9th National Convention on Statistics (NCS)
EDSA Shangri-la Hotel
1 Garden Way, Ortigas Center, Mandaluyong City
October 4-5, 2004

Geographic Distribution of the Poor: Is Poverty Contaminating?

Erniel B. Barrios and Ohmar Z. Landagan

For additional information, please contact:

Author's Name:	Erniel B. Barrios and Ohmar Z. Landagan
Designation:	Associate Professor and Instructor
Agency:	School of Statistics, University of the Philippines
Address:	Diliman, Quezon City, Philippines
Telefax:	(632) 9280881
E-mail:	ernielb@yahoo.com and ozlandagan@up.edu.ph

Geographic Distribution of the Poor: Is Poverty Contaminating?

Erniel B. Barrios and Ohmar Z. Landagan

The Family Income and Expenditure Survey in 1994, 1997 and 2000 are linked at the provincial level. Various specifications of spatio-temporal models are fitted to search for evidence of spatial association among poverty indicators. There is indeed an evidence of geographic clustering of provinces according to various poverty indicators, especially for household income and expenditures.

Keywords: *poverty indicators, geographic characteristics of poverty, spatio-temporal interaction, panel data*

I. Introduction

The first goal listed in the Millennium Development Goals of the United Nations is to "Eradicate Poverty and Hunger". Under this goal, the targets are to halve the people earning under $1 a day and those suffering from hunger in 25 years since the declaration in 1990. The UNDP developed the framework for the Common Country Assessment (CCA) whose by-products are intervention policies that are tailor-fitted to the needs of the country. Poverty interventions need not be universal across countries. In the context of development anthropology, even within a country, poverty alleviation strategies should be developed in a participatory approach to ensure that the programs would indeed address the needs of the stakeholders. Glewe (1991) emphasized the need to understand the determinants of poverty as well as the tracking of its changes over time as this is a vital input to policy analysis and in the formulation of poverty reduction strategies.

There is a vast literature on various determinants of poverty. Fofack (2002) demonstrated the interaction of spatial location, burden of age dependency, human and physical assets and household amenities in shaping up of the household welfare in Burkina Faso. The sectors most vulnerable to rural poverty in India were identified by Jha(2002) and these included: those engaged in agriculture (specially the agricultural labor and those with small land area for cultivation), the scheduled caste and scheduled tribes (accounting for 42% of the rural poor), the women, the young, and those that has limited access to productive resources (including finance). The NEC, NSO and IFPRI (2001) investigated the socio-demographic variables and some community characteristics (may serve as proxy to spatial determinants) in understanding the structure and dynamics of poverty in Malawi.

Geda, et.al. (2001) emphasized the role of education, household size and being in agriculture as important determinants of poverty in Kenya. Jalan and Ravallion (1998) differentiated transient and chronic poverty in China and pointed out that household wealth and geographic characteristics generally affects chronic as well as transient poverty. While demographic as well education and health status affect chronic poverty, it is not as important in explaining transient poverty.

There is a growing literature exploring on the geographic dimensions of poverty. More information is generated that clarify the dynamics of poverty over space. It is thus, necessary for policies and interventions on poverty to consider site-specific attributes to foster a rapid impact on poverty alleviation. Thus, the purpose of this paper is to use spatio-temporal models in assessing the geographic dimensions of poverty. Evidence on the contamination or clustering of poverty among neighboring areas will be collected through an appropriate statistical model.

II. Spatio-Temporal Models

Limited marginal data over space or time often limits the information generated from modeling. One thrust in the past decade in statistics has been focused on the spatio-temporal interaction to address the inadequacy of the information on time series or on cross-sections. Spatio-temporal interactions would borrow information for the same location at some other time or borrow at the same time from some other locations. The works on panel data analysis has intensively explored on various ways of postulating models considering the nature of the data available. The field has been enrich substantially by the work of environmental statisticians whose concentration was on spatial distributions but because of the availability of monitoring data, also considered temporal distribution as well. The spatio-temporal models has so far been used to resolve a wide-array of issues foremost in environmental applications, medical/diseases mapping and many other concerns that benefits from the information on geographic distribution of certain indicators.

Lawson and Denison (2002) edited a volume that explored various themes of spatial cluster modeling. As an illustration, methods are applied in a wide range of disciplines including medicine, the social sciences and economics. Kent and Mardia (2002) explored the rationale for spatial-temporal data and presented the methodological sketch in analyzing such data taking into consideration spatio-temporal interactions. According to the motivation, of the model, Kent and Mardia (2002) identified four classes of models under spatio-temporal interactions, namely: extension of time series methods to space, extension of random fields and imaging techniques (including smoothing) to time, interaction of time and space, and physical models. Objectives are either for dimension-reduction of for prediction. The size of the data can be of various magnitudes requiring appropriate methodologies to vary extensively according to the nature of the available data indexed over time and space.

A sample of methodologies developed for various data situations are presented. Brown, et.al. (2000) investigated a physical dispersion model and showed that it has a non-separable covariance function. Model components are sequentially added generating a “blurring” effect to the space and time components. Cressie and Huang (1999) on the otherhand, derived a new approach that allows one to obtain many classes of nonseparable, spatio-temporal stationary covariance functions.

Recognizing the size of the data from monitoring systems that are in place for quite sometime, Wilke and Cressie (1999), developed an approach to spatio-temporal modeling that achieves dimension-reduction and uses a statistical models that is temporally dynamic and spatially descriptive. The result is a spatio-temporal generalization of the Kalman filter but has the flexibility of entertaining non-dynamic spatial component.

Wilke and Royle (1999) used spatio-temporal models in determining an optimal design for environmental monitoring networks. Some of the recent works using spatio-temporal models include Korie, et.al. (2000) and Cressie and Majure (1997) using environmental data.

Cressie and Majure (1997) considered the decomposition of the target indicator (log-nitrate concentration) Z(s,t) observed at location s and at time t. The decomposition postulates: $Z(s,t)=\mu(s,t)+\delta(s,t)$. The first component $\mu(s,t)$ represents the large-scale variation modelled as: $\mu(s,t)=x(s,t)'\beta$, where x(.) includes variables explanatory variables, indicator of seasonality, time events and spatial locations, β is the usual parameter vector. The second component $\delta(s,t)$ represents the small-scale variation, usually modelled after fitting the large-scale variation. The model for the small-scale variation is based on the computed residuals and usually considers temporal autocorrelation and/or spatial autocorrelation. The models proposed in subsequent sections follow similar argument as Cressie and Majure (1997).

III. The Data and Modeling Strategies

The family income and expenditure survey data from 1994, 1997 and 2000 are aggregated at the provincial level. The number of observations for 82 provinces (including the breakdown of NCR into 4 districts) for three years is 241 since there are provinces created only within the period 1994-2000.

Dependent Variables

This paper do not intend to contribute in the growing debate on how poverty should be measured hence, several indicators are analyzed. The three direct indicators of poverty includes: total income (average for the province), total expenditure (average for the province) and savings (computed at the household level and averaged for the province). Some inequity indicators were also included: proportion of food expenditures to total expenditures and proportion of income in rural to urban areas.

Independent Variables

Indicators of the demographic profile, economic profile, living conditions and assets and geographic attributes are postulated as determinants of poverty indicators.

For the demographic profile, the following indicators were included: proportion of male-headed households, proportion of households headed by <40 years old, proportion headed by married persons, proportion of nuclear families, average household size.

For the economic profile: proportion of households with employed head, proportion of households where wife is employed, average number of employed persons per household, and proportion of households engaged in agriculture.

For the living conditions and possession of assets: proportion of single detached houses, proportion of houses with strong walls and strong roof, proportion of households using water-sealed toilet, proportion of households with electricity, proportion of households with television, and proportion of households with refrigerator.

For the geographic indicators, neighborhoods are defined in terms of membership to the same region only. It will be more comprehensive though (and tedious too) to define neighbors through physical proximity and accessibility of the provinces. Data on adjacent provinces, accessibility routes, and even the ethno-linguistic relations of the provinces can truly enhance the notion of neighborhood and hence, the spatial relationship among the provinces. For this study, the hypothesis of spatial relationship is explored using two approaches. The first approach defines regional dummy variables. The second approach computes the average level of the dependent variable for the region the use it as a measure of proximity of provinces. Provinces from the same region assumes the same average for the dependent variables, hence, they are statistically postulated to be neighbors.

Notes on Modeling

For the total income and total expenditures, linear models were postulated. For savings, proportion of food to total expenditures and proportion of rural to urban income, extreme skewness is expected, thus, a logistic link function is proposed.

For the temporal pattern, both independence and first order autocorrelation (since there are only 3 time points) were postulated.

Fixed effects and random effects assumptions were also investigated.

The prototype models are as follows: $Z(s,t) = \mu(s,t) + \delta(s,t)$ where s is a province and t can either be 1994, 1997 or 2000. The function μ considers the simultaneous interaction of space (provinces) and time (FIES years). The function δ on the other hand, assumes independence in space and time. Another scenario for δ assumes first-order autocorrelation with fixed or random component among the provinces ($\delta(s,t) = \nu_s + \varepsilon_{st}$ where ν_s is a fixed or random parameter for province s, $\varepsilon_{st} = \rho\varepsilon_{st-1} + \eta_{st}$, η_{st} is a white noise).

IV. Results and Discussions

In general, the use of regional average as a measure of spatial distance yields better-fitting model than using regional dummy variables. While dummy variables can clearly differentiate regional affiliation of provinces, regional averages would further differentiate one region to another, separating them far enough to distinctively recognize one region over another.

Random effect models are also superior to fixed effect models. This is a possible evidence of dynamism of poverty situation among the provinces over time. The effect of a province on the poverty indicator is random since the effect of the alleviation effort at one time of measurement may vary at another point of measurement. This also indicates that intervention efforts in a province are not consistently stimulating poverty alleviation in one direction. The strategies applied during the period 1994-2000 need not be sustainable.

Models with autocorrelated errors are comparable to those where independence is assumed. The 3-observation period for poverty indicators may not be sufficient to assess the temporal dimensions of the indicator.

Total Income

Average household income per province is explained primarily by the spatial distance measure (average regional household income) ($p<0.000$). This is an evidence of spatial proximity of provinces in the same region. Of the total household income in province, 59% can be attributed to the average regional household income, a characteristic shared by provinces in the same region. This means that once the regional affiliation of a province is known, the household income may also be partly known.

Average household income may also admit possession of assets like refrigerator ($p<0.000$), economic activity like being in the agriculture sector ($p<0.009$), and living condition like strong wall materials ($p<0.033$).

The temporal autocorrelation of the household income is estimated at 0.11, indicating the weak dependence of current household income to the last measurement. This further supports the claim above that the movement of poverty alleviation result is not consistently moving along the same path. The current poverty alleviation strategies do not consistently push poverty status along the same direction.

The variation in household income is explained prominently by provincial variation (47%). This further proves the geographic clustering of provinces relative to total household income.

Total Expenditures

The same determinants/correlates of household incomes also explain household expenditures. The measure of spatial proximity (average household expenditure in the region) still dominated all other determinants of poverty.

The temporal autocorrelation of the household expenditure is estimated at 0.31, higher than the one estimated for income but still indicating a weak dependence of current household expenditures to the last measurement. This also supports the claim above that the movement of poverty alleviation result is not consistently moving along the same path.

The variation in household expenditure is explained prominently by provincial variation (45%). This further proves the geographic clustering of provinces relative to total household income.

Savings

With a logistic link function, the household savings per province is explained primarily by regional dummy variables, proxy indicators of spatial distance. Savings is also affected by the proportion of employed wife in a province, the proportion of households in a residence with strong wall and roof materials and with water-sealed toilet.

The temporal autocorrelation of the household savings is estimated at 0.16, indicating the weak dependence of current household savings to the last measurement. The saving behavior of the household also exhibited inconsistent pattern during the three-measurement point analyzed. The variation in household saving is explained by 25% OF provincial variation.

Proportion of Food to Total Expenditures

With a logistic link function, the proportion of food expenditure to total household expenditure per province is still explained by regional dummy variables, proxy indicators of spatial distance. Food expenditure is also affected by the proportion of employed wife in a province, proportion of male-headed households, proportion with electricity and proportion with refrigerator.

The temporal autocorrelation of the food to total expenditure ratio is negligible, indicating the absence of a single direction of the ratio over the three measurement period.

Proportion of Rural to Urban Income

There is a strong evidence of geographic clustering of the proportion of rural to urban income as exhibited by the regional dummy variables. In fact, only the regional dummy variables appeared to be significant determinants of this indicator. There is also a weak autocorrelation among the three-year data points indicating perhaps that there has been minimal movement on the ratio of the 9-year period.

V. Conclusions and Recommendations

There is an evidence of spatial clustering among the provinces with reference to selected poverty indicators. The implication is that targeted intervention is feasible and need not be different among provinces. Provinces within a region (perhaps even beyond) exhibit similar picture of poverty situation and thus, a strategy for alleviation may be adopted for a group of provinces rather than tailor-fitting it for all provinces.

There is a minimal autocorrelation for the 3-period observations points on poverty indicators. This could mean any of the following: inadequacy of the observations to establish the dependence structure, that the programs geared towards alleviation are not consistent (or sustainable) to foster a rapid change in the indicators, or that current programs are not necessarily learning for the mistakes and success of past programs.

A spatio-temporal interaction is an important component that needs to be integrated in modeling a system that is dynamic spatially and temporally. It is important however, to carefully plan how the spatial relations can be defined in terms of proxy indicators to facilitate the proper representation of the spatial relations. It is no longer proper to assume in modeling that sample points (province) are uncorrelated since the interaction of socio-cultural situations would necessarily link and group together adjacent communities.

It is recommended that the definition of neighborhood be improved to better approximate the distance. It is also important to enhance/explore other models in studying the small-scale variations (errors), perhaps an additive modeling strategy can be adopted to facilitate an interactive model development. The data may be extended from 1988, 1991 for a better understanding of the autocorrelation structure.

VI. Appendices

Table 1: Coefficient of Variation for Models for Total Income

Assumption of Small-Scale Variation	Spatial Indicators	Type of Effect	R^2
Temporally Independent	Regional Dummy Variable	Fixed	85%
Temporally Independent	Regional Average	Fixed	88%
AR(1)	Regional Dummy Variable	Fixed	17%
AR(1)	Regional Average	Fixed	73%
AR(1)	Regional Dummy Variable	Random	85%
AR(1)	Regional Average	Random	87%

Table 2: Coefficient of Variation for Models for Total Expenditures

Assumption of Small-Scale Variation	Spatial Indicators	Type of Effect	R^2
Temporally Independent	Regional Dummy Variable	Fixed	85%
Temporally Independent	Regional Average	Fixed	88%
AR(1)	Regional Dummy Variable	Fixed	35%
AR(1)	Regional Average	Fixed	82%
AR(1)	Regional Dummy Variable	Random	85%
AR(1)	Regional Average	Random	88%

Table 3: Coefficient of Variation for Models for Savings (Logit Link)

Assumption of Small-Scale Variation	Spatial Indicators	Type of Effect	R^2
Temporally Independent	Regional Dummy Variable	Fixed	33%
AR(1)	Regional Dummy Variable	Fixed	1%
AR(1)	Regional Dummy Variable	Random	37%

Table 4: Coefficient of Variation for Models for Proportion of Food to Total Expenditures (Logit Link)

Assumption of Small-Scale Variation	Spatial Indicators	Type of Effect	R^2
Temporally Independent	Regional Dummy Variable	Fixed	75%
AR(1)	Regional Dummy Variable	Fixed	85%
AR(1)	Regional Dummy Variable	Random	75%

Table 5: Coefficient of Variation for Models for Proportion of Rural to Urban Income (Logit Link)

Assumption of Small-Scale Variation	Spatial Indicators	Type of Effect	R^2
Temporally Independent	Regional Dummy Variable	Fixed	31%
AR(1)	Regional Dummy Variable	Fixed	15%
AR(1)	Regional Dummy Variable	Random	34%

References:

Brown, PE, Karesen KF, Roberts, GO, and Tonellato, S, 2000, "Blur-Generated Non-Separable Space-Time Models", *J. of the Royal Statistical Society B*, Vol. 62, Part 4, pp. 847-860.

Cressie, N and Huang, HC, 1999, "Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions", *J. of the American Statistical Association*, Vol. 94, No. 448, pp. 1330-1340.

Cressie, N and Majure, JJ, 1997, "Spatio-Temporal Statistical Modeling of Livestock Waste in Streams", *J. of Agricultural, Biological, and Environmental Statistics*, Vol. 2, No.1, pp. 24-47.

Fofack, H, 2002, "The Nature and Dynamics of Poverty Determinants in Burkina Faso", *WB Policy Research and Working Paper Series*

Geda, A, de Jong, N, Mwabu, G and Kimenyi, M, 2001, "Determinants of Poverty in Kenya: A Household Level Analysis", *Working Papers, Institute of Social Studies, The Hague-The Netherlands*

Glewwe, P, 1991, "Investigating the Determinants of Household Welfare in Cote d'Ivoire", *J. of Development Economic*, Vol. 35, No.2, pp. 307-337.

Jalan, J and Ravallion, M, 1998, "Determinants of Transient and Chronic Poverty: Evidence from Rural China", *WB Policy Research and Working Paper Series*

Jha, R, 2002, "Rural Poverty in India: Structure, Determinants and Suggestions for Policy Reforms", *Australia South Asia Research Centre, Australia National University*

Kent, JT and Mardia, KV, 2002, "Chapter 12: Modelling Strategies for Spatial-Temporal Data" in Lawson, A.B. and Denison, D.G.T., 2002, **Spatial Cluster Modelling**, Chapman & Hall/CRC, USA

Korie, S, Perry, JN, Mugglestone, MA, Clark, SJ, Thomas, CFG, Mohamad Roff, MN, 2000, "Spatiotemporal Associations in Beetle and Virus Count Data", *J. of Agricultural, Biological, and Environmental Statistics*, Vol. 5, No.2, pp. 214-239.

Lawson, A.B. and Denison, D.G.T., 2002, **Spatial Cluster Modelling**, Chapman & Hall/CRC, USA

National Economic Council(NEC), National Statistics Office(NSO), International Food Policy Research Institute(IFPRI), 2001, "The Determinants of Poverty in Malawi, 1998"

Wilke, CK and Cressie, N, 1999, "A Dimension-Reduced Approach to Space-Time Kalman Filtering", *Biometrika*, Vol. 86, No. 4, pp. 815-829.

Wilke, CK and Royle, JA, 1999, “Space-Time Dynamic Design of Environmental Monitoring Networks”, *J. of Agricultural, Biological, and Environmental Statistics*, Vol. 4, No.4, pp. 489-508.