An Empirical Evaluation of Poverty Mapping Methodology: Explicitly Spatial versus Implicitly Spatial Approach

Susan Olivia

John Gibson

Aaron Smith

Scott Rozelle

And

Xiangzheng Deng

A Contributed Paper to the Australian Agricultural & Resource Economics Society's Annual Conference, Cairns, February 11-13, 2009.

Draft for Comments
January 2009

An Empirical Evaluation of Poverty Mapping Methodology: Explicitly Spatial versus Implicitly Spatial Approach

Susan Olivia[1]*, John Gibson[2], Aaron Smith[1,] Scott Rozelle[3],and Xiangzheng Deng[4]

Abstract

Poverty maps provide information on the spatial distribution of welfare and can predict poverty levels for small geographic units like counties and townships. Typically regression methods are used to estimate coefficients from the detailed information in household surveys, which are then applied to the more extensive coverage of a census. One problem with standard regression techniques is that they do not take into account the 'spatial dependencies' that often exist in the data. Ignoring spatial autocorrelation in the regression providing the coefficient estimates could lead to misleading predictions of poverty, and estimates of standard errors. Household survey data usually lack exact measures of location so it is not possible to fully account for this spatial autocorrelation. In this paper, we use data from Shaanxi, China with exact measures of distance between each household to explicitly model this spatial autocorrelation. We also investigate which set of augmenting variables (i) census means or (ii) environmental variables mainly from satellite imagery have the most impact in soaking up unwanted spatial autocorrelation.

JEL: C31, C42, O53, P36
Keywords: China, Poverty, Small Area Estimation, Survey Methods, Spatial Models

[*] Corresponding author: olivia@primal.ucdavis.edu. [1] University of California, Davis; [2] University of Waikato, New Zealand; [3] Stanford University; [4] Chinese Academy of Sciences.

Acknowledgments: We are grateful to the Research and Education Advanced Network New Zealand Capability Fund for financial support and to Albert Park, Hongbin Li and Shuming Bao for assistance. All remaining errors are our responsibility.

1. Introduction

In most countries, poverty is concentrated in certain areas. Since resources available for poverty reduction are usually limited, public spending for poverty reduction requires careful targeting to prevent both leakages of benefits to the non-poor and missing those that are poor. Geographical targeting, where place of residence is the main eligibility criterion, may be one of the most effective ways in reaching the poor (Baker and Grosh, 1994). Unfortunately, methods that rely on traditional sources of data such as national-level household surveys are unable to provide targeting maps at a level of disaggregation that is sufficient to capture heterogeneity related to local spatial variability. For example, China's rural household survey samples 80,000 households but yields poverty estimates that are representative only for each province *(n = 31)*. But policymakers wish to know, not only which provinces are the poorest, but also which areas *within* a province are the poorest. Previous research has shown that geographic targeting is most effective when the geographic units are relatively small, such as a village, county or district (Bigman and Fofack, 2000; Elbers, Fujii, Lanjouw and Özler and Yin, 2007). While surveys lack sufficient sample points, census data can be easily disaggregated to fine levels but lacks details on income or consumption which are needed for measuring poverty.

To enable finer geographic targeting, poverty analysts have recently experimented with techniques for combining the detailed information from household surveys with the more extensive coverage of census data (*aka* small area estimation). The methodology is developed by Elbers, Lanjouw and Lanjouw (2003), hereafter denoted as ELL, and has been applied to a substantial number of countries, including Albania, Brazil, Cambodia, Indonesia, Madagascar, Mexico, Morocco, and South Africa. In some cases, the poverty maps are used by governments to target financial resources to particularly needy areas[1]. In the approach introduced by ELL, household survey data are used to estimate a model of consumption, with the explanatory variables restricted to those that are also available from a recent census (*'*the first stage model'*)*. The coefficients from this estimated model are then combined with the overlapping variables from the census (which cover all

[1] See http://www.worldbank.org/poverty for a list of applications that apply ELL's (2003) poverty mapping technique.

households), and consumption and income levels are predicted for each household in the census ('the second stage'). Using such data, we can then predict the odds of being poor for each census household and add these up to yield estimated poverty rates for disaggregated (small) geographic units. These welfare indicators are then plotted on a map, which is conventionally called a poverty map.

However, in a recent report of an expert review panel entitled "Evaluation of World Bank Research, 1998 – 2005", the ELL's (2003) poverty mapping methodology received severe critical comments (Banerjee et. al., 2006). The report claims that while the method has been a popular tool in many countries in recent years, it is increasingly understood that there are problems with the methods, or at the very least better ways to improve the precision of the predictions than the methods typically used. Above all, the major shortcoming of using ELL technique to generate poverty maps is that it cannot explicitly account for *spatial autocorrelation* that often exist in the data. This *spatial autocorrelation* can arise either because nearby locations have unobserved factors in common (e.g. deteriorating environmental conditions) or because of interaction between one household and another (e.g. poverty rate in one area is directly affected by poverty in nearby areas). The first model, of unobserved common factors, is known as a *spatial error* model while the second, of neighbour's interactions, is a *spatial lag* model. If this autocorrelation is ignored, the calculated standard errors will overstate the true precision of the local-level estimates of poverty (Tarozzi and Deaton, 2007). In response to this criticism, Elbers et al. (2008) point out that one may reduce the impact of the correlations to negligible levels by introducing a variety of cluster means measuring household and individual characteristics calculated from the census into the first stage model.

While the ELL poverty mapping technique attempts to deal with spatial autocorrelation, it necessarily does so in a way which does not rely on knowing the location of either sample or census households. In this paper, we use data from Shaanxi, China explicitly model the spatial autocorrelation in the first stage regression of the poverty mapping methodology.[2] Shaanxi is selected because it is an area of high poverty in China, with an incidence of rural poverty in 2000 that was 2.9 times as high as national average. Furthermore, it has one of the slowest rates of poverty reduction in rural China since 1981 (Chen and Ravallion, 2008) and has considerable environmental heterogeneity (Huang et al., 2007).

A key feature of our analysis is that we georeference the existing rural household survey data so that we can measure the exact distance between each households and use this information to estimate the first stage regression model using the spatial framework. The aim of such modelling is to see how good is the ELL approach to dealing with spatial autocorrelation compared with using spatial methods that utilize household location. These comparisons may matter since there are unpleasant consequences of modelling spatial effects in the wrong way. For example, ignoring a spatial error structure can cause inference problems while ignoring spatial lags can bias coefficient estimates since the omitted autocorrelated in the lag model enters through the systematic part of the model (Anselin, 1988). We will also investigate which set of augmenting variables (i) census means or (ii) environmental variables mainly from satellite imagery have the most impact in soaking up unwanted spatial autocorrelation.

The rest of the paper is organised as follows. Section 2 summarizes the ELL (2003) methodology and discusses the spatial approach used in the paper. In addition, it also discusses tests for the presence of spatial autocorrelation in the data. Section 3 describes the data used in the study. Section 4 presents the results. The final section concludes.

[2] It is not appropriate to carry on our spatial framework model into the second stage as we do not have the spatial information of the census households. Furthermore, given that our observations from the household survey are small, we will not be able to get a good estimate of poverty at the township level.

2. The implicitly spatial approach and the explicitly spatial approach

2.1. Overview of the ELL methodology

The ELL's methodology combines the strength of both the detailed information about living standards available in the household survey and the more extensive coverage of the census to derive spatially disaggregated welfare indicators. In the first stage, a model of (log) per capita consumption expenditure y_i is estimated using household survey data:

$$\ln\, y_i = \mathbf{x}_i \boldsymbol{\beta} + u_i \qquad (1)$$

where $\mathbf{x}_i$ is the vector of explanatory variables for the ith household and is restricted to those variables that can also be found in the census, $\boldsymbol{\beta}$ is a vector of parameters and u_i is the error term satisfying $E[u_i \mid x_i] = 0$. This error term can be decomposed into two independent components: a cluster specific effect η_c and a household specific effect ε_{ci}. This complex error structure allows for both spatial autocorrelation (that is, a 'location effect' common to all households in the same area) and heteroskedasticity (non-constant variance) in the household component of the error term.

In the second stage of the analysis, the estimated regression coefficients from equation (1) are applied to data from the 2000 Population Census using the characteristics included in the vector $\mathbf{x}_i$ to obtain predicted consumption for each household within the micro census. While it is possible to directly predict consumption by simply combining the characteristics for census household j, $\mathbf{x}_j^c$ with $\hat{\boldsymbol{\beta}}$ from equation (1), a more refined methodology is needed to account for the complex nature of the disturbance term (Elbers et al., 2003). Specifically, estimates of the distribution for both η and ε are obtained from the residuals of equation (1) and from an auxiliary equation that explains the heteroskedasticity in the household-specific part of the residual. A simulated value of expenditure for each household is then based on both predicted log expenditure, $\mathbf{x}_j'^c \tilde{\boldsymbol{\beta}}$ and random draws from the estimated distributions of the disturbance terms. For each simulation distributional statistics, including the poverty measures, are calculated. These simulations are repeated 100 times. For any given location, the mean across the 100 simulations of statistics such as the headcount poverty rate and the average predicted expenditure level provides the point estimates of those statistics for that location, while the standard deviations serve as estimates of the standard errors.

In this paper, we mainly focus on the first stage of the analysis and compare two different methods: the implicit non-spatial (traditional) method and the spatial econometrics approach (which has not been adequately considered in previous studies) to account for the unobserved spatial correlation - rather than the second stage of the analysis due to the fact that we do not have spatial information for the households in the population census.[3]

2.2. A spatial regression approach

As discussed earlier, a major weakness of the conventional statistical method in the poverty mapping methodology is that it does not explicitly account for spatial dependencies that often exist in the data, especially in the first stage of the regression model. Ignoring spatial autocorrelation in the first stage regression could lead to misleading estimates of the parameters. If this were the case, such analysis could result in a large proportion of poor households being excluded say from the allocation of transfers, while a number of non-poor households might be deemed as potential beneficiaries.

A key issue in adjusting for this spatial autocorrelation is that some structure has to be imposed on the data. A *spatial weight matrix W,* is one way of imposing the required structure on the study of spatial autocorrelation. This is an $N \times N$ positive and symmetric matrix which exogenously determines for each observation (row) which locations (columns) belong in its neighborhood. For non-neighbors, $w_{ij}=0$, while for neighbors the weights are either $w_{ij}=1$ (binary weights) or a function of something else, such as: $w_{ij} = 1/d_{ij}$ where d_{ij} is the distance between observations i and j (inverse distance weights). Who is a neighbor may be defined either by a distance criteria, especially with point data, or by whether they share a common border and/or vertex (contiguity) for areal

[3] In their paper, Tarozzi and Deaton (2007) point out that the ELL methodology could significantly over-state the precision of local-level estimates of poverty, if underlying assumptions of spatial homogeneity do not hold. Our paper differs from that of Tarozzi and Deaton, in a way that we focus on the failure of spatial homogeneity assumption in the first stage instead of the second stage of the analysis.

data (Wilhelmsson, 2002). The diagonal elements of the weights matrix are conventionally set to zero, and typically standardized such that the elements of a row sum to one (Anselin and Bera, 1998). Hence, the spatial weight matrix allows all of the interactions between observation *i* and each of its neighbors to be parameterized in the form of a weighted average. Specifically, for some random variable of interest *z*, each element of the spatially lagged variable *Wz* equals $\sum_j w_{ij} z_j$ which is a weighted average of the *z* values in the neighborhood of point *i*.

The spatial weight matrix is used by both main approaches for incorporating spatial effects into regression models: the *spatial lag* model and the *spatial error* model. Spatial lag dependence refers to a situation in which the dependent variable in one area is affected by the dependent variable in nearby areas. For instance, if the dependent variable is income or poverty, it is likely that the level of economic activity in one area is directly affected by the level of economic activity in neighbouring areas through migration or trade-investment linkages.

Formally, the spatial lag model is defined as:

$$Y = \rho WY + X\beta + \varepsilon \tag{2}$$

where Y is an $N\times 1$ vector of observations on the dependent variable, WY is the spatially lagged dependent variable, X is an $N\times k$ matrix of explanatory variables, ε is a vector of errors, β is the vector of regression parameters and ρ is the spatial autoregressive parameter.

Although equation (2) looks like a dynamic model from time-series analysis, one key difference causes OLS (the conventional regression model) to always be an inconsistent estimator of the spatial lag model. In the time-series context, if there is no serial correlation in the errors, ε_t there will be no correlation between y_{t-1} and ε_t and OLS will be a consistent estimator. In contrast, $(WY)_i$ is always correlated with both ε_i and the error term at all other locations. Hence, OLS is not consistent for the spatial lag model (Anselin, 1988).

In contrast to the spatial lag model, the spatial error model is defined as:

$$\begin{aligned} Y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + \mu \end{aligned} \tag{3}$$

where λ is the spatial autoregressive coefficient, μ is a vector of errors that are assumed to be independently and identically distributed and the other variables and parameters are as defined in equation (2). In this model, the error for one observation depends on a weighted average of the errors for neighboring observations, with λ measuring the strength of this relationship. This can happen if there are variables that are not included in the regression model but do have an effect on the dependent variable (omitted variable bias problem) and they are spatially correlated. For example, the quality of local government and environment factors affect income and poverty, but it is difficult to include in a regression model. Because the quality of local government and environment is likely to be spatially correlated, the error term in each area is likely to be correlated with those in nearby area. This consequently violates one of the underlying assumptions of the OLS regression model that the disturbance terms for each observation are not correlated with one another. In this case, the estimates of the coefficient are no longer efficient and can cause inference problems.

It is clear that both equations (2) and (3) are restricted versions of a more general spatial autoregressive model (SAC) with autoregressive disturbances:

$$\begin{aligned} Y &= \rho W_1 Y + X\beta + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + u \end{aligned} \tag{4}$$

It may therefore seem preferable to always begin with a model like equation (4) and test in a general-to-specific way to see if either equation (2) or equation (3) are data-acceptable. Indeed, equation (4) could always be the starting point for cross-sectional regressions because the standard OLS regression model:

$$Y = X\beta + \varepsilon \tag{5}$$

is just a special case with ρ=λ=0. However, spatial models are much more computationally demanding and for most econometric software there are limits on the sample sizes that they can accommodate (due to the need to form a weights matrix of order $N \times N$). Moreover, they have to be estimated by methods such as instrumental variables and maximum likelihood that require additional assumptions.

Spatial models embedded in equation (4) can be thought as random effects model, where part of the effects are restricted to be correlated in space and the remaining error is assumed to be uncorrelated with the right-hand-side variables (Case, 1991). Thus, equation (4) can be rewritten as:

$$Y = X\beta + u + (\rho W_1 Y + \lambda W_2 \varepsilon) \tag{6}$$

where the term in bracket is identical for every observation say within a township. That is, model as shown in equation (4) assigns to the random effect of each observation within a township: ρ times the average value of Y in townships surrounding the observation's township plus λ times the average error in townships surrounding the observation's township. [4]

It is possible to use Lagrange Multiplier (LM) tests for spatial autocorrelation, which only need the restricted model to be estimated. Therefore it is common in the spatial econometrics literature to start with an OLS model and use the residuals from that model to test against spatial alternatives. In addition to these LM tests, Moran's I test, which has some parallels with the Durbin-Watson statistic, is also widely used (Anselin and Bera, 1998). For a row-standardized spatial weight matrix, Moran's I can be expressed as:

$$I = \frac{\mathbf{e}'\mathbf{W}\mathbf{e}}{\mathbf{e}'\mathbf{e}} \tag{7}$$

where $\mathbf{e}$ is a vector of OLS residuals and $\mathbf{W}$ is the spatial weight matrix. Moran's I is asymptotically normally distributed with mean $-1/(N-1)$ and its statistical significance can be evaluated from a standardized normal table. A feature of Moran's I is that the alternative hypothesis does not specify the process generating the autocorrelated disturbances. However, there is a simple intuition for Moran's I because for any variable

[4] Although one could argue that the β parameters in equation (4) can be estimated using a least squares dummy variable estimator of the corresponding fixed effect model: $Y = X\beta + D\delta + u$ where D is a matrix of dummy variables. However, the fixed effect model is only useful for poverty mapping if one has enough observations in the first stage model to get a good estimate of the fixed effect model, which is rarely the case.

z in deviation from mean form, I is equivalent to the slope coefficient in a linear regression of **Wz** on **z** (Anselin, 1995).

The LM tests are based on explicitly specified alternative hypotheses. For testing OLS against the spatial error model (λ=0) the test statistic is:

$$LM_{\lambda} = \left[\mathbf{e}'\mathbf{W}\mathbf{e}/\hat{\sigma}^{2}\right]^{2} / T \tag{8}$$

where $T = tr(W' + W)W$ and LM_{λ} is distributed as χ^2 with 1 degree of freedom. For testing OLS against the spatial lag model (ρ=0) the test statistic is:

$$LM_{\rho} = \left[\mathbf{e}'\mathbf{W}\mathbf{Y}/\hat{\sigma}^{2}\right]^{2} / T_{1} \tag{9}$$

where $T_1 = (\mathbf{WX}\hat{\beta})'\mathbf{M}(\mathbf{WX}\hat{\beta}) / \hat{\sigma}^2 + T$ and $\mathbf{M} = \mathbf{I} - \mathbf{X(X'X)}^{-1}\mathbf{X}'$. One difficulty with both LM_{λ} and LM_{ρ} is that they each have power against the other alternative. In other words, when testing λ=0, LM_{λ} responds to nonzero ρ and when testing ρ=0, LM_{ρ} responds to nonzero λ. To test in the possible presence of both spatial error and spatial lags, Anselin et al. (1996) develop specification tests for spatial lags that are robust to ignored spatial errors and tests for spatial errors that are robust to ignored spatial lags. These tests denoted LM_{λ}^{*} and LM_{ρ}^{*} should be used when both LM_{λ} and LM_{ρ} are statistically significant.

All five of the spatial autocorrelation tests described here are used in the current study. Depending on the outcome of the specification tests, the regression model for (log) per capita consumption expenditure will be re-estimated in either the spatial lag or spatial error framework.

3. Data

The data come from four sources: (i) the 2001 Rural Household and Income Expenditure Survey conducted by the China's National Bureau of Statistics; (ii) GPS data on the location of households in the RHIES (conducted by ourselves); (iii) the 2000 Population Census; and (iv) environmental variables from the satellite remote sensing data for Shaanxi.

The 2001 Rural Household Income and Expenditure Survey (RHIES), as its name implies, collected information on the income and expenditure of households. Apart from this, the survey also collected information on household characteristics, employment, seasonal labor migration, agricultural production, dwelling characteristics, ownership of durable goods and fixed assets and access to public infrastructure. The RHIES used a random multi stage sampling of 1,990 households in Shaanxi, with 34 counties selected in the first stage and 182 townships selected from amongst these counties. Each of these townships has many constituents villages, so typically one and occasionally two villages per township were selected by the RHIES. During the final stage, 10 households from each selected village are selected randomly into the sample. For the analysis reported here, we obtained from the statistics bureau a randomly chosen sub-sample of just under 60 percent of these RHIES households ($n = 1068$), located in 22 counties. Figure 1 shows the location of the counties and townships in the sub-sample that we use.

Like the household survey in many other developing countries, the Chinese version did not geo-reference households in part because of lack of information about the benefits. To precisely measure the distance between sampled households, we retrospectively (in 2007) established the exact coordinates of the primary dwelling of each sample household. To carry out this work, we first obtained a list of the village and household names of the 2001 RHIES sample from the State Statistical Bureau (SSB) in Shaanxi. This step of our task required the close cooperation between our colleagues in China and the personnel in the provincial SSB since this information is not part of the standard RHIES data in order to protect confidentiality. [5]

After the counties, towns, villages and households were identified, the second step of our data collection effort involved recruiting teams of enumerators and training them in the operation of the GPS equipment. All of the enumerators were masters students from the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of

[5] In our agreement with the statistical authorities, we only used the data for obtaining the GPS coordinates and then we eliminated this identifying information from our data set.

Science. Once each team entered a village that was part of the RHIES sample, the first task was to find the exact location of each household from the list of RHIES respondents provided by the statistics bureau. This was done by getting one or more villagers who were familiar with the households in their villages to guide the team. Because villages in China are relatively small (on average, the distance between the furthest two households within a single village was only 0.5 kilometres), most villagers could easily take us to all of the households. In each case, the guide was able to tell us if the household had moved their residence between the time of the original survey (2001) and the time of our fieldwork for the GPS data collection (2007).

The environmental component of this research uses a variety of spatially referenced variables that provides information on land cover, rainfall, temperature, elevation and terrain slope for Shaanxi, which can be considered part of what Ravallion (1998) calls geographic capital. The land cover data are from satellite remote sensing data provided by the US Landsat TM/ETM images which have a spatial resolution of 30 by 30 meters. These data have been interpreted, involving considerable ground-truthing and aggregated into 1 kilometer by 1 kilometer at the county level by Chinese Academy of Sciences – CAS (Liu et al., 2003a and 2003b). These data have previously been used by Deng et al. (2002, 2003 and 2008). A hierarchical classification system of 25 land-cover classes was applied to the data and the total land area of each county were aggregated from the 25 classes of land cover in this study. The data for measuring *rainfall* (measured in millimetres per year) and *temperature* (measured in degrees centigrade per year) are from the CAS data centre but were initially collected and organized by the Meteorological Observation Bureau of China from more than 600 national climatic and meteorological data centres.

The *elevation* and *terrain slope* variables, which measure the nature of the terrain of each county, are generated from China's digital elevation model data set that are part of the basic CAS data base. A variable to measure the share of plain area is also is created by dividing the land area in a county that has a slope that is less than eight degrees by the total land area of the county. Information on the properties of soil also is part of our set of geographic and climatic variables from the CAS data center. Originally collected by a special nationwide research and documentation project (the *Second Round of China's National Soil Survey*) organized by the State Council and run by a consortium of universities, research institutes and soils extension centres, we use the data to specify two variables: the loam and organic content of the soil (measured in percent).

In addition, a variable that measures the density of a county's highway network is also included in this study. This variable is based on a digital map of transportation networks that exist in each county. It was developed by CAS and the measure includes all highways, national expressways, provincial-level roads and other more minor roads in the mid-1990s. The variable (henceforth—highway density) is measured as the total length of all highways in a county divided by the land size of the county.

4. Results

4.1. Results of Testing for Spatial Autocorrelation Effects

Results are presented initially for tests of spatial autocorrelation in household per capita consumption in rural Shaanxi. The spatial autocorrelation in a dependent variable need not cause any concern if it is removed by the covariates in the regression model. Therefore, the results of testing the residuals of an OLS consumption model are also reported.

A spatial weights matrix is needed to test for spatial autocorrelation and in turn requires a measure of distance between household. Latitude and longitude coordinates for each household were used to calculate this and then Moran's Index of spatial autocorrelation in household per capita consumption was calculated for varying neighbourhood sizes of 1 – 50 km (Figure 2). When households within a one kilometer radius of location i are considered as the neighborhood, Moran's I = 0.24 and is statistically significant ($p<0.01$). In other words, a regression of spatially weighted average per capita consumption within this neighborhood, W_s on per capita consumption of each household s, would have a statistically significant coefficient of 0.24. The strength of the spatial autocorrelation declines sharply as the neighborhood is defined to include a larger area. For all neighborhood sizes considered, Moran's I is still statistically significant.

To see whether this spatial autocorrelation is also transmitted to the residuals of an OLS regression, a (log) per capita consumption model (first stage model) was estimated using household-level variables typically found in poverty mapping studies in the literature. These variables included type and size of dwelling, household size and age composition, proportion of household members with primary, secondary and higher education, indicator for households who engaged in non-agricultural sector and households who use LPG as main cooking fuel. The first stage model of consumption, which is estimated for 1,068 rural households from the sample survey, is reported in Table 1. The model suggests that per capita consumption is higher for households with larger dwellings (as a proxy for housing quality and wealth), with a greater number of their members engaged in the non-agricultural sector. Having access to safe drinking water as well as having sanitary facility in the house also leads to a higher level of consumption. On the other hand, consumption is lower for households with a greater proportion of kids aged 6 years and below, greater proportion of youths aged 7 – 15 years, greater proportion of adults and greater proportion of elderly in the household.

Column (2) of Table 1 reports the first stage model of consumption based on household characteristics as well as the township level means of the household level variables from the census. The use of census means in the survey model of consumption has been recommended by Elbers et al. (2003) as a way to proxy for location-specific correlates of consumption, which can help to make the cluster specific variance η_c smaller and improve precision of the second stage predictions. This model has an R^2 of 0.23, as compared with 0.20 for the model in column (1) that is without the census means but otherwise has the same variables. However, many of the added variables are statistically insignificant, with the exception of the proportion of households in the township living in brick dwelling, having three generations living under the same roof and higher proportion of adults in the household. The coefficients on most of the household variables that were already in the model generally maintain their size and significance.

Most applications of ELL's (2003) method do not include any environmental variables and instead rely mainly on census and survey variables (see Table 2). However, there are number of geographic variables that may help to explain the spatial patterns in poverty in rural Shaanxi. To capture excluded location effects and other elements of geographical capital, we augment the model in column (1) with environmental variables. Inclusion of environmental variables raises the value of R^2 of the consumption model from 0.20 to 0.25 and these variables are jointly statistically significant with a *F-statistic* of 9.65, suggesting that consumption is highly related to the characteristics of the environment of where people live. The environmental variables show that consumption is lower for households in areas on steep slopes, with higher temperature and soils with higher percentage of organic matter. Soil with lower percentage of loam is correlated with lower consumption. On the other hand, consumption is higher for households in areas with higher annual rainfall and higher density of highways.

Column (4) of Table 1 reports the results of augmenting the model with household characteristics with environmental variables and means of the census variables. Inclusion of these variables raise the R^2 from 0.20 to 0.29. Most of the household characteristics in this model maintain the same sign as they had in the model estimated only on household variables (i.e. the model reported in column (1)). However, the inclusion of the location variables (both environmental and census means) reduces the size and significance of the coefficient dummy for households with access to safe drinking water and number of household members engaged in non-agricultural activities, alters the significance of the coefficients on household members with vocational degree and above. The inclusion of the location variables also alters the coefficients on some of the environmental variables, in particular strengthening the positive (negative) impact that living in areas with higher rainfall (higher percentage of organic matter) has on the value of household consumption.

How reliable are the OLS results reported in Table 1 in terms of ignored spatial autocorrelation? Tests using the methods described in Section 2 were used with two different types of weights – binary and inverse distance. In this paper, the spatial weights matrix is defined in a way such that households within the same township are considered as neighbours; else they are considered as non-neighbours. Note that the minimum feasible distance (to prevent 'islands' with no neighbours) between two households averages at 0.9 km.

According to these tests, there is substantial evidence of misspecification in the OLS results (see Table 3). The Moran's *I* statistics across four models reported in Table 1 are statistically significant suggesting that the models have spatially correlated errors. The strength of the spatial autocorrelation in the error terms decline as we include locational conditions such as census means of household characteristics and environmental variables in the model. That neighbouring observations have related residuals and that the correlation of the residuals between observations i and j, for $i \neq j$ are quite strong especially at the rather small distance range can also be seen in Figures 3-6, which show that including environmental variables helps in reducing the spatial error correlation. Including township averages also reduces the spatial error correlation but not to a negligible level as suggested by Elbers et al. (2003). Taken together, these indicators suggest that modelling a spatial component in the first stage regression model of the poverty mapping methodology is appropriate.

Both the LM_{λ} and the LM_{ρ} tests are statistically significant, so it is necessary to use the specification tests for spatial lags that are robust to unaccounted for spatial errors, and the tests for spatial errors that are robust to ignored spatial lags. According to these robust tests there is less evidence in favour of the spatial lag model (i.e., the values of LM_{λ}^{*} are almost always above the threshold for statistical significance while those for LM_{ρ}^{*} are sometimes below the threshold), with the exception of the model that incorporating environmental variables, when the inverse distance weight matrix is used.

4.2 Spatial Regression Results

In light of the above results about the misspecification when OLS is used, a variety of spatial lag and spatial error models were estimated. A comparison of the maximized log likelihoods of the resulting models indicated that there was better performance when the spatial weight matrix was based on inverse distance rather than a simple 0/1 set of weights.

Tables 4 and 5 contain a comparison of the previously presented OLS results on three types of unrestricted exogenous variables (household characteristics, census means of household characteristics, environmental variables) from the results of the spatial error and spatial lag models respectively (based on inverse distance weights and a neighbourhood size of 0.9 km).

As can be seen from Table 4, the spatial parameters corresponding to the spatial error (λ) and spatial lag (ρ) are statistically significant across 4 sets of estimation. According to the maximum likelihood estimates, in the spatial error specification λ =0.34 with a standard error of 0.04. Table 4 also show that location attributes can go a considerable way towards reducing spatial autocorrelation, but not fully eliminate the spatial autocorrelation in the data. Once we move from the model that includes only household characteristics to model with environmental variables and census means, the spatial parameter is reduced to λ =0.21 with a standard error of 0.04. In other words, the spatially weighted residual per capita consumption within a 0.9 km is significantly associated with the residual of per capita consumption for a particular household even after controlling for household characteristics and limited set of location attributes (census means and environmental variables – Column 4 of Table 4).

The results reported in Table 4 also allows one to assess which set of augmenting variables: census means or environmental variables have the most impact in soaking up unwanted spatial autocorrelation in the poverty mapping methodology. According to our results, including environmental variables in the first stage model of consumption reduces the spatial correlation by almost 20 percent ($\lambda_{env}=0.27$ cf. $\lambda_{hh}=0.34$) , while the reduction is smaller when we augment the model with census means ($\lambda_{census\,means}=0.29$ cf. $\lambda_{hh}=0.34$) . However, when both environmental variables and census means are used in the model, the results confirm that including these variables further reduce the unwanted spatial autocorrelation by almost 38 percent ($\lambda_{env+census\,means}=0.21$ cf. $\lambda_{hh}=0.34$) . The above patterns also hold when binary weights matrix is used in the model (Appendix Table 1).

This result suggesting including both census mean variables and environmental variables in the first stage regression model can reduce the impact of the correlations to negligible levels. However, integrating both census means and environmental variables did not soak up all the spatial autocorrelation, implying that there is a need to explicitly model the spatial autocorrelation, if not OLS model is likely to be mis-specified. There are two supporting evidence that evidence from the explicitly spatial models are superior to the OLS estimates. First, the increase in the likelihood functions over OLS when we allow for spatial correlation in the dependent variable and in errors range from 16 to 50 points. Secondly, when the spatial error model is used, standard errors are generally smaller than those for the OLS model.

The results are largely the same when the spatial lag model is used. The spatial lag parameters (ρ) are statistically significant across four models. The value of the correlation coefficient say using only the household characteristics ρ = 0.32 indicates that on average, a 10 percentage point increase in per capita consumption in a particular location will result in a 3.2 percentage point increase in the per capita consumption in a neighbouring location *ceteris paribus.* This seems to suggest a strong evidence of spill over effects in rural Shaanxi. The same conclusion is reached when we estimate the spatial model using the binary weight for neighbourhood of 0.9km (Appendix Table 2).

5. Conclusions

In this paper, we take an explicit spatial econometric approach in estimating specifications that incorporate spatial dependence in the first stage of consumption model of the poverty mapping exercises. The significance of the spatial parameters indicates that spatial dependencies should be incorporated in the poverty mapping methodology. Ignoring a spatial error structure can cause inference problems while ignoring spatial lags can bias coefficient estimates since the omitted autocorrelation in the lag model enters through the systematic part rather than the random part of the model. Given that the ELL's (2003) poverty mapping methodology depends on a model specification that is carefully chosen such that the explanatory variables on the right hand side of the consumption function (the first stage model) need to be restricted to those variables that are also available from a recent census, along with aggregated level variables to capture latent cluster-level effects. Thus, if the spatial autocorrelation exists in the data, the estimation methodology that does not explicitly take this autocorrelation into account, could significantly over-state the precision of local-level estimates of poverty in the second stage of the analysis. This may result in a large proportion of poor households being excluded from say the allocation of transfers while a number of non-poor households might be deemed as potential beneficiaries. Our results show that both census means and environmental variables can go a considerable way towards removing spatial autocorrelation, however, these locational control variables did not soak up all of the unwanted spatial autocorrelation. The results suggest that in order to explicitly model the spatial effects, analysts need to know actual distance between households, which in this sense are supportive of the growing use of GPS in household surveys. There are several other improvements in both analysis and survey implementation that can result from the more accurate location and distance data that GPS allows (Gibson and McKenzie, 2007).

References

Ahmad, Y. and C. Goh (2007), 'Poverty maps of Yunnan province: Uses and lessons for scaling-up in China', in T. Bedi, A.Coudouel, and K. Simler (eds.), *More than a pretty picture: Using poverty maps to design better policies and interventions.* Washington, D.C.: World Bank.

Alderman, H., M. Babita, G. Demombynes, N. Makhatha, and B. Özler, B. (2003), 'How low can you go? Combining census and survey data for mapping poverty in South Africa', *Journal of African Economics* **11**(2): 169 – 200.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Kluwer Academic Publishers: Dordrecht.

Anselin, L., Bera, A., Florax, R. and Yoon, M. (1996). 'Simple diagnostic tests for spatial dependence', *Regional Science and Urban Economics* **26**(1): 77 – 104.

Anselin, L. and Bera, A. (1998). 'Spatial dependence in linear regression models with an introduction to spatial econometrics' in Ullah, A. and Giles, D. (eds), *Handbook of Applied Economic Statistics.* Marcel Dekker: New York.

Baker, J. and M. Grosh (1994), 'Poverty reduction through geographic targeting: How well does it work?', *World Development* **22**(7): 983 – 995.

Benson, T., J. Chamberlin and I. Rhinehart (2005), 'An investigation of the spatial determinants of the local prevalence of poverty in rural Malawi', *Food Policy* **30**(5-6): 532 – 550.

Bigman, D. and Fofack, H. (2000), 'Geographical targeting for poverty alleviation: An introduction to the special issue', *World Bank Economic Review* **14**(1): 129 – 145.

Case, A. (1991). 'Spatial patterns in household demand', *Econometrica* **59**(4): 953 – 965.

Chen, S. and M. Ravallion (2008), 'China is poorer than we thought, but no less successful in the fight against poverty', World Bank Policy Research Working Paper No. 4621. Washington, D.C: World Bank.

Dasgupta, S., U. Deichmann, C. Meisner, and D. Wheeler (2005), 'Where is the poverty-environment nexus? Evidence from Cambodia, Lao PDR and Vietnam', *World Development* **33**(4): 617 – 638.

Elbers, C., J. Lanjouw, and P. Lanjouw (2003), 'Micro-level estimation of poverty and inequality', *Econometrica* **71**(1): 355 – 364.

Elbers, C., P. Lanjouw, J. Mistiaen, B. Özler, and K. Simler (2003), 'Are neighbors equal?', UNU-WIDER Discussion Paper No. 52. Helsinki: United Nations University.

Elbers, C., Lanjouw, P., and Leitye, P. (2008). 'Brazil within Brazil: Testing the poverty map methodology in Minas Gerais', *World Bank Policy Research Working Paper No. 4513.* World Bank: Washington, D.C.

Elbers, C., Fujii,T., Lanjouw, P., Özler, B. and Yin, W. (2007), 'Poverty alleviation through geographic targeting: How much does disaggregation help?', *Journal of Development Economics* **83**(1): 198 – 213.

Fujii, T. (2004), 'Commune-level estimation of poverty measures and its application in Cambodia', UNU-WIDER Discussion Paper No. 48. Helsinki: United Nations University.

Gibson, J., G. Datt, B. Allen, V. Hwang, M. Bourke and D. Parajuli (2005), 'Mapping poverty in rural Papua New Guinea', *Pacific Economic Bulletin* **20**(1): 27 – 43.

Gibson, J. and McKenzie, D. (2007). “Using global positioning systems in household surveys for better economics and better policy”, *World Bank Research Observer* **22**(2): 217 – 241.

Healy, A. and S. Jitsuchon (2007), ‘Finding the poor in Thailand’, *Journal of Asian Economies,* **18**(5): 739 – 759.

Hentschel, J., J. Lanjouw, P. Lanjouw, and J. Poggi (2000), ‘Combining census and survey data to trace the spatial dimensions of poverty: A case study of Ecuador’, *World Bank Economic Review* **14**(1): 147 – 65.

Hoogeveen, J. (2005), ‘Measuring welfare for small but vulnerable groups: Poverty and disability in Uganda’, *Journal of African Economies,* **14**(4): 603 – 631.

Huang, Q., R. Wang, Z. Ren, J. Li, and H. Zhang (2007), ‘Regional ecological security assessment based on long periods of ecological footprint analysis’, *Resource, Conservation and Recycling* **51**(1): 24 – 41.

López-Calva, L., L. Rodríguez-Chamussy, and M. Székely (2007), ‘Poverty maps and public policy: Lessons from Mexico’ in Bedi, T., Coudouel, A. and Simler, K. *More than a pretty picture: Using poverty maps to design better policies and interventions.* World Bank: Washington, D.C.

Minot, N. and B. Baulch (2005), ‘Poverty mapping with aggregate census data: What is the loss in precision?’, *Review of Development Economics,* **9**(1): 5 – 24.

Mistiaen, J., B. Özler, T. Razafimanantena and J. Razafindravonona (2002), ‘Putting welfare on the map in Madagascar’ World Bank Africa Region Working Paper Series No. 34. World Bank: Washington, D.C.

Simler, K. and V. Nhate (2005), ‘Poverty, inequality and geographic targeting: Evidence from small-area estimates in Mozambique’ IFPRI FCND Discussion Paper No. 192. Washington, D.C.: IFPRI.

Suryahadi, A., W. Widyanti, D. Perwira, S. Sumarto, C. Elbers, and M. Pradhan (2003), ‘Developing a poverty map for Indonesia: An initiatory work in three provinces’ SMERU Research Paper May 2003. Jakarta: SMERU.

Tarozzi, A. and Deaton, A. (2007). “Using census and survey data to estimate poverty and inequality for small areas”, *mimeo* Department of Economics, Duke University.

Wilhelmsson, M. (2002). ‘Spatial models in real estate economics’, *Housing Theory and Societ*y, **19**(1): 92 – 101.

Figure 1. Sampled Counties and Townships in the Rural Household Income and Expenditure Survey for Shaanxi

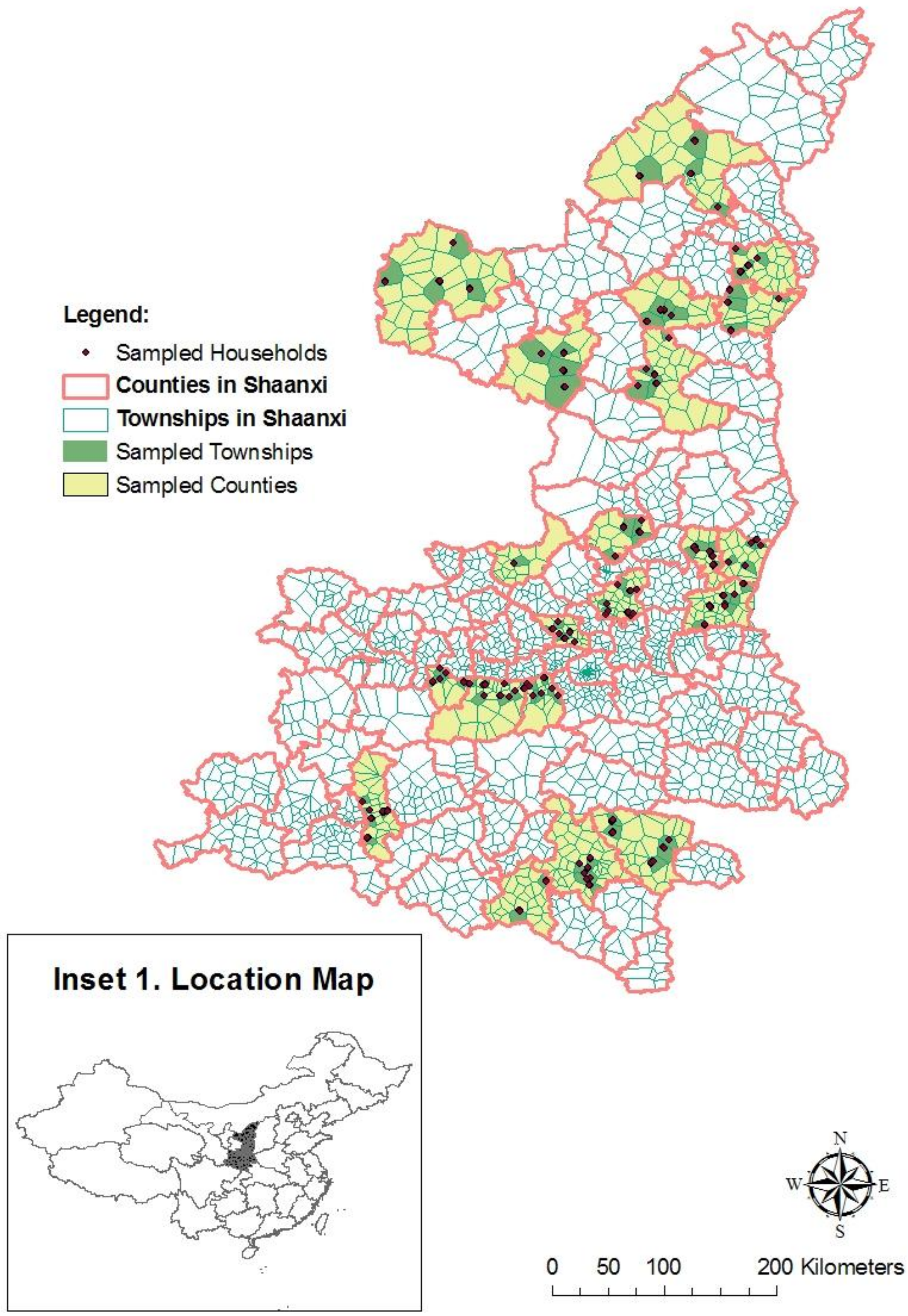

Figure 2. Spatial Correlation in Per Capita Household Consumption

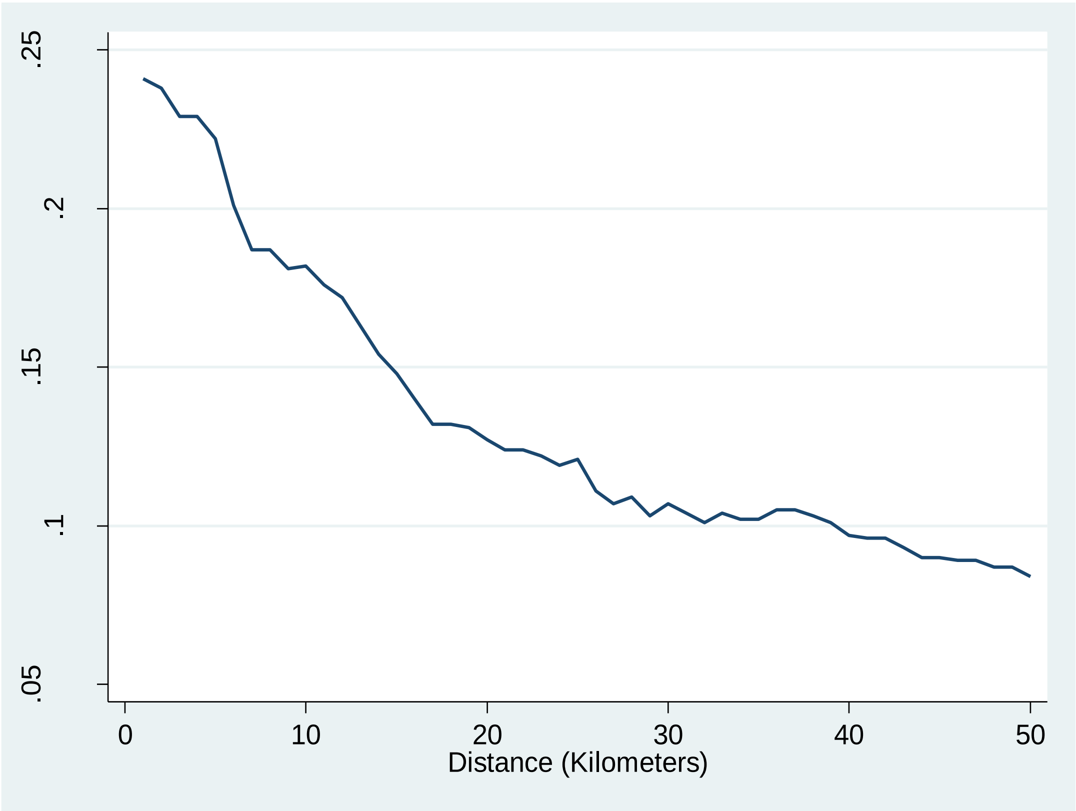

Figure 3. Residuals Correlation Plot based on the Model per Capita Households on Household Characteristics

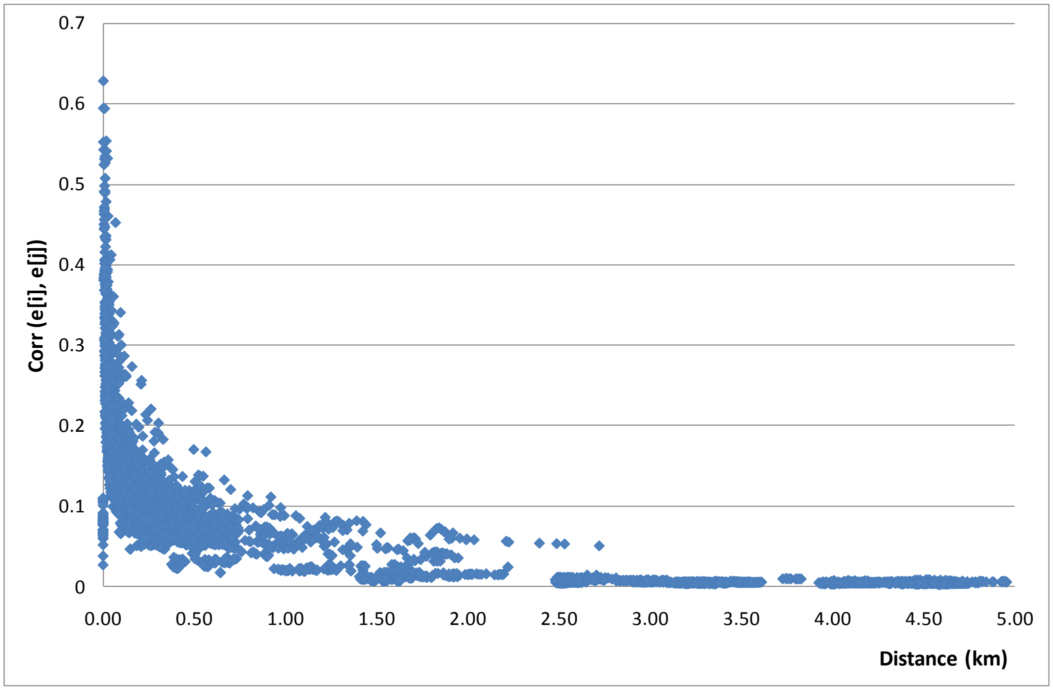

Figure 4. Residuals Correlation Plot based on the Model per Capita Households on Household Characteristics & Township Level Means

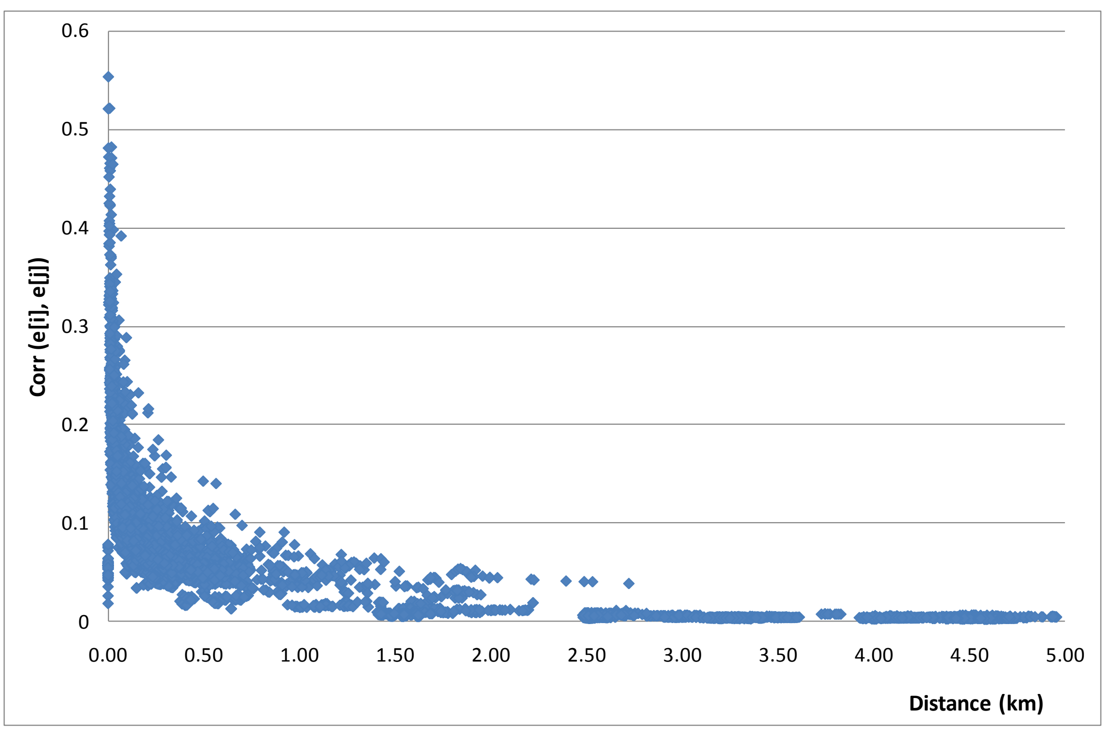

Figure 5. Residuals Correlation Plot based on the Model per Capita Households on Household Characteristics & Environmental Variables

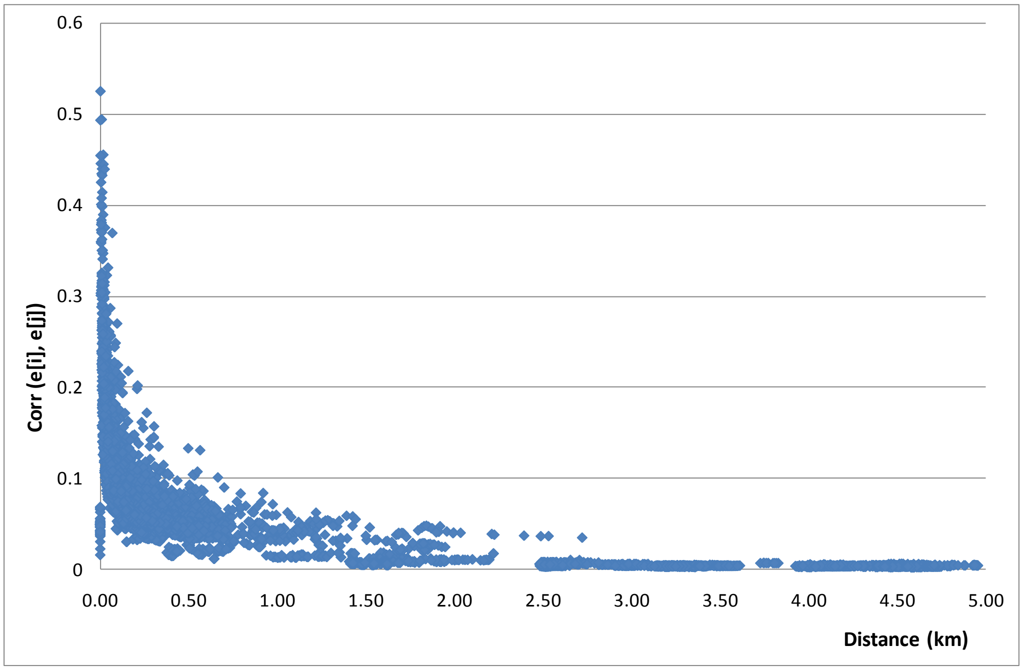

Figure 6. Residuals Correlation Plot based on the Model per Capita Households on Household Characteristics & Environmental Variables & Township Level Means

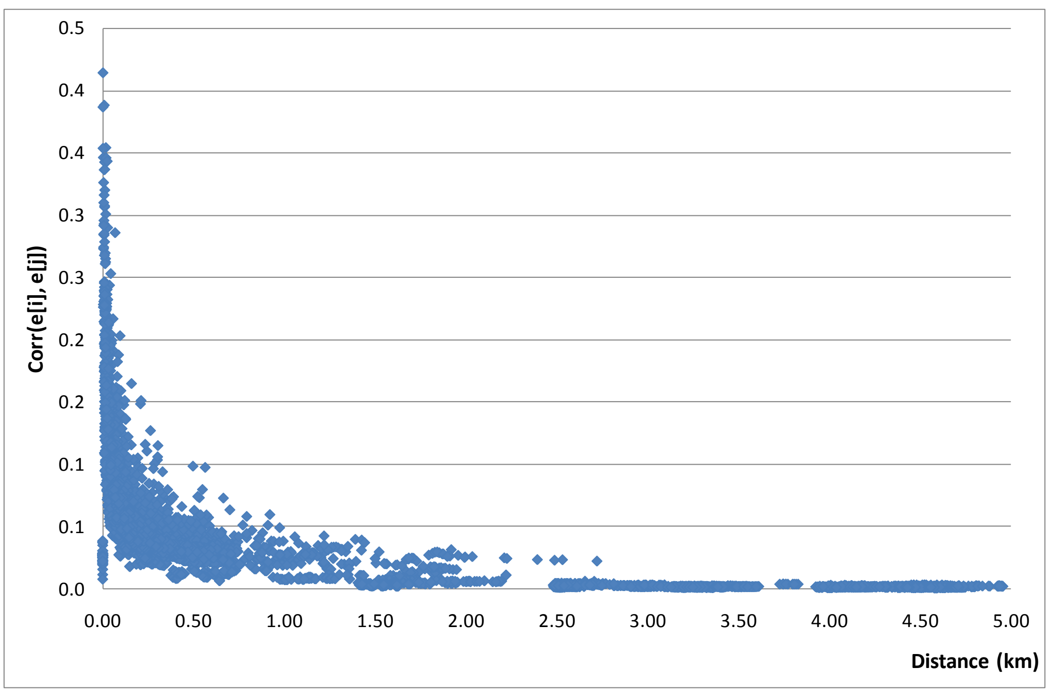

Table 1. First Stage Regression Model of Per Capita Expenditure (OLS)

	(1)	(2)	(3)	(4)
Household Level Characteristics				
# HH members age < 6	-0.278***	-0.287***	-0.290***	-0.289***
	(0.043)	(0.042)	(0.042)	(0.041)
# HH members age 7 - 15 years	-0.109***	-0.117***	-0.128***	-0.119***
	(0.024)	(0.024)	(0.024)	(0.024)
# HH members age 16 - 60 years	-0.082***	-0.089***	-0.100***	-0.086***
	(0.028)	(0.029)	(0.028)	(0.028)
# HH members age > 60 years	-0.180***	-0.176***	-0.208***	-0.200***
	(0.036)	(0.036)	(0.036)	(0.035)
# HH members completed primary school	-0.120***	-0.100***	-0.098***	-0.094***
	(0.035)	(0.035)	(0.034)	(0.034)
# HH members completed junior high school	-0.052	-0.042	-0.030	-0.044
	(0.032)	(0.033)	(0.032)	(0.032)
# HH members completed senior high school	0.033	0.027	0.054	0.028
	(0.045)	(0.047)	(0.045)	(0.046)
# HH members completed vocational degree	0.144	0.190	0.124	0.118
	(0.124)	(0.123)	(0.121)	(0.120)
# HH members with college degree and above	0.194	0.214	0.247	0.275
	(0.188)	(0.187)	(0.184)	(0.182)
# HH members engaged in non-agricultural activities	0.125***	0.127***	0.125***	0.118***
	(0.030)	(0.031)	(0.030)	(0.031)
Housing area (meter square)	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
HH uses LPG as main cooking fuel (dummy=1; 0 otherwise)	0.277	0.324	0.213	0.201
	(0.238)	(0.239)	(0.235)	(0.237)
House made of brick (dummy=1; 0 otherwise)	0.067	-0.003	0.086*	0.021
	(0.042)	(0.046)	(0.044)	(0.046)
Household has access to safe drinking water (dummy=1; 0 otherwise)	0.183***	0.159***	0.110**	0.087*
	(0.042)	(0.044)	(0.045)	(0.045)
Households with sanitary equipment (dummy=1; 0 otherwise)	0.128***	0.227***	0.228***	0.230***
	(0.048)	(0.053)	(0.054)	(0.062)

Census Means at Township Level			
# of kids in the household	0.265		-0.147
	(0.245)		(0.264)
# of youths in the household	-0.258*		-0.444***
	(0.131)		(0.136)
# of adults in the household	0.472***		0.026
	(0.154)		(0.166)
# of elderly in the household	0.214		0.416
	(0.255)		(0.263)
# HH members completed primary school	-0.113		-0.082
	(0.145)		(0.158)
# HH members completed junior high school	-0.197		0.094
	(0.129)		(0.152)
# HH members copmleted senior high school	0.008		-0.214
	(0.260)		(0.282)
# HH members completed vocational degree and above	0.193		-0.983
	(0.647)		(0.666)
# HH members with college degree and above	0.108		0.282
	(0.569)		(0.569)
Housing area (meter square)	-0.002**		-0.003***
	(0.001)		(0.001)
House made of brick (dummy=1; 0 otherwise)	0.224**		0.439***
	(0.093)		(0.118)
# labor force engaged in non-agriculture sector	0.077		0.183*
	(0.100)		(0.108)
Married household head (dummy = 1; 0 otherwise)	0.702		0.427
	(0.455)		(0.468)
3 generations living under the same roof (dummy = 1; 0 otherwise)	-4.943**		-3.795
	(2.319)		(2.473)
Environmental Variables			
Total area of land		0.056	0.163**
		(0.054)	(0.081)
Elevation (log)		0.078	0.140
		(0.106)	(0.113)
Density of highway (log)		0.030***	0.049***
		(0.007)	(0.008)
% loam in the soil		0.007	0.009

			(0.005)	(0.006)
Annual rainfall (log)			0.597***	0.952***
			(0.125)	(0.177)
Slope (log)			-0.142**	-0.115*
			(0.055)	(0.064)
% organic matter in soil texture			-0.433***	-0.685***
			(0.093)	(0.111)
Temperature			-0.069***	-0.070***
			(0.017)	(0.020)
% plain area			0.098	0.144
			(0.090)	(0.102)
Constant	6.747***	5.514***	2.918**	-0.786
	(0.092)	(0.434)	(1.270)	(1.485)
Log-likelihood function	-1060.12	-1034.16	-1025.88	-997.94
Number of observations	1,068	1,068	1,068	1,068
R2	0.196	0.234	0.246	0.285

Note: standard errors in ( ); *** significant at 1%; **significant at 5%;
*significant at 1%

Table 2. Selected Applications of Elbers, Lanjouw and Lanjouw's (2003) Method

Author(s)	Country Studies	Main Data Sources
Mistiaen, Özler, Razafimanantena and Razafindravonona (2002)	Madagascar	▪ 1993/1994 Household Survey ▪ 1993 Population Census
Alderman, Babita, Dembynes, Makhatha, and Özler (2003)	South Africa	▪ 1995 Household Survey and Expenditure Survey ▪ 1996 Population Census
Suryahadi, Widyanti, Perwira,Sumarto, Elbers and Pradhan (2003)	Indonesia	▪ 1999 Consumption Module and Core Socio-Economic Survey ▪ 2000 Population Census ▪ 1999 Village Census
Fujii (2004)	Cambodia	▪ 1997 Socioeconomic Survey ▪ 1998 Population Census
Benson, Chamberlin and Rhinehart (2005)	Malawi	▪ 1997/1998 Integrated Household Survey ▪ 1998 Population and Housing Census
Gibson, Datt, Allen, Hwang, Bourke, and Parajuli (2005)	Papua New Guinea	▪ 1996 Household Survey ▪ 2000 National Census ▪ PNG Resource Inventory System ▪ Mapping Agricultural System Project
Hoogeveen (2005)	Uganda	▪ 1992 Integrated Household Survey ▪ 1991 Population and Housing Census
Minot and Baulch (2005)	Vietnam	▪ 1998 Living Standards Survey ▪ 1999 Population and Housing Census
Simler and Nhate (2005)	Mozambique	▪ 1996/1997 National Household Survey on Living Conditions ▪ 1997 Population Census
Ahmad and Goh (2007)	China (Yunnan Province)	▪ 2000 Urban and Rural Household Surveys ▪ 2000 Population Census
Healy and Jitsuchon (2007)	Thailand	▪ 2000 Socio-Economic Survey ▪ 2000 Population and Housing Census
López-Calva, Rodríguez-Chamussy and Székely (2007)	Mexico	▪ 2000 Household Survey ▪ 2000 Population Census

Table 3. Specification Tests for Spatial Autocorrelation in the OLS residuals of the per Capita Household Consumption Model.

Type of weighting matrix	Moran's *I*	LM_λ	LM^*_λ	LM_ρ	LM^*_ρ
Inverse distance weights					
OLS w/ HH characteristics	11.11***	116.07***	10.06***	106.50***	0.49
OLS w/ HH characteristics & census means	9.69***	72.79***	3.30*	70.549***	1.06
OLS w/ HH characteristics & environmental variables	8.73***	62.51***	0.80	65.839***	4.131**
Binary weights					
OLS w/ HH characteristics	14.05***	182.24***	23.12***	159.46***	1.08
OLS w/ HH characteristics & census means	12.87***	121.38***	10.47***	111.414***	0.79
OLS w/ HH characteristics & environmental variables	11.41***	102.39***	3.57*	103.49***	4.67**

Note: ***=p<0.01, **=p<0.05, *=p<0.10.

Table 4. Spatial Error Estimates of the per capita Household Consumption Model

	(1)	(2)	(3)	(4)
Household Characteristics				
# HH members age < 6	-0.260***	-0.267***	-0.271***	-0.274***
	(0.040)	(0.040)	(0.040)	(0.040)
# HH members age 7 - 15 years	-0.113***	-0.116***	-0.122***	-0.117***
	(0.023)	(0.023)	(0.023)	(0.023)
# HH members age 16 - 60 years	-0.084***	-0.087***	-0.094***	-0.085***
	(0.027)	(0.027)	(0.027)	(0.027)
# HH members age > 60 years	-0.188***	-0.184***	-0.201***	-0.197***
	(0.034)	(0.034)	(0.034)	(0.034)
# HH members completed primary school	-0.108***	-0.099***	-0.098***	-0.096***
	(0.032)	(0.033)	(0.032)	(0.032)
# HH members completed junior high school	-0.043	-0.041	-0.033	-0.041
	(0.031)	(0.031)	(0.031)	(0.031)
# HH members completed senior high school	0.033	0.028	0.043	0.028
	(0.044)	(0.044)	(0.044)	(0.044)
# HH members completed vocational degree	0.181	0.196*	0.167	0.151
	(0.118)	(0.118)	(0.117)	(0.116)
# HH members with college degree and above	0.227	0.233	0.247	0.267
	(0.178)	(0.178)	(0.177)	(0.176)
# HH members engaged in non-agricultural activities	0.105***	0.109***	0.109***	0.106***
	(0.031)	(0.031)	(0.031)	(0.031)
Housing area (meter square)	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
HH uses LPG as main cooking fuel (dummy=1; 0 otherwise)	0.269	0.282	0.227	0.214
	(0.242)	(0.241)	(0.239)	(0.237)
House made of brick (dummy=1; 0 otherwise)	0.052	0.013	0.066	0.025
	(0.045)	(0.047)	(0.046)	(0.046)
Household has access to safe drinking water (dummy=1; 0 otherwise)	0.183***	0.161***	0.118**	0.092*
	(0.057)	(0.057)	(0.057)	(0.054)
Households with sanitary equipment (dummy=1; 0 otherwise)	0.123**	0.200***	0.203***	0.209***
	(0.060)	(0.063)	(0.063)	(0.068)

Census Means at Township Level			
# of kids in the household	0.181		-0.211
	(0.328)		(0.322)
# of youths in the household	-0.245		-0.434***
	(0.176)		(0.166)
# of adults in the household	0.450**		0.015
	(0.202)		(0.200)
# of elderly in the household	0.198		0.411
	(0.340)		(0.320)
# HH members completed primary school	-0.111		-0.086
	(0.191)		(0.191)
# HH members completed junior high school	-0.197		0.095
	(0.167)		(0.184)
# HH members copmleted senior high school	0.002		-0.212
	(0.345)		(0.341)
# HH members completed vocational degree and above	0.035		-1.079
	(0.861)		(0.806)
# HH members with college degree and above	0.098		0.263
	(0.761)		(0.695)
Housing area (meter square)	-0.002		-0.003**
	(0.001)		(0.001)
House made of brick (dummy=1; 0 otherwise)	0.214*		0.437***
	(0.123)		(0.144)
# labor force engaged in non-agriculture sector	0.101		0.200
	(0.132)		(0.130)
Married household head (dummy = 1; 0 otherwise)	0.616		0.397
	(0.607)		(0.569)
3 generations living under the same roof (dummy = 1; 0 otherwise)	-4.424		-3.646
	(3.090)		(3.017)
Environmental Variables			
Total area of land		0.056	0.170*
		(0.071)	(0.097)
Elevation (log)		0.075	0.149
		(0.140)	(0.138)
Density of highway (log)		0.029***	0.049***
		(0.009)	(0.010)
% loam in the soil		0.006	0.008
		(0.007)	(0.007)
Annual rainfall (log)		0.577***	0.955***

			(0.161)	(0.215)
Slope (log)			-0.139*	-0.113
			(0.073)	(0.077)
% organic matter in soil texture			-0.410***	-0.681***
			(0.121)	(0.135)
Temperature			-0.064***	-0.068***
			(0.022)	(0.025)
% plain area			0.095	0.149
			(0.118)	(0.124)
Constant	6.744***	5.621***	3.032*	-0.900
	(0.098)	(0.573)	(1.650)	(1.799)
λ	0.343***	0.293***	0.276***	0.213***
	(0.035)	(0.037)	(0.038)	(0.040)
Number of observations	1,068	1,068	1,068	1,068
Log-likelihood function	-1012.62	-1002.15	-997.97	-982.53

Note: standard errors in ( ); *** significant at 1%; **significant at 5%; *significant at 1%

Table 5. Spatial Lag Estimates of the per Capita Household Consumption Model

	(1)	(2)	(3)	(4)
Household Characteristics				
# HH members age < 6	-0.258***	-0.274***	-0.275***	-0.279***
	(0.040)	(0.040)	(0.040)	(0.040)
# HH members age 7 - 15 years	-0.105***	-0.115***	-0.124***	-0.118***
	(0.022)	(0.023)	(0.023)	(0.023)
# HH members age 16 - 60 years	-0.074***	-0.087***	-0.094***	-0.085***
	(0.026)	(0.027)	(0.026)	(0.027)
# HH members age > 60 years	-0.184***	-0.182***	-0.204***	-0.200***
	(0.034)	(0.034)	(0.034)	(0.034)
# HH members completed primary school	-0.111***	-0.097***	-0.094***	-0.093***
	(0.033)	(0.033)	(0.032)	(0.033)
# HH members completed junior high school	-0.051*	-0.038	-0.031	-0.040
	(0.030)	(0.031)	(0.030)	(0.031)
# HH members completed senior high school	0.028	0.035	0.051	0.034
	(0.043)	(0.045)	(0.043)	(0.044)
# HH members completed vocational degree	0.177	0.199*	0.149	0.137
	(0.117)	(0.117)	(0.116)	(0.116)
# HH members with college degree and above	0.248	0.250	0.271	0.287
	(0.178)	(0.178)	(0.176)	(0.176)
# HH members engaged in non-agricultural activities	0.099***	0.103***	0.103***	0.101***
	(0.028)	(0.029)	(0.029)	(0.030)
Housing area (meter square)	0.002***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
HH uses LPG as main cooking fuel (dummy=1; 0 otherwise)	0.233	0.288	0.205	0.201
	(0.225)	(0.227)	(0.225)	(0.228)
House made of brick (dummy=1; 0 otherwise)	0.050	0.006	0.063	0.021
	(0.040)	(0.044)	(0.042)	(0.044)
Household has access to safe drinking water (dummy=1; 0 otherwise)	0.107***	0.112***	0.075*	0.065
	(0.040)	(0.042)	(0.043)	(0.044)
Households with sanitary equipment (dummy=1; 0 otherwise)	0.091**	0.171***	0.183***	0.201***
	(0.046)	(0.051)	(0.052)	(0.060)

Census Means at Township Level			
# of kids in the household	0.124		-0.178
	(0.233)		(0.254)
# of youths in the household	-0.161		-0.344***
	(0.126)		(0.132)
# of adults in the household	0.362**		0.031
	(0.147)		(0.160)
# of elderly in the household	0.170		0.324
	(0.243)		(0.255)
# HH members completed primary school	-0.109		-0.071
	(0.138)		(0.153)
# HH members completed junior high school	-0.192		0.049
	(0.123)		(0.147)
# HH members completed senior high school	-0.048		-0.222
	(0.248)		(0.272)
# HH members completed vocational degree and above	-0.044		-0.879
	(0.617)		(0.642)
# HH members with college degree and above	-0.003		0.139
	(0.542)		(0.549)
Housing area (meter square)	-0.002**		-0.003**
	(0.001)		(0.001)
House made of brick (dummy=1; 0 otherwise)	0.167*		0.342***
	(0.089)		(0.115)
# labor force engaged in non-agriculture sector	0.069		0.161
	(0.095)		(0.104)
Married household head (dummy = 1; 0 otherwise)	0.333		0.240
	(0.436)		(0.453)
3 generations living under the same roof (dummy = 1; 0 otherwise)	-4.291*		-3.436
	(2.209)		(2.385)
Environmental Variables			
Total area of land		0.048	0.121
		(0.052)	(0.078)
Elevation (log)		0.055	0.114
		(0.101)	(0.109)
Density of highway (log)		0.020***	0.038***
		(0.007)	(0.008)
% loam in the soil		0.007	0.008
		(0.005)	(0.006)
Annual rainfall (log)		0.375***	0.735***
		(0.123)	(0.175)

Slope (log)			-0.082	-0.083
			(0.054)	(0.062)
% organic matter in soil texture			-0.339***	-0.563***
			(0.090)	(0.110)
Temperature			-0.048***	-0.056***
			(0.017)	(0.020)
% plain area			0.051	0.111
			(0.086)	(0.098)
Constant	4.656***	4.232***	2.400**	-0.287
	(0.236)	(0.444)	(1.217)	(1.435)
ρ	0.322***	0.281***	0.269***	0.209***
	(0.034)	(0.036)	(0.036)	(0.038)
Number of observations	1,068	1,068	1,068	1,068
Log-likelihood function	-1015.21	-1002.67	-996.56	-981.79

Note: standard errors in ( ); *** significant at 1%; **significant at 5%; *significant at 1%

Appendix Table 1. Spatial Error Estimates of the per capita Household Consumption Model (Based on Binary Weight Matrix)

	(1)	**(2)**	**(3)**	**(4)**
Household Characteristics				
# HH members age < 6	-0.262***	-0.268***	-0.271***	-0.273***
	(0.039)	(0.040)	(0.040)	(0.039)
# HH members age 7 - 15 years	-0.118***	-0.120***	-0.122***	-0.120***
	(0.023)	(0.023)	(0.023)	(0.023)
# HH members age 16 - 60 years	-0.093***	-0.095***	-0.094***	-0.092***
	(0.026)	(0.027)	(0.027)	(0.027)
# HH members age > 60 years	-0.198***	-0.195***	-0.201***	-0.203***
	(0.034)	(0.034)	(0.034)	(0.033)
# HH members completed primary school	-0.097***	-0.091***	-0.098***	-0.090***
	(0.032)	(0.032)	(0.032)	(0.032)
# HH members completed junior high school	-0.034	-0.032	-0.033	-0.035
	(0.031)	(0.031)	(0.031)	(0.031)
# HH members completed senior high school	0.037	0.034	0.043	0.032
	(0.043)	(0.044)	(0.044)	(0.044)
# HH members completed vocational degree	0.164	0.179	0.167	0.146
	(0.115)	(0.115)	(0.117)	(0.114)
# HH members with college degree and above	0.250	0.252	0.247	0.275
	(0.175)	(0.175)	(0.177)	(0.174)
# HH members engaged in non-agricultural activities	0.102***	0.104***	0.109***	0.104***
	(0.031)	(0.031)	(0.031)	(0.031)
Housing area (meter square)	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
HH uses LPG as main cooking fuel (dummy=1; 0 otherwise)	0.254	0.264	0.227	0.209
	(0.242)	(0.241)	(0.239)	(0.238)
House made of brick (dummy=1; 0 otherwise)	0.041	0.007	0.066	0.018
	(0.046)	(0.048)	(0.046)	(0.047)
Household has access to safe drinking water (dummy=1; 0 otherwise)	0.190***	0.167***	0.118**	0.098*
	(0.063)	(0.063)	(0.057)	(0.059)
Households with sanitary equipment (dummy=1; 0 otherwise)	0.125*	0.195***	0.203***	0.201***
	(0.065)	(0.067)	(0.063)	(0.072)

Census Means at Township Level			
# of kids in the household	0.257		-0.149
	(0.372)		(0.357)
# of youths in the household	-0.254		-0.447**
	(0.200)		(0.185)
# of adults in the household	0.473**		0.025
	(0.228)		(0.223)
# of elderly in the household	0.231		0.437
	(0.385)		(0.356)
# HH members completed primary school	-0.146		-0.115
	(0.216)		(0.212)
# HH members completed junior high school	-0.227		0.075
	(0.189)		(0.205)
# HH members copmleted senior high school	0.017		-0.194
	(0.391)		(0.379)
# HH members completed vocational degree and above	0.055		-1.069
	(0.976)		(0.896)
# HH members with college degree and above	0.042		0.204
	(0.865)		(0.774)
Housing area (meter square)	-0.002		-0.003**
	(0.002)		(0.002)
House made of brick (dummy=1; 0 otherwise)	0.229*		0.448***
	(0.139)		(0.160)
# labor force engaged in non-agriculture sector	0.101		0.199
	(0.149)		(0.145)
Married household head (dummy = 1; 0 otherwise)	0.626		0.372
	(0.689)		(0.633)
3 generations living under the same roof (dummy = 1; 0 otherwise)	-4.798		-3.828
	(3.506)		(3.361)
Environmental Variables			
Total area of land		0.056	0.181*
		(0.071)	(0.107)
Elevation (log)		0.075	0.151
		(0.140)	(0.154)
Density of highway (log)		0.029***	0.050***
		(0.009)	(0.011)
% loam in the soil		0.006	0.009
		(0.007)	(0.008)
Annual rainfall (log)		0.577***	0.944***

			(0.161)	(0.239)
Slope (log)			-0.139*	-0.119
			(0.073)	(0.086)
% organic matter in soil texture			-0.410***	-0.667***
			(0.121)	(0.149)
Temperature			-0.064***	-0.071***
			(0.022)	(0.027)
% plain area			0.095	0.159
			(0.118)	(0.138)
Constant	6.751***	5.612***	3.032*	-0.933
	(0.101)	(0.648)	(1.650)	(1.998)
λ	0.434***	0.386***	0.276***	0.300***
	(0.037)	(0.040)	(0.038)	(0.045)
Number of observations	1,068	1,068	1,068	1,068
Log-likelihood function	-1015.56	-1009.35	-1001.47	-988.24

Note: standard errors in ( ); *** significant at 1%; **significant at 5%;
*significant at 1%

Appendix Table 2. Spatial Lag Estimates of the per capita Household Consumption Model (Based on Binary Weight Matrix)

	(1)	**(2)**	**(3)**	**(4)**
Household Characteristics				
# HH members age < 6	-0.259***	-0.276***	-0.275***	-0.280***
	(0.040)	(0.040)	(0.040)	(0.040)
# HH members age 7 - 15 years	-0.104***	-0.115***	-0.124***	-0.119***
	(0.022)	(0.022)	(0.023)	(0.023)
# HH members age 16 - 60 years	-0.073***	-0.089***	-0.094***	-0.088***
	(0.026)	(0.027)	(0.026)	(0.027)
# HH members age > 60 years	-0.185***	-0.185***	-0.204***	-0.202***
	(0.034)	(0.034)	(0.034)	(0.034)
# HH members completed primary school	-0.107***	-0.094***	-0.094***	-0.091***
	(0.032)	(0.033)	(0.032)	(0.032)
# HH members completed junior high school	-0.052*	-0.036	-0.031	-0.039
	(0.030)	(0.031)	(0.030)	(0.031)
# HH members completed senior high school	0.024	0.036	0.051	0.034
	(0.042)	(0.044)	(0.043)	(0.044)
# HH members completed vocational degree	0.176	0.196*	0.149	0.137
	(0.116)	(0.116)	(0.116)	(0.115)
# HH members with college degree and above	0.256	0.253	0.271	0.287
	(0.176)	(0.176)	(0.176)	(0.174)
# HH members engaged in non-agricultural activities	0.096***	0.100***	0.103***	0.100***
	(0.028)	(0.029)	(0.029)	(0.030)
Housing area (meter square)	0.002***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
HH uses LPG as main cooking fuel (dummy=1; 0 otherwise)	0.208	0.268	0.205	0.195
	(0.223)	(0.225)	(0.225)	(0.227)
House made of brick (dummy=1; 0 otherwise)	0.049	0.009	0.063	0.021
	(0.039)	(0.044)	(0.042)	(0.044)
Household has access to safe drinking water (dummy=1; 0 otherwise)	0.091**	0.100**	0.075*	0.059
	(0.040)	(0.042)	(0.043)	(0.043)
Households with sanitary equipment (dummy=1; 0 otherwise)	0.089**	0.162***	0.183***	0.192***
	(0.045)	(0.050)	(0.052)	(0.060)

Census Means at Township Level			
# of kids in the household	0.172		-0.101
	(0.231)		(0.253)
# of youths in the household	-0.136		-0.316**
	(0.124)		(0.132)
# of adults in the household	0.344**		0.038
	(0.146)		(0.159)
# of elderly in the household	0.154		0.292
	(0.240)		(0.253)
# HH members completed primary school	-0.123		-0.081
	(0.136)		(0.152)
# HH members completed junior high school	-0.207*		0.022
	(0.122)		(0.146)
# HH members copmleted senior high school	-0.059		-0.213
	(0.245)		(0.270)
# HH members completed vocational degree and above	-0.037		-0.774
	(0.610)		(0.639)
# HH members with college degree and above	-0.035		0.079
	(0.536)		(0.546)
Housing area (meter square)	-0.002*		-0.002**
	(0.001)		(0.001)
House made of brick (dummy=1; 0 otherwise)	0.158*		0.310***
	(0.088)		(0.115)
# labor force engaged in non-agriculture sector	0.053		0.138
	(0.094)		(0.103)
Married household head (dummy = 1; 0 otherwise)	0.269		0.182
	(0.431)		(0.450)
3 generations living under the same roof (dummy = 1; 0 otherwise)	-4.311**		-3.358
	(2.185)		(2.369)
Environmental Variables			
Total area of land		0.048	0.111
		(0.052)	(0.078)
Elevation (log)		0.055	0.098
		(0.101)	(0.109)
Density of highway (log)		0.020***	0.034***
		(0.007)	(0.008)
% loam in the soil		0.007	0.008
		(0.005)	(0.006)
Annual rainfall (log)		0.375***	0.645***

			(0.123)	(0.176)
Slope (log)			-0.082	-0.076
			(0.054)	(0.061)
% organic matter in soil texture			-0.339***	-0.509***
			(0.090)	(0.110)
Temperature			-0.048***	-0.053***
			(0.017)	(0.020)
% plain area			0.051	0.099
			(0.086)	(0.098)
Constant	4.122***	3.780***	2.400**	-0.073
	(0.252)	(0.449)	(1.217)	(1.426)
ρ	0.403***	0.365***	0.269***	0.286***
	(0.036)	(0.039)	(0.036)	(0.043)
Number of observations	1,068	1,068	1,068	1,068
Log-likelihood function	-1019.59	-1009.06	-999.84	-987.36

Note: standard errors in ( ); *** significant at 1%; **significant at 5%;
*significant at 1%