# 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

**Nr. 173**

## Spatial inequalities explained – Evidence from Burkina Faso

**Johannes Gräb, Michael Grimm**

**July 2008**

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

# Spatial inequalities explained — Evidence from Burkina Faso*

Johannes Gräb†

*University of Göttingen*

*Center for Statistics and Department of Economics*

and

Michael Grimm‡

*Institute of Social Studies, The Hague*

*DIAL Paris, DIW Berlin, University of Göttingen*

The literature shows that regional disparities in growth and poverty are often relatively high, that these regional disparities do not necessarily disappear as the economies grow and develop and that these disparities are itself in many cases an important driver of the overall performance of an economy. In this paper we make use of the advantage of a multilevel random coefficient model to explain spatial disparities in incomes among Burkinabè households. Our findings show that it is not a geographical concentration of people with poor endowments that make areas poor in Burkina Faso. Household income disparities are largely driven by differences in neighborhood endowments and to a smaller extent by provincial or regional characteristics. We conclude that the policy should target small scale geographical units, such as villages. Providing infrastructure, enhancing the functioning of labor markets and fostering demand for education can compensate for climatical disadvantages.

**Key words:** Spatial inequality, poverty, multilevel modeling, decomposition, Sub-Saharan Africa.

***JEL* codes:** C21, I32, O12, R12.

*We thank the Institut National de la Statistique et de la Démographie (INSD) in Burkina Faso for providing the household survey data. We are also grateful to Hartmut Janus and Bakary Kindé for their great hospitality and assistance during the data collection process in Burkina Faso. The paper benefited from discussions with Stephan Klasen, Jann Lay, Andy McKay, and Stefan Sperlich. Financial support for this research by the German Research Foundation is greatly acknowledged.

† University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany. Email: jgraeb@uni-goettingen.de.

† Institute of Social Studies, Kortenaerkade 12, P.O.Box 29776, 2502LT The Hague, The Netherlands. Email: grimm@iss.nl.

# 1 Introduction

Country case studies on growth and poverty frequently show that regional disparities are relatively high, that these regional disparities do not necessarily disappear as the economies grow and develop and that these disparities are itself in many cases an important driver of the overall performance of an economy, i.e. regional inequality affects adversely the growth trajectory of a country. The 'Operationalizing Pro-Poor Growth Project', for instance, coordinated by British, French and German donors and the World Bank and covering 14 countries shows various cases in point (see Besley and Cord (2007); Grimm et al. (2007)). Often such regional inequalities are closely linked to key policy choices (e.g. trade policy) and patterns of public spending. But in most cases lagging regions also suffer under infrastructure bottlenecks, adverse agroclimatic conditions, import competition and limited scope for non-agricultural activities.

Burkina Faso is one among many Sub-saharan African countries where the regional pattern of living standards is particularly puzzling. Some of the observed inequality can be related to cotton production which is the main export commodity of the Burkinabè economy. However, despite the cotton boom, some cotton provinces did grow slower than other non-cotton regions, and in particular the traditionally poor and arid North of the country knew a quite good development during the past decade. Hence, it is not obvious to what extent agro-climatic conditions and trade exposure are responsible for a region's living standard and, hence, whether poverty alleviation policies should rather target region and provinces instead of villages and households.

Standard poverty assessments usually address such issues simply by undertaking rather descriptive analyses of growth patterns across regions and by performing decompositions of inequality indices by regional units. However, such decompositions make it very difficult to disentangle what is due to heterogeneity in household characteristics and what is due to heterogeneity in area-specific characteristics or endowments. In other words poor areas could simply be poor because households with poor endowments are geographically concentrated.

To deal with this problem, Ravallion and Wodon (1999) use two consecutive cross-sections of household survey data for Bangladesh to run separate regressions for each year and for each of the urban and rural sectors. They include a wide range of household characteristics and attribute the remaining part of the observed variance to geographic effects. They then undertake a number of robustness checks to exclude that there is a bias due to omitted household characteristics which are spatially correlated. The authors conclude that there are sizeable spatial differences in the returns to given household characteristics, i.e. the same household might be poor in one but not in the other region.

Another approach was chosen by Jalan and Ravallion (2002) and later by De Vreyer et al. (2008). They use several waves of panel-data to implement a quasi-differencing method to identify the impact of locally determined geographic and socioeconomic variables on household's consumption growth while removing unobserved household and community fixed effects. These authors find, for rural China and Peru respectively, robust evidence of geographic poverty traps and highlight in particular the socio-economic features of villages and the provision of public goods, such as rural roads, as important area-specific determinants.

Benson et al. (2005) have used alternatively spatially regression and geographically weighted regression techniques to allow regression error terms to be spatially correlated and to assess the degree to which determinants of poverty and the prevalence of poverty vary across space. For rural Malawi the authors find not much evidence for local poverty traps, characterized for instance by low agricultural productivity, and emphasize that the determinants of poverty vary spatially in their effects across the country. However, they find some evidence that regions which more opportunities for non-agricultural earnings and more markets, public infrastructure and services show less poverty.

While all these studies suggest that poverty reduction efforts have to be targeted at the sub-national level, none of these studies can provide a quantification of the respective weights of household determinants and the various sub-national levels in the variance of living standards across households. In this paper we suggest a methodology to address this issue. We construct a multi-level random coefficient model able to differentiate between spatial inequality due to differences in household characteristics and inequality due to area-specific characteristics. The model allows to quantify and identify simultaneously the relative impact and the main drivers of each administrative level on household income disparities.[1]

The main idea of that approach is that all these factors have, given their specific natural aggregation level, a different variance, and can therefore be separated even if they are strongly correlated. Put differently, multi-level modeling allows for instance to separate household characteristics from community characteristics, because both 'vary' on different levels. In a simple household level regression, their correlation would make it impossible to get efficient estimates, in particular if many of the relevant factors on both levels are unobserved. Moreover, within each level the variance in living standards can be decomposed to the extent that specific variables on that level, such as land and human capital endowments at the household level or climatic conditions and public services at the community level are observed.

To implement that approach for Burkina Faso, we build a very detailed and exhaustive data set combining household living standard measurement survey data, population census data, agricultural productivity survey data and a number of statistics collected at the provincial level.

The reminder of our paper is organized as follows. In Section 2 we provide a brief description of the extent of spatial inequality and its evolution over time in Burkina Faso. In Section 3 we present our data and describe in detail our empirical strategy. In Section 4 we present the results of various decompositions using our multilevel random coefficient model. In Section 5 we conclude.

[1]Similar techniques have been applied by Bolstad and Manda (2001), Ecob (1996) and Van De Poel et al. (2007) to study spatial inequality in child mortality and health.

## 2 Regional growth and inequality in Burkina Faso

Burkina Faso is one of the poorest countries in the world. In 2005, GDP per capita was estimated at only PPP US$ 1,213 and according to the Human Development Index, the country was ranked 176th out of 177 countries (UNDP, 2007). It is a landlocked country in the middle of West-Africa with a population of roughly 13,4 million. It has a very low human capital base and only very few natural resources. The country depends highly on cotton exports, which account for almost 60 percent of total export earnings, as well as on international aid. More than 80 percent of the Burkinabè population lives in rural areas working predominantly in the agricultural sector, which, again, suffers from very limited rainfall and recurrent severe droughts. The country experienced sustained growth with moderate poverty reduction during the last 15 years however accompanied by important variations over time spatial disparities, which cannot easily be explained (Grimm and Günther (2007)).

If income levels and growth rates as well as poverty shares are compared across Burkina's 13 regions (see Table 1),[2] one gets the impression that the Western regions, where the majority of the Burkinabè cotton is produced - Hauts Bassins, Mouhoun and Cascades - are richer than the remaining regions (abstracting from the two urban centers Ouagadougou and Bobo-Diolassou). However, in terms of growth in the subsequent period, the non-cotton and initially very poor Eastern regions - Sahel, Est and Centre-nord - performed better than all cotton regions, despite the very favorable development of cotton exports and the widespread belief that cotton exports were the driver of Burkina Faso's growth. In terms of poverty, Hauts-Bassins has still, given its relatively high income level (by Burkinabè standards) moderate poverty without however any significant poverty reduction since 1994. Mouhon, another of the important cotton regions, had ever and has still very high poverty levels. The cotton region Cascade halved poverty between 1994 and 2003 (Grimm and Günther (2007)).

[insert Table 1]

Given the somehow puzzling descriptive statistics on the regional level, we further disaggregate the data to see if the observed pattern of economic growth and poverty reduction is similar on the provincial level. The rationale for this procedure is our belief that certain factors, like the cultivation of cotton or livestock production, do indeed have a significant impact on income but that these activities are specific to provinces or districts but not to regions as a whole. The economic performance of all 45 Burkinabè provinces can be best presented using geographical maps.

Figure 1 indicates two important issues. First, neither does economic growth occur on some widespread regional level nor does there seem to be a high regional concentration of poverty. The intensity of growth and poverty rather varies across provinces over the whole country. Second, the provinces with the highest poverty incidence are not the same over time. Similar to what Benson et

[2] This household survey data is presented in detail in Section 3.

al. (2005) have found for rural Malawi, there do not seem to be spatial poverty traps in Burkina Faso.

[insert Figure 1]

If we disaggregate our data by the 135 districts (*Départements*) which are covered by the surveys[3] and plot household expenditures per capita in 1994 against growth of household expenditures per capita over the period 1994 to 2003, the data even suggest convergence in living standards across these local units.

However such kind of $\beta$-convergence might be exaggerated if expenditures per capita are measured with error (see e.g., Sala-iMartin (1996)). Although we provide below some evidence why such convergence could have occurred, we do not find robust empirical evidence for these channels and we cannot rule out that measurement error plays an important role. First, because we do not find evidence for $\sigma$ convergence, which would be immune to the measurement error problem (see e.g., Sala-iMartin (1996)). Second, we find a much smaller $\beta$-convergence coefficient if we regress the growth rate of expenditures from 1998 to 2003 on expenditure levels in 1994, which again could be sign of measurement error. However one should note that 1998 is a very particular year.

[insert Figure 2]

Hence, the question arises how disparities in expenditure, or more general in income levels, and the growth processes across areas can be explained. In particular, it is important to find out what the (relative) impact is of household characteristics on the one hand and area-specific, more macro-level factors, like geographic endowments, infrastructure and public services on the other hand. How is it possible that cotton exports seem to have boosted the economy without having boosted the regions where cotton is produced? What boosted the development in the lagging regions and what is their potential given their a priori less 'agriculture-friendly' geography? Is the effect of relevant factors the same across spatial units or does it vary significantly across the country?

Answers to this kind of questions have not yet been given for Burkina Faso, but seem crucial to appropriately target poverty alleviation strategies. The only study we have found that did research in that direction for the case of Burkina Faso is the one by Bigman et al. (2000). Similar to our study, these authors build a very detailed data set combining information of the household, village, department and provincial level. In a first step they estimate a prediction model for household consumption, using the household data and the community data from all other sources. In a second step, the authors use the prediction model to predict poverty at the village level for all villages inside and outside the household survey sample. The authors conclude that differences in the incidence of poverty among regions would be primarily due to differences in agro-climatic conditions, whereas differences in the incidence of poverty among villages within the same region would often reflect past policy biases that led to differences in the quality of roads or public services.

[3] In total Burkina Faso has 301 districts (*Départements*).

# 3 Data and empirical strategy

## 3.1 Data

Burkina Faso is composed by 13 agro-climatic regions which in turn are organized in 45 provinces, 301 districts (départements), 26 cities and towns (population > 5,000) and roughly 9,000 villages. According to the last census in 2006 the urbanization rate was about 16 percent and the average population density 48.4 persons per km$^2$. The two major cities are Ouagadougou, the capital, with a population of roughly 1.1 million and Bobo-Dioulasso with a population of about 0.4 million. The third city, Koudougou only has a population of 83.4 thousand.[4] The variables we use have been collected from a large number of sources and on different levels of that organizational structure. However, it was very difficult to find and get access to data on agro-climatic characteristics, infrastructure and public services and if it existed to match these data to other sources. This seems to be a problem in many least developed countries and may explain why only very few attempts have been made so far to analyze the effects of area-specific characteristics on households' living standards.

First, household data is drawn from three nation-wide representative household surveys, the Enquête Prioritaires (EP), conducted in 1994 (EP I), 1998 (EP II) and 2003 (EP III) covering around 8,500 different households in each year. These surveys were conducted by the Institut National de la Statistique et de la Démographie (INSD) with technical and financial support of the World Bank. These surveys contain relatively detailed information on household's socio-demographic characteristics, education, employment, agricultural and non-agricultural activities as well as consumption, income and some assets. However, except in 1998 these surveys were not linked with any village survey. The questionnaires only contain some questions regarding the time needed to reach the next primary and secondary school, the next health center, market and drinking water point. A more detailed description of these data sets can be found in Grimm and Günther (2007).

Given the usual low quality of income data in poor rural settings, we use household expenditure per capita as an indicator of households' living standards. Expenditures were deflated over time and space using appropriate price deflators. A critical issue in our study are of course the deflators used to correct for price differences across space. For this purpose we use deflators provided by the INSD in each survey year for Burkina Faso's 13 regions (based on price data collected on 37 different regional markets). Details about the construction of our expenditure aggregate and the price deflators we used can again be found in Grimm and Günther (2007).

Second, we can draw data on the community (or cluster) level from several sources. Each survey contains questions regarding access to roads, markets, and health institutions and alike. In 1998 a specific community survey was added to the usual household survey which collected further community data for 325 of the 425 communities covered by the survey. In addition, community characteristics were constructed by aggregating household characteristics

[4] Statistics taken from INSD, see http://www.insd.bf.

at the community level. It is important to note that each survey does however not cover exactly the same communities. Thus, a community panel cannot be constructed.

Third, data on landsize, fertilizer use and the use of modern production technologies in agriculture are drawn from a yearly agricultural survey called Enquête Agricole. This survey is conducted by the Ministry of Agriculture in collaboration with INSD. Since the data set uses a different survey design than the EPs, we merged the information to the other data sources on the provincial level, the smallest common regional unit. Landsize, fertilizer use and information about modern production technologies are therefore provincial averages.

Fourth, data on agro-climatic conditions such as monthly rainfall for the period 1993-2006 on the provincial level, and monthly minimum and maximum temperatures on the regional level were obtained from the Directorate of Meteorology (Direction de la Météorologie).

Fifth, data on the provision of public services and infrastructure and population densities, also at the provincial level, were obtained from the Ministry of Infrastructure (Direction Génerale de l'Amenagement du Territoire).

Hence, we have a data set which is organized in four levels: the household, the community (or cluster), the province and the region. Table 2 provides an overview of all used variables, their sources and their means and standard deviations.

[insert Table 2]

## 3.2 Empirical strategy

### 3.2.1 Multilevel modeling

To analyze the determinants of income levels and the contribution of these determinants to income disparities on different levels, we use a multilevel (or hierarchical or mixed) regression model.[5] Multilevel models are especially used in social science, sociology and health research to specify the effect of social context on individual level outcomes.[6] In economics, multilevel applications have not been as popular as in other research fields. Since the strength of multilevel models lies not in the estimation of causal relationships but rather in out of sample predictions. Hence economists rely on these models in particular to perform small area estimations, for instance to construct a poverty map (see Elbers et al. (2003) and Jiang and Lahiri (2006)).[7]

Another reason which might explain the low use of multilevel models is that these models are only consistent if OLS is consistent. In the context of hierarchical data this would mean that area effects should be independent of the

[5] For a comprehensive overview of the statistical theory underlying multilevel modeling and of various illustrative applications, see e.g. Goldstein (2003) and Hox (1995)

[6] For a good overview of applications in that area, see DiPrete and Forristal (1994).

[7] A paper which deals with causal multilevel models is, for example, Aassve and Arpino (2007)

covariates and any unobserved individual effects. Obviously, this independence assumption is seldom justified. However, this assumption is also violated when OLS is applied to this kind of hierarchical data. In the absence of panel data, as in our case, one could estimate a OLS model with dummy variables for each higher level unit to hold the independence assumption. However, this would impose severe restrictions on the model. Due to the few observations that we observe per first level unit (maximum 20), introducing dummy variables would lead to a significant over-parametrization (Lombardía and Sperlich, 2007).

Moreover, as it will be discussed thoroughly below, multilevel models offer itself several unique advantages, as their flexibility to allow for combining nested data from different sources, the possibility to partition variation across levels or to model variation of effects across areas. We will construct our multilevel model such that we can benefit from these advantages while using our large data set to control as much as possible for unobserved heterogeneity.

In the following we will differentiate between multilevel random intercept and multilevel random coefficient models as well as between fixed and random effects. In a multilevel random intercept model, a single error term for each level is introduced into the model (see equation 5). In a multilevel random coefficient model, in addition to random intercepts random effects are introduced into the model, i.e. coefficients are allowed to vary by higher level units (see equation 6). Fixed effects are hereafter denoted as coefficients which are directly estimated by the model. Contrary, for random effects only the variance and its standard error is estimated.

### 3.2.2 Efficient Estimation

We built our data set using several different and independent data sets. Variables are observed on multiple, nested levels. For instance, households' composition is observed on the household level, population density on the provincial level and temperature on the regional level. Clustering stemming from this nested structure requires to account for intra-group correlations. Under the assumption that individuals and households on the same level are more alike than individuals and households from different levels, within group residuals must be correlated. The classical linear regression model however assumes residuals to be independent among individuals by modeling the unexplained variability solely as the variance of the residual. Applying standard OLS regression to nested data leads to an overestimation of standard errors and, hence, statistical inference can be wrong. In a data set consisting of multiple levels, i.e. multiple 'populations', unexplained variability should be decomposed into the variability on all nested levels. This is exactly done by the multilevel structure and it allows to obtain efficient estimates (see Goldstein (2003)).

### 3.2.3 Variance partitioning

In a multilevel random intercept model, the decomposition of the error term allows to assess how much of the total variance in the dependent variable is attributable to higher levels and how much of this variance on higher levels can be explained by differences of characteristics on lower levels.[8] Put differently, the decomposition allows to draw conclusions on the explanatory power of the regressors with respect to the variation at the different levels (see Borgoni et al. (2002)). For instance, we can ask the question whether the observed spatial pattern in income levels across regions can rather be explained by differences in regional variables, like geography and institutions, by differences in community characteristics like access to certain public goods or rather by differences in household characteristics, like household size and education. This is a major conceptual advantage of a multilevel model. If we ran a household income regression with explanatory variables on higher levels, but without a multilevel structure, significant coefficients of these variables are likely to pick up variation which is at least partly due to omitted household level variables. In contrast, if we introduce a random variability coefficient on each level, we can test the explanatory power of level-specific variables on each level separately. Whenever an introduced variable reduces the variance of the level-specific error term, we can conclude that this variable explains a part of the variance in incomes (see Ecob (1996)).

### 3.2.4 Area-specific returns

A multilevel model designed as a multilevel random coefficient model ('RC' hereafter), allows to take into account a possible variation in the factor coefficients across spatial units. Finding significant variation in the effects of individual characteristics across spatial units suggests that area modifies the association between individual characteristics and income (see Merlo et al. (2005b)). In our case, for instance, it will be interesting to see whether effects associated with education, cotton cultivation or household composition are constant across spatial units.

### 3.2.5 Covariance structure of random effects

Finally, the RC model allows us to investigate the covariance structure of the random intercepts and random slope coefficients. For instance, it might be that communities with lower average income levels (a lower intercept) have higher returns associated with education or cotton cultivation. A significant negative correlation could for example explain the convergence described in Section 2.

### 3.2.6 Strengths of a multilevel model

Based on these methodological considerations, we belief that a multilevel model is particularly suitable for explaining spatial inequalities. Our methodology is capable of decomposing spatial inequality into the contribution of household and spatial characteristics, of identifying the key spatial determinants of inequality

[8] As emphasized by Goldstein et al. (2002), a straightforward interpretation of the variance partition coefficient or intra-class correlation coefficient is only possible in a random intercept model. In a random coefficient model the variance partition coefficient depends on the level of the covariates.

and of tracking variations in returns across space, thereby preserving simultaneously the advantages of the methods proposed by Ravallion and Wodon (1999), Jalan and Ravallion (2002) and Benson et al. (2005). Beyond the geographical analogue of the Oaxaca-Blinder decomposition proposed by Ravallion and Wodon (1999), our decomposition methodology allows to attribute weights to the contribution of the various sub-national levels on inequality. In addition to the identification of higher level variable effects on household income, as modeled by the GMM-type approach by Jalan and Ravallion (2002), our model differentiates between significant higher level effects explaining higher level inequality and significant higher level effects just picking up omitted household characteristics. Our methodology does also have some drawbacks. In the absence of panel data for Burkina Faso, we cannot exclude that we face with some of our explanatory variables an endogeneity bias. However, the methodology we propose is just as applicable to panel data as is is to cross sectional data. It should also be noted that our method does only control for unobserved heterogeneity on each level in case the independence assumption between unobserved characteristics and the regressors holds. To reduce at least the bias stemming from unobserved heterogeneity, we we build a large dataset in order to control for as many variables as possible.

### 3.2.7 The models to estimate

Our multilevel RC model can formally best be described by beginning with a two level random coefficient model with only one explanatory variable. The idea of the model is, that the regression coefficient on the first level is treated as a random variable at the second level.

The model equation reads:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}. \tag{1}$$

The regression coefficients $\beta_{0j}$ and $\beta_{1j}$ can be expressed as:

$$\beta_{0j} = \gamma_{00} + U_{0j} \tag{2}$$

$$\beta_{1j} = \gamma_{10} + U_{1j} \tag{3}$$

Finally, the combined model can be expressed as consisting of a fixed part (first term) and a random part (second term):

$$Y_{ij} = (\gamma_{00} + \gamma_{10}X_{ij}) + (U_{0j} + U_{1j}X_{ij} + \varepsilon_{ij}) \tag{4}$$

With such a model it is straightforward to check for significant variation of the random intercepts and slope coefficients within each nested group. Moreover, it is possible, as described above, to investigate the covariance of the intercepts and slopes.

We will use a four level model. In a first step, we will investigate the intraclass correlation coefficients and its proportional change when covariates on different levels are subsequently included.

$$Y_{ijkl} = (\gamma_{0000} + \sum_{p=1}^{P} \gamma_{p000} X_{pijkl} + \sum_{q=1}^{Q} \gamma_{0q00} C_{qjkl} + \sum_{r=1}^{R} \gamma_{00r0} P_{rkl} + \sum_{m=1}^{M} \gamma_{000m} R_{ml}) \quad (5)$$
$$+(U_{0jkl} + V_{00kl} + W_{000l} + \varepsilon_{ijkl}),$$

where $i$ stands for individuals, $j$ for communities, $k$ for provinces and $l$ for regions. $X$, $C$, $P$ and $R$ are vectors of individual, community, provincial and regional characteristics, respectively.

In a second step, we will augment our multilevel model by allowing coefficients of individual characteristics to vary across communities and by modeling the covariances of the random effects on that level.

$$Y_{ijkl} = (\gamma_{0000} + \sum_{p=1}^{P} \gamma_{p000} X_{pijkl} + \sum_{q=1}^{Q} \gamma_{0q00} C_{qjkl} + \sum_{r=1}^{R} \gamma_{00r0} P_{rkl} + \sum_{m=1}^{M} \gamma_{000m} R_{ml}) \quad (6)$$
$$+(U_{0jkl} + V_{00kl} + W_{000l} + \sum_{p=1}^{P} U_{pjkl} X_{pijkl} + \varepsilon_{ijkl})$$

We will estimate the model using Stata and its implemented mixed model command 'xtmixed'. The estimation procedure is based on an iterative generalised least squares approach (discussed in Goldstein (2003)). This procedure starts with the estimation of the fixed effects coefficients using ordinary least squares. The resulting residuals are stored. Afterwards, an iterative procedure begins, starting with a generalised least squares regression in a first step. Then, in a second step the residuals of this regression are used to compute the variance of the random coefficients. These steps are then iterated.

We estimate our model for three points in time: 1994, 1998 and 2003. This will also allow to get some insights into the dynamics of spatial inequality and its determinants.

# 4 Results: Sources of spatial inequality

## 4.1 Model $M0$: Intra-level correlations

For each year for which we estimate our model, we begin by a four level null model where we introduce nothing but a random intercept on the community, the provincial and the regional level. Using a likelihood ratio test we check whether the three level model, nested in the four level model, performs better (see Goldstein (2003)). Since this is not the case for any of the three years under consideration, we will use a four level model in the following.

Our base model, $M0$, reads:

$$Y_{ijkl} = \gamma_{0000} + U_{0jkl} + V_{00kl} + W_{000l} + \varepsilon_{ijkl}, \quad (7)$$

where $Y_{ijkl}$ stands for household expenditure per capita.

The results of model $M0$ for each year are shown in Tables 3 - 5. As expected, without controlling for the effect of covariates, the variation of the intercepts is highly significant at all levels. A good indicator to measure the contribution of the variance at each level to the total variance is the variance partition coefficient, also called the intra-class correlation coefficient (icc). This coefficient measures the degree to which observations in the same unit of a level, e.g. the same community, are dependent. As in an ANOVA model, the total variance of the dependent variable can be decomposed as the *sum* of the variance on each level, i.e. as the sum of the level-4, the level-3, the level-2 and the level-1 variances. The decomposition of the variance by level then reads:

$$\mathrm{var}(Y_{ij}|X_{ij}) = \mathrm{var}(U_{0j00}) + \mathrm{var}(W_{00k0}) + \mathrm{var}(W_{000l}) + \mathrm{var}(\varepsilon_{ijkl}). \tag{8}$$

Equation (8) can be rewritten as follows:

$$\mathrm{var}(Y_{ij}|X_{ij}) = \sigma^2_{u_0} + \sigma^2_{v_{00}} + \sigma^2_{w_{000}} + \sigma^2_{\varepsilon}. \tag{9}$$

For instance, the intra-class correlation coefficient $\rho$, for the year 1998 and the community level, is equal to

$$\rho = \frac{\sigma^2_{u_0}}{\sigma^2_{u_0} + \sigma^2_{v_{00}} + \sigma^2_{w_{000}} + \sigma^2_{\varepsilon}} = \frac{.221}{.221 + .025 + .078 + .51} \approx 27 \text{ percent.}$$

In words, in 1998 26.5 percent of the total variance is situated at the community level. In this case, the intra-class correlation coefficient measures the correlation of the residual of the response variable of households stemming from the same community. The high icc of the community level, which is almost as high in 1994 (19.2 percent) and 2003 (20.5 percent) depicts two things. First, it underlines the importance of using a multilevel approach to get efficient estimates. Second, it suggests strong community effects which are, however, relatively stable over time. The latter finding is particularly interesting in our case, since it means that the more households' incomes within a community are alike, the more likely is it that incomes are directly related to the contextual environment of the communities (see Merlo et al. (2005a)). The finding of a high community effect of around 20 to 25 percent is in line with the only other study we know which uses a multilevel model to assess the importance of community effects on household incomes (see Aassve and Arpino (2007)).

[insert Tables 3 - 5]

The icc for all years and levels are shown in Table 6. Clearly, most of the variance exists at the household level. It should be emphasized, however, that household expenditure data in developing countries is usually measured with error, given that it is generally very difficult to get precise information on expenditures if simple recall questions are used. Our model attributes the total variance which is due to measurement error in the expenditure data to the household level component. If we were able to account for these errors, the contribution of the household level variance to the total variance would probably be significantly lower, and, in consequence, the contribution of the higher levels higher. The contribution of the variance on the provincial and regional level is relatively low. We conclude that differences in household incomes are mainly driven by household and community (or cluster) characteristics and only to a minor extent by provincial and regional characteristics. In Burkina Faso regions rather than provinces follow agro-climatic zones, this can explain why regions make a higher contribution than provinces.

[insert Table 6]

The finding of a significant contribution of higher level characteristics on income does not necessarily have to be the result of differences in higher level characteristics itself. For instance, differences between communities might result from differences in household characteristics between communities, i.e. similar households are spatially concentrated. To see whether this is the case, we have to test the proportional change in the icc after accounting for household characteristics (to control for differences in household characteristics between higher levels). However, it should be noted, that household characteristics might lie in the causal pathway between area characteristics and household income. Including household characteristics will probably lead to an understatement of the importance of area characteristics. Hence, it is important to carefully discuss the household level variables and the potential influence of area characteristics on these variables. Therefore, we put special emphasis on the results of our null-model, where we find a significant contribution of unobserved large scale regional variables. Candidates for such unobservables are variables related to geographic and institutional factors as well as infrastructure endowments.

## 4.2 Model $M1$: The role of household characteristics

In the second step, we introduce on the household level explanatory variables to the random intercept model of equation (5). but we do not yet allow coefficients to vary across spatial units. We call this model '$M1$'. Our main concern is about two questions: First, what are the key household characteristics determining per capita income disparities? Second, to what extent are household level characteristics responsible for the spatial variation observed on higher levels as well as the proportional change of the icc? The results are presented - for each year separately - in Tables 3 - 5. Since we use maximum likelihood techniques for estimation, we rely on the Akaike Information criterion (AIC) to select the best model. We estimated many other versions of $M1$ with a much larger set of potentially important explanatory variables, but present here only those models with the lowest AIC and where all insignificant variables have been dropped.

All household variables have the expected sign and are in line with standard regression results. In particular, household composition has a considerable effect on income levels. In terms of per capita incomes, smaller households seem to be significantly better off in all years under consideration. The dependency ratios, measured via the children (0-6 years) per adult ratio, the youth (7-14 years) per adult ratio and the elderly (55 years and older) per adult ratio do all have a significant effect. While young household members lower per capita income in all years, the old-age dependency ratio is insignificant in 1994 and 2003 (thus dropped from the regression for those years) and negative in the drought year 1998 when food prices were extremely high.

Age of the household head has a significant negative effect on household income in all years. The household head being a male adult does not seem to play a major role concerning income since its effect is only significantly positive in 2003. The education of the household head is, as expected, very important in all years. Households with a literate head and households with a higher percentage of literate adults have on average a higher household income. Ethnicity has no influence on household income. Religion does. Belonging to one of the two world religions, Islam and Christianity, has a positive, but only hardly significant effect on income.

The effect of cotton farming differs across periods. Cotton farmers were better off in 1998 and 2003. In 1994 cotton did not yet have a significant effect. This is plausible, since the 'cotton boom' set in after the devaluation in 1994, enhanced by a very favorable evolution of cotton prices and accompanied by a substantial expansion of land used for cultivation. Farmers who were also engaged in livestock herding which is often done to diversify risk, and hence, to lower the vulnerability to external shocks, were significantly better off in 2003. However, a deeper analysis of this issue would require to take into account the possible endogeneity, since richer farmers are more likely to be engaged in livestock herding than poorer farmers. For the latter, the income constraint does not allow to buy any livestock. Obviously, it is now interesting to see whether for all these household characteristics the effects differ across communities.

For all years the community and regional variance component fell after the incorporation of household level explanatory variables. For the provincial component the direction of the change was unstable over the years which is not surprising given the small size and low significance of the provincial random intercept. The extent of the proportional change of the icc for the community and regional variance component is surprisingly stable across survey years (see Table 7). Controlling for household level characteristics reduces the icc of the community by around 50 percent.

[insert Table 7]

If we were neglecting any unobserved individual characteristics we could conclude that 50 percent of the community level variation in income levels was due to household characteristics while the rest was due to community characteristics. Clearly, this is an unrealistic assumption. We have to consider that household characteristics are itself influenced by higher level factors. Levels as well as returns to education, cotton farming and livestock herding might be influenced by neighborhood characteristics, which would lead to an underestimation of area importance. Hence, our result described above seems plausible. On the regional level the inclusion of household level variables was also non-ambiguous. In 1998 and 2003 observed household characteristics were able to explain around 60 percent of total unexplained regional variance. In 1994, this change was significantly lower since only 40 percent could be explained. Given that we controlled for individual characteristics as much as possible from the Burkinabè household surveys (see Table 2), we conclude also for regions that large scale variables have a non-negligible impact on household level income.

### 4.3 Model $M2$: The role of community characteristics

To test for the meaningfulness of our results which indicated a high relevance of community (or cluster) characteristics, we will check the proportional change of the icc after the incorporation of community characteristics (Model $M2$). The remaining significant variation of the community level random intercept could be either due to unobserved household characteristics leaving the community icc more or less unchanged or due to neighborhood-specific characteristics lowering the community icc towards zero. Again, we use the AIC as a model selection criterion and present only our final models $M2$ with the best fit in Tables 3 - 5. All community variables which we initially included in model $M2$ are listed in Table 2.

If neighborhood matters, the question is of course which are the factors which have the highest importance. Tables 3 - 5 depict a distinct pattern across all years. Urban communities with a high ethnic fragmentation[9], a high percentage of literate household heads, i.e. adults per household, and access to electricity are better off, ceteris paribus. Besides the direct effect of having a literate adult in the household, there seems to exist a contextual or spill-over effect of better educated on less educated individuals within communities. However, access to primary and secondary schools - as measured by the time needed to reach them - does not turn out to be significant. Education is the only household characteristic which appears to have some spill-over effects. Except for youth per adult in 1994, all aggregates of household level characteristics turn out to be insignificant. This is also true for communities with a higher percentage of cotton farmers, even though cotton farmers themselves are better off in 1998 and 2003, and cotton is always seen as a factor with some contextual effect in a neighborhood.

Since we did not have a direct measure of electricity in a community, we declared a community to have access to electricity if at least one household had access. Electricity might be a good proxy for infrastructure, such as access to roads, in a community, since power transmission lines are usually found along (gravel) roads. Since at the community level we only have information on electricity but not on other infrastructure such as roads, we interpret the positive effect of electricity carefully as a general positive effect of neighborhood infrastructure on household income. Though, like access to schools, access to health-centers and access to markets turns out to be insignificant. The effect of these kind of public services might be, at least to some extent, captured by the significant positive effect of urban communities since all these services are usually provided in urban neighborhoods.

[9]Ethnic fragmentation is measured as the variance of the shares of each ethnicity in a community

In 1998, the household survey was accompanied by a community survey for 325 out of the 425 clusters. This much larger community level dataset in 1998 can however only be examined at the cost of loosing a fourth of all households in the sample. Hence, we report regression results using data for the community survey separately in model $M*$ in table 4. Of all community survey variables listed in table 2 only the access to a road and to a hospital and high malaria incidence in a cluster determine household income. Malaria seems to lower income levels significantly by deteriorating the productivity of workers in Burkina Faso. Contrary, the possibility of visiting a nearby official hospital to seek medical examination as well as the access to a road foster household income. These results strongly confirm our findings and suggestions from the regressions for all years of model $M2$ which used community level data drawn from the household survey. Beyond the positive effect of a community to be urban, access to markets and educational institutions do not seem to play a major role in determining household income. Access to roads however, as already suggested by by the positive effect of electricity in model $M2$ which we thought to be highly correlated with road access, seems crucial in enhancing income generating potential by fostering mobility of goods as well as of labor.

After accounting for community factors the community icc reduces significantly in all years (see Table 7). Around 60 percent of the remaining unexplained community level variation was explained by those factors in 1994 and 1998. In 2003 it was still more than 40 percent. We have only a modest database on community level variables which is drawn from the household level questionnaire. However this small set of variables is capable of explaining a significant part of the remaining unexplained between neighborhood differences. Hence, in addition to simply specifying some significant relationship between contextual variables and household income as done above, we conclude that these variables are actually responsible for a large part of the community level disparity. The remaining unexplained community variation cannot not be dissolved with our data at hand.

The question remains whether provincial and regional income disparities are actually driven by differences in provincial and regional endowments or if they are mainly driven by differences in community characteristics between these areas. Table 7 shows that around 60 percent of the remaining regional level variation in 1994, 80 percent in 1998 and 40 percent in 2003 can be explained by differences in community endowments. After the consideration of household and community level determinants, less than 5 percent in 1994 and 1998 and less than 12 percent in 2003 of the remaining total unexplained variation is situated at the provincial and regional level together. It should be emphasized, once again, that household as well as community factors are likely to lie in the pathway of macro factors, and, hence, we risk to understate the influence of variables on higher aggregation levels. Moreover, likelihood ratio tests show that both levels still have a significant impact.

### 4.4 Model $M3$: The role of provincial and regional characteristics

In model $M3$ we incorporate provincial and regional level variables. However, except for rainfall in (the drought) year 1998, all provincial and regional variables turned out to be insignificant. Population density, the density of tarred and gravel roads, the average maximum temperature or the variation of rainfall did not show a significant effect, once household and community level characteristics were included. The remaining unexplained variation could not be lowered in any of the three years under consideration. This result might seem surprising, but it is in fact quite consistent with other findings in the literature. Jalan and Ravallion (2002) and Benson et al. (2005) do also not find a significant effect of population density on household income. Benson et al. (2005) even confirm our result of a missing effect of access to roads which is according to Jacoby (2000) the result of a low infrastructure elasticity of poverty. The finding that rainfall played an important role in the drought year 1998, while it did not have a significant impact in the other two survey years is also consistent with other studies. Benson et al. (2005) find the amount of rainfall in Malawi only to be significant when it is exceptionally high, while Dercon (2004) refers to the negative long-term impact of rainfall shocks, such as severe droughts, in Ethiopia. Farmers in Burkina Faso - as elsewhere in developing countries - seem to be able to cope with very low rainfall as long as there is at least some rain.

Our results are also in line with those by Bigman et al. (2000), the only other study investigating the determinants of spatial income disparities in Burkina Faso. The authors of that study conclude that regional inequality is driven by agro-climatic conditions and disparities between villages are driven by differences in infrastructure. However, compared to Bigman et al. (2000), we stress the importance of community characteristics even more. Our analysis suggests that a large part of regional disparity is actually driven by differences in community characteristics between these regions. Hence, we think the actual impact of agro-climatic conditions is lower than suggested by Bigman et al. (2000). We do, however, not negate an impact of climate, especially rainfall, on spatial income disparities. Findings of insignificant climatical effects in a Sahel country like Burkina Faso have to be interpreted with caution. Even though climatical variables are known to vary strongly on a small scale geographical unit, they are usually measured on a provincial or regional level due to an insufficient geographical distribution of monitoring stations. In addition, the distribution of rainfall over time is hard to capture.[10] Therefore, missing variation in observations rather than in actual patterns might be responsible for insignificant climatical variables. However, we get this result of strong community effects even though climate - which can just not be measured in an appropriate way - plays an important role in Burkina. In our view, policies which aim to equalize access to infrastructure could reduce regional inequalities quite substantially.

[10] We included the amount and variation of rainfall over the year as well as over the harvest and the pre-harvest period in our models.

## 4.5 Model $M4$: Variations in household level effects across communities

In a next step, we allow household level variables to differ in their impact across communities. Thus, in addition to random intercepts, we now also add random coefficients - see equation 6 - at the community level. Covariances of random effects are modeled unstructured, i.e. all variances-covariances are distinctly estimated. We use an iterative procedure to test for significant variance-covariances of all significant household level variables of model $M2$. We use likelihood-ratio tests by estimating the deviance for the model without the specific random effect and for the model with the specific random effect. We keep those random effects in model $M4$ whenever the test-statistic - the difference between the deviances of the two models - is significant, i.e. if we get a $\chi^2$ below 5% (Goldstein, 2003). In addition, variances and covariances are regarded as insignificant when their standard error is larger than their estimate (Tseloni, 2006). All estimates and their standard errors for model $M4$ are shown in Tables 8 - 10.

[insert Tables 8 - 10]

The results of our analysis are, once again, relatively homogeneous across time. We find indeed, that returns associated with education, household size and dependency ratios (children per adult and youth per adult) vary significantly across communities in all three years. On the other hand, returns associated with age and gender of the household head, with cotton farming and livestock herding do not vary significantly across Burkinabè communities in either year.

The variation of returns across communities is not only statistically but also economically meaningful. The fixed effect estimate of the variable 'literate head' of .26 in 1994 states that households with a literate head have on average around 26 percent more per capita income compared to households with an illiterate head. The variance of the random effect of the household head variable states however that this return differs significantly between communities. The effect varies between $-23$ $(.26-2*\sqrt{.055}*100)$ and 75 percent $(.26+2*\sqrt{.055}*100\%)$ between the 2.5th and 97.5th quantile of Burkinabè communities. In 1998 ($-31$ to 85 percent) and 2003 ($-29$ to 79 percent), the variation of educational returns is also highly significant across Burkinabè communities.

With an additional household member, per capita income declines between 0 and 8 percent in 95 percent of the Burkinabè communities in 1994. In 1998 households per capita income changes from 1 to $-9$ percent and in 2003 from 1 to $-12$ percent due to an additional member. In 1994, a doubling in the number of youth per adult in a family changes per capita income between 3 and $-51$ percent and in 2003 between $-4$ and $-36$ percent. In 1998, the variation in returns to youth per adult was insignificant. However, the effect of a doubling in the number of children per adult changes income between 9 and $-56$ percent.

We conclude that neighborhood has an influence on returns associated with household characteristics, in particular with education. From a policy point of view, it is important to find out what drives these neighborhood effects. In the case of returns to literacy, it might be channeled by unobserved factors like labor market characteristics or the access to modern (agricultural) production technologies. These factors will rather be found in better developed communities with high population density. However, higher returns to education could also be the result - decreasing marginal returns to education assumed - of higher marginal effects in some poor and remote communities. While the former case would rather lead to divergence of communities, the latter could lead to convergence. Note that since the regression coefficients associated with the household characteristics, which are random variables at the community level, are directly determined by the observed community level factors (see equation 3) further regression analysis it not feasible.

To get further insights, we can calculate the best linear unbiased predictors (BLUP) of the random effects and look if variations in returns across communities follow a distinct pattern across the 13 agro-climatic regions in Burkina Faso. Based on these calculations, Figures 9 to 11 show the mean of the BLUP across regions for each year. Figures 9 to 11 do not show any North-South or East-West pattern in returns to literacy across the 13 regions in any year. The same is true for the household size and dependency ratios (see figures 3 to 8). We conclude that returns to these factors are driven by small scale community characteristics but not by any regional factor.

[insert Figures 3 - 11]

[insert Table 11 and Figure 12]

Moreover, we can examine the covariance of random effects and random intercepts. For the returns to education the covariance between its random effect and the random community intercept turns out to be insignificant in 1994 and positively significant in 1998 and 2003. On average returns to education are higher in richer communities, ceteris paribus. This indicates the impact of unobserved community factors on educational returns. Labor markets are usually better developed in richer communities in a sense that they are offering more opportunities for a better educated and trained work force. Moreover, modern agricultural inputs which may require skilled labor are rather found in richer communities. Hence, there is little evidence for higher returns to education in poorer communities. This is probably due to a only weakly competitive labor market and the general low demand for skilled labor in rural areas of poor countries such as Burkina Faso. Therefore, we conclude that disparities in returns to education cannot explain convergence across departments.

Regarding the effect of household size, the covariance with the community intercept is significantly negative in all years. The same is true for the effect of the dependency rates, children and youth per adult. This is an interesting result, stating that an additional household member, at working age or not, lowers per capita income even more in otherwise richer communities. This is a finding which is usually not found in the literature. Garrett and Ruel (1999) for instance found a higher negative effect for rural than for urban households on calorie intake. In the Burkinabè context, the pressure on urban labor markets might explain this high negative effect. On the other hand, subsistence farmers might have a higher facility to feed an additional person.[11]

## 5 Concluding remarks

The objective of this paper was to explain spatial disparities in income among households. Poverty maps of developing countries often show areas - on community as well as on regional levels - with high poverty prevalence. The question is if income differentials between areas are actually driven by contextual differences or rather by a geographical segregation of people with similar characteristics. If it is geography that matters, then the question is which geographical level emits the largest impact. While a strong community impact would rather hint at a large influence of infrastructure, a strong regional impact would suggest a large impact of climate and institutions.

Multilevel models offer several options to gather statistically correct and economically meaningful information on these questions. Modeling the unexplained variation separately for each level allows to take into account the intra-class residual correlation to get efficient estimates. The contribution of each level to the total unexplained variance, the intra-class correlation coefficient, informs about the impact of each geographical level on household income. The proportional change of this coefficient after the iterative inclusion of variables on all levels, allows to gain insight on the contribution of variables on each level to income disparities on the same or higher levels. Finally, the modeling of random coefficients on some area-level informs about the extent to which area modifies returns on household characteristics. Looking at the covariances of random effects and intercepts and at a geographical clustering of random effects even allows to draw some conclusions about the channels by which area modifies returns.

Our findings show that it is not a geographical concentration of people with poor endowments that make areas poor in Burkina Faso. Put differently, area matters and contextual differences of communities, especially in education and infrastructure, and to a smaller extent of regions drive household income disparities. After taking into account the large set of household level characteristics, around 50 percent of the unexplained community and regional variation remains. Differences in community characteristics between regions account for the largest part of any remaining unexplained regional variation. Less than 5 percent in 1994 and 1998 and less than 12 percent in 2003 of the total unexplained variation on the regional level remains after inclusion of household and community factors.

Random coefficient models show that returns to literacy, household size and dependency ratios are area-specific. In contrast, returns to cotton farming and livestock herding are constant across Burkinabè communities. Significant positive covariances between random literacy effects and community intercepts

[11] Our finding of a negative covariance between intercepts and household size/dependency ratios do also hold when Ouagadougou is omitted from the regression.

show that higher returns to education are found where neighborhoods are better off.

We conclude that household income disparities in Burkina Faso are largely driven by differences in endowments of the neighborhood people live in. A clear policy implication emerges. Policy should target small scale geographical units, such as villages. By providing infrastructure, enhancing the functioning of labor markets and fostering demand for education, policy can compensate for climatical disadvantages between regions and significantly reduce poverty.

Clearly, our model is constrained by the very limited availability and the modest quality of data on all levels. In Burkina Faso, as wells as in many other developing countries, community surveys accompanying the household surveys are missing. Geo-referenced data to compensate for this lack is also most often missing. Small-scale area data is however a key in determining the effects of infrastructure on welfare outcomes. In future, household survey design should take this into account.

## References

**Aassve, A. and B. Arpino**, "Dynamic Multi-Level Analysis of Households' Living Standards and Poverty: Evidence from Vietnam," Econ Working Papers WP-2007-10 February 2007.

**Benson, T., J. Chamberlin, and I. Rhinehart**, "An investigation of the spatial determinants of the local prevalence of poverty in rural Malawi," *Food Policy*, 2005, *30* (5-6), 532–550.

**Besley, T. and L. Cord**, *Delivering on the promise of pro-poor growth: Insights and lessons from country experiences*, New York: Palgrave Macmillan, 2007.

**Bigman, D., S. Dercon, D. Guillaume, and M. Lambotte**, "Community Targeting for Poverty Reduction in Burkina Faso," *World Bank Economic Review*, 2000, *14* (1), 167–193.

**Bolstad, W.M. and S.O. Manda**, "Investigating Child Mortality in Malawi Using Family and Community Random Effects: A Bayesian Analysis," *Journal of the American Statistical Association*, 2001, *96* (453), 12–19.

**Borgoni, R., U.-C. Ewert, and A. Fürnkranz-Prskawetz**, "How important are household demographic characteristics to explain private car use patterns? A multilevel approach to Austrian data," MPIDR Working Papers WP-2002-006, Max Planck Institute for Demographic Research, Rostock, Germany February 2002.

**Dercon, S.**, "Growth and Shocks: Evidence from rural Ethopia," *Journal of Development Economics*, 2004, *74* (2), 309–329.

**DHS+, Measure**, "Burkina Faso Enquête Démographique et de Santé. Rapport de synthèse," Technical Report, ORC Macro International Inc. 2004.

**DiPrete, T.A. and J.D. Forristal**, "Multilevel Model: Methos and Substance," *Annual Review of Sociology*, 1994, *20*, 331–357.

**Ecob, R.**, "A Multilevel Modelling Approach to Examining the Effects of Area of Residence on Health and Functioning," *Journal of the Royal Statistical Society*, 1996, *159* (1), 61–75.

**Elbers, C., J.O. Lanjouw, and P. Lanjouw**, "Micro-level Estimation of Poverty and Inequality," *Econometrica*, 2003, *71*, 355–364.

**Garrett, J. L. and M. T. Ruel**, "Are Determinants of Rural and Urban Food Security and Nutritional Status Different? Some Insights from Mozambique," *World Development*, 1999, *27* (11), 1955–1975.

**Goldstein, H.**, *Multilevel Statistical Models*, Vol. 3, London: Edward Arnold, 2003.

**_ , W. Browne, and J. Rasbash**, "Partitioning Variation in Multilevel Models," *Understanding Statistics*, 2002, *1* (4), 223–231.

**Grimm, M., A. McKay, and S. Klasen**, *Determinants of Pro-Poor Growth: Analytical Issues and Findings from Country Cases*, London: Palgrave-Macmillan, 2007.

**_ and I. Günther**, "Growth and Poverty in Burkina Faso. A Reassessment of the Paradox," *Journal of African Economies*, 2007, *16*, 70–101.

**Hox, J.J.**, *Applied Multilevel Analysis*, Amsterdam: TT-Publikaties, 1995.

**Jacoby, H. G.**, "Access to Markets and the Benefits of Rural Roads," *The Economic Journal*, July 2000, *110*, 713–737.

**Jalan, M. and J. Ravallion**, "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China," *Journal of Applied Econometrics*, 2002, *17* (4), 329–346.

**Jiang, J. and P. Lahiri**, "Mixed Model Prediction and Small Area Estimation," *Test*, 2006, *15* (1), 1–96.

**Lombardía, M.J. and S. Sperlich**, "Multi-level Regression Between Fixed Effects and Mixed Effects Models," Technical Report ZfS 14.2007, Center for Statistics 2007.

**Merlo, J., M. Yang, B. Chaix, J. Lynch, and L. Rastam**, "A brief conceptual tutorial on multilevel analysis in social epidemiology: interpreting neighbourhood differences and the effect of neighbourhood characteristics on individual health," *Journal Epidemiol. Community Health*, 2005, *59*, 1022–1029.

**_ , _ , _ , _ , and _ ,** "A brief conceptual tutorial on multilevel analysis in social epidemiology: investigating contextual phenomena in different groups of people," *Journal Epidemiol. Community Health*, 2005, *59*, 729–736.

**Poel, E. Van De, O. A. O'Donnell, and E. Van Doorslaer**, "What Explains the Rural-Urban Gap in Infant Mortality - Household or Community Characteristics?," Technical Report 2007.

**Ravallion, M. and Q. Wodon**, "Poor Areas, or Only Poor People?," *Journal of Regional Science*, November 1999, *39* (4), 689–711.

**Sala-iMartin, X. X.**, "The classical approach to convergence analysis," *Economic Journal*, 1996, *106*, 1019–1036.

**Tseloni, A.**, "Multilevel modelling of the number of property crimes: household and area effects," *Journal of the Royal Statistical Society A*, 2006, *169* (2), 205–233.

**UNDP**, *Human Development Report 2007/2008* UNDP 2007.

**Vreyer, P. De, J. Herrera, and S. Mesplé-Somps**, "Consumption growth and spatial poverty traps: an analysis of the effects of social services and community infrastructures on living standards in rural peru," in F. Nowak-Lehmann and S. Klasen, eds., *Poverty, Inequality, and Policy in Latin America*, MIT Press (forthcoming), 2008.

Table 1: Descriptive Regional Growth and Poverty Statistics

| | 1994 | | | 1998 | | | | 2003 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Pop.-share | PC Expen-diture* | P0 | Pop.-share | PC Expen-diture* | Growth 94-98 | P0 | Pop.-share | PC Expen-diture* | Growth 98-03 | Growth 94-03 | P0 | Change in P0 94-03 (%points) |
| Burkina Faso | 100 | 100 | 55.51 | 100 | 82 | -17.5 | 61.81 | 100 | 108 | 31.5 | 8.5 | 47.20 | -8.3 |
| *Eastern regions* | 32.9 | 69.3 | 68.44 | 23.9 | 70.9 | 63.0 | 72.18 | 22.6 | 89.2 | 65.3 | 39.5 | 46.90 | -21.5 |
| Sahel | 5.5 | 74.2 | 62.88 | 6.4 | 67.5 | -9.0 | 59.32 | 5.8 | 124.6 | 84.5 | 67.9 | 36.89 | -26.0 |
| Est | 8.8 | 81.1 | 64.53 | 8.6 | 63.4 | -21.9 | 66.74 | 8.5 | 100.6 | 58.7 | 24.0 | 42.05 | -22.5 |
| Centre-nord | 8.8 | 69.7 | 65.04 | 8.9 | 51.2 | -26.6 | 78.35 | 8.3 | 104.8 | 104.7 | 50.3 | 36.01 | -29.0 |
| Nord | 9.8 | 55.6 | 78.09 | 9.6 | 54.8 | -1.4 | 79.89 | 8.6 | 66.2 | 20.8 | 19.1 | 68.97 | -9.1 |
| *Western regions* | 20.5 | 89.7 | 51.77 | 21.5 | 89.0 | -18.2 | 59.13 | 22.8 | 82.6 | 22.1 | -0.1 | 51.71 | -0.1 |
| Cascades | 2.2 | 85.4 | 58.34 | 3.0 | 94.6 | 10.7 | 48.16 | 3.6 | 124.3 | 31.4 | 45.5 | 38.35 | -20.0 |
| Hauts Bassins** | 8.1 | 95.1 | 40.18 | 8.0 | 75.7 | -20.5 | 54.80 | 6.9 | 105.2 | 39.0 | 10.5 | 41.43 | 1.3 |
| Mouhoun | 10.2 | 86.2 | 59.52 | 10.6 | 65.7 | -23.8 | 65.46 | 12.2 | 70.4 | 7.1 | -18.4 | 61.55 | 2.0 |
| *Cenral regions* | 35.1 | 86.4 | 60.45 | 44.5 | 81.8 | -18.8 | 65.15 | 42.5 | 79.2 | 27.4 | -0.6 | 55.49 | -5.0 |
| Sud-ouest | 4.9 | 108.7 | 54.12 | 4.2 | 62.0 | -43.0 | 64.20 | 4.9 | 75.9 | 22.3 | -30.2 | 57.94 | 3.8 |
| Centre-ouest | 10.2 | 90.2 | 61.89 | 10.7 | 76.1 | -15.6 | 61.83 | 8.6 | 108.8 | 42.9 | 20.6 | 42.13 | -19.8 |
| Plateau | 5.0 | 78.6 | 63.28 | 5.6 | 56.9 | -27.7 | 67.67 | 6.1 | 78.3 | 37.6 | -0.5 | 60.53 | -2.7 |
| Centre-est | 8.0 | 81.3 | 57.42 | 8.0 | 71.2 | -12.5 | 70.30 | 8.3 | 88.3 | 24.0 | 8.5 | 56.35 | -1.1 |
| Centre*** | 2.0 | 67.4 | 63.55 | 2.0 | 73.8 | 9.5 | 52.35 | 1.8 | 68.8 | -6.8 | 2.0 | 66.48 | 2.9 |
| Centre-sud | 5.0 | 80.2 | 64.56 | 4.4 | 55.4 | -31.0 | 67.33 | 4.3 | 65.1 | 17.6 | -18.8 | 65.89 | 1.3 |
| Urban Centers | 11.6 | 246.7 | 10.38 | 10.1 | 282.4 | -5.6 | 21.69 | 12.2 | 218.4 | 1.7 | -4.0 | 16.38 | 6.0 |
| Ougadougou | 8.2 | 258.8 | 8.44 | 7.3 | 255.6 | -1.2 | 20.51 | 8.3 | 270.2 | 5.7 | 4.4 | 13.64 | 5.2 |
| Bobo | 3.4 | 217.8 | 15.03 | 2.8 | 173.9 | -20.2 | 24.74 | 3.8 | 164.2 | -5.6 | -24.6 | 22.37 | 7.3 |

*Source: EP1, EP2, EP3, estimations by the authors*

**average per capita expenditure in Burkina Faso 1994 = 100*

*** without Bobo*

**** Without Ouagadougou*

Figure 1: Growth and Poverty Incidence on Provincial Level

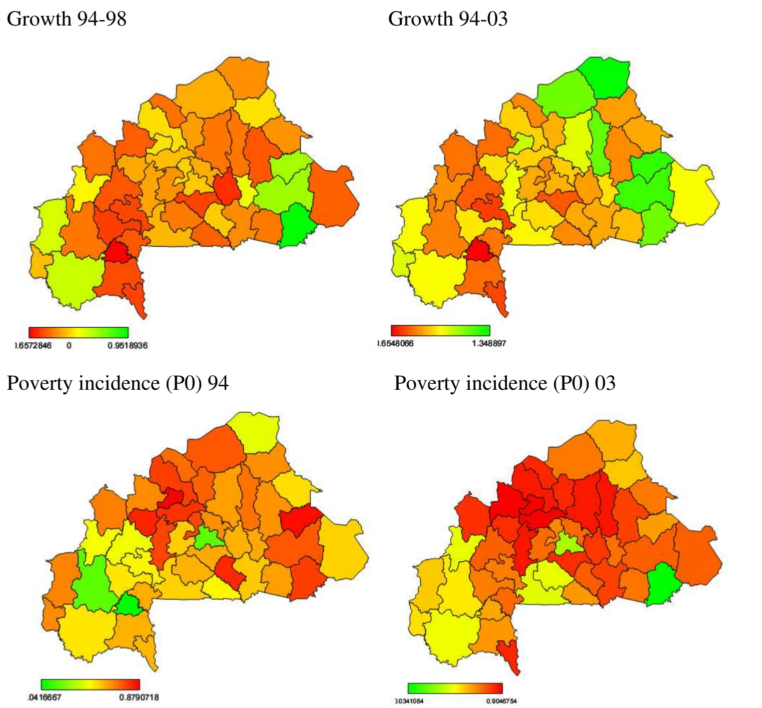


Figure 2: Convergence in Burkina Faso, initial per capita income and growth on the department level (135 observations), 1994-2003

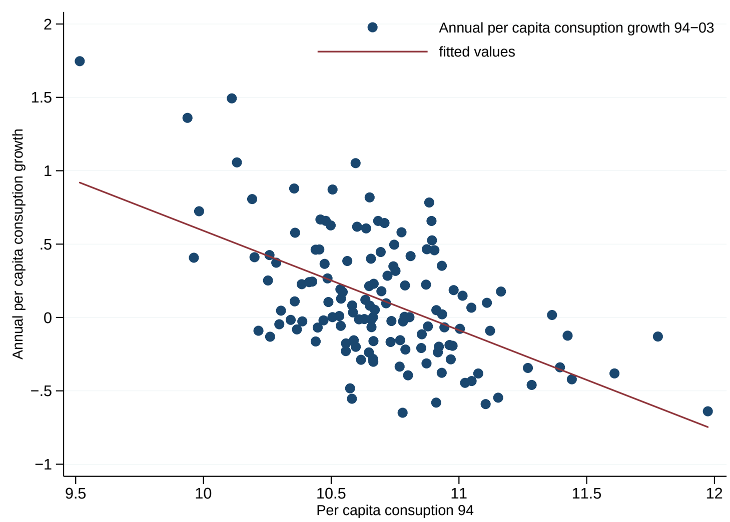

Table 2: Determinants of spatial inequality

| **Variable** | | **Descriptive Statistics** | | | | | | **Source** | **Aggregation** |
|---|---|---|---|---|---|---|---|---|---|
| Name | Label | 1994 | | 1998 | | 2003 | | | **level** |
| | | Mean | Std dev | Mean | Std dev | Mean | Std dev | | |
| Hhsize | HH size | 7.53 | 5.50 | 7.50 | 5.18 | 6.36 | 4.07 | Enq. Prioritaire | Houeshold |
| Children Adult | Children (0-6) per adult | 0.54 | 0.51 | 0.50 | 0.49 | 0.49 | 0.48 | Enq. Prioritaire | Houeshold |
| Youth Adult | Youth (7-14) per adult | 0.51 | 0.54 | 0.51 | 0.55 | 0.47 | 0.54 | Enq. Prioritaire | Houeshold |
| Elderly Adult | Elderly (55+) per adult | 0.13 | 0.30 | 0.13 | 0.30 | 0.12 | 0.30 | Enq. Prioritaire | Houeshold |
| Age | Age of HH head | 45.79 | 15.21 | 46.08 | 15.01 | 44.23 | 15.17 | Enq. Prioritaire | Houeshold |
| Sex | Sex of HH head | 1.09 | 0.29 | 1.09 | 0.28 | 1.09 | 0.29 | Enq. Prioritaire | Houeshold |
| Literate Head | Literate HH head | 0.25 | 0.43 | 0.24 | 0.42 | 0.26 | 0.44 | Enq. Prioritaire | Houeshold |
| Literate Adult | % of literate adults in hh | 0.36 | 0.49 | 0.33 | 0.46 | 0.37 | 0.47 | Enq. Prioritaire | Houeshold |
| Cotton | HH primarily engaged in cotton farming | 0.06 | 0.23 | 0.12 | 0.33 | 0.13 | 0.34 | Enq. Prioritaire | Houeshold |
| Livestock | HH engaged in some livestock herding | 0.56 | 0.50 | 0.63 | 0.48 | 0.65 | 0.48 | Enq. Prioritaire | Houeshold |
| Muslim | HH head is Muslim | 0.58 | 0.49 | 0.54 | 0.50 | | | Enq. Prioritaire | Houeshold |
| Christian | HH head is Christian | 0.24 | 0.43 | 0.25 | 0.43 | | | Enq. Prioritaire | Houeshold |
| Mossi | HH head is Mossi | 0.49 | 0.50 | 0.50 | 0.50 | | | Enq. Prioritaire | Houeshold |
| ZD Religion | Variation of religous groups in community* | 0.74 | 0.52 | 0.78 | 0.50 | | | Enq. Prioritaire | Community |
| ZD Ethnicity | Variation of ethnicity in community* | 2.05 | 1.88 | 2.29 | 2.79 | | | Enq. Prioritaire | Community |
| ZD Cotton | % of HHs primarily engaged in cotton | 0.06 | 0.14 | 0.12 | 0.25 | 0.13 | 0.25 | Enq. Prioritaire | Community |
| ZD Livestock | % of HHs engaged in some livestock | 0.56 | 0.31 | 0.63 | 0.34 | 0.65 | 0.32 | Enq. Prioritaire | Community |
| ZD Literate | % of literate adults in community | 0.36 | 0.32 | 0.33 | 0.29 | 0.37 | 0.30 | Enq. Prioritaire | Community |
| ZD Literate Head | % of literate HH heads in community | 0.25 | 0.23 | 0.24 | 0.23 | 0.26 | 0.23 | Enq. Prioritaire | Community |
| ZD Hhsize | Avg HH size in community | 7.53 | 2.19 | 7.51 | 2.39 | 6.36 | 1.56 | Enq. Prioritaire | Community |
| ZD Children Adult | Avg number of Childrens per adult | 0.54 | 0.17 | 0.50 | 0.17 | 0.49 | 0.17 | Enq. Prioritaire | Community |
| ZD Youth Adult | Avg number of youth per adult | 0.51 | 0.16 | 0.51 | 0.17 | 0.47 | 0.16 | Enq. Prioritaire | Community |
| ZD Elderly Adult | Avg number of elderly per adult | 0.13 | 0.09 | 0.13 | 0.10 | 0.12 | 0.08 | Enq. Prioritaire | Community |
| Electricity | 1 HH in community has electr. | 0.24 | 0.43 | 0.28 | 0.45 | 0.32 | 0.47 | Enq. Prioritaire | Community |
| ZD Urban | Urban community | 0.32 | 0.47 | 0.31 | 0.46 | 0.31 | 0.46 | Enq. Prioritaire | Community |
| Primary Access | Next primary school within 30 min | 0.92 | 0.27 | 0.89 | 0.31 | 0.94 | 0.23 | Enq. Prioritaire | Community |
| Secondary Access | Next secondary school within 30 min | 0.53 | 0.50 | 0.46 | 0.50 | 0.49 | 0.50 | Enq. Prioritaire | Community |

| **Variable** | | **Descriptive Statistics** | | | | | | **Source** | **Aggregation** |
|---|---|---|---|---|---|---|---|---|---|
| Name | Label | 1994 | | 1998 | | 2003 | | | **level** |
| | | Mean | Std dev | Mean | Std dev | Mean | Std dev | | |
| Healthcenter Access | Next health center within 30 min | 0.83 | 0.37 | 0.77 | 0.42 | 0.74 | 0.44 | Enq. Prioritaire | Community |
| Market Access | Next market within 30 min | 0.95 | 0.22 | 0.92 | 0.28 | 0.90 | 0.31 | Enq. Prioritaire | Community |
| Road | Access to road | | | 0.65 | 0.48 | | | Enq. Communit. | Community |
| Next Road | Distance to next road in km | | | 9.30 | 13.52 | | | Enq. Communit. | Community |
| Next Tarred Road | Distance to next tarred road in km | | | 79.50 | 87.67 | | | Enq. Communit. | Community |
| Freshwater | Access to fresh water point | | | 0.95 | 0.22 | | | Enq. Communit. | Community |
| Market | Access to market | | | 0.55 | 0.50 | | | Enq. Communit. | Community |
| Next Market | Distance to next market in km | | | 3.72 | 6.02 | | | Enq. Communit. | Community |
| School | Access to school | | | 0.67 | 0.47 | | | Enq. Communit. | Community |
| Formation | Access to formation center | | | 0.06 | 0.24 | | | Enq. Communit. | Community |
| Hospital | Access to hospital | | | 0.33 | 0.47 | | | Enq. Communit. | Community |
| Pharmacie | Access to pharmacie | | | 0.31 | 0.46 | | | Enq. Communit. | Community |
| Next Hospital | Distance to next hospital in km | | | 6.84 | 7.36 | | | Enq. Communit. | Community |
| Next Pharmacie | Distance to next pharmacie in km | | | 8.07 | 9.18 | | | Enq. Communit. | Community |
| Malaria | Malaria most frequent disease | | | 0.72 | 0.45 | | | Enq. Communit. | Community |
| Rain Mean | Avg rainfall in region | 82.67 | 17.87 | 70.88 | 17.65 | 82.65 | 24.12 | Direc. of Meteo. | Province |
| Rain Var | Variation of rainfall | 12623 | 4678 | 7712 | 3376 | 10233 | 3627 | Direc. of Meteo. | Province |
| Pop Density | Population density | 24.39 | 58.35 | 18.88 | 50.44 | 21.00 | 59.44 | Minis. of Infra. | Province |
| Landsize | Avg size of cultivated land per hh in ha | | | 4.03 | 1.68 | 5.41 | 2.08 | Enq. Agricole | Province |
| Fertilizer | Use of fertilizer | | | 0.33 | 0.25 | 0.15 | 0.14 | Enq. Agricole | Province |
| Modernequipment | Use of modern agricult. equipment | | | 0.70 | 0.20 | 0.69 | 0.28 | Enq. Agricole | Province |
| Tempmax | Avg max temperature | 34.46 | 0.76 | 35.21 | 0.99 | 35.76 | 1.02 | Direc. of Meteo. | Region |
| Tarred Size | Density of tarred roads (km/km2) | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | Minis. of Infra. | Region |

**Measured as the variance of the shares in a community*

Table 3: Models - 1994 - Fixed effects

| | M0 | | M1 | | M2 | | M4 | |
|---|---|---|---|---|---|---|---|---|
| **Household level** | | | | | | | | |
| HHsize | | | -0.040 | *** | -0.040 | *** | -0.042 | *** |
| Children Adult | | | -0.054 | *** | -0.053 | *** | -0.042 | *** |
| Youth Adult | | | -0.060 | *** | -0.055 | *** | -0.238 | *** |
| Elderly Adult | | | | | | | | |
| Age | | | -0.006 | *** | -0.006 | *** | -0.004 | *** |
| Sex | | | | | | | | |
| Literate Head | | | 0.038 | *** | 0.036 | *** | 0.260 | *** |
| Literate Adult | | | 0.440 | *** | 0.367 | *** | 0.337 | *** |
| Cotton | | | | | | | | |
| Livestock | | | | | | | | |
| Muslim | | | | | | | | |
| Christian | | | 0.056 | *** | 0.048 | ** | 0.018 | |
| Mossi | | | | | | | | |
| **Community level** | | | | | | | | |
| ZD Religion | | | | | | | | |
| ZD Ethnicity | | | | | 0.039 | *** | 0.030 | *** |
| ZD Cotton | | | | | | | | |
| ZD Livestock | | | | | | | | |
| ZD Literate Adult | | | | | 0.549 | *** | 0.390 | *** |
| ZD Literate Head | | | | | | | | |
| ZD Hhsize | | | | | | | | |
| ZD Children Adult | | | | | | | | |
| ZD Youth Adult | | | | | -0.092 | ** | -0.029 | |
| ZD Elderly Adult | | | | | | | | |
| Electricity | | | | | 0.169 | *** | 0.176 | *** |
| ZD Urban | | | | | 0.164 | *** | 0.144 | *** |
| Primary Access | | | | | | | | |
| Secondary Access | | | | | | | | |
| Healthcenter Access | | | | | | | | |
| Market Access | | | | | | | | |
| **Provincial level** | | | | | | | | |
| Landsize | | | | | | | | |
| Rain | | | | | | | | |
| Pop. Density | | | | | | | | |
| Tarred Road | | | | | | | | |
| Size | | | | | | | | |
| **Regional Level** | | | | | | | | |
| Ltempmax | | | | | | | | |
| Constant | 11.080 | *** | 11.590 | *** | 11.350 | *** | 11.330 | *** |
| AIC | 19423 | | 17065 | | 16780 | | 16187 | |
| LR test | 0.000 | | - | | - | | 0.000 | |
| Obs | 8595 | | 8595 | | 8595 | | 8595 | |

Table 4: Models - 1998 - Fixed effects

| | M0 | | M1 | | M2 | | M3 | | M4 | | M* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Household level** | | | | | | | | | | | | |
| HHsize | | | -0.040 | *** | -0.038 | *** | -0.039 | *** | -0.044 | *** | -0.038 | *** |
| Children Adult | | | -0.234 | *** | -0.229 | *** | -0.230 | *** | -0.238 | *** | -0.175 | *** |
| Youth Adult | | | -0.180 | *** | -0.169 | *** | -0.169 | *** | -0.163 | *** | -0.148 | *** |
| Elderly Adult | | | 0.084 | *** | 0.091 | *** | 0.090 | *** | 0.093 | *** | 0.112 | *** |
| Age | | | -0.006 | *** | -0.006 | *** | -0.006 | *** | -0.006 | *** | -0.006 | *** |
| Sex | | | | | | | | | | | | |
| Literate Head | | | 0.328 | *** | 0.277 | *** | 0.278 | *** | 0.270 | *** | 0.244 | *** |
| Literate Adult | | | 0.270 | *** | 0.253 | *** | 0.252 | *** | 0.214 | *** | 0.151 | *** |
| Cotton | | | 0.055 | * | 0.084 | *** | 0.102 | *** | 0.107 | *** | 0.098 | *** |
| Livestock | | | | | | | | | | | | |
| Muslim | | | 0.040 | ** | 0.036 | ** | 0.036 | ** | 0.052 | *** | 0.060 | *** |
| Christian | | | | | | | | | | | | |
| Mossi | | | | | | | | | | | | |
| **Community Level** | | | | | | | | | | | | |
| ZD Religion | | | | | | | | | | | | |
| ZD Ethnicity | | | | | 0.010 | ** | 0.011 | ** | 0.010 | ** | 0.010 | * |
| ZD Cotton | | | | | | | | | | | | |
| ZD Livestock | | | | | | | | | | | | |
| ZD Literate Adult | | | | | | | | | | | | |
| ZD Literate Head | | | | | 0.672 | *** | 0.753 | *** | 0.482 | *** | 0.345 | *** |
| ZD Hhsize | | | | | | | | | | | | |
| ZD Children Adult | | | | | | | | | | | | |
| ZD Youth Adult | | | | | | | | | | | | |
| ZD Elderly Adult | | | | | | | | | | | | |
| Electricity | | | | | 0.184 | *** | 0.211 | *** | 0.170 | *** | 0.120 | ** |
| ZD Urban | | | | | 0.129 | ** | 0.196 | *** | 0.213 | *** | 0.345 | *** |
| Primary Access | | | | | | | | | | | | |
| Secondary Access | | | | | | | | | | | | |
| Healthcenter Access | | | | | 0.054 | * | 0.059 | * | 0.051 | * | | |
| Market Access | | | | | | | | | | | | |
| Malaria | | | | | | | | | | | -.068 | *** |
| Hospital | | | | | | | | | | | .058 | ** |
| Road | | | | | | | | | | | .047 | ** |
| **Provincial level** | | | | | | | | | | | | |
| Landsize | | | | | | | | | | | | |
| Rain | | | | | | | 0.203 | * | 0.227 | * | 0.229 | * |
| Pop. Density | | | | | | | | | | | | |
| Tarred Road | | | | | | | | | | | | |
| Size | | | | | | | | | | | | |
| **Regional Level** | | | | | | | | | | | | |
| Ltempmax | | | | | | | 1.600 | | 1.470 | | 1.390 | |
| Constant | 10.860 | *** | 11.480 | *** | 11.390 | *** | 4.620 | | 5.000 | | 5.28 | |
| AIC | 19284 | | 16474 | | 16197 | | 16200 | | 15988 | | 11714 | |
| LR test | 0.000 | | - | | - | | - | | 0.000 | | - | |
| Obs | 8477 | | 8477 | | 8477 | | 8477 | | 8477 | | 6277 | |

Table 5: Models - 2003 - Fixed effects

| | M0 | | M1 | | M2 | | M4 | |
|---|---|---|---|---|---|---|---|---|
| **Household level** | | | | | | | | |
| HHsize | | | -0.054 | *** | -0.054 | *** | -0.060 | *** |
| Children Adult | | | -0.227 | *** | -0.219 | *** | -0.206 | *** |
| Youth Adult | | | -0.190 | *** | -0.178 | *** | -0.183 | *** |
| Elderly Adult | | | | | | | | |
| Age | | | -0.004 | ** | -0.004 | ** | -0.004 | ** |
| Sex | | | -0.060 | | -0.064 | | -0.074 | |
| Literate Head | | | 0.272 | *** | 0.232 | *** | 0.252 | *** |
| Literate Adult | | | 0.268 | *** | 0.2538 | *** | 0.212 | *** |
| Cotton | | | 0.079 | *** | 0.117 | *** | 0.105 | *** |
| Livestock | | | 0.034 | *** | 0.083 | *** | 0.096 | *** |
| Muslim | | | | | | | | |
| Christian | | | | | | | | |
| Mossi | | | | | | | | |
| **Community level** | | | | | | | | |
| ZD Religion | | | | | | | | |
| ZD Ethnicity | | | | | | | | |
| ZD Cotton | | | | | | | | |
| ZD Livestock | | | | | -0.347 | *** | -0.370 | *** |
| ZD Literate Adult | | | | | -0.260 | * | -0.290 | ** |
| ZD Literate Head | | | | | 0.827 | *** | 0.706 | *** |
| ZD Hhsize | | | | | | | | |
| ZD Children Adult | | | | | | | | |
| ZD Youth Adult | | | | | | | | |
| ZD Elderly Adult | | | | | | | | |
| Electricity | | | | | 0.138 | *** | 0.146 | *** |
| ZD Urban | | | | | | | | |
| Primary Access | | | | | | | | |
| Secondary Access | | | | | | | | |
| Healthcenter Access | | | | | | | | |
| Market Access | | | | | 0.088 | ** | 0.075 | * |
| **Provincial level** | | | | | | | | |
| Landsize | | | | | | | | |
| Rain | | | | | | | | |
| Pop. Density | | | | | | | | |
| Tarred Road | | | | | | | | |
| Size | | | | | | | | |
| **Regional Level** | | | | | | | | |
| Ltempmax | | | | | | | | |
| Constant | 11.150 | *** | 11.780 | *** | 12.080 | *** | 11.870 | *** |
| AIC | 19143 | | 16305 | | 16132 | | 15976 | |
| LR test | 0.000 | | - | | - | | 0.000 | |
| Obs | 8488 | | | | 8488 | | 8488 | |

Table 6: Intraclass correlation coefficients (ICC)

| | 1994 | | | 1998 | | | | 2003 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **M0** | **M1** | **M2** | **M0** | **M1** | **M2** | **M3** | **M0** | **M1** | **M2** |
| Region | 9.6% | 7.8% | 3.5% | 9.3% | 5.0% | 1.2% | 1.4% | 11.1% | 6.2% | 3.9% |
| Province | 4.1% | 2.1% | 1.0% | 3.0% | 3.1% | 2.1% | 2.0% | 3.3% | 6.3% | 7.8% |
| Community | 21.9% | 15.9% | 7.4% | 26.5% | 18.9% | 9.0% | 9.3% | 20.5% | 15.1% | 9.0% |
| Households | 64.4% | 74.2% | 88.1% | 61.2% | 72.9% | 87.7% | 87.3% | 65.2% | 72.5% | 79.3% |

Table 7: Proportional change of ICC

| | 1994 | | 1998 | | | 2003 | |
|---|---|---|---|---|---|---|---|
| | **M1** | **M2** | **M1** | **M2** | **M3** | **M1** | **M2** |
| Region | -42.5% | -62.1% | -65.4% | -79.6% | 15.5% | -62.4% | -42.3% |
| Province | -64.0% | -60.4% | -33.7% | -45.0% | -3.2% | 29.0% | 13.2% |
| Community | -49.1% | -61.1% | -54.6% | -60.5% | 3.4% | -50.3% | -45.8% |
| Households | -19.1% | -0.2% | -23.9% | -0.2% | 0.0% | -25.1% | -0.2% |

Table 8: Models - 1994 - Random effects

| | **M0** | | **M1** | | **M2** | | **M4** | |
|---|---|---|---|---|---|---|---|---|
| | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. |
| **Variances** | | | | | | | | |
| var(region) | 0.075 | 0.047 | 0.043 | 0.023 | 0.016 | 0.008 | 0.015 | 0.008 |
| var(province) | 0.032 | 0.018 | 0.012 | 0.007 | 0.005 | 0.003 | 0.006 | 0.003 |
| var(community) | 0.171 | 0.014 | 0.087 | 0.008 | 0.034 | 0.004 | 0.091 | 0.012 |
| var (household) | 0.502 | 0.008 | 0.406 | 0.006 | 0.406 | 0.006 | 0.360 | 0.006 |
| var(hhsize) | | | | | | | 0.000 | 0.000 |
| var(youth adult) | | | | | | | 0.018 | 0.006 |
| var(literate head) | | | | | | | 0.055 | 0.013 |
| **Covariances** | | | | | | | | |
| cov(hhsize, youth adult) | | | | | | | 0.000 | 0.001 |
| cov(hhsize, lit. head) | | | | | | | -0.002 | 0.001 |
| cov(youth ad, lit. head) | | | | | | | -0.007 | 0.007 |
| cov(hhsize, cons) | | | | | | | -0.005 | 0.001 |
| cov(youth ad, cons) | | | | | | | -0.018 | 0.007 |
| cov(literate head, cons) | | | | | | | 0.009 | 0.009 |

Table 9: Models - 1998 - Random effects

| | M0 | | M1 | | M2 | | M3 | | M4 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. |
| **Variances** | | | | | | | | | | |
| var(region) | 0.078 | 0.048 | 0.027 | 0.020 | 0.005 | 0.004 | 0.006 | 0.005 | 0.005 | 0.004 |
| var(province) | 0.025 | 0.016 | 0.017 | 0.010 | 0.009 | 0.005 | 0.009 | 0.005 | 0.009 | 0.009 |
| var(community) | 0.221 | 0.018 | 0.100 | 0.009 | 0.040 | 0.004 | 0.041 | 0.004 | 0.104 | 0.104 |
| var (household) | 0.510 | 0.008 | 0.388 | 0.006 | 0.387 | 0.006 | 0.387 | 0.006 | 0.357 | 0.357 |
| var(hhsize) | | | | | | | | | 0.001 | 0.000 |
| var(Children adult) | | | | | | | | | 0.027 | 0.008 |
| var(literate head) | | | | | | | | | 0.085 | 0.016 |
| **Covariances** | | | | | | | | | | |
| cov(hhsize,Children adult) | | | | | | | | | 0.002 | 0.001 |
| cov(hhsize, lit. head) | | | | | | | | | -0.003 | 0.001 |
| cov(Children ad, lit. head) | | | | | | | | | -0.026 | 0.008 |
| cov(hhsize, cons) | | | | | | | | | -0.006 | 0.001 |
| cov(Children ad, cons) | | | | | | | | | -0.031 | 0.009 |
| cov(lit. head, cons) | | | | | | | | | 0.038 | 0.011 |

Table 10: Models - 2003 - Random effects

| | M0 | | M1 | | M2 | | M4 | |
|---|---|---|---|---|---|---|---|---|
| | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. | Est. | Std.Err. |
| **Variances** | | | | | | | | |
| var(region) | 0.085 | 0.045 | 0.032 | 0.021 | 0.019 | 0.013 | 0.019 | 0.0137 |
| var(province) | 0.025 | 0.013 | 0.032 | 0.012 | 0.037 | 0.011 | 0.041 | 0.0122 |
| var(community) | 0.157 | 0.013 | 0.078 | 0.007 | 0.042 | 0.005 | 0.066 | 0.0114 |
| var (household) | 0.502 | 0.008 | 0.376 | 0.006 | 0.375 | 0.006 | 0.350 | 0.0059 |
| var(hhsize) | | | | | | | 0.001 | 0.000 |
| var(youth adult) | | | | | | | 0.005 | 0.004 |
| var(literate head) | | | | | | | 0.074 | 0.014 |
| **Covariances** | | | | | | | | |
| cov(hhsize, youth adult) | | | | | | | 0.002 | 0.0007 |
| cov(hhsize, lit. head) | | | | | | | -0.003 | 0.0011 |
| cov(youth ad, lit. head) | | | | | | | -0.012 | 0.0069 |
| cov(hhsize, cons) | | | | | | | -0.004 | 0.0011 |
| cov(youth ad, cons) | | | | | | | -0.013 | 0.0061 |
| cov(literate head, cons) | | | | | | | 0.016 | 0.0094 |

Figure 3: HHsize-94

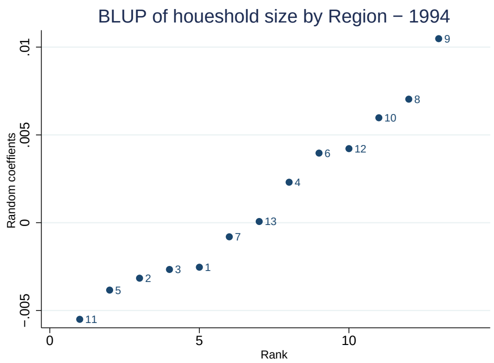


Figure 4: HHsize-98

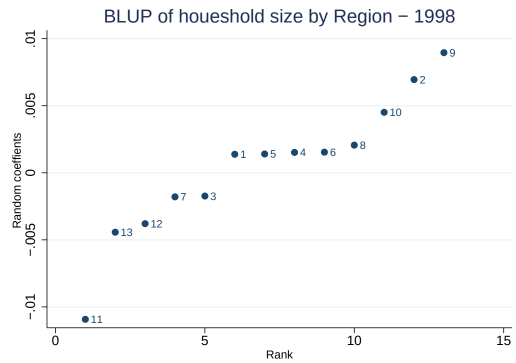


Figure 5: HHsize-03

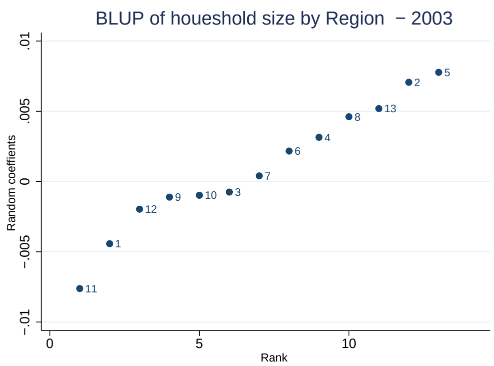


Figure 6: Youth per adult-94

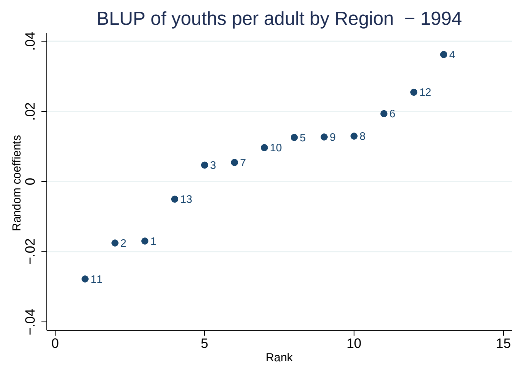


Figure 7: Children per adult-98

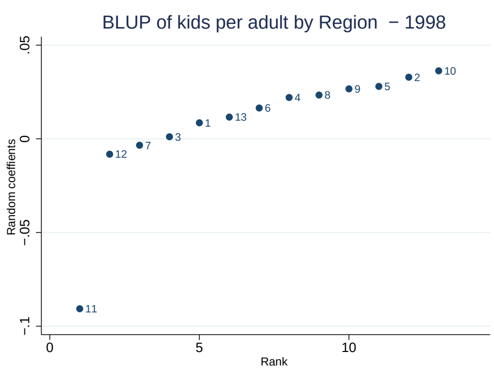


Figure 8: Youth per adult-03

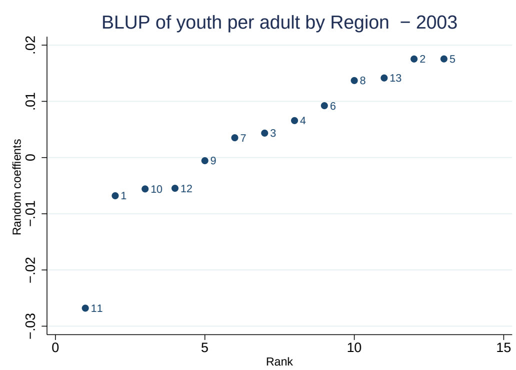

Figure 9: Education-94

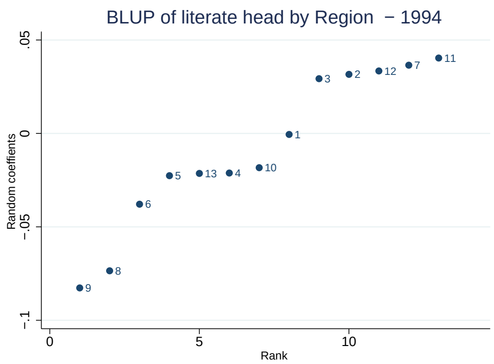


Figure 10: Education-98

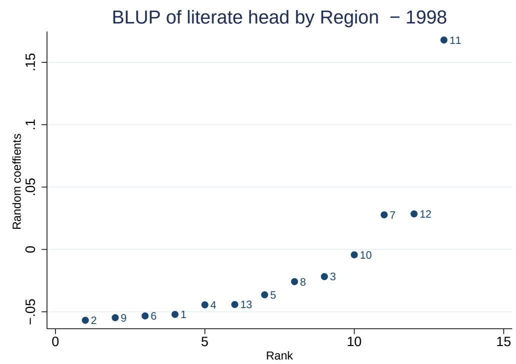


Figure 11: Education-03

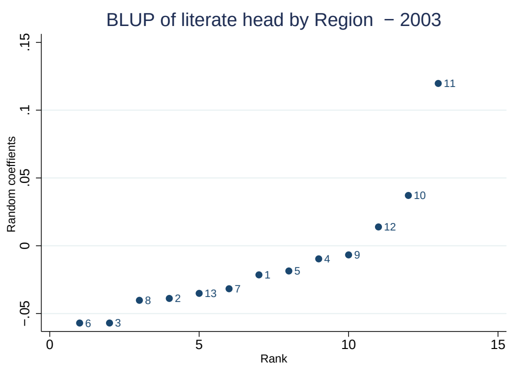


Table 11: Regional Codes

| **Code** | **Region** |
|---|---|
| 1 | Hauts Bassins |
| 2 | Boucle de Mouhoun |
| 3 | Sahel |
| 4 | Est |
| 5 | Sud-Ouest |
| 6 | Centre Nord |
| 7 | Centre Ouest |
| 8 | Plateau Central |
| 9 | Nord |
| 10 | Centre Est |
| 11 | Centre (Ougadougou) |
| 12 | Cascades |
| 13 | Centre Sud |

Figure 12: Map of Burkina Faso


*Source: DHS+ (2004)*