Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)

Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

124

Consumption growth and spatial poverty traps: an analysis of the effect of social services and community infrastructures on living standards in rural Peru

Philippe De Vreyer, Javier Herrera,
Sandrine Mesplé-Somps

October 2005

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173

e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

Consumption growth and spatial poverty traps: an analysis of the effect of social services and community infrastructures on living standards in rural Peru.

Philippe De Vreyer*, Javier Herrera†, Sandrine Mesplé-Somps‡

June 2005

*Member of CADRE, University of Lille 2, and associate with DIAL. Email: devreyer@dial.prd.fr

†UR DIAL-IRD, Paris. E-mail: herrera@dial.prd.fr

‡UR DIAL-IRD, Paris. Email: mesple@dial.prd.fr Address: DIAL, 4 rue d'Enghien, 75010 Paris. We wish to thank Jean-Louis Arcand and seminar participants at the universities of Caen, Cergy Pontoise, Clermont-Ferrand and Nantes together with our colleagues at DIAL for their remarks. The usual disclaimer applies.

Abstract:

We test the effect of local geographic endowment of capital on household growth in living standards in rural Peru, using a four years unbalanced panel data set. Our theoretical model of household consumption growth allows for the effect of community variables to modify the returns to augmented capital in the household production function. Three different sources of data are used: the ENAHO 1997-2000 household surveys, the population census of 1993 and the district infrastructure census of 1997. Altogether the addition of these different data sources makes an unusually rich data set, at least when considered with developing country standards. As in Jalan and Ravallion (2002), we use a quasi-differencing method to identify the impact of locally determined geographic and socioeconomic variables, while removing unobserved household and community level fixed effects. GMM are then used to estimate the model parameters. Several significant interesting results appear, confirming that private consumption growth depends on local geographic variables.

JEL Classification : C33 - H23 - I18 - I32 - I38 - 012

1 Introduction.

The fight against poverty is at the top of the World's economic agenda, as resumed by the so-called Millenium Development Goals. The objective, fixed at the 1995 Human Development World Summit in Copenhague, is to cut the 1990 rate of absolute poverty by a factor two in 2015. In order to achieve this goal, there is a large agreement among economists that two conditions must be met: first, a minimum pace of growth in the mean per capita consumption must be attained and, second, the benefits of that growth have to be sufficiently well distributed in the direction of the poor. Recently several studies have examined the relationship between growth at the aggregate level and distribution in the population. The study by Dollar and Kraay (2000), for instance, concludes that "growth is good for the poor" because, on average, the rate of increase into the mean welfare level of the poorest part of the population is the same as that of the population at large. However, as pointed by Ravallion (2001), such studies on aggregate data hide more things than they reveal. In particular they do not account for the potential spatial heterogeneity of growth, that is the fact that in a given country some regions do not grow as fast as others. Persistently slow growth areas can often be identified with persistenly poor areas, at least in relative terms. Such areas have been a concern in many countries, including those undergoing sustained aggregate economic growth. Examples include China, the eastern Outer Islands of Indonesia, parts of northeastern India, northwestern and southern rural areas of Bangladesh, much of northern Nigeria, the northeast of Brazil (see Ravallion, 1998).

Peru is yet another example as the prevalence of poverty varies considerably across regions: the *sierra* and *selva* have poverty rates that are nearly twice and extreme poverty rates that are about seven times that of the coast.[1,2] More than half of the extreme poor reside in the rural sierra, though it has less than a quarter of the national population. This region is characterized by subsistence agriculture, scarce arable land, fractured topography and high production risks. The Fondo Nacional de Compensación Social (FONCODES) has built a poverty map of Peru, based on an index of unsatisfied-basic needs, constructed at the district level. This map, reproduced in Escobal and Torero (2000), shows evidence of large welfare disparities across the country, but a heavy concentration of poor people in the most geographically adverse zones, as in provinces located in the *sierra* and *selva*.[3] However, it seems that there is a high regional heterogeneity of poverty and growth. Inequality is found higher within the three geographic areas than between them, with more than 85% of the generalized entropy indices due to within geographic area inequality. Turning to growth, Escobal and Torero (2000) evaluated the rate of increase in per capita consumption at the provincial level combining information from household surveys and census data between 1972, 1981 and 1993. They noted a high degree of disparity of the per capita expenditure growth rate between provinces. They also found that provinces with the highest, or the lowest, consumption growth rates tend to be clustered. The regional heterogeneity in poverty and growth rates combines with an apparent high degree of poverty persistence. According to Herrera (2001), three quarters of the poor in 1997 remained poor in 1998 and about 60% of

[1] The figures in this paragraph are taken from World Bank (2002). In this paper, a household is defined as poor if its per consumption is lower than the cost of a minimum basket of goods and services and extremely poor if it is lower than the cost of a minimum basket of food, necessary to maintain adequate caloric intake. See also INEI et al. (2001), and Herrera (2001).

[2] Following the Spanish tradition, the regions of the country have been classified into three distinct zones: the *costa* (coast or plains), the *sierra* (Andean mountain range) and the *selva* (jungle or Amazon). The Coast, which represents around 11% of the territory, and about 49% of the population, has a cold humid desert climate. Lima, capital of the country, is located in this area and represents 30% of the Peruvian population. The Andes constitutes 31% of Peruvian territory, where 42% of the population live, essentially in rural areas. In this region, altitude is superior to 2300 meters in the majority of cases. Only 8% of Peruvian population live in the Amazon region, which represents 58% of the territory. Most of this area is covered by dense forest.

[3] More generally, poverty rates are significantly higher in rural than in urban areas: 78.4% of households are poor in rural areas, against "only" 42% in urban areas. For extreme poverty these rates are 51.3% and 9.9% respectively.

them were still poor in 1999.

The question therefore is to determine why are there areas with persistenly low levels of income or consumption. The particular setting of Peru makes likely that heterogeneity in geographical capital is correlated with that of living standards in different regions of the country. This does not have to be the entire story though, and one cannot exclude the possibility that households with similar characteristics tend to concentrate. Moreover, "pure" geographic endowments like ecological conditions, climate, altitude or latitude are not the sole geographic capital. The supply of local public goods and infrastructure, or the local endowments of private goods, are community-level variables that constitute the environment of people and that can impact their productivity. Our purpose is then to determine whether and which components of "geographic capital" have a non zero impact on the marginal productivity of capital and thus help in determining growth in living standards in rural areas of Peru.

In section 2 the various models that can explain spatial poverty traps and several identification problems are discussed. In section 3, a model of consumption growth that allows for the effect of community variables to modify the returns to augmented capital in the household income generating function is presented. The way to control for latent heterogeneity is also exposed. The model is very similar to that of Jalan and Ravallion (2002). Detailled presentation of the data is given in section 4. The presence of geographical externalities is tested using a 4 years household panel, from 1997 to 2000 (ENAHO survey), the population census of 1993 and the district infrastructure census of 1997. Altogether the addition of these different data sources makes an unusually rich data set, at least when considered with developing country standards. In particular, the panel dimension of the data allows to purge the estimation from any household and community unobservables that could bias our results. Econometric estimation results are analyzed in section

5. They confirm that private consumption growth depends on local geographic variables. For instance, it appears that living in areas with a relatively high density of population has a positive effect on the household consumption growth, whereas the prevalence of some infectious diseases has a negative impact.

2 How to explain and identify spatial poverty traps?

Schematically two models compete to explain spatial poverty traps. With free household mobility the spatial concentration of poverty can arise because people with similar characteristics concentrate. If these people were to move to other areas they would experience the same growth in their living standards, holding everything else equal (this is what Ravallion, 1998, terms the *individualistic model*). The alternative explanation is that, with no mobility, spatial poverty traps occur because in some areas the "geographic capital" is lower or less efficient than in others, and because such capital has a positive impact on the marginal productivity of private inputs. In this case otherwise identical households do not experience the same growth in their living standards, if they live in areas with different endowments of geographic capital (this is called the *geographic model*). Free mobility, that is mobility without any cost, is an ideal situation that one is unlikely to find in a low-income country. In Bangladesh for instance, Ravallion and Wodon (1997) find that *"sizable geographic differences in living standards persist when one takes account of the spatial concentration of households with readily observable non geographic characteristics conducive to poverty. The same, observationally equivalent, household is poor in one place but not another."* What is remarkable in this example is that this occurs even though there are no administrative restrictions to migration, and very few physical ones, and the vast majority of the country population shares the same ethnicity, language and religion. Just the direct costs of migration - small in absolute terms but prohibitively high relatively to their ressources - prevent poor people from migrating to areas in which they would enjoy higher living standards. In Peru, like in Bangladesh, migration is “free” but, unlike that of Bangladesh, the geography raises physical barriers to the mobility of households. Thus, high transportation costs, lack of information on opportunities outside the area of residence, ethnic fragmentation, ill-functioning markets for land and so on are as many impediments to the migration of poor households.

From an empirical point of view, making the distinction between the *individualistic* and the *geographic* models is not easy, because with free mobility of people or households, it is not difficult to imagine cases where an apparent effect of geographic capital in fact results from the concentration of households with similar characteristics. Suppose for instance that people with high endowments of private capital concentrate in areas with a given range of average temperature. One would then observe that people living under temperate climatic conditions have higher living standards than others, but it would be wrong to attribute this difference to the climate. The issue is the potential endogenous location of individuals and households. It turns out to be particularly accute in static models of living standard levels (Ravallion, 1998). As noted by Ravallion (1998), one way of dealing with this is to estimate a switching regression which determines which region the household is located in. But such regression would likely be plagued by endogeneity and identification problems of its own. In other words this might be asking too much to the data. Another strategy is to use an estimation method that permits to control for the effects of unobserved household characteristics and that could bias the coefficients of the geographic variables. This can be done provided that panel data are available. This is the approach followed in this paper and that is developed in the next section.

3 A simple model of consumption growth.

3.1 Theoretical model.

As in Jalan and Ravallion (2002), we extend the Ramsey model of consumption planning to the case of a household facing geographic externalities in its income generating process. The household, h, finances its consumption entirely from its current income, which is produced according to a production function that admits as arguments the level of productive capital, K, and a vector G of community level variables that might have a positive or a negative effect on the returns to capital:

$$y_{ht} = F(K_{ht}, G_{ht}) \tag{1}$$

We depart from Jalan and Ravallion in assuming that K is the level of "augmented capital" in the sense that it includes physical as well as human capital. The reason for this choice will become clear when we will turn to the specification of the capital marginal productivity in the econometric application. Thus y_{ht} is Becker's full income or, in other words, the potential income of the household. It is the income that the household could obtain if it were using entirely both physical and human capital to produce income. The household is assumed to have no access to the credit market for capital, an assumption that seems reasonable given the particular context of rural Peru. Potential income can be used by the household either to increase its capital stock (by accumulating physical assets or by investing in the human capital of its members) or to consume:

$$F(K_{ht}, G_{ht}) = K_{h,t+1} - (1-\delta)K_{ht} + c_{ht} \tag{2}$$

with δ the rate of depreciation of augmented capital. The household is assumed to have perfect foresight and to maximize the actualized value of its utility flows at date 0 under the budget constraint:

$$(3) \qquad \max_{\{c_{ht}\}_{t=0,..,+\infty}} \sum_{t=0}^{+\infty} \beta^t.u_{ht}$$

$$s.t. \quad K_{h,t+1} = (1-\delta)K_{ht} + F(K_{ht}, G_{ht}) - c_{ht} \text{ for all } t \geq 0$$

This yields the following set of first order conditions:

$$(F'_K(K_{ht}, G_{ht}) + 1 - \delta).u'_{ht}(c_{ht}) = u'_{h,t-1}(c_{h,t-1}) \text{ for all } t \geq 1$$

They show that an increase into the marginal productivity of capital induces an increase into consumption if the marginal utility of consumption is decreasing. The particular feature of this model is that geographic externalities can influence consumption growth rates through effects on the marginal productivity of capital.

In order to get an estimable form of this equation we follow Jalan and Ravallion in assuming that the instantaneous utility function is of the isoelastic form:

$$u_{ht}(c_{ht}) = \frac{(c_{ht})^{1-\delta}}{1-\delta}$$

where $\delta \in [0, 1[$. This yields, after reporting the corresponding marginal utilities into the first order equation, linearizing and writing the marginal productivity of augmented capital in a reduced form:

$$(4) \qquad \Delta \ln c_{ht} = \ln c_{ht} - \ln c_{ht-1} = \alpha + x'_{ht}\beta + z'_h\gamma$$

where x_{ht} and z_h are vectors of specific household-community time dependant and independent variables that modify the marginal productivity of capital.[4] Note that the marginal productivity of capital does not depend on the level of capital, that is constant returns to scale are assumed. But recall that K is, in our case, augmented capital, in the sense that it includes the human capital stock of the household members, so that imposing constant returns to scale is not as restrictive as it might seem.[5]

3.2 The econometric model.

In order to allow for unobserved heterogeneity we complete equation (4) by adding to the deterministic part a stochastically determined error term: v_{ht}. In this paper we are particularly interested in determining the effect of community specific variables on the marginal productivity of capital. In order to do so, we have to precisely control for the effect of community and household unobserved specific effects that our model cannot account for and that one cannot hope to fully capture in the available data. As these unobserved variables are likely to be correlated with our included explanatory variables, lack of control of their effects will result in biased OLS estimates of the β and γ coefficients vectors. The usual cure for such unobserved effects is to work with the first differenced version of the base model. But, in our case, this would result into the dropping of the time invariant variables, a most undesirable consequence given our purposes. However, as noted in Jalan and Ravallion (2002), the existence of economy-wide factors suggests that the impact of observed and unobserved heterogeneity on the marginal productivity of capital is not necessarily constant over time. For instance a well maintained irrigation

[4] Reporting the marginal utilities in the first order equation yields: $F'_K(K_{ht}, G_{ht}) + 1 - \delta = c_{ht-1}^{-\delta}/c_{ht}^{-\delta}$. Taking the logarithm on both sides and assuming that $F'_K < 1$ we obtain: $\Delta \ln c_{ht} \equiv (F'_K - \delta)/\delta$. Now assuming a linear specification for the marginal utility of capital we get: $\Delta \ln c_{ht} = \alpha + x'_{ht}\beta + z'_h\gamma$

[5] The AK model in the endogenous growth litterature bears on the same assumptions. See Barro and Sala I Martin (1995) for details.

network is likely to increase the productivity of farmers in the corresponding area, but this could matter more in bad (dry) years than in good (rainy) ones. In other words, economy-wide shocks do not necessarily have the same impact on all households and it is a reasonable assumption to allow the effect of these shocks to vary with unobserved household heterogeneity. We thus follow Holtz-Eakin, Newey and Rosen (1988), Ahn, Lee and Schmidt (2000) and Jalan and Ravallion (2002) and decompose the error term as follows:

$$v_{ht} = \theta_t \omega_h + \mu_{ht} \tag{5}$$

Equation (4) is written as:

$$\Delta \ln c_{ht} = \ln c_{ht} - \ln c_{ht-1} = \alpha + x'_{ht}\beta + z'_h\gamma + \theta_t\omega_h + \mu_{ht} \tag{6}$$

where μ_{ht} is assumed to be an i.i.d. variable with zero mean and ω_h is a household specific effect (also with zero mean) which is not assumed orthogonal to the regressors and that modifies the impact of external shocks, θ_t, on consumption growth. Now, lagging equation (6) by one period, multiplying the resulting equation by $r_t = \theta_t/\theta_{t-1}$, and substracting it from (6) we get:

$$\Delta \ln c_{ht} = \alpha(1-r_t)+r_t\Delta \ln c_{ht-1}+(x_{ht}-r_t x_{ht-1})'\beta+(1-r_t)z'_h\gamma+\mu_{ht}-r_t\mu_{ht-1} \tag{7}$$

The determinant advantage of this modelling strategy is that in the preceding equation the coefficients of the time invariant variables are identified, provided r_t is not found equal to one.

This specification is tested in Jalan and Ravallion (2002) for China. However it assumes that external shocks are identical for every household of the economy. But in a country like Peru, that presents a wide disparity of ecological conditions and, in particular, in rural areas, this does not seem a very reasonable assumption. Thus we choose to relax this hypothesis and to give more flexibility to the error term decomposition, by allowing inter-regional variation of the r_t ratio. We experimented with several regional classifications. The best results are obtained with a six natural regions classification, defined according to altitude and the position with respect to the *Cordillera de los Andes*.[6] We also allow the constant term to change with the year of observation.

The model is estimated by the Generalized Method of Moments on two consecutive growth periods. A four steps estimation method is employed. In the first step initial values of the coefficients are obtained by least squares, without imposing cross-equations restrictions. Estimated coefficients are then employed to provide starting values for the model estimation by maximum likelihood. The resulting estimates are then used to construct a starting value of the weighting matrix of the GMM criterion. Using this matrix we obtain the one step GMM estimator. The residuals of this estimation are then employed to obtain a White-heteroscedastic consistent weighting matrix that we use to compute the two step GMM estimator.

[6]Early versions of the model were estimated with r_t constant, then with r_t taking a different value for each of the three regions traditionaly distinguished in Peru (see note 2). In both cases the model was rejected on the basis of the overidentification test (see *infra*). Better, encouraging, results were obtained with an *ad hoc* grouping of the departments in six regions, but the best results were obtained on the basis of the classification established by the geographer Javier Pulgar Vidal in 1946. A total of eight natural regions can be distinguished in Peru, depending on altitude and the side (coast or Amazon) of the *Cordillera de los Andes: Chala* (coastal region, between 0 and 500 meters), *Yunga* (between 500 and 2300 meters if on the coast, and between 1000 and 2300 meters if on the Amazon side of the *Andes*), *Quechua* (altitude between 2300 and 3500 meters), *Suni* (between 3500 and 4000 meters), *Puna* (between 4000 and 4800 meters), *Janca* (between 4800 and 6746 meters), *Rupa-Rupa* (Amazon side, between 400 and 1000 meters of altitude) and *Omagua* (Amazon side, between 80 and 400 meters). Accordingly each so defined region has its own homogeneous climatic characteristics, with the corresponding population settlement and economic specializations. A total of six regions have been kept in this survey because there are no observations for the *Janca* region in our sample (this regions has only about 47,000 inhabitants) and because the *Suni* and *Puna* regions have been joined, since we have only a very small number of households from the *Puna* region.

3.3 Determining the instrumental set of variables.

In equation (7) one of our regressors, namely $\Delta \ln c_{ht-1}$, is correlated with the error term $\mu_{ht} - r_t\mu_{ht-1}$, so that instrumentation of this variable is required. A natural choice of instrument is the growth or level of log-consumption with an appropriate lag. As measurement error on consumption growth is a possibility that we cannot reasonably exclude, the year t-2 consumption level cannot be retained as an instrument for $\Delta \ln c_{ht-1}$, and we have to rely on year t-3 observations, so that a minimum of four years of observations are *a priori* necessary in order to properly identify our model if lagged consumption is the only available instrument.[7] However, one can imagine to use the log-income level observed in year t-2 as an instrument, if one is willing to assume that measurement errors on income are independant from those on consumption. Under this assumption, the model can be estimated using two consecutive periods and the use of GMM estimation techniques allows each equation to be instrumented with a different set of instruments. Moreover, one can extend the list of potential instruments and include capital stock variables as measured at the beginning of the observation period and the household and community variables, either fixed or measured at the beginning of the corresponding observation period. For instance, in estimating the determinants of consumption growth between year t-1 and year t, households characteristics as observed in years t-1, t-2 and t-3 are potential valid instruments. However, even though extending the set of instruments never lessens efficiency in infinite samples, in finite samples this could result in very poor estimator properties (Wooldridge 2002). For this reason we tried to restrict the set of instrumental variables to a minimum.

In order to test our specification, we follow Arellano and Bond (1991) and rely on a sargan test to validate our instrument set. Under the null hypothesis of zero

[7] Since $\Delta \ln c_{ht-1} = \ln c_{ht-1} - \ln c_{ht-2}$, any measurement error on this variable either comes from the measurement of c_{ht-1} or c_{ht-2} or both. As measurement error on c_{ht-2} impacts μ_{ht-1}, $\ln c_{ht-2}$ cannot be used to instrument $\Delta \ln c_{ht-1}$.

correlation between the estimated residuals and the instrumental variables, the sargan statistic follows a chi-square with a number of degrees of freedom equal to that of overidentifying restrictions.[8]

4 Data and choice of variables.

To test the hypothesis of poverty traps due to geographic or man-made infrastructure endowment in Peru, we use household panel data from 1997 to 2000. This panel has been constructed from surveys conducted by the INEI between 1997 and 2000 at the national level (ENAHO surveys). Surveys were carried out in the last quarter of 1997, 1998, 1999 and 2000. They are representative of three rural areas (*Selva, Sierra and Costa*) and four urban areas (*Selva, Sierra, Costa* and *Lima city*). Household surveys include information on housing, demographics, education, health, expenditure, income, and employment. Expenditure and income data were calculated taking into account the inflation rates between 1997 and 2000, and regional price spreads.[9]

Our theoretical model is not restricted to farm household behavior and it can be representative of non-farm household behavior as well. We nethertheless choose to restrict the sample to rural areas, the reason being that working with rural households makes easier the computation of the pertinent community variables that measure geographic externalities. Urban households are likely to have better access than rural households to facilities that are not located in their district, and the value of the community variables computed for urban households at the district level might not properly reflect the extent of geographic externalities.

[8] The use of the Sargan statistic as a test of over-identification bears upon the assumption of zero second order autocorrelation between the model error terms. This can be tested provided at least five years of observations are available. Unfortunately only four years are available in the present case.

[9] See INEI (2001a) for the details of the method used to calculate and deflate expenditure and income data.

Our data suffers from substantial attrition, since it covers 1162 rural households over the first three years, but only 492 rural households over the complete period.[10] This attrition is essentially due to the fact that the survey has not been implemented during the last year in an effective way. Worried by the possibility of attrition bias in our estimation, we compared the distributions of the log-consumption level in 1999 for households present both in 1999 and 2000 (4 years panel) and for observations present only from 1997 to 1999 (3 years panel). We also compared the distribution in the 4 years panel with that of the total sample of observations included in either one of the panels (unbalanced panel). The result of this comparison is shown in figure 1 where the results of kernel density estimates are reproduced. The distributions appear to be very close to each other and this is confirmed by the results of a Kolmogorov-Smirnov test that we run on the hypothesis that the distributions of the 3 years and 4 years panels are equal (see table 1). The null hypothesis cannot be rejected with a p-value of 0.479. We also compared the distributions for panel households and for the total sample of households present each year of observations (since panel households are only a sub-set of the entire sample of households surveyed by the ENAHO). The results are shown in figure 2 and table 1. Once again we found no significant difference between the distributions.

Districts (*districtos*) are the smallest administrative division of Peru. Peru is then composed of 1886 districts distributed in 212 provinces (*provincias*) and 25 counties (*departamentos*). Geographic externalities are tested at the district level. Whereas the population census of 1993 (*IX Censo de Poblacion y IV de Vivienda 1993*) gives information on the average demographic and socioeconomic characteristics of people living in each district, the district infrastructure census of 1997 (*Encuesta Nacional de Municipalidades e Infraestructura Socio-economica Distrital, 1997*) collects information on the availability of public infrastructure. From

[10] This is also the case for urban households. There are 1809 observations over the first three years, but only 716 over the complete period.

the ENAHO household surveys it is also possible to calculate geographic level variables. Altogether, the three data sources that we use in this paper provide us with a wealth of information on the geographic capital at the district level.

The list of explanatory variables includes, at the household level and for estimating the determinants of consumption growth between years t-1 and t, a set of dummies controling for the sex, age and employment status of the head at the beginning of year t, together with the proportion of children less than five years of age and the proportion of adults more than 65. As one of our assumptions is that there are constant returns to scale to augmented capital in the household production function, we choose not to include the household size neither the proportion of children of working age nor the proportion of other adults as explanatory variables, since these variables are proxies for the level of productive human capital in the household. This assumption will be checked by adding these variables, together with other proxies for productive capital, such as household owned assets, the household's head education level and the connection of the household to electricity, public water and public sewage, into the list of explanatory variables and checking that their coefficients are not found different from zero. The proportion of children less than five years old is included in order to account for eventual opportunity costs borned by active adults when caring for these children. The proportion of adults more than 65 has a rather different status. We include it as a way to control for potential opportunity costs of caring for the elderly, but also, and mainly, because, as one of our geographic explanatory variable is the proportion of old people in 1993 at the district level, we think it is important to control for that proportion at the household level, in order to exclude the possibility that the geographic-level variable captures the effect of the corresponding omitted household level variable.

A set of three household dummies is included in order to account for the household participation to one or more anti-poverty programmes in the areas of nutrition, health and education during year t. A fourth dummy is included that takes the unit value when the household members have heard of an infrastructure public programme in their district. These programmes designed to reduced extreme poverty are quite numerous in Peru, and, according to the World Bank (World Bank, 2002), spending on them increased substantially between 1992 and 2002.[11] As many of these programmes are geographically targeted, we cannot properly identify the effect of local geographic capital on consumption growth without holding account of possible non random selection of households among their beneficiaries. Consequently we run a set of probit regressions and use the results to construct the appropriate Heckman-type correction terms.[12] The independant variables of these probits are household and community level exogenous variables, particularly five key geographic poverty indicators that have been used, among others, to design the targets of these programmes. For infrastructure, *having heard* of a public programme does not mean participation, but only that the district in which the household lives beneficiates from such a program. As for the other programmes, we control for the possible endogenous selection of the districts by running a probit, but the list of independant covariates includes only community level variables, with the key household variables (e.g. household consumption in 1997) being replaced by their district means.

The list of geographic level explanatory variables includes "pure" geographic

[11] In the recent years, in spite of the recession, resulting in a drop of per capita GDP of 0.77% a year and in a fiscal retrenchment, the share of social expenditure to GDP has not declined. On the contrary, public expenditures on education, health and water grew between 1997 and 2000, while the budget for defense and national security has been reduced. The budget for social assistance has decreased in 1998 from its level of 1997, but increased again in 1999. There is little doubt that the capacity of the government to maintain the budget dedicated to fight poverty can be related to the limited increase in the prevalence of extreme poverty that grew only slowly, increasing 1.3% to its current rate of 24.4% between 1997 and 2001, and several key social indicators, including infant and maternal mortality rates, improved significantly. However recent evaluations have raised questions on the targeting efficiency of anti-poverty programmes (Paxon, Schady, 1999, INEI, 2000a, 2000b, Schady, 2002, Alderman and Stifel, 2003)

[12] Specifically we follow Maddala (1983) and suppose that under programme participation, the household consumption growth rate is written $y_i = y_{i1} = \alpha_1 + x_i'\beta + u_{i1}$ whereas under non-participation we have $y_i = y_{i2} = \alpha_2 + x_i'\beta + u_{i2}$. Household participation is commanded by the following latent variable $I_i^* = z_i'\gamma + \varepsilon_i$ and we allow the correlation between ε_i and u_{i1} and u_{i2} to differ. Under these assumptions $E(y_i) = x_i'\beta + (\alpha_1 - \alpha_2)\Phi(z_i'\gamma) + (\sigma_{u_1\varepsilon} - \sigma_{u_2\varepsilon})\phi(z_i'\gamma)$ and the difference $\alpha_1 - \alpha_2$ measures the average programme impact holding everything else equal.

variables, such as altitude, the distance to equator, the distance to the provincial capital; demographic indicators like population density[13], the percentage of people of catholic or evangelist confession in the district or the percentage of people with spanish as a native language; public infrastructure variables, such as the road network density - calculated as the ratio of the total kilometers of roads to the district area - the proportion of paved roads, the percentage of people connected to electricity or public water and sewer; finally socio-economic variables such as the district proportion of people with a tertiary education level, the proportion of people suffuring from a digestive illness (typhoid or diarrhoea), the part of the active population working in the primary sector, the proportion working as independent workers, the proportion of old people in the population, the unemployment rate and the number of doctors per inhabitant.

Some variables in this list deserve some further comments. Altitude and the distance to equator are justified as explanatory variables due to the particular geographical setting of Peru. The country area lays between 400 kms and 2000 kms from equator and, as shown by Bloom and Sachs (1998), living under a tropical climate can be significantly detrimental to economic activity. However, the disagreements of life under the tropics can be largely reduced first by the effect of oceanic influence and second by altitude. As Peru is crossed by the *Cordillera de los Andes* it offers a wide variety of climates (see Escobal and Torero 2000 for a more thorough description of Peru's ecological variety) and we include altitude[14] and distance to equator in order to control, albeit imperfectly, for this heterogeneity. Percentages of people of catholic or evangelist confession, together with the proportion with spanish as a mother tongue, are included in order to account for the potential effect of a public

[13] Unfortunately, this variable is calculated using the entire surface of each district, but not only areas where human settlement is possible.

[14] We introduce also a dummy variable for districts located above 2000 meters, in an attempt to control for the significant change of ecological system at this level of altitude.

policy bias to the detriment of minorities[15] that is not captured by our included geographic infrastructure variables. These variables could also proxy for local ethnic fragmentation as it has been recently shown that local public good availabilities are inversely related to it (Alesina et al., 1999, Vidgor, 2001, Miguel and Gugerty, 2002). Population density, the decomposition of the district population according to the level of education or the kind of work may measure externalities due to linkages between economic units. For instance, the concentration of people - with or without specific characteristics - may improve individual productivity because agglomeration encourages information spillovers or because a high level of activity brings efficiency (Romer, 1986, Durlauf, 1994). TABLE 2 ABOUT HERE

Descriptive statistics of the variables are given in table 2. The means and standard deviations of the household level variables are computed taking the household as the unit of observation, whereas those of the geographic variables are computed over the sampled districts. A few points are worth mentionning. First we observe that, as could be expected, all through the panel the mean age of the household head, as well as the proportion of elderly people tend to increase, whereas the proportion of young children tends to decrease together with the average size of households. Second, it can be seen that the percentage of household heads working in the public sector goes down. This is coherent with the fact that, during the nineties, the substantial downsizing of the public sector led to a drop in the public sector employment resulting in a significant number of new job-seekers. In the majority of cases, they found work in the generally lower-paying informal sector (IMF, 2001).[16]

[15] Unfortunately, the ethnic origin of the population is not available in the 1993 census population, nor in the 1997-2000 ENAHO households surveys.

[16] Moreover, between 1997 and 1999, real wage in the public sector has decreased by 11%, whereas formal sector and informal sector real wages increased by 14% and 8% respectivelly (IMF, 2001).

5 Results.

Results are presented in table 3. In the first column the results obtained when no account is held of the potentially endogenous household selection in anti-poverty programmes are reported. Column two presents the estimates obtained when Heckman-type correction terms are added to the list of covariates. Finally, in column three, the results that we get when adding the household level proxies for productive capital are shown. At the bottom of each column, the value of the Sargan overidentification test is reported, together with the normal approximation of the statistic that can be employed when the number of degrees of freedom is large. We shall first comment results concerning the model identification and instrumentation, then turn to the estimated coefficients.

The validity of our instrumentation procedure bears upon the value of the Sargan statistic.[17] In all three cases this value is found well below the critical value at the 5%, or even the 1%, level. This means that we can be confident that, first, the quasi-differentiation of our model indeed has removed any household or community unobserved specific effect and that, second, instrumentation of the lagged-consumption growth is correct. This conclusion is reinforced by the fact that, in all three regressions, controls for the household's head professional activity, sex and age are included, together with the proportion of children less than 5 and the proportion of adults more than 65 in the household and that none of these variables have coefficients statistically different from zero (results not shown). In case the model quasi-differentiation did not remove all unobserved household specific effects, one would expect such effects to be correlated with one or more of these variables and their coefficients to be, spuriously, found different from zero. This is not what

[17] The set of non included instruments is as follows : number of years since administrative creation of the district, longitude, nine housing quality variables in 1997, household log-income in 1997 (for growth periods 1998-1999 and 1999-2000) and household log-consumption in 1997 (for growth period 1999-2000 only).

we find.

Turning now to the identification issue remember that coefficients of the geographic time-invariant variables are identified provided that the r_t ratios are found different from one. We find that this is always the case in the *Chala*, *Yunga*, *Quechua* and *Omagua* regions and also, in year 1999, for the *Rupa-Rupa* region (at the 1% level). Only the *"Suni+Puna"* region and, for year 2000, the *Rupa-Rupa* region do not check this rule, meaning that observations from these regions do not contribute to the identification of the time-invariant variables coefficients. As only about 14% of our sampled observations belong to the *"Suni+Puna"* region, this means that overall our time-invariant variables coefficients are correctly identified (albeit maybe altitude, see *infra*).

Comparison of results in columns two and three shows that estimated coefficients are not significantly modified when proxies for household productive assets are included. Moreover, coefficients of these variables are not found significant, thus confirming the validity of our constant returns to scale assumption (results not reported) and strengthening our conclusion on the efficiency of the quasi-differencing procedure. Such results are close to those reported by Jalan and Ravallion (2002) for China, where fixed productive assets and cultivated land per capita do not have any significant impact on consumption growth. However, expenditure on agricultural inputs have a significant but negative effect and household size has a positive effect in the case of China (in the present case, the coefficient of this variable is found very close to zero and unsignificant).

Comparing results from columns one and two shows that holding account of the household endogenous selection process in anti-poverty programmes only changes the degree of precision of the estimated coefficients of the programme dummies. With no correction for endogenous selection, we find that education and health oriented anti-poverty programmes have a positive and significant effect on consumption growth. For health this effect is rather important, since participation to one of these programmes adds about 12 percentage points to the consumption growth rate. However it looses in significance when we correct for endogenous selection and for education the effect vanishes. Nutrition programmes do not seem to have any significant impact on consumption growth. The vanishing of the education coefficient conforms to what could be expected, given that in the present period the effect of enrolling in an education oriented programme is ambiguous. On the one hand, as some of these programmes have the effect of reducing the direct costs of schooling, one can expect programme enrolment to have a direct positive effect on consumption. On the other hand, if children have to attend school to receive the benefits, the total effect on consumption is ambiguous because, while attending school, children are not taking part in the household productive activities. Overall the total effect is likely to be small and the positive and significant coefficient reported in the first column of table 3 could result from the household endogenous selection, as is suggested by the unsignificant coefficient reported in the second column. As for infrastructure programmes their long-lasting impacts cannot manifest in our data. However they can have a direct and immediate effect on household consumption through an increase into local employment opportunities. Our results do not confirm this possibility but this could come from the fact that among the households that have heard of an infrastructure programme, only a subset of them are likely to directly benefit from it. To resume, these results suggest the effectiveness of health anti-poverty programmes, a lack of effect of nutrition programmes and a process of endogenous selection of households among the beneficiaries of education oriented programmes.[18] TABLE 3 ABOUT HERE

[18]The impact of some education oriented social programmes on educational outcomes has been analysed by Paxson and Schady (1999). For instance, they show that in districts which received FONCODES support, education expenditure increased school attendance for young children, but no evidence that these programmes affect the probability of being at the right school level, and weak evidence that it decreased the average time it takes children to go to school. Alderman and Stifel (2003) evaluate the "Vaso de Leche" (glass of milk) feeding programme. They find that the

Turning now to the core of our results a large proportion of infrastructure and socio-economic geographic variables are found with a non zero coefficient of the expected sign, confirming the existence of externalities phenomena due to neighbourhood endowments of physical and human capital or geographic characteristics. Not all results do confirm to what could have been expected however.

Distance to equator is positively related to consumption growth. As this variable is expressed in thousand kilometers, its coefficient means that, *ceteris paribus*, moving south by one thousand kilometers adds 11 percentage points to consumption growth. This could be expected given that, in Peru, the degree of humidity diminishes with increasing latitudes and is consistent with the "geographic" point of view developed by Bloom and Sachs (1998). However, altitude is not found to have a direct impact, in opposition to what has been found by Escobal and Torero (2000), but this could result, first, from our control for community and household unobserved specific effects and, second, from the fact that observations from the *"Suni+Puna"* region do not contribute to the identification of time-invariant variables, since the r_t ratio is never found different from one for this region (see *supra*). As it lies entirely above 3500 meters of altitude, this could explain why we do not find any significant impact of this variable.

At the district level, road network density, the percentage of paved roads and the proportion of households connected to the public water network are not found to have any effect. On the other hand, the proportion of people connected to the public sewage system is found positive and significant and the distance to the province capital has an estimated coefficient negative and marginally significant: ten more kilometers reduce consumption growth by one half a percentage point. Surprisingly the proportion of households connected to electricity is found to have a negative and significant effect. The zero coefficient of the variable measuring the

programme is relatively well targeted to the poor, but no econometric evidence that its nutritional objectives are achieved.

household access to the public water network might result from the inclusion in the regression of the proportion of individuals presenting a digestive illness in the area, since these diseases are frequently the results of a low quality of the drinking water (as we shall see *infra* this variable is found to have a negative coefficient). Overall these results suggest that the supply of public facilities (albeit maybe water) do not play an important role in explaining the spatial heterogeneity of consumption growth in Peru.

Much different are the results obtained for the district level socio-economic and demographic characteristics. Relatively high growth areas appear to be those with a high population density (an increase in 100 inhabitants per square kilometers translates into an increase of about 2.4 percentage point in the growth rate), high proportions of catholics and evangelists and (marginally) with spanish as native language. A low proportion of old people, a low proportion of the active population working as self-employed and a low unemployment rate (all variables measured in 1993) are also linked to high growth rates. These effects are quantitatively important: indeed, since the dependant variable is a percentage, the coefficients of the variables measuring proportions can be directly interpreted as elasticities. Thus a one per cent increase into the proportion of people more than 65 years old translates into a 1.8% decrease in the consumption growth rate, *ceteris paribus.* The effects of population density, proportions of elderly and self-employed and of the unemployment rate are consistent with the existence of agglomeration and pecuniary externalities. The positive coefficients for the proportions of catholics, evangelists and spanish speaking people have a rather different status: they suggest the existence of a negative bias detrimental to minorities, especially to Indian communities. This negative effect can be due to a public policy bias, as mentioned above, but also to a private segmented labor market meaning that Indian people have no access to jobs with possible high pay-rises or without opportunities to accumulate knownhow. In the case of Peru, this negative bias detrimental to Indian communities could also result from their specific pre-colonial organization and/or post-colonial administration.

Finally we find that the district proportion of people with a digestive illness (typhoid, diarrhoea...) in the previous year has a large negative and strongly significant effect on consumption growth. Once again this effect is quantitatively important: a one per cent increase in the prevalence of these diseases reduces the consumption growth rate by about the same proportion. This is consistent with the estimated positive effect of the household participation in health anti-poverty programmes and suggests that much could be accomplished in this domain, though it is not clear what should be done, since in our results the number of physicians per inhabitant has a zero effect. These results are similar with those found in China where infant mortality rate and medical personnel per capita have a significant impact on farm-household productivity, and those found by Murrugarra et al. (1999) and Cortez (1999) on wages and productivities in rural and urban areas in Peru.

6 Conclusion.

The aim of this paper is to test the effect of local geographic capital endowments on consumption growth in Peru, using a micro model of household behavior that allows for the effect of community variables to modify the returns to augmented capital in the household income generating function. Estimation results depend crucially on the control for community and household unobserved specific effects and tend to be consistent with the hypothesis that local geographic endowments have a non zero effect on consumption growth, a prediction of the *geographic model*. Somewhat unexpectedly, given the heterogeneity of the Peruvian geographic and the obvious difficulty to live in some areas, it appears that most socio-economic variables have significant coefficients, but not all pure geographic characteristics nor local public goods.

These results have several important analytical and policy implications. First, it pinpoints the weakness of models that only consider income dynamics purely in terms of individual household characteristics. Income dynamics are also explained by "geographic" endowments. Second, the way in which geographic capital affects consumption is complex. Spatial poverty traps are linked more strongly to socio-economic features of villages and provision of public goods rather than to purely geographic attributes. Lower endowments have negative externalities adversely affecting the returns to households assets and therefore their consumption growth.

This adverse impact of spatial factors bears also crucial policy implications. It leads to stress the need to combine policies focused on income transfers and assets reinforcement (particularly human capital) with policies that favour mobility across regional markets. In this sense, reduction of transaction costs plays an important role (access to markets, information on market opportunities etc.). Households in poverty trap areas will then more easily take advantage of growth opportunities offered by more dynamic markets across local communities.

Targeting is the other aspect of anti-poverty policies that may be affected when the dynamic and spatial dimensions are taken into account. The existence of poverty traps implies that chronic and transient poverty may be distinguished and have different determinants which in turn implies specific policy contents. Dynamic targeting also implies identifying factors associated with vulnerability in order to prevent household falling into poverty (transient or permanent) after a shock. Since externalities are mostly linked to provision of public goods and agglomeration effects, medium-term anti-poverty policy will necessarily have a public investment component. Besides, anti-poverty policies may not necessarily target poor households or villages but may also focus on bridging poor villages with more dynamic regional markets.

Although we have considered regional fixed effects and taken into account unobservable individual effects, an explicit and more complete treatment of covariant shocks is needed. In the same vein, we have made the hypothesis that institutions are identical inside Peru. We have not explored at all the impact of different levels of institutional development and complexity which may be determinant in the efficiency of public policies at the local level. Neither have we tackled with the difficult issue of ethnic discrimination embedded in public policies and market results. These issues are potentially linked since in Peru the colonization process has framed the social stratification and the settlement of rural communities. In particular, communities were managed differently according to their ethnic composition and the period of settlement. It is possible that this kind of difference between communities persists nowadays and has an impact on the growth process. These issues are all part of a future research agenda.

References

Ahn, S., Lee, Y. & Schmidt, P. (2001), 'Gmm estimation of linear panel data models with time-varying individual effects', *Journal of Econometrics* **101**, 219–255.

Alderman, H. & Stifel, D. (2003), 'The "glass of milk" subsidy program and malnutrition in peru', *Policy Research Working Paper* **3089**.

Alesina, A., Baqir, R. & Easterly, W. (1999), 'Public goods and ethnic divisions', *Quarterly Journal of Economics* **114**(4), 1243–1284.

Arellano, M. & Bond, S. (1991), 'Some tests of specification for panel data: Monte carlo evidence and an application to employment equations', *Review of Economic Studies* **58**, 277–297.

Barro, R. J. & Sala-i Martin, X. (1995), *Economic Growth*, McGraw-Hill, Inc.

Bloom, D. E. & Sachs, J. D. (1998), 'Geography, demography, and economic growth in africa', *Harvard Institute for Economic Development Working Paper* .

Cortez, R. (1999), 'Salud y productividad en el peru, un analisis empirico por genera y region', *Documento de trabajo* **R-363**.

Dollar, D. & Kraay, A. (2000), 'Growth is good for the poor', *Development Research Group, Working Paper* .

Durlauf, S. N. (1994), 'Spillovers, stratification, and inequality', *European Economic Review* **38**, 836–845.

Escobal, J. & Torero, M. (2000), 'Does geographic explain differences in economic growth in peru?', *Research Network Working paper* **R-404**, 64p.

Herrera, J. (2001), 'Poverty dynamics in peru 1997-99', *Document de travail* **2001/09**.

Holtz-Eakin, D., Newey, W. & Rosen, H. (1988), 'Estimating vector autoregressions with panel data', *Econometrica* **56**, 1371–1395.

IMF (2001), 'Peru: Selected issues', *Country Report* **01/51**.

INEI (2000*a*), 'Impacto de los programas de apoyo en las condiciones de vida', p. 124p.

INEI (2000*b*), 'Impacto de los programas de apoyo social 1997-1998: el valor de la necesidades basicas', p. 82p.

INEI, CIES & IRD (2001), 'Nuevas estimationes de la pobreza en el peru, 1997-2000', p. 60p.

Jalan, J. & Ravallion, M. (2002), 'Geographic poverty traps? a micro model of consumption growth in rural china.', *Journal of Applied Econometrics* **17**, 329–346.

Miguel, E. & Gugerty, M. (2002), 'Ethnic diversity, social sanctions and public goods in kenya', *University of California and University of Washington* .

Murrugarra, E. & Valdivia, M. (1999), 'The returns to health for peruvian urban adults: differentials across genders, the life-cycle and the wage distribution', *Documento de trabajo* **R-352**.

Paxson, C. & Schady, N. (1999), 'Do scholl facilities matter ? the case of the peruvian social fund (foncodes)', *Policy Research Working Paper* **2229**.

Ravallion, M. (1998), Poor areas, *in* A. Ullah, ed., 'Hanbook of Applied Economic Statistics', Marcel Dekker, Inc., New York, Basel, Hong-Kong, pp. 63–91.

Ravallion, M. (2001), 'Growth, inequality and poverty: Looking beyond averages.', *Policy Research Working Paper* **2558**, 31p.

Ravallion, M. & Wodon, Q. (1997), 'Poor areas or only poor people?', *Policy Research Working Paper* **1798**.

Romer, P. (1986), 'Increasing returns and long-run growth', *Journal of Political Economy* **94(5)**, 1002–1037.

Schady, N. (2002), 'Picking the poor: indicators for geographic targeting in peru', *Review of Income and Wealth* **48(3)**, 417–433.

Vigdor, J. L. (2001), 'Community composition and collective action: Analyzing initial mail response to the 2000 census', *Working Papers Series* **SAN01-05**.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross-section and Panel Data*, MIT Press.

World-Bank (2002), 'Peru country assistance strategy', *Memorandum of the President of the International Bank for Reconstruction and Development and the International Finance Corporation to the Executive Directors* .

Table 1: Kolmogorov-Smirnov tests of distributions equality

Null hypothesis tested	**P-value**
Distribution of household log-consumption in:	
- 3 years and 4 years panel are equal in 1999	0.479
- Total sample and panel are equal in	
- 1997	0.064
- 1998	0.122
- 1999	1.000
- 2000	0.985

Comparing kernel density estimates of log-consumption
Balanced and unbalanced panels - rural areas - year 1999

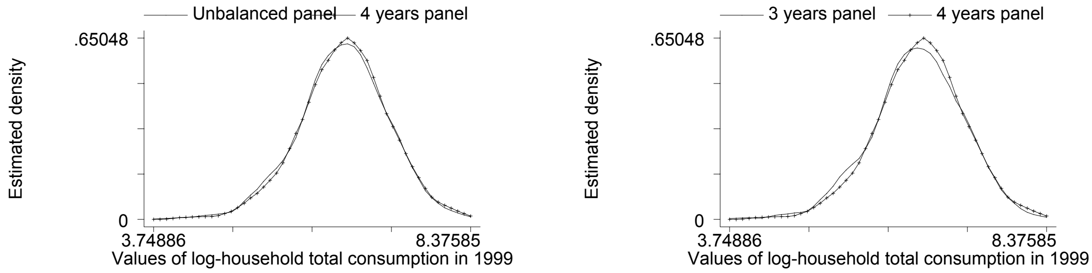
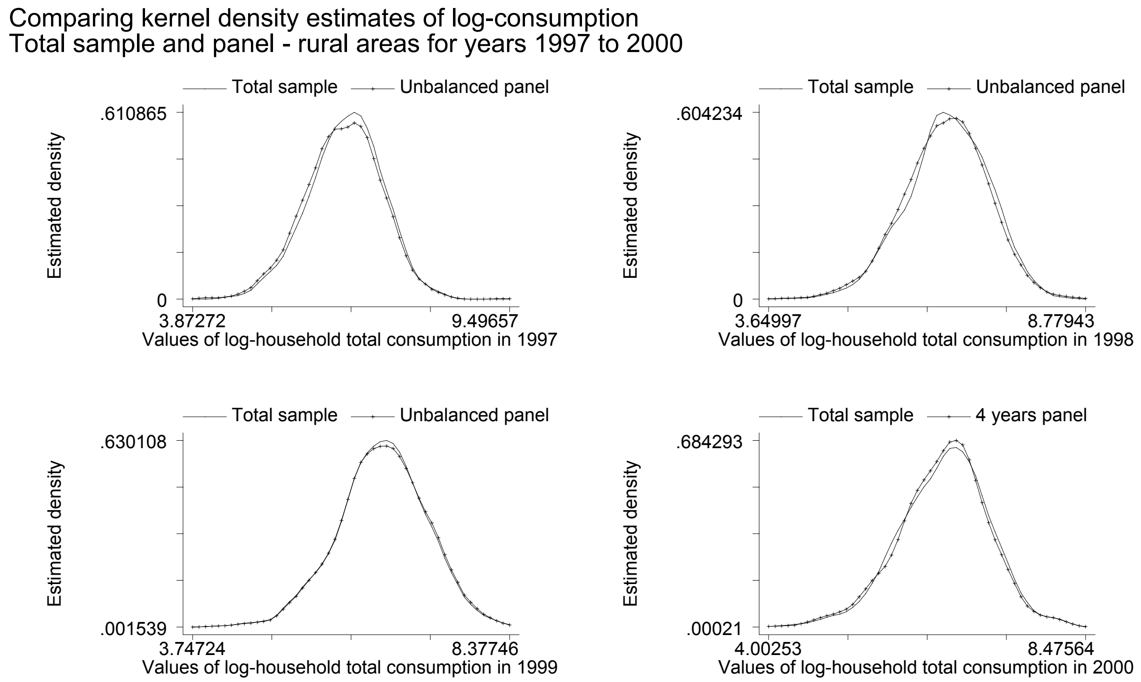


Figure 1

stata™



Comparing kernel density estimates of log-consumption
Total sample and panel - rural areas for years 1997 to 2000
Total sample
Unbalanced panel
.610865
0
3.87272
9.49657
Estimated density
Values of log-household total consumption in 1997
Total sample
Unbalanced panel
.604234
0
3.64997
8.77943
Estimated density
Values of log-household total consumption in 1998
Total sample
Unbalanced panel
.630108
.001539
3.74724
8.37746
Estimated density
Values of log-household total consumption in 1999
Total sample
4 years panel
.684293
.00021
4.00253
8.47564
Estimated density
Values of log-household total consumption in 2000

Figure 2

Table 2: Descriptive statistics (Means Standard Deviations)

Household level variables	1997	1998	1999	2000
Consumption (ln)	6.60 (0.6)	6.56 (0.7)	6.52 (0.6)	6.48 (0.6)
Sexe of Hh. head (male=1)	0.85 (0.4)	0.84 (0.4)	0.82 (0.4)	0.81 (0.4)
Age of the Hh. head				
Less than 26	0.04 (0.2)	0.03 (0.2)	0.02 (0.1)	0.03 (0.2)
Between 26 and 35	0.19 (0.4)	0.17 (0.4)	0.16 (0.4)	0.16 (0.4)
Between 36 and 55	0.44 (0.5)	0.45 (0.5)	0.47 (0.5)	0.45 (0.5)
More than 55	Ref.	Ref.	Ref.	Ref.
Activity of Hh. heads				
Self-employed	0.67 (0.5)	0.68 (0.5)	0.66 (0.5)	0.67 (0.5)
Executive	0.09 (0.3)	0.10 (0.3)	0.12 (0.3)	0.11 (0.3)
Wage earner - private	0.18 (0.4)	0.16 (0.4)	0.15 (0.4)	0.15 (0.4)
Wage earner - public	0.03 (0.2)	0.03 (0.2)	0.02 (0.1)	0.01 (0.1)
Other	Ref.	Ref.	Ref.	Ref.
% of 0-5 y-o. chid in the hh.	14.2 (16.1)	12.4 15.1)	9.3 (12.8)	8.56 (12.4)
% of elderly in the hh (>= 65 y-o.)	4.40 (15.7)	5.2 (16.9)	6.04 (18.9)	5.53 (16.6)
Hh size	5.30 (2.4)	5.26 (2.4)	5.20 (2.4)	5.01 (2.3)
Hh head with tertiary level of edu	0.03 (0.2)	0.03 (0.2)	0.03 (0.2)	0.02 (0.1)
Hh who own TV set or radio	0.84 (0.4)	0.83 (0.4)	0.81 (0.4)	0.85 (0.4)
Hh who own a vehicle	0.23 (0.4)	0.24 (0.4)	0.24 (0.4)	0.27 (0.4)
Hh connected to pub. water	0.18 (0.4)	0.26 (0.4)	0.28 (0.5)	0.32 (0.5)
Hh connected to pub. sewage	0.05 (0.2)	0.05 (0.2)	0.04 (0.2)	0.03 (0.2)
Hh connected to electricity	0.24 (0.4)	0.27 (0.4)	0.29 (0.5)	0.30 (0.5)
Hh part. to Nutrition anti poverty progr.		0.47 (0.5)	0.48 (0.5)	0.50 (0.5)
Hh part. to Health anti poverty progr.		0.32 (0.5)	0.32 (0.5)	0.25 (0.4)
Hh part. to Educ. anti poverty progr.		0.62 (0.5)	0.63 (0.5)	0.62 (0.5)
Hh know. the existence of Infrast. prog.		0.60 (0.5)	0.82 (0.4)	0.77 (0.4)

Table 2 (end)

Pure geographic variables				
Altitude a)	1697 (1409)			
Altitude > 2000 a)	0.42 (0.5)			
Distance to equator a)	1125.1 (439)			
Distance to provincial capital a)	31.4 (46.1)			
Road network (1000 km/km2) a)	4.79 (13.9)			
% of paved roads a)	14.7 (26.2)			
Nber of Phys. (per 10 000 inhab.) a)	0.78 (1.5)			
% of hh con. to public water a)	35.6 (32.7)			
% of hh con. to public sewage a)	14.2 (21.3)			
% of hh con. to electricity a)	39.9 (35.6)			
Population density (inhab./ km2) b)	56.2 (76.1)			
Urbanization rate b)	40.9 (26.5)			
% of catholics b)	86.9 (8.2)			
% of evangelists b)	9.0 (6.3)			
% with spanish as native lang. b)	61.7 (29.3)			
% of people more than 65 y-o. b)	4.9 (2.2)			
% working in primary sector b)	56.9 (20.0)			
% working as executive b)	3.4 (2.4)			
% working as self-employed b)	41.8 (14.0)			
% working as manual worker b)	20.3 (14.2)			
% working as clerk b)	10.6 (7.6)			
Unemployment rate b)	5.1 (5.1)			
	1997	1998	1999	2000
% with primary education c)	44.7 (9.6)	48.1 (10.3)	51.1 (11.3)	51.1 (14.9)
% with secondary education c)	20.1 (10.9)	20.1 (10.9	22.0 (11.1)	22.2 (11.7)
% with tertiary education c)	6.7 (4.9)	4.4 (4.8)	4.5 (5.3)	4.7 (5.9)
% illiterate c)	30.5 (13.1)	31.3 (13.1)	30.2 (13.3)	26.6 (11.7)
% with digestive illness c)	3.4 (4.5)	4.1 (4.3)	4.1 (4.5)	2.7 (2.7)

Standard deviation in brackets a) Source: *Encuesta Nacional de Municipalidades e Infraestructura Socio-economica Distrital, 1997*
b) Source: *IX Censo de Poblacion y IV de Vivienda 1993* c) Source: *ENAHO 1997- 2000-IV, computation by the authors.*

Table 3: Selected results (two step GMM estimation, with White-corrected standard errors)

Variable	**Model 1**	**Model 2**	**Model 3**
Household enrollment in anti-poverty programmes in year t (dummy variables)			
Nutrition	0.032	0.006	-0.042
Education	0.049*	0.047	0.031
Health	0.116***	0.172*	0.150
Infrastructure	-0.007	-0.036	-0.038
"Pure" geographic variables			
Altitude (unit is 1000 meters)	0.059*	0.036	0.046*
Altitude>2000	-0.066	-0.056	-0.065
Distance to equator (unit is 1000 kms)	0.060*	0.113***	0.134***
Distance to provincial capital (unit is 10 kms)	-0.002	-0.005*	-0.007*
Infrastructure variables			
Road network density	0.818	0.923	0.768
% of paved roads	0.029	-0.012	-0.025
Number of physicians per inhabitant	0.025	0.034	0.014
% of hh connected to public water	-0.019	-0.010	0.052
% of hh connected to public sewage	0.110	0.211**	0.161
% of hh connected to electricity	-0.087**	-0.116**	-0.051
Other geographic demographic and socio-economic characteristics			
Population density (1000s of inh. per km^2)	0.356**	0.236*	0.309**
% of catholics	1.086***	0.901***	0.946***
% of evangelists	1.257**	1.322***	1.220***
% with spanish as native language	0.093*	0.054	0.104*
% of people more than 65 years old	-2.318***	-1.748***	-1.955***
% working in the primary sector	0.021	0.000	0.020
% working as self-employed	-0.213**	-0.239***	-0.251***
Unemployment rate	-0.584*	-0.666**	-0.718**
% with tertiary education in year t-1	-0.017	-0.367	-0.605*
% with digestive illness in year t-1	-1.082***	-0.972***	-0.779**

Table 3 (end)

	Model 1	**Model 2**	**Model 3**
r_t *ratios for each natural region*			
r_{1999} (Chala)	0.030^{s}	-0.352^{s}	-0.054^{s}
r_{1999} (Yunga)	$-0.329^{s,*}$	$-0.342^{s,*}$	-0.125^{s}
r_{1999} (Quechua)	-0.334^{s}	$-0.551^{s,**}$	$-0.536^{s,**}$
r_{1999} (Suni+Puna)	1.265^{**}	1.319^{***}	1.134^{***}
r_{1999} (Rupa-Rupa)	-0.249^{s}	$-0.522^{s,*}$	-0.285^{s}
r_{1999} (Omagua)	$-0.368^{s,*}$	$-0.795^{s,***}$	$-0.825^{s,***}$
r_{2000} (Chala)	$-0.815^{s,***}$	$-0.681^{s,***}$	$-0.606^{s,***}$
r_{2000} (Yunga)	$-0.535^{s,***}$	$-0.341^{s,***}$	$-0.277^{s,***}$
r_{2000} (Quechua)	-0.227^{s}	$-0.432^{s,**}$	$-0.79^{s,**}$
r_{2000} (Suni+Puna)	0.970	1.131^{**}	0.998^{**}
r_{2000} (Rupa-Rupa)	0.688	0.454	0.067^{t}
r_{2000} (Omagua)	-0.367^{s}	-0.167^{s}	-0.028^{s}
Sargan statistic	89.4	101.8	108.2
Degrees of freedom	80	92	101
Normal approximation	0.76	0.74	0.53
Number of observations			
from 1997 to 1999	1162	1162	1162
from 1997 to 2000	492	492	492

Model 1: Heckman's type correction terms not included
Model 2: Heckman's type correction terms included
Model 3: Model 2 with proxies for household owned assets included
*, **, ***: significant at the 10%, 5% and 1% level, respectively
s: different from 1 at the 1% level; t: different from 1 at the 10% level. In all regressions, controls for the household head's professional activity, sex and age are included, together with the proportion of children less than 5 and the proportion of adults more than 65 in the household. None of these variables have coefficients statistically different from zero. Other unreported results are the values of the intercept coefficients for years 1998 to 2000 and, in model 3, the coefficients of the household's proxies for productive capital (all unsignificant).