.2873	0.3552	-0.0034	-0.3481
	(0.4281)	(0.3907)	(0.3049)	(0.3426)	(0.4468)
Basic needs	-0.0917 *	-0.0881 *	-0.0671 *	-0.0442 *	-0.0164
	(0.0169)	(0.0148)	(0.0111)	(0.0125)	(0.0174)
Household size	-0.0955 *	-0.0964 *	-0.1199 *	-0.1224 *	-0.1247 *
	(0.0060)	(0.0054)	(0.0046)	(0.0058)	(0.0085)
Schooling years (household head)	0.0371 *	0.0413 *	0.0356 *	0.0354 *	0.0347 *
	(0.0049)	(0.0044)	(0.0033)	(0.0038)	(0.0052)
Schooling years (other members)	0.05 *	0.0428*	0.0371 *	0.0346 *	0.0346 *
	(0.0053)	(0.0047)	(0.0036)	(0.0041)	(0.0056)
Potential labor experience (household head)	0.0053 *	0.0059 *	0.0047 *	0.0057 *	0.0049 *
	(0.0012)	(0.0011)	(0.0008)	(0.0009)	(0.0011)
Household head gender	-0.0775	-0.0135	-0.024	-0.0198	-0.0307
	(0.0431)	(0.0375)	(0.0287)	(0.0320)	(0.0439)
Number of migrants	0.0245	0.0132	0.0135	0.0097	0.0154
	(0.0126)	(0.0112)	(0.0087)	(0.0100)	(0.0134)
Spell of illness (household head)	-0.0216	-0.0046	0.0134	0.0164	0.0299 ~
	(0.0126)	(0.0111)	(0.0084)	(0.0093)	(0.0125)
Savings	0.0231 *	0.0234 *	0.0311 *	0.0325 *	0.0316 *
	(0.0016)	(0.0064)	(0.0029)	(0.0026)	(0.0025)
Value of durable goods	0.0004	0.0034 ~	0.023 *	0.0309 *	0.0342 *
	(0.0005)	(0.0014)	(0.0005)	(0.0004)	(0.0004)
Observations	3623	3623	3623	3623	3623
Pseudo R-squared	0.2673	0.2764	0.3095	0.3294	0.3454

Group of variables	Joint test: All coefficients equal to zero (Pr>Fstat)				
Geography	0.000	0.000	0.000	0.005	0.421
Localization	0.039	0.076	0.095	0.213	0.792
Infrastructure	0.000	0.000	0.000	0.005	0.477
Private assets	0.000	0.000	0.000	0.000	0.000

Note: Standard deviation in parenthesis and p<0.01=*, p<0.5=~

Source: Authors' calculation based on LSMS 1994.

In addition, our variable that captures the impact of the access to public infrastructure also seems to have a stronger effect on the poorer households. The basic needs variable is negative and significant for the first percentiles and loses its significance for the 90th percentile.

The variables measuring the impact of private assets, mainly schooling years and potential labor experience, are significant and seem to be similar among poor and rich households. On the other hand, the two variables that we use as a proxy for wealth, savings and value of durable assets, become bigger and more significant the richer the household.

Lastly, as mentioned in Section 3, following equation (5) we develop a micro model for consumption growth allowing for constraints on factor mobility and externalities, whereby geographic factors, in the specific region or neighboring regions, can influence the productivity of a household's own capital. For this purpose we develop two households panels, one for 1991-1994 and the other for 1994-1997, and we explain the changes in the difference of logs of expenditure using geographical variables, public infrastructure variables and private assets of the households. The results are shown in Table 13.

As with our previous findings, geographical variables do seem to be significant. Altitude is negative and significant in the last panel. Temperature also reveals its negative effect when its level is too high or too low (the coefficient for temperature is positive while the coefficient for its square term is negative and significant). The absence of public assets, measured through unsatisfied basic needs, also seems to be very important in explaining changes in expenditure differentials between households. Furthermore, private assets, measured by schooling years, again showed themselves to be significant and positive.

Lastly, the lagged expenditure is negative and significant. This can be explained by the reduction in inequality, especially in the period of 1991-1994, for which the Gini coefficient is reduced from 0.369 to 0.364. On the other hand, when recovering the implied λ there is a clear indication of convergence. In this respect, it is important to mention that there is much debate about the possible evidence of convergence and there is not yet a consensus on what is the best method to use for measuring it.[14]

[14] Furthermore, Quah (1993) and Friedman (1992) question the methodology of estimating the convergence rate using the growth and the lagged expenditure variables. They argue that this methodology suffers from the Galton Fallacy.

Table 13

Panel Data Analysis of Per capita Expenditure Growth Rate: 1991-94, 1994-97

(OLS estimation with robust errors including geographic variables)

	Periods			
Variables (final period)	1991-94		1994-97	
Intercept	2.792	*	2.893	*
	(0.266)		(0.306)	
Schooling years (household head)	0.045	*	0.043	*
	(0.004)		(0.004)	
Age (household head)	0.006	*	0.009	*
	(0.001)		(0.001)	
Household head gender (male=1)	-0.115	*	-0.167	*
	(0.037)		(0.048)	
Unsatisfied basic needs	-0.053	*	-0.162	*
	(0.018)		(0.019)	
Altitude	0.536		-0.974	*
	(0.176)		(0.184)	
Temperature	0.047		0.056	~
	(0.025)		(0.025)	
Temperature squared	-0.001	*	-0.002	~
	(0.001)		(0.001)	
Expenditure (initial period)	-0.542	*	-0.578	*
	(0.024)		(0.029)	
Number of observations	1212		900	
R-squared adjusted	0.3136		0.4097	
Gini (initial period)	0.369		0.358	
Gini (final period)	0.364		0.400	
Annual growth rate (%)	10.8		2.3	

Note: Standard deviation in parenthesis and p<0.01=*, p<0.5=~

Gini coefficients and growth rates calculations are based on per capita expenditure

Source: Authors' calculation based on 1991, 1994 and 1997 LSMS.

5.4 Breakdown of Regional Per Capita Expenditure

To disentangle the effect of geography on regional expenditure and expenditure growth we have applied the breakdown technique described in Section 3 to the household-level estimation performed for per capita expenditure and shown in Table 11. For this breakdown we have assumed that parameters are stable across the three main geographic areas: coast, highland and jungle. This initial breakdown is shown in Table 14. In the first column we see that most of the difference in log per capita expenditure between the highland and the coast can be accounted for by the differences in infrastructure endowments and private assets. In other words, once the main geographic variables are accounted for (altitude, temperature and surface characteristics), only private assets and infrastructure endowments are needed to explain regional expenditure differences. Similarly, the second column shows the breakdown of the differences in log per capita expenditure between the jungle area and the coast, showing again that once main geographic variables are accounted for most of the regional expenditure differences can be explained by infrastructure endowment and private asset composition.

Table 14

Decomposition of Regional Per Capita Expenditure Differences

(Log differences)

Group of variables	Highland-Coast	Jungle-Coast
Geography	**-0.163**	**0.031**
Altitude	-0.036	-0.004
Temperature	-0.235 *	0.173 *
Temperature squared	0.117	-0.121
Igneous rock	0.015 ~	-0.004 ~
Sedimentary rock	-0.004	-0.009
Land depth	-0.022	-0.005
Location	**0.050**	**0.039**
Urbanization	0.055	0.038
Distance to province capital	-0.005	0.001
Geography*location	**0.081** ~	**0.007** ~
Urbanization*altitude	0.081 ~	0.007 ~
Infrastructure	**-0.024** ~	**-0.064** ~
Per inhabitant schools in town	0.024	0.023
Per inhabitant medical centers in town	0.010	0.009
Basic needs	-0.058 *	-0.095 *
Private assets	**-0.185** *	**-0.258** *
Household size	-0.031 *	-0.064 *
Schooling years (household head)	-0.061 *	-0.065 *
Schooling years (other members)	-0.069 *	-0.102 *
Potential labor experience	-0.013 *	-0.024 *
Household head gender	0.000	-0.001
Number of migrants	-0.009 ~	-0.005 ~
Spell of illness (household head)	0.000	0.000
Savings	0.002 *	0.000 *
Value of durable goods	-0.003	0.004
Explained	-0.241	-0.244
Residual	0.024	0.077
Total	**-0.217**	**-0.167**

Note: *=p<.01, ~=p<=.05, +=p<.1.

Source: Authors' calculation based on 1994 LSMS.

Obviously, the fact that geography has no additional impact on regional per capita expenditure differences has to do with the fact that key infrastructure variables such as school and medical facilities, access to electricity, water and sanitation, as well as private assets, have dampened the effect of geography on regional expenditure differentials. To see this, Table 15 performs the same breakdown exercise introducing each set of variables sequentially. First, only geography variables are entered in the model, and the breakdown exercise is conducted only with these variables. In this case, geography is highly significant in explaining per capita expenditure differentials between the highland and coastal areas of Peru, as well as between the jungle and coastal areas. Geography remains highly significant even after we introduce location variables and their cross-products into the analysis. However, once infrastructure variables come into play in the analysis, the impact of geography disappears, as the coefficients associated with these types of variables are shown to be jointly non-significant. This could be because, in the models without infrastructure, the geography variables were choosing their effect and therefore when improving our specification the effect of these variables disappears.

Table 15

Decomposition of Regional Per Capita Expenditure Differences, by Model

Group of variables	Highland-Coast				Jungle-Coast			
	1	1+2	1+2+3	1+2+3+4	1	1+2	1+2+3	1+2+3+4
(1) Geography	-0.239 *	-0.162 ~	-0.283 ~	-0.163	-0.152 *	-0.084 ~	-0.052 ~	0.031
(2) Location		-0.181	0.024	0.05		-0.123	0.021	0.039
(3) Geo*location		0.093 *	0.137 *	0.081 ~		0.008 *	0.012 *	0.007 ~
(4) Infrastructure			-0.118 *	-0.024 ~			-0.237 *	-0.064 ~
(6) Private assets				-0.185 *				-0.258 *
Explained	-0.239	-0.250	-0.240	-0.241	-0.152	-0.199	-0.256	-0.244
Residual	0.022	0.033	0.023	0.003	-0.015	0.032	0.089	0.072
Total	-0.217	-0.217	-0.217	-0.217	-0.167	-0.167	-0.167	-0.167

Source: Authors' calculation based on 1994 LSMS.

The same type of breakdown can also be carried out with the per capita expenditure growth equations that we reported in Table 8. In this case, per capita growth rate differentials between highland and coastal regions and between jungle and coastal regions can be broken down into their main determinants: geographical differences, infrastructure differences and asset endowment differences, as reported in Table 16. Here, as was the case with the previous result, geography does not appear to significantly contribute to growth differentials, once infrastructure differences and private asset endowment differences are accounted for. In this case, however, only private asset endowment differentials seem to play an important role in explaining differential growth patterns between highland, jungle and coastal regions.

As was the case in the analysis of differential expenditure levels across regions, the role of geography variables seems to be shadowed by the presence of infrastructure and private asset endowments. To see whether this is the case, Table 17 shows the same breakdown exercise, introducing each set of variables sequentially. First, only geography variables are entered in the model, and the decomposition exercise is conducted only with these variables. In this case geography is highly significant in explaining per capita expenditure growth differentials. However once infrastructure variables are introduced into the analysis, the significance of geography disappears and does not reappear as the remaining variables are introduced. It must be noted that the analysis remains valid even if we correct for possible spatial autocorrelation due to possibly omitted non-geographic spatially correlated variables.

Table 16

Decomposition of Regional Per Capita Expenditure Differences

(Growth rates differences at Province level)

Group of variables	Highland-Coast	Jungle-Coast
Geography	**0.2126**	**0.1296**
Altitude level	0.1182	0.0055
Latitude	-0.0280	0.0471
Longitude	0.0437	0.0396
Soil slope	0.0518	-0.0159
Soil depth	-0.0020	0.0379
Igneous rock	-0.0329 *	0.0222 *
Metamorphic rock	0.0300	0.0399
Temperature	0.0319	-0.0467
Infrastructure	**-0.0431**	**-0.0920**
Basic needs	-0.0431	-0.0920
Geography*Infrastructure	**-0.0125**	**-0.0041**
Altitude*Basic needs	-0.0125	-0.0041
Private assets	**-0.3430** *	**-0.0031** *
School attendance rate	-0.1335 *	-0.0663 *
Female household head (%)	-0.0739 ~	0.0147 ~
Working children (%)	0.0278 ~	0.0090 ~
Household size	-0.0689	0.0580
Household size growth a/	-0.0881 +	-0.0133 +
Number of migrants	-0.0063	-0.0051
Total explained	-0.1860	0.0304
Residual	0.1048	0.0989
Total	**-0.0812**	**0.1293**

a/ Instruments variables are shown in the Appendix.

Note: *=p<.01, ~=p<=.05, +=p<=0.1

Source: Authors' calculation based on 1972 and 1993 Population and House Censuses.

Table 17

Decomposition of Regional Per capita Growth Expenditure Differences, by model
(At province level)

Group of variables	Highland-Coast 1	1+2	1+2+3	1+2+3+4	1+2+3+4 a/	Jungle-Coast 1	1+2	1+2+3	1+2+3+4	1+2+3+4 a/
(1) Geography	-0.163 ~	-0.113	-0.047	0.158	0.213	0.023 ~	0.154	0.136	0.126	0.130
(2) Infrastructure		-0.108 *	-0.075 ~	-0.043	-0.043		-0.229 *	-0.161 ~	-0.091	-0.092
(3) Geo*infrastructure			-0.093	0.004	-0.013			-0.031	0.001	-0.004
(4) Private assets				-0.327 *	-0.343 *				-0.025 *	-0.003 *
Explained	-0.163	-0.221	-0.215	-0.208	-0.186	0.023	-0.075	-0.056	0.012	0.030
Residual	0.082	0.139	0.134	0.127	0.105	0.106	0.205	0.185	0.118	0.099
Total	-0.081	-0.081	-0.081	-0.081	-0.081	0.129	0.129	0.129	0.129	0.129

a/ Modelling first-order spatial error autocorrelation.
Note: *=p<.01, ~=p<=.05, +=p<=0.1.
Source: Authors' calculation based on 1972 and 1993 Population and House Censuses.

6. Conclusions

Peru's enormous geographic diversity makes it an extremely interesting case study for analyzing whether geography has a causal role in determining how household welfare evolves over time. We know that there are huge welfare disparities across Peru, and there is a heavy concentration of very poor people throughout the most geographic adverse regions, as in the *sierra* and *selva.* Although these welfare disparities can be attributed to geography they can also be related, at least in part, to a significant dispersion in access to infrastructure and other public assets. Therefore, there is no clear evidence that regional income differences can only be explained by geography or if they had been hampered (or facilitated) by local or neighboring natural or manmade geographical endowments.

Despite the fact that there have been many efforts to link Peru's geographical diversity to key issues such as settlement location or construction of administrative or political regions, very little has been done to analyze the links between this geographic diversity and development, economic growth or poverty.

To reduce this gap, our research strategy consisted of describing how geography might play a fundamental role in regional economic growth and what relationship there is between geographic variables and expenditure levels and growth across regions within Peru. To formally answer whether geography is a determinant of the evolution of welfare over time, we developed a micro model of consumption which not only took in the local effect of geographic variables but also included public and private assets as variables that could reduce the potentially adverse effect of geography. For this purpose we used national census data for 1972, 1981 and 1993, the LSMS surveys for 1991, 1994, 1996, 1997, information from the district-level infrastructure census, geographical datasets, and information from the third National Agrarian Census of 1994. This cross-sectional analysis helped us in attempting to understand whether geographic externalities arising from local or neighboring public assets, or local endowments of private goods, entail that living in or near a well-endowed area implies that a poor household can eventually escape poverty.

We have shown that what seems to be sizable geographic differences in living standards in Peru can be almost fully explained when one takes into account the spatial concentration of households with readily observable non-geographic characteristics, in particular public and private assets. In other words, the same observationally equivalent household has a similar expenditure level in one place as in another with different geographic characteristics such as altitude or temperature. This does not mean, however that geography is not important, but that its influence on expenditure level and growth differential comes about through a spatially uneven provision of public infrastructure. Furthermore, when we measure the expected gain (or loss) in consumption from living in a geographic region (i.e., coast) as opposed to living in another geographic region (i.e., highlands), we found that most of the difference in log per capita expenditure between the highland and the coast can be accounted for by the differences in infrastructure endowments and private assets. This could be an indication that the availability of infrastructure could be limited by geography and therefore the more adverse geographic regions are the ones with less access to public infrastructure.

Another interesting result is that, despite the fact that in our models of expenditure growth we included all relevant geographic variables, as well as infrastructure and private assets variables, the residuals continue to show spatial autocorrelation. This fact suggests the idea that there may be non-geographic non-observables that may be affecting the provincial expenditure pattern. This is consistent with Ravallion and Wodon (1997) when they show that sizable geographic differences in living standards can persist even if we take into account the spatial concentration of households with readily observable non-geographic characteristics conducive to poverty.

It is important to note that there appear to be non-geographic, spatially correlated omitted variables that need to be taken into account in our expenditure growth model. Therefore policy programs that use regional targeting do have a rationale even if geographic variables do not explain the bulk of the difference in regional growth, once we have taken into account differentials in access to private and public assets.

Lastly, an issue that we had not taken into account, and which could be very important for future research, is the fact that adverse geographic externalities can provide incentives to migration. This is something that we did not control for in this study. The migration effect could be twofold. On the one hand, it could be the reason why households with fewer private assets are those that choose to locate in the more adverse geographical regions. On the other hand, it could be very important for policymaking in developing infrastructure, in the sense that certain investments in infrastructure, such as education, are mobile with migration, while others are not. Therefore, it could be more profitable to invest in mobile infrastructure in the more adverse geographic regions, to give the individuals the necessary tools to migrate from these regions and therefore increase their probability of escaping a poverty trap.

7. Bibliography

Anselin, L. 1980. *Estimation Methods for Spatial Autoregressive Structures*. Regional Science Dissertation and Monograph Series. 8. Ithaca, United States: Cornell University.

----. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

----. 1995a. "Local Indicators of Spatial Association-LISA." *Geographical Analysis.* 27 (2): 93-115.

----. 1995b. *SpaceStat Version 1.80, User's Guide*. Morgantown, United States: West Virginia University, Regional Research Institute.

Anselin, L., Varga, A. and Zoltan, A. 1996. "Local Geographic Spillovers Between University Research and High Technology Innovations: A Spatial Econometric Approach." Research Paper 9606. Morgantown, United States: West Virginia University, Regional Institute and Department of Economics.

Bloom, D. E. and Sachs, J. 1998. "Geography, Demography, and Economic Growth in Africa." Cambridge, United States: Harvard University, Harvard Institute for International Development.

Case, A.C. 1992. "Neighborhood Influence and Technological Change." *Regional Science and Urban Economics.* 22 (0): 491-508.

Cliff, A.D. and Ord, J.K. 1981. *Spatial Processes.* London, United Kingdom: Pion Limited.

----. 1972. "Testing for Spatial Autocorrelation Among Regression Residuals." *Geographical Analysis*. 4: 267-284.

Davis, R. D. and Weinstein, D.E. 1997. "Economic Geography and Regional Production Structure: An Empirical Investigation." NBER Working Paper 6093. Cambridge, United States: National Bureau of Economic Research.

Deaton, A. 1997. *The Analysis of Household Surveys. A Microeconometric Approach to Development Policy*. Washington, DC, United States: World Bank/Johns Hopkins University Press.

Dubin, R. 1992. "Spatial Autocorrelation and Neighborhood Quality." *Regional Science and Urban Economics*. 22 (3): 433-52.

Dumais, G., Ellison, G. and Glaeser, E.L. 1997. " Geographic Concentration as a Dynamic Process." NBER Working Paper No. 6270. Cambridge, United States: National Bureau of Economic Research.

Eichengreen, B. 1998. "Geography as Destiny: A Brief History of Economic Growth." *Foreign Affairs*. 77 (2): 128-33.

Engerman, S. L. and Kenneth L. Sokoloff, K.L. 1998. “Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States.” In: S. Haber, editor. *How Latin American Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914.* Stanford, United States: Stanford University Press.

Escobal, J., Saavedra, J. and Torero, M. 1998. “Los Activos de los Pobres en el Perú.” Research Network Working Paper R-361. Washington, DC, United States: Inter-American Development Bank.

Friedman, M. 1992. “Communication. Do Old Fallacies Ever Die?” *Journal of Economic Literature*. 30: 2129-2132.

FONCODES. 1995. “El Mapa de la Inversión Social: FONCODES frente a la Pobreza 1991-1994.” Lima, Peru: UNICEF.

Gallup, J. L., Sachs, J., and Mellinger, A.D. 1998. “Geography and Economic Development.” Cambridge, United States: Harvard University, Harvard Institute for International Development. Mimeographed document.

Gallup, J.L. 1998. “Agriculture Productivity and Geography.” Cambridge, United States: Harvard University, Harvard Institute for International Development.

Getis, A., and Ord, K. 1995a. “The Analysis of Spatial Association by Use of Distance Statistics.” *Geographical Analysis*. 24: 189-206.

----. 1995b. “Local Spatial Autocorrelation Statistics: Distributional Issues and an Application.” *Geographical Analysis*. 27 (4): 286-306.

Hall, R. E and Jones, C. 1998. “Why do Some Countries Produce So much More Output per Worker than Other?” Stanford, United States: Stanford University. Mimeographed document.

----. 1997. “Levels of Economic Activity Across Countries.” Paper presented at the American Economic Association Annual Meeting, New Orleans, United States.

Hentschel, J., Lanjouw, J.O., Peter Lanjouw, P. *et al.* 1998. “Combining Census and Survey Data to Study Spatial Dimensions of Poverty: A Case Study of Ecuador.” Washington, DC, United States: World Bank. Mimeographed document.

Heshmati, A. and Kumbhakar, S.C. 1997. “ Estimation of Technical Efficiency in Swedish Crop Farms: A Pseudo Panel Data Approach.” *Journal of Agricultural Economics*. 48 (1): 22-37.

Hordijk, L. 1974. “Spatial Correlation in the Disturbance of a Linear Interregional Model.” *Regional Science and Urban Economics*. 4: 117-40.

Hordijk, L and Nijkamp, P. 1977. “Dynamic Models of Spatial Autocorrelation.” *Environment and Planning A*. 9: 505-519.

Hordijk, L. and Paekinck, J. 1976. “Some Principles and Results in Spatial Econometrics.” *Recherches Economiques de Louvain*. 42: 175-97.

Jalan, J. and Ravallion, M. 1998. “Geographic Poverty Traps?” IED Discussion Paper Series 86. Boston, United States: Boston University, Institute for Economic Development.

Kim, S. 1997. “Regions, Resources, and Economic Geography: Sources of U.S. Regional Comparative Advantage, 1880-1987.” NBER Working Paper 6322. Cambridge, United States: National Bureau of Economic Research.

Knight, J. and Song, L. 1993. “The Spatial Contribution to Income Inequality in Rural China.” *Cambridge Journal of Economics*. 17 (2): 195-213.

Landes, D.S. 1998. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York, United States: W.W. Norton.

Lyons, T. 1991. “Interprovincial Disparities in China: Output and Consumption, 1952-1987.” *Economic Development and Cultural Change*. 39 (3): 471-506.

Moran, P.A.P. 1950. “Notes on Continuous Stochastic Processes.” *Biometrika*. 37: 17-23.

Moreno, R. and Trehan, B. 1997. “Location and the Growth of Nations.” San Francisco, United States: Federal Reserve Bank of San Francisco, Research Department. Mimeographed document.

Nass, C. Garfinkle, D. 1992. “Localized Autocorrelation Diagnostic Statistic (LADS) for Spatial Models: Conceptualization, Utilization, and Computation.” *Regional Science and Urban Economics*. 22: 333-346.

Peñaherrera, C.1986. “El desarrollo de la geografía en el Perú.” In: E. Yepes, editor. *Estudios de historia de la ciencia en el Perú: ciencias sociales*. Volume 2. Lima, Peru: Sociedad Peruana de Historia de la Ciencia y la Tecnología.

Pulgar Vidal, J. 1946. *Geografía del Perú: las ocho regiones naturales*. Tenth edition, 1986. Lima, Peru: Editorial PEISA.

Quah, D. 1993. “Galton’s Fallacy and Tests of the Convergence Hypothesis.” *Scandinavian Journal of Economics*. 95 (4): 427-443.

Ravallion, M. and Jalan, J. 1996. “Growth Divergence Due to Spatial Externalities.” *Economic Letters*. 53 (2): 227-232.

Ravallion, M. and Wodon, Q. 1997. “Poor Areas, or Only Poor People?” Policy Research Working Paper 1798. Washington, DC, United States: World Bank.

Verbeek, M. 1992. “Pseudo Panel Data.” In: L. Matyas and P. Sevestre, editors. *The Econometrics of Panel Data: Handbook of Theory and Applications.* Dordrecht, The Netherlands: Kluwer Academic Publishing.

World Bank. 1999. Poverty and Social Development in Peru: 1994-97. Washington, DC, United States: World Bank.

Data Sources

At household level

- Living Standard Measurement Surveys 1985-86 and 1994, Cuanto Institute.

At province level

- Population and Household Censuses 1972, 1981 and 1994, Instituto Nacional de Estadística e Informática: population and household characteristics.

- Third National Agrarian Census 1994, Instituto Nacional de Estadística e Informática: agricultural variables, cattle and land.

- Basic Needs Map 1994. Instituto Nacional de Estadística e Informática: basic needs and health variables

- Social Investment Map 1994, FONCODES: poverty index and its components, living standard.

Geographic variables

- Arc data online in: http://www.esri.com/data/online/esri/wothphysic.html. This information was afterwards overlaid on a map of Peru at provincial and district levels. The score for each province or district was selected according to the position of its centroid on the thematic map: earthquake zones, precipitation, soils and vegetation.

- Natural Resources in Peru 1995, Instituto Nacional de Recursos Naturales: bioclimate and land potential scores.

- Social Investment Map 1994, FONCODES: altitude and geographic location.

Annex 1: Measuring Geographical Association: Theoretical Framework

The importance of spatial relationships began to be recognized in the seventies with the works of Cliff and Ord (1972) in the United Kingdom, and Hordijk (1974), Hordijk and Pelinck (1976) and Hordijk and Nijkamp (1977). These studies created a great interest in the development of a methodology for the study of observations distributed in a specific geographical location and gave birth to what is called "spatial econometrics."

Spatial autocorrelation says that what is observed in one place is in part determined by what is occurring in the other spatial locations. So, any observation of a variable *y* in *i* (where *i* is an element of a population *S*), is related formally through a function *f* to the magnitudes of the variable in other spatial units in the system.

(a.1) $$y_i = f\ (y_1, y_2, \ldots y_{i-1}, y_{i+1}, \ldots y_n)$$

There are a large number of tests to detect the presence of spatial correlation (Anselin, 1988), but those that are most used are the "Moran Statistic" (I) and the G-statistics (Getis and Ord, 1992).

The Moran Statistic

Formally, Moran's *I* is:

(a.2) $$I = \frac{N}{S_o} * \frac{\sum_i \sum_j w_{ij}(x_i - m).(x_j - m)}{\sum_i (x_i - m)^2}$$

where N is the number of observations, x and x_j are observations for location i and j (with mean m), and w_{ij} is the element in the spatial weight matrix corresponding to the observation pair *i*,*j*. The *W* used here can be characterized: $W=\{w_{ij}\}$ such that $w_{ij}=1$ if *i* and *j* are neighbors, $w_{ij}=0$ otherwise, and $w_{ii}=0$ for all *i*. The rows of *W* are then normalized such that each observation's neighbors have the same amount of influence, that is $\sum_j w_{ij} = 1$, for all *i*. In addition it will be assumed that each neighbor of a given farm carries equal weight, $w_{ij=}\ w_{ik}$ for non-zero elements (neighbors) k and *j* for farmer *i*. If more information were available about the amount of influence each household exercises, this could be incorporated into the *W* matrix (regarding the different structures, see Anselin, 1988).

S_o is a scaling constant:

(a.3) $$S_o = \sum_i \sum_j w_{ij}$$

i.e., the sum of all weights. For a row standardized spatial matrix, which is the preferred way to implement the test and the way it is done in this paper, S_o equals N (since each row sums to 1), and the statistic simplifies to the ratio of a spatial cross product to a variance:

(a.4) $$I = \frac{\sum_i \sum_j w_{ij}(x_i - m).(x_j - m)}{\sum_i (x_i - m)^2}$$

Moran's *I* is similar but not equivalent to a correlation coefficient and is not centered around 0. In fact, the theoretical mean of Moran's *I* is $-1/N$-1. In other words, the expected value is negative and is only a function of the sample size (*N*). Note however, that this mean will tend to zero as the sample size increases.

Instead of using the *I* statistics by themselves, inference is typically based on a standardized z-value. This is computed by subtracting the theoretical mean and dividing the result by the theoretical standard deviation.

(a.5) $$z_I = \frac{(I - E(I))}{SD(I)}$$

where E(I) is the theoretical mean and SD(I) is the theoretical standard deviation. For a technical discussion and detailed expressions for the moments see Cliff and Ord (1973, 1981). The most common approach is to assume that the variable in question follows a normal distribution. Based on asymptotic considerations (i.e., by assuming that the sample may became infinitely large) the z-value, using the proper measures for mean and standard deviation, follows a standard normal distribution (i.e., normal distribution with mean 0 and variance 1). Significance of the statistic can then be judged by comparing the computed z-value to its probability in a standard normal table (see Case, 1987).

Deriving the G and G Statistic*

The Getis and Ord (1992) statistic is used as a validation of the Moran *I*. Getis and Ord introduced a family of statistics, *G*, that can be used as measures of spatial association in a number of circumstances (see Getis and Ord, 1992, and chapter 23 in the *SpaceStat Tutorial* for general background).

Formally, the *G* statistic, for a chosen critical distance *d*, *G(d)*, is defined as:

(a.6) $$G(d) = \frac{\sum_i \sum_j wij(d) x_i x_j}{\sum_i \sum_j x_i x_j}$$

where x_i is the value observed at location *i*, and w_{ij} (*d*) stands for an element of the symmetric (non-standardized) spatial weights matrix for distance *d*. The numerator of the statistic is similar to that of Moran's *I*, but its denominator is different. Its significance is assessed by means of a standardized z-value, obtained in the usual fashion. The mean and variance of the *G(d)* statistic can be derived under a randomization assumption and the z-value can be shown to tend to a standard normal variable in the limit. (See Getis and Ord, 1992, for detailed derivations.)

For each observation *i*, the G_i and G_i* statistics indicate the extent to which that location is surrounded by high values or low values for the variable under consideration, for a given distance *d*. Formally, the G_i *and* G_i* statistics are defined as:

$$G_i = \frac{\sum_j wij(d) x_j}{\sum_j x_j} \qquad \text{(a.7)}$$

where $w_{ij}(d)$ are the elements from the contiguity matrix for distance *d*. The G_i and G_i* measures differ with respect to the number of observations that are included in the computation of the denominator. For the G_i statistic, $j \neq i$ while for the G_i* statistic $j = i$ is included in the sum. In other words, the G_i* measure provides a measure of spatial clustering that includes the observation under consideration, while the G_i measure does not.

Inference about the significance of the G_i and G_i* statistics is based on a standardized z-value, which is computed by substituting the theoretical mean and dividing by the theoretical standard deviation (for more details see Getis and Ord, 1992).

A positive and significant z-value for a *G*, G_i or G_i* statistic indicates spatial clustering of high values, whereas a negative and significant z-value indicates spatial clustering of low values. Note that this interpretation is different from that of the more traditional measures of spatial autocorrelation, such as the Moran *I*, where spatial clustering of like values, either high or low, is indicated by positive autocorrelation.

Local Indicators of Spatial Association (LISA)

In Anselin (1995), a local indicator of spatial association (LISA) is defined, showing how it allows for the breakdown of global indicators, such as Moran's *I*, into the contribution of each observation. The LISA statistics serve two purposes. On one hand, they may be interpreted as indicators of local pockets of nonstationarity, or hot spots, similar to the G_i and G_i^* of Getis and Ord (1992). On the other hand, they may be used to assess the influence of individual locations on the magnitude of the global statistic and to identify outliers, as in Anselin's Moran scatterplot (1993a). Both of these uses will help in determining which locations have the greatest correlation with their neighbors.

The LISA for a variable y_i, observed at location *i*, can be expressed as a statistic L_i, such that:

$$L_i = f(y_i, y_{J_i}) \qquad \text{(a.8)}$$

Where *f* is a function (possibly including additional parameters), and the y_{Ji}, are the values observed in the neighborhood J_i of *i*.

Similar to the rationale behind the significance tests for G_i and G_i^* statistics of Getis and Ord (1992), the general LISA can be used as the basis for a test on the null hypothesis of no local spatial association. However, in contrast to what holds for the G_i and G_i^* statistics, general results on the distribution of a generic LISA may be hard to obtain.

As a special case of the local Gamma,[15] a local Moran statistic for an observation *I* may be defined as:

$$I_i = z_i \sum_j w_{ij} z_j \qquad (a.9)$$

where, analogous to the global *Moran's I*, the observations z_i, z_j are in deviations from the mean, and the summation over *j* is such that only neighboring values *j* element of J_i are included. For ease of interpretation the weights w_{ij} may be in row-standardized form, though it is not necessary, and by convention, w_{ii}=0.

It can be easily seen that the corresponding global statistic is indeed the familiar *Moran's I*. The sum of the local Morans is:

$$\sum_i I_i = \sum_i z_i \sum_j w_{ij} z_j , \qquad (a.10)$$

The moments of I_i under the null hypothesis of no spatial association can be derived using the principles outlined by Cliff and Ord (1981, pp.42-46) and a reasoning similar to that employed by Getis and Ord (1992, pp. 1990-92).

A test for significant local spatial association is based on these moments, although, as mentioned by Anselin (1995), the exact distribution of such a statistic is still unknown.

[15] See Anselin (1995).

Annex 2: Data Description

A.2.1. Province-Level Per Capita Expenditure Estimates

To estimate per capita expenditure at the province level for Census years 1972, 1981 and 1993, we estimated a household-level expenditure equation based on the information available in the LSMS surveys for 1985-86 and 1994. Following Escobal *et al*. (1998) we regress per capita expenditure on private and public assets, allowing interactions between them. A more detailed discussion of these estimations can be found in Escobal *et al*. (1998).

Table A.2.1 shows the results of this procedure. The endogenous variable in each equation was the per capita expenditure in constant Nuevos Soles of 1994. From the coefficients obtained in Table A.2.1, we simulated the province-level per capita expenditure using the province-level variables obtained from the Census data, and the means of the household surveys whenever there was not a counterpart variable in the census. For 1972 and 1981 we used the parameters of LSMS 1985-86 and for 1993 the calculations of LSMS 1994, due to the proximity of the sample surveys and Census dates.

The province-level variables used in all Census years were: household size, percentage of houses without access to potable water, without drainage, without electricity, total illiteracy rate, schooling attendance rate, percentage of child laborers and percentage of population living in urban areas. Additionally, for 1993 we included the percentage of the non-professional, economically active population, percentage of households headed by women, and college attendance rate. We complete the set of variables (to estimate province-level expenditures) using sample average values of the LSMS by regions. As we mention above, LSMS data are divided into geographical regions to improve the quality of the sampling. These regions were included in the regression as dummy variables associated with location: northern coast, central highland, and greater Lima, for example.

Per capita expenditure at the province level in each Census year was adjusted to reproduce the Aggregate Consumption growth rate of National Accounts within those years. Using 1981 as an anchor, we changed slightly the intercept coefficients of the other regressions to re-estimate the projected variables. Thus, we replace the OLS estimated coefficients 6.690 with 6.350 and 7.695 with 7.595 for 1993 and 1972, respectively. In this way the growth rate of projected per capita expenditure (weighted by population in each year) is equal to the macroeconomic statistics. The coefficients reported in Table A.2.1 display the new values for the intercepts.

Finally, the number of provinces has not remained constant in the last 30 years. In 1972 the number of provinces was 150, in 1981, 153, and in 1993, 188. We therefore had to homogenize the province areas and shapes through time. With this purpose we decided to use the political-administrative division of Peru in 1993 because the Geographical Information System (GIS) was developed following the 1993 Census. To impute the values in 1972 for new provinces we repeated the "original" province information in each of its new regions or areas. For 1981 we had district-level data and, since the creation of a new province is basically a new clustering of districts, we aggregate those district values to create data for the new provinces.

Table A.2.1
Determinants of (Log) Per Capita expenditure
(OLS estimation with robust errors)

	Census year					
	1972 a/		1981 a/		1993 b/	
Variables	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
Intercept	7.6959	(0.1954)	7.7777	(0.3271)	6.3502	(0.1377)
Access to credit	0.1384	(0.0399)	0.1351	(0.0364)	0.0826	(0.0366)
Access to drinking water	-0.1051	(0.0589)	-0.1316	(0.0535)		
Access to electricity	0.0846	(0.0541)	0.0788	(0.0497)	0.0021	(0.0004)
Access to in-house drainage services	0.1165	(0.1455)	0.1032	(0.1030)	0.0016	(0.0009)
Cattle	0.1288	(0.0827)	0.1368	(0.0800)	0.0913	(0.0788)
Durable goods	0.0680	(0.0092)	0.0681	(0.0087)	0.0051	(0.0046)
Fertilizers usage	0.1619	(0.0436)	0.1839	(0.0414)	0.1056	(0.0327)
Household head gender	0.0278	(0.0627)	-0.0035	(0.0523)		
Household members with secondary education (%)			0.0031	(0.0023)		
House with inadequate floor	-0.0042	(0.0009)	-0.0038	(0.0008)	-0.0021	(0.0003)
Household size	-0.2760	(0.0341)	-0.3361	(0.0306)	-0.3253	(0.0283)
Illiteracy rate	-0.0017	(0.0008)	-0.0012	(0.0008)	-0.0016	(0.0007)
School attendance (children)	0.0010	(0.0006)	0.0006	(0.0006)		
Land size	0.0432	(0.0503)	0.0185	(0.0413)		
Number of migrants (household members)	-0.0061	(0.0410)	-0.0039	(0.0409)	0.1359	(0.0261)
Number of rooms in the house	0.0050	(0.0015)	0.0041	(0.0013)	0.0562	(0.0108)
Non-professional labor force			0.0002	(0.0028)		
Potential work experience	-0.0001	(0.0065)	0.0002	(0.0057)	0.0153	(0.0058)
Savings	0.0772	(0.0343)	0.0471	(0.0349)	0.0775	(0.0359)
Schooling attendance rate					0.0004	(0.0004)
Schoolling years (household head)	0.0167	(0.0119)	0.0168	(0.0114)	0.0310	(0.0073)
Schoolling years (other members)	0.0372	(0.0188)	0.0388	(0.0160)	0.0326	(0.0070)
Seeds usage	0.1419	(0.0366)	0.1390	(0.0335)	0.0798	(0.0322)
Social networks	0.2282	(0.0601)	0.2197	(0.0620)	0.0862	(0.1102)
Spell of illness (household head)	0.0153	(0.0299)	0.0268	(0.0299)	-0.0516	(0.0326)
Urban zone	0.0064	(0.0021)	0.0092	(0.0034)	0.0176	(0.1592)
Working children (%)	-0.0014	(0.0005)	-0.0013	(0.0005)		
Northern coast	-0.1374	(0.0334)	-0.1408	(0.0321)	-0.0460	(0.0257)
Central coast	-0.1991	(0.0375)	-0.2033	(0.0393)	-0.0304	(0.0332)
Southern coast	-0.0352	(0.0595)	-0.0552	(0.0642)	-0.0939	(0.0490)
Northern highlands	-0.5987	(0.0541)	-0.5789	(0.0508)	0.1185	(0.0358)
Central highlands	-0.3599	(0.0379)	-0.3670	(0.0374)	-0.0564	(0.0267)
Southern highlands	-0.7135	(0.0365)	-0.0413	(0.0356)	-0.0769	(0.0287)
Northern high altitude jungle	-0.4818	(0.0579)	-0.4313	(0.0583)	-0.2987	(0.0488)
Central high altitude jungle	-0.4875	(0.0547)	-0.4324	(0.0509)	-0.2745	(0.0501)
Low altitude jungle					-0.2327	(0.0561)
Durable goods (squared)	-8.59E-04	(0.0003)	-8.07E-04	(0.0002)	-7.72E-06	(0.0000)
Household size (squared)	0.0120	(0.0024)	0.0156	(0.0021)	0.0153	(0.0020)
Number of migrants (household members) squared	0.0002	(0.0072)	-0.0019	(0.0073)		
Potential work experience (squared)	1.07E-05	(0.0001)	-3.00E-05	(0.0001)	-1.63E-04	(0.0001)
Savings (squared)	0.0002	(0.0003)	0.0004	(0.0003)	-0.0015	(0.0007)
Schoolling years (other members, squared)	-0.0020	(0.0022)	-0.0034	(0.0021)		
Spell of illness (household head) squared					0.0002	(0.0063)
Durable goods*social networks	-0.0060	(0.0022)	-0.0035	(0.0021)	0.0007	(0.0037)
Household size*potential work experience	0.0001	(0.0003)	0.0004	(0.0003)	0.0001	(0.0002)
Household size*savings	-0.0065	(0.0033)	-0.0053	(0.0036)	-0.0032	(0.0017)
Household size*spell of illness	0.0011	(0.0078)	0.0020	(0.0084)	0.0076	(0.0135)
Number of migrants*durable goods	-0.0002	(0.0005)	-0.0003	(0.0006)	0.0005	(0.0009)
Number of migrants*land size	0.0296	(0.0319)	0.0227	(0.0354)	0.0596	(0.0506)
Number of migrants*savings	0.0043	(0.0023)	0.0040	(0.0026)	-0.0004	(0.0030)
Potential work experience*durables goods	-0.0001	(0.0001)	-0.0001	(0.0001)	0.0000	(0.0001)
Potential work experience*number of migrants	-0.0003	(0.0006)	0.0001	(0.0006)	-0.0017	(0.0006)
Potential work experience*savings	-0.0005	(0.0004)	-0.0004	(0.0004)	0.0002	(0.0004)
Potential work experience*spells of illness	-0.0001	(0.0006)	-0.0003	(0.0006)	0.0007	(0.0006)
Savings*durable goods	-5.06E-05	(0.0002)	-2.19E-05	(0.0002)	-2.12E-04	(0.0001)
Schooling years (household head)*durable goods	-0.0001	(0.0003)	-0.0003	(0.0003)	-0.0006	(0.0003)

continued...

Table A.2.1
Determinants of (Log) per-capita expenditure
(OLS estimation with robust errors)

	Census year					
	1972 a/		1981 a/		1993 b/	
Variables	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
						...conclusion
Schooling years (household head)*land size	-0.0113	(0.0120)	-0.0053	(0.0102)	0.0092	(0.0089)
Schooling years (household head)*potential work experience	-0.0001	(0.0002)	0.0000	(0.0002)	-0.0002	(0.0002)
Schooling years (household head)*potential work experience	0.0023	(0.0019)	0.0027	(0.0020)	-0.0067	(0.0016)
Schooling years (household head)*savings	-0.0044	(0.0016)	-0.0044	(0.0017)	0.0003	(0.0013)
Schooling years (household head)*spells of illness	-0.0026	(0.0023)	-0.0013	(0.0022)	0.0056	(0.0017)
Spell of illness*durable goods	0.0005	(0.0007)	0.0002	(0.0007)	-0.0001	(0.0006)
Spell of illness*number of migrants	-0.0024	(0.0044)	-0.0028	(0.0045)	-0.0014	(0.0057)
Spell of illness*savings	0.0042	(0.0024)	0.0024	(0.0026)	-0.0006	(0.0033)
Urban zone*household head gender	-7.85E-05	(0.0007)	1.95E-04	(0.0006)		
Urban zone*land size	0.0007	(0.0013)	0.0001	(0.0012)		
Urban zone*savings (squared)	-6.82E-06	(0.0000)	-8.07E-06	(0.0000)	1.29E-03	(0.0006)
Urban zone*schooling years (household head, squared)	7.18E-05	(0.0001)	4.79E-05	(0.0001)	6.57E-03	(0.0066)
Urban zone*schooling years (other member)	-0.0001	(0.0002)	-0.0002	(0.0002)	-0.0015	(0.0079)
Urban zone*schooling years (other member, squared)	2.20E-05	(0.0000)	3.07E-05	(0.0000)		
Urban zone*access to credit	0.0004	(0.0005)	0.0004	(0.0004)	0.0560	(0.0540)
Urban zone*access to drinking water	0.0009	(0.0007)	0.0010	(0.0006)		
Urban zone*access to electricity	-1.31E-04	(0.0007)	-4.18E-05	(0.0006)	-7.86E-04	(0.0006)
Urban zone*access to in-house drainage services	-0.0003	(0.0015)	-0.0001	(0.0011)	-0.0006	(0.0009)
Urban zone*cattle	-0.0009	(0.0013)	-0.0004	(0.0012)	-0.0223	(0.1018)
Urban zone*durable goods	-0.0003	(0.0001)	-0.0003	(0.0001)	0.0519	(0.0056)
Urban zone*durable goods (squared)	6.12E-06	(0.0000)	5.38E-06	(0.0000)	-3.06E-04	(0.0000)
Urban zone*fertilizers usage	-0.0011	(0.0008)	-0.0011	(0.0008)	-0.1592	(0.0816)
Urban zone*household size	0.0009	(0.0004)	0.0013	(0.0003)	0.0609	(0.0326)
Urban zone*household size (squared)	-0.0001	(0.0000)	-0.0001	(0.0000)	-0.0054	(0.0024)
Urban zone*illiteracy rate	7.28E-06	(0.0000)	6.38E-06	(0.0000)	7.38E-04	(0.0010)
Urban zone*number of migrants	0.0001	(0.0001)	0.0001	(0.0001)		
Urban zone*number of migrants (squared)	-0.0001	(0.0004)	-0.0003	(0.0004)		
Urban zone*number of room in the house	-2.31E-05	(0.0000)	-3.27E-05	(0.0000)	-0.0004	(0.0122)
Urban zone*pesticides usage	0.2702	(0.0764)	0.3074	(0.0659)	0.1272	(0.0326)
Urban zone*potential work experience	0.0001	(0.0001)	0.0001	(0.0001)	-0.0032	(0.0059)
Urban zone*potential work experience (squared)	-7.84E-07	(0.0000)	-1.12E-06	(0.0000)	0.0001	(0.0001)
Urban zone*savings	0.0006	(0.0003)	0.0008	(0.0003)	-0.0535	(0.0255)
Urban zone*schoolling attendance rate					0.0006	(0.0005)
Urban zone*seeds usage	-0.0024	(0.0008)	-0.0017	(0.0007)	0.0109	(0.0830)
Urban zone*social networks	-0.0009	(0.0005)	-0.0011	(0.0005)	0.0554	(0.0770)
Urban zone*spells of illness	0.0003	(0.0002)	0.0001	(0.0002)		
Urban zone*Urban zone*inadequate floor	4.02E-05	(0.0000)	3.51E-05	(0.0000)	0.0004	(0.0005)
Urban zone*working children	2.04E-05	(0.0000)	1.62E-05	(0.0000)	-0.0989	(0.0863)
Number of observation	4949		4949		3623	
R-squared	0.7546		0.7612		0.8596	

a/ Based on 1985-86 LSMS.

b/ Based on 1994 LSMS.

Note: Standard deviation in parenthesis and p<0.01=*, p<0.5=~

Authors' calculation based on LSMS 1985-86 and 1994.

Annex 3: Results of Spatial Autocorrelation at the Province Level

Table A.3.1
Spatial Correlations at Province Level

Variables	Moran Index	Z-Value	Geary Index	Z-Value
South latitude	0.9302	20.21 *	0.057	-18.76 *
North longitude	0.8870	19.27 *	0.093	-18.04 *
Precipitation	0.7573	16.47 *	0.259	-14.73 *
Household size 1993	0.7495	16.30 *	0.241	-15.10 *
Temperature (average)	0.7486	16.29 *	0.256	-14.79 *
Temperature (min.)	0.7469	16.25 *	0.255	-14.83 *
Temperature (max.)	0.7422	16.15 *	0.265	-14.62 *
Altitude of the district capital (meters over sea level)	0.6693	14.57 *	0.322	-13.47 *
% household head that are female 1993	0.6560	14.28 *	0.325	-13.43 *
Inadequate floor	0.6518	14.19 *	0.339	-13.16 *
Soil depth	0.6422	13.99 *	0.328	-13.37 *
Total illiteracy rate 1981	0.6352	13.83 *	0.356	-12.82 *
Overcrowded houses 1993	0.6286	13.69 *	0.339	-13.15 *
Household size 1981	0.6130	13.35 *	0.377	-12.39 *
Per capita expenditure in 1981	0.6084	13.26 *	0.399	-11.95 *
Perimeter of the province	0.6032	13.14 *	0.390	-12.12 *
Illiteracy rate of women 1993	0.6030	13.14 *	0.389	-12.16 *
Igneous rock	0.5994	13.06 *	0.389	-12.14 *
Total illiteracy rate 1993	0.5977	13.02 *	0.397	-11.99 *
Illiteracy rate of women 1981	0.5948	12.96 *	0.386	-12.20 *
Malnutrition rate 1993	0.5871	12.80 *	0.389	-12.14 *
Schooling years 1993	0.5833	12.71 *	0.396	-12.02 *
Potential bioclimate score	0.5798	12.64 *	0.412	-11.68 *
Forestry land potential score	0.5798	12.64 *	0.425	-11.43 *
% Urban population in 1993	0.5781	12.60 *	0.437	-11.19 *
Soil slope	0.5750	12.53 *	0.395	-12.02 *
Population 1993	0.5740	12.51 *	0.440	-11.13 *
Forestry potential bioclimate score	0.5738	12.51 *	0.432	-11.30 *
Natural Resources score	0.5721	12.47 *	0.413	-11.67 *
Total area of the province	0.5712	12.45 *	0.351	-12.91 *
Living Standard 1993 according to FONCODES	0.5609	12.23 *	0.436	-11.22 *
% households without electric appliances 1993	0.5577	12.16 *	0.426	-11.41 *
Illiteracy rate of men 1993	0.5558	12.12 *	0.441	-11.12 *
Rural basic needs: household head with low schooling 1993	0.5536	12.07 *	0.419	-11.55 *
Number of rooms per house 1993	0.5521	12.04 *	0.424	-11.45 *
Urban basic needs: household head with low schooling 1993	0.5392	11.76 *	0.464	-10.66 *
Urban basic needs: house with inadequate characteristics 1993	0.5382	11.74 *	0.450	-10.95 *
FONCODES Poverty Index 1996	0.5372	11.72 *	0.459	-10.75 *
Total illiteracy rate 1972	0.5352	11.68 *	0.453	-10.87 *
Total land potential score	0.5344	11.66 *	0.447	-11.01 *
Change per capita expenditure 81-93	0.5267	11.49 *	0.462	-10.71 *
Per capita expenditure in 1993	0.5265	11.49 *	0.457	-10.81 *
Household size 1972	0.5183	11.31 *	0.471	-10.52 *
School attendance 1993	0.5074	11.07 *	0.475	-10.44 *
Child mortality rate 1993	0.5070	11.07 *	0.481	-10.31 *
Rate of migration 1988-93	0.5056	11.04 *	0.514	-9.66 *
FONCODES Poverty Ranking 1996	0.5023	10.96 *	0.491	-10.12 *
School attendance 1981	0.5004	10.92 *	0.481	-10.33 *
Elementary school attendance 1981	0.4940	10.78 *	0.496	-10.02 *
School attendance 1972	0.4861	10.61 *	0.493	-10.08 *
Agriculture land potential score	0.4833	10.55 *	0.490	-10.15 *
Agriculture potential bioclimate score	0.4825	10.54 *	0.501	-9.93 *
Climate II zones	0.4731	10.33 *	0.511	-9.72 *
Unsatisfied basic needs	0.4590	10.03 *	0.546	-9.02 *
Cattle potential bioclimate score	0.4446	9.72 *	0.539	-9.17 *
Per capita expenditure in 1972	0.4399	9.62 *	0.608	-7.80 *

...continued

Table A.3.1
Spatial Correlations at Province Level

Variables	Moran Index	Z-Value	Geary Index	Z-Value
% households without electric appliances 1972	0.4398	9.61 *	0.548	-9.00 *
Access to drinking water 1993	0.4376	9.57 *	0.558	-8.79 *
Earthquake zone	0.4306	9.42 *	0.558	-8.79 *
Metamorphic rock	0.4221	9.23 *	0.564	-8.66 *
Change in female household headed 1972-1993	0.4214	9.22 *	0.596	-8.04 *
Climate I zones	0.4173	9.13 *	0.583	-8.30 *
% of children who works 1981	0.4124	9.02 *	0.571	-8.53 *
Total fallow cropping land	0.4111	8.99 *	0.639	-7.18 *
Access to electricity 1993	0.4081	8.93 *	0.584	-8.28 *
Change in illiteracy rate 72-93	0.3858	8.45 *	0.640	-7.16 *
School attendance 72-93	0.3857	8.44 *	0.614	-7.67 *
Cattle land potential score	0.3842	8.41 *	0.614	-7.67 *
Change in illiteracy rate 72-81	0.3814	8.35 *	0.614	-7.68 *
Access to electricity 1981	0.3812	8.35 *	0.600	-7.95 *
Access to sanitation services 1993	0.3811	8.35 *	0.608	-7.79 *
Types of natural resources	0.3804	8.33 *	0.615	-7.65 *
Change in per capita expenditure 72-81	0.3762	8.24 *	0.619	-7.59 *
% of rural population 1993	0.3643	7.98 *	0.658	-6.80 *
Access to sanitation services 1981	0.3577	7.84 *	0.625	-7.46 *
% of permanent crops sell consumed	0.3540	7.76 *	0.635	-7.26 *
Access to drinking water 1981	0.3505	7.68 *	0.622	-7.51 *
Rural population 1993	0.3501	7.68 *	0.658	-6.79 *
Change in women illiteracy rate 72-93	0.3486	7.65 *	0.675	-6.47 *
School attendance 81-93	0.3329	7.31 *	0.643	-7.11 *
Non-irrigated land for temporal crops	0.3327	7.30 *	0.663	-6.70 *
Total fallow land	0.3285	7.21 *	0.731	-5.34 *
Economic Active Population without profession 1993	0.3281	7.20 *	0.674	-6.48 *
Total temporal irrigated land	0.3241	7.11 *	0.680	-6.36 *
% household head that are female 1972	0.3238	7.11 *	0.669	-6.59 *
Annual per capita income 1981	0.3234	7.10 *	0.655	-6.86 *
Inadequate ceiling 1993	0.3173	6.97 *	0.665	-6.67 *
Total grassland	0.3148	6.91 *	0.742	-5.13 *
% of children that work 1993	0.3050	6.70 *	0.639	-7.18 *
Temporal crops sold in farm	0.2984	6.56 *	0.683	-6.30 *
Number of medics 1993	0.2895	6.37 *	0.730	-5.38 *
Number of rural houses 1993	0.2883	6.34 *	0.722	-5.53 *
Total agricultural land	0.2880	6.34 *	0.728	-5.41 *
Total used land crop	0.2784	6.13 *	0.718	-5.61 *
Change in access to drinking water 81-93	0.2755	6.07 *	0.717	-5.63 *
Total agrarian units	0.2751	6.06 *	0.731	-5.36 *
Change in access to sanitation services 72-93	0.2713	5.97 *	0.790	-4.17 *
Change household size 81-93	0.2700	5.95 *	0.695	-6.06 *
Change in access to drinking water 72-93	0.2698	5.94 *	0.781	-4.35 *
Change in access to sanitation services 81-93	0.2691	5.93 *	0.723	-5.51 *
Number of rural towns	0.2685	5.91 *	0.786	-4.26 *
Land of crops sold in farm	0.2677	5.90 *	0.704	-5.88 *
Inadequate ceiling 1981	0.2643	5.82 *	0.727	-5.43 *
Total harvested land	0.2639	5.81 *	0.737	-5.22 *
Temporal crops self consumed	0.2622	5.78 *	0.790	-4.18 *
Hospital beds per thousand inhabitants 1981	0.2600	5.73 *	0.751	-4.95 *
Change in non-durable goods 72-93	0.2557	5.64 *	0.799	-4.00 *
Change in women illiteracy rate 81-93	0.2514	5.54 *	0.788	-4.21 *
Economic Active Population without profession 81-93	0.2491	5.50 *	0.711	-5.75 *
Change in per capita expenditure 72-93	0.2460	5.43 *	0.731	-5.35 *
Change in working children 81-93	0.2443	5.39 *	0.716	-5.65 *
Change access to drinking water 81-93	0.2325	5.14 *	0.807	-3.84 *

...continued

Table A.3.1
Spatial Correlations at Province Level

Variables	Moran Index	Z-Value	Geary Index	Z-Value
				conclusion
Land used to sell in market	0.2313	5.11 *	0.712	-5.72 *
Change working children 81-93	0.2306	5.10 *	0.754	-4.90 *
Change household size	0.2271	5.02 *	0.750	-4.96 *
Number of towns	0.2254	4.98 *	0.869	-2.60 *
Temporal crops sold in market	0.2184	4.83 *	0.766	-4.65 *
Change in access to electricity 72-81	0.2110	4.67 *	0.844	-3.10 *
Permanent crop land	0.2095	4.64 *	0.844	-3.09 *
Change in women illiteracy rate 72-81	0.2032	4.50 *	0.818	-3.62 *
% children who work 1972	0.2016	4.47 *	0.794	-4.09 *
Access to drinking water 1972	0.2006	4.45 *	0.793	-4.11 *
Access to electricity 1972	0.1955	4.34 *	0.797	-4.03 *
Change in schooling attendance 72-81	0.1938	4.30 *	0.806	-3.86 *
Land used for self consumed purposes	0.1920	4.26 *	0.878	-2.44 ~
Total no agricultural land	0.1872	4.16 *	0.828	-3.42 *
Population with college studies	0.1817	4.04 *	0.835	-3.27 *
Total land	0.1763	3.92 *	0.837	-3.24 *
Total land used	0.1711	3.81 *	0.807	-3.84 *
Change in access to sanitation services 81-72	0.1709	3.81 *	0.941	-1.18
Permanent crops sold in market	0.1659	3.70 *	0.858	-2.82 *
Permanent crops sold in farm	0.1650	3.68 *	0.948	-1.03
Inadequate ceiling 72-93	0.1620	3.61 *	0.856	-2.87 *
Change in access to electricity 72-93	0.1612	3.60 *	0.897	-2.05 ~
Economic Active Population without profession 1981	0.1556	3.47 *	0.812	-3.73 *
Hospital beds per thousandinhabitants 1993	0.1447	3.24 *	0.870	-2.59 *
Change in % of children who work 72-81	0.1434	3.21 *	0.862	-2.75 *
Number of urban towns	0.1153	2.60 *	1.256	5.09 *
Urban population 1993	0.1127	2.55 ~	1.282	5.61 *
Urban houses 1993	0.1122	2.54 ~	1.285	5.66 *
Total houses 1993	0.1020	2.32 ~	1.291	5.78 *
Total occupied houses 1993	0.1016	2.31 ~	1.291	5.79 *
Rural town density 1993	0.0999	2.27 ~	0.864	-2.71 *
Access to sanitation services 1972	0.0786	1.81	0.921	-1.56
Change in access to drinking water 72-81	0.0624	1.46	1.029	0.58
Change in illiteracy rate 1981-93	-0.0688	-1.37	1.060	1.20
Rural house density 1993	0.0554	1.31	0.900	-1.98 ~
Hospital beds 1993	0.0395	0.97	0.933	-1.34
Temporal crops used as seeds	0.0383	0.94	0.912	-1.74
Rural population density 1993	0.0309	0.78	0.934	-1.32
Land whose main output was used as seeds	0.0281	0.72	0.918	-1.63
Change in household size 72-81	0.0188	0.52	0.988	-0.24
Forestry land	0.0157	0.45	1.081	1.61
Permanent crop whose main output was used as seeds	0.0036	0.19	1.026	0.51
Towns density 1993	0.0019	0.16	0.870	-2.58 ~
Urban towns density 1993	-0.0106	-0.12	0.825	-3.49 *
Houses density 1993	-0.0094	-0.09	0.818	-3.62 *
Population density 1993	-0.0093	-0.09	0.818	-3.61 *
Vegetation zones typolgy	-0.0020	0.07	0.796	-4.06 *
Urban houses density 1993	-0.0081	-0.06	0.817	-3.63 *
Urban population density 1993	-0.0070	-0.04	0.816	-3.66 *

Note: $p<0.01$=*, $p<0.5$=~

Source: Authors' calculation based on National Census of Populations 1972, 1981 and 1993.

Annex 4: Determinants of Household Growth

Table A.4.1

Determinants of Household Size Growth: 1972-93

(At province level)

Variables at initial period	Coefficient
Non school-attendance rate	0.0048
	(0.004)
Household size	-0.6410 *
	(0.045)
Female household head (%)	-0.0086
	(0.005)
Woman illiteracy rate	-0.0050 ~
	(0.002)
Household members <1 year (proportion of household size)	33.7979 *
	(9.368)
Household members 1-4 years (proportion of household size)	-1.4481
	(1.703)
Household members 5-9 years (proportion of household size)	-3.6317
	(4.165)
Household members 10-14 years (proportion of household size)	5.9568
	(3.850)
Household members 15-19 years (proportion of household size)	1.0405
	(1.945)
Latitude	-0.0580 *
	(0.009)
Total illiteracy rate	0.0134 *
	(0.004)
Intercept	1.6245 *
	(0.462)
Observations	190
Adjusted R-squared	0.7733

Note: Standard errors in parenthesis. *=p<.01, ~=p<=.05.

Source: Authors' calculation based on 1972 and 1993 Population and House Censuses.

Annex 5: Alternative Methods for correcting for Spatial Autocorrelation

Table A.5.1

Comparing Methods: Determinants of Per Capita Expenditure Growth Rate: 1972-93

(OLS estimations with robust standard errors, at province level)

Variables at initial period	Models (1)	(2)	(3)
Intercept	-0.0277	-1.2303	-0.3270
	(1.385)	(7.537)	(1.706)
Altitude	0.2616	0.1863	0.4580
	(0.385)	(0.171)	(0.389)
Latitude	-0.0231	-0.0815	-0.0170
	(0.019)	(0.083)	(0.019)
Longitude	-0.0182	-0.0788	-0.0171
	(0.015)	(0.068)	(0.015)
Soil slope	0.0033	0.0154	0.0035
	(0.002)	(0.010)	(0.002)
Soil depth	0.002	0.0102	0.0023
	(0.002)	(0.010)	(0.002)
Igneous rock	-0.3197 *	-1.2763 *	-0.2757
	(0.100)	(0.463)	(0.106)
Metamorphic rock	-0.1318	-0.6157	-0.1362
	(0.122)	(0.540)	(0.122)
Temperature	-0.0114	-0.0382	-0.0082
	(0.009)	(0.039)	(0.009)
Basic needs	-0.0222	-0.1029	-0.0225
	(0.017)	(0.073)	(0.016)
High*basic needs	0.0045	-0.0347	-0.0149
	(0.090)	(0.358)	(0.080)
School attendance rate	0.0143 *	0.0649 *	0.0144
	(0.003)	(0.014)	(0.003)
Household headed by women (%)	-0.0109 ~	-0.0574 ~	-0.0134
	(0.005)	(0.024)	(0.005)
Working children (%)	0.0533 *	0.2151 *	0.0462
	(0.020)	(0.082)	(0.018)
Household size	0.0783	0.4608	0.1057
	(0.133)	(0.573)	(0.128)
Household size growth a/	-0.2624	-1.0146	-0.2208
	(0.140)	(0.606)	(0.136)
Number of migrants	0.0171	0.0588	0.0101
	(0.029)	(0.128)	(0.029)
Spatial autocorrelation		0.1702 *	0.2305
		(0.000)	(0.102)
Number of observations	190	190	190
Adjusted R-squared	0.486	0.528	0.526

a/ Instrumental variables are shown in Annex 2.

Note: Standard deviation in parenthesis and p<0.01=*, p<0.5=~

Model 1: Geography + infraestructure.+Geo*infra+private assets

Model 2: Geography + infraestructure.+Geo*infra+private assets, modelling first-order spatial error autocorrelation. (GMM).

Model 3: Geography + infraestructure.+Geo*infra+private assets, modelling first-order spatial error autocorrelation (ML).

Source: Authors' calculation based on 1972 and 1993 Population and House Censuses.