Regional Income Inequality and Economic Growth: A Spatial Econometrics Analysis for Provinces in the Philippines

Valerien O. Pede*, Adam H. Sparks, and Justin D. McKinley

Social Sciences Division
International Rice Research Institute
DAPO Box 7777, Metro Manila, Philippines

(筹备版本。请不要引用)

Abstract

This paper revisits the inequality-growth relationship using data at the sub-national (provincial) level in the Philippines over the period 1991-2000. A conditional convergence growth model is considered where the growth of per capita income depends on inequality and other growth factors. The contribution of each province to the overall inequality obtained from the Theil index is considered. Results indicate that inequality has a positive and significant effect on per capita income growth. However, the magnitude of the inequality effect is not stable across regions. Geographically Weighted Regression estimates show that the magnitude of the inequality growth relationship varies over a range of 0.72 to 3.36. Other results are also noteworthy in this study. Per capita income grows faster in provinces that contribute more to the overall inequality. Provinces with higher poverty incidence tend to grow less and human capital appears to be a significant booster to per capita income growth. Additionally, urban provinces tend to grow faster than the rural ones.

JEL codes: R11, R12, O15, C21

Key words: clusters, growth, inequality, spatial econometrics

Contributed paper prepared for presentation at the 56th AARES annual conference, Fremantle, Western Australia, February 7-10, 2012.

Copyright 2012 by Authors names. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Introduction

Dealing with income inequality among regions and individuals has remained an enduring development challenge. In many countries of the world it is common to observe that the major portion of wealth is concentrated in the hands of a minority, while the vast majority of the population is living in poverty. These facts raise several questions. Does income inequality hurt a nation’s economic growth? Is increased income inequality good or bad for an economy? Is income inequality something to be encouraged, or not?

Southeast Asia is no exception to this observation. For instance, since 1985, the Philippines’ richest quintile of the population has consistently commanded more than 50% of total family income while the poorest quintile at less than 5% (ADB, 2009). Evidence of inequality can be traced back even further, Africa (2011) reports that since 1961, the top 50% of families in the Philippines have represented approximately 80% of the share of income; leaving just 20% of income for the bottom 50% of families.

While Singapore, Hong Kong, Taiwan, and South Korea have experienced high economic growth and rapid industrialization to become “tiger economies,” the Philippines has been referred to as “East Asia’s stray cat,” because of its failure to grow like its neighbors (Vos and Yap, 1996). This is a particularly interesting position for the Philippines to be in considering its history. Up until the early 1970’s, the Philippines had the second highest per capita earnings in Asia, second only to Japan (Galang, 1996). Since then, the economic conditions have changed significantly and there are now mixed messages being expressed in Filipino newspapers. The Manila Standard Today positively reports that the per capita earnings in the Philippines have
surpassed 2,000 USD/year in 2011. This is a milestone that has been perceived as crucial for neighboring Thailand, who now possesses per capita earnings that are three times greater than that of the Philippines. This is all in spite of the fact that Thailand’s per capita earning were less than the Philippines up until the 1970s (Dela Cruz, 2011). These gains in per capita earning are not being distributed evenly through the population. The Philippines currently has the highest income inequality in Southeast Asia with a Gini coefficient of 44% according to the Philippine Star (Xinhua, 2011).

In the pursuit of how to deal with income inequality a number of studies have investigated how it relates to economic indicators such as per capita income or economic growth. The debate started with a path-breaking publication from Kuznets (1955), who found that there is an inverted U-shape relationship between inequality and per capita income. Following Kuznets, numerous studies have investigated the relationship between income inequality and economic growth. However, conflicting findings have always emerged from these studies. A negative relationship is claimed in some studies (Alesina and Rodrik, 1994; Persson and Tabellini, 1994; Clarke, 1995; Deininger and Squire, 1998), while other studies find that inequality is positively related to economic growth (Li and Zou, 1998; Forbes, 2000; Bell and Freeman, 2001; Siebert, 1998).

Theories supporting a positive link between income inequality and growth are summarized in Aghion and Howitt (1998). Aghion first states, given that the rich have a higher marginal propensity to save than poor, more unequal economies will tend to grow faster than economies characterized by a more equitable income distribution. Second, due to large sunk costs required for setting up new industries or implementing new ideas, it is more efficient that wealth be concentrated in the hands of few people (individuals or a family for example). Third, providing
incentives to workers will reduce differences in income and favors redistribution, but doing so lowers the rate of growth because of the trade-off between equity and efficiency.

Perotti (1996) also summarizes the theories supporting a positive relationship between income inequality and growth in four approaches: the endogenous fiscal policy approach; the socio-political instability approach; the borrowing and investment in education approach; and the joint education/fertility approach. Aghion and Howitt (1998) also enumerate three main reasons why inequality may have a direct negative effect on economic growth. First, they argue that redistribution enhances investment opportunities in the absence of well-functioning capital markets, and helps to raise aggregate productivity and growth. Indeed, the poor have a relatively higher marginal productivity of investment compared to the rich. Therefore, when income redistribution happens, income differences are narrowed and this will enhance productivity and promote growth. Second, inequality worsens borrower’s incentives to invest in productive activities. Wealth redistribution increases the ability of individuals to invest and thereby promotes growth whenever the positive incentive effect outbalances the potentially negative incentive effect on lender’s effort. Their third reason is linked to the macroeconomic volatility effect that inequality may provoke. Individuals have different attitudes toward risk, and they also have different access to investment opportunities. Consequently, this creates separation between investors and savers that will give rise to volatility in term of investment rate and interest rate.

Differences in the method used to measure inequality or in the econometrics estimation method can result in large differences in the estimated inequality growth link. For instance, Panizza (2002) shows that the relationship between inequality and growth is not robust, demonstrating that no relationship was detected on US states data when using fixed effects and Generalized Method of Moment (GMM). Partridge (2005) relates the mixed findings to differing short- and
long-term responses. Using U.S. state level data, and accounting for short- and long-term responses, he observes that inequality is positively related to growth, but short run income distribution response is unclear. Mixed results are also obtained when differentiation is made between types of regions. For instance, Fallah and Partridge (2007) re-examined the inequality-growth relationship and observed opposite signs for urban and rural samples. In order to shed some light on the ambiguity related to the correlation between inequality and economic growth, De Dominicis et al. (2008) use meta-analysis techniques, their conclusions point to the dependence of the correlation on estimation methods, data quality and sample coverage. They observed that the use of a fixed effects model and regional dummies tends to indicate a positive relationship between growth and inequality on pooled data. Also, the negative effect of inequality on economic growth tends to be more accentuated in developing countries than in developed countries. The measures of inequality, the length of growth period, and data quality also tend to have important implication on the form of the relationship between growth and inequality.

Income distribution and inequality in the Philippines has been a popular topic for researchers for quite some time (see Paukert et al., 1981; Blejer and Guerrero, 1990; Estudillo, 1997; Rodriguez, 1998; and Hossain et al., 2000). Additionally, the impact of income inequality on economic growth has also been a popular topic more recently (see Balisacan and Fuwa, 2003; Balisacan and Fuwa, 2004b, Felipe and Sipin, 2004). However, in past literature, the role of space has been largely disregarded. Balisacan and Fuwa (2004a) discuss changes in spatial income inequality, but fail to use spatial econometric techniques to discuss the role of income inequality on economic growth. The spillover of economic activities across regions creates a spatial interdependence. We hypothesize this spatial dependence between regions could influence the
inequality-growth link. For instance, knowledge spillovers across regions could induce convergence towards equality.

This paper revisits the inequality-growth relationship in the context of the Philippines using data at the provincial level over the period 1991 to 2000. The goal is to investigate how income inequality in the Philippines affects economic growth. A conditional convergence growth model is considered. Spatial econometrics techniques are used to estimate the inequality-growth model in order to account for technological dependence and spillover effects from neighbors that might affect the growth process.

The rest of the paper is organized as follows. The next section presents the econometric methodology and estimation procedure. The following section describes the data used in the econometric estimations. Then, results are presented and discussed. The last section concludes the paper.

**Econometric methodology**

We first consider a conditional growth model in linear form given as:

\[ y = X\beta + \epsilon, \]  

where \( y \) is an \( N \times 1 \) spatial data series representing the growth of per capita income over the period 1991 - 2000; \( X \) is an \( Nxk \) matrix of explanatory variables, \( \epsilon \) is a vector of innovations and \( N \) represent the number of observations or spatial. The linear model as specified in [1] may be estimated using Ordinary Least Square (OLS). However, when spatial dependence is present,
OLS estimation of [1] will yield biased or inefficient coefficients, whether the spatial dependence operates in the dependent variable or in the disturbances.¹

A general form illustrating the consideration of spatial dependence in equation [1] could be illustrated by a spatial autoregressive model given as:

\[ y = \rho Wy + X\beta + \varepsilon \]

\[ \varepsilon = \lambda W\varepsilon + \mu , \]

where \( \rho \) and \( \lambda \) are scalar lag and error and spatial autoregressive parameters, \( W \) is an exogenously determined weight matrix that illustrate the spatial structure of units, is a vector of independently and identically distributed disturbances. All other symbols are defined as before. Depending on the values taken by the spatial parameter \( \rho \) and \( \lambda \), two nested models could be obtained from [1]. A spatial lag model is obtained when the parameter \( \lambda \) is equal to zero. However, a spatial error model is obtained when the parameter \( \rho \) is equal to zero.

The model in equation [2], commonly called SARAR, may be estimated using Maximum likelihood (MLE) or General Method of Moments (GMM) (see Kelejian and Prucha, 1998). Given the presence of spatial autoregressive component in the model of equation [2], a correct interpretation of the estimated coefficients involve computing the measure of direct, indirect and total effects. These computations are extensively explained in LeSage and Pace (2009). The direct effect characterizes the average impact of a change in the explanatory variable in each of the spatial units on the dependent variable at the same location. The indirect effect characterizes

¹ OLS estimation can still be valid when the spatial dependence is modeled in the independent variables X. this is referred as cross-regressive model.
the average impact of a change in the explanatory variable in each location on the dependent variable in different locations. The total effect represents the sum of direct and indirect effects.

The reduced form of the equation in [2] is given as:

\[ y = (I - \rho W)^{-1}X\beta + (I - \lambda W)^{-1}\mu \]  

[3]

where \( I \) represent an \( N \times N \) identity matrix. The marginal effect of a change in an explanatory variable \( X_i \) is given as:

\[ \frac{\partial y}{\partial X_i} = (I - \lambda W)^{-1}\beta_i, \]  

[4]

where \( \beta_i \) represents the coefficient associated to the variable \( X_i \).

The models in equation [1] and [2] all assume that the estimated coefficients are global, but it may well be that the estimated relationships are not stable and vary across space. Many of the previous studies on the inequality-growth link had made similar assumptions. However, it has well been demonstrated that the relationship between growth and inequality may not be stable.

For robustness check, we consider an alternative estimation procedure that allows parameter variation across space: the so-called Geographically Weighted Regression (GWR). Brundson et al. originally introduced the GWR technique. Commonly, regressions coefficients are assumed to be global or fixed across all spatial units. But this may not always be the case, as some economic phenomena may induce variation across location in terms of impacts or effects being investigated. For instance, parameters describing the same relationship may show different magnitudes or signs across the spatial units. Reasons for why one might expect spatial non-
stationary in some relationships/effects are discussed in Fotheringham et al. (2002). The GWR model, as described in Fotheringham et al. (2002), is expressed as:

\[ y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{K} \beta_k(u_i, v_i) x_{ik} + \epsilon_i, \]  

where \((u_i, v_i)\) represents the coordinates of the \(i\)th point in space, and \(\beta_k(u_i, v_i)\) is a realization of the continuous function \(\beta_k(u,v)\) at point \(i\), and is an error term. The weighted least square estimates are given as:

\[ \hat{\beta}(u_i, v_i) = (X'W(u_i, v_i)X)^{-1} X'W(u_i, v_i)y, \]

where all symbols are defined as before except that although the notation \(W\) is similar to the weights matrix in spatial process models (defined in equation [2]), the weight matrix in GWR has zeros everywhere except for some of the diagonal elements, whereas the traditional weights matrix has zeros on the diagonal and non-zeros in some of the off-diagonal positions.

The estimation procedure starts with the specification of the weighting function and then the choice of the circle of influence or “bandwidth.” Two types of weighting functions are commonly used: the Gaussian distance-decay weighting function and the bi-square function. The Gaussian distance-decay weighting function is given as:

\[ w_{ij} = \exp \left[ -\frac{1}{2} \left( \frac{d_{ij}}{b} \right)^2 \right], \]  

[7]
where $i$ is the data point for which the parameters are being estimated, $j$ represents any other point in space, $d_{ij}$ is the distance between $i$ and $j$, and $b$ is the bandwidth over which the spatial interaction extends. The bi-square function is specified as:

$$
 w_{ij} = \begin{cases} 
 1 - \left( \frac{d_{ij}}{b} \right)^2 & \text{if } d_{ij} < b \\
 0 & \text{otherwise.}
\end{cases} \quad [8]
$$

The choice of the bandwidth constitutes an important step in the estimation procedure. The bandwidth may be selected by using the least squares criterion, which boils down to minimizing the sum of squared errors given as:

$$
 SS = \sum (y_i - \hat{y}_i(h))^2, \quad [9]
$$

where $y_i$ represent the value of the observation at point $i$ and $\hat{y}_i(h)$ the predicted value from the GWR evaluated at the bandwidth $h$. Obviously, the drawback of this optimization is that when the bandwidth tends to zero, the predicted values are close to the actual observation. Therefore, the sum of square errors tends to a minimum value of zero. This optimization will therefore suggest $h = 0$ as optimal solution or result in computational errors. This problem can be avoided by omitting the i-th observation when computing the GWR estimate of $\hat{y}_i$, and subsequently minimize the resulting adjusted sum of squared errors. Alternative methods for selecting the bandwidth are based on the Akaike Information Criterion and the Schwartz Information Criterion, or on the use of cross-validation techniques (Fotheringham et al., 2002).
Data and estimation procedures

In this paper, we consider economic growth data over the period 1991-2000 on 80 provinces in the Philippines. Per capita income and human capital data are obtained from the National Statistics Office (NSO). Using regional consumer price indexes (CPI), the per capita income were all converted into the year 2000 Philippine Peso. Human capital (education) variable is defined as the proportion of population with post-secondary (undergraduate and graduate) and college degree and higher. Data on poverty incidence at the provincial level are obtained for the year 1997. The contribution of each province to the overall inequality is computed using the Theil inequality formula. It is expressed as follow:

\[ T = \sum_{i=1}^{N} Y_i \log NY_i, \]  

[10]

where \( Y_i \) represents the share of income of region i relative to the nation, \( N \) is the number of provinces. The term in the summation represents the contribution of province i to the overall inequality.

We consider a conditional growth model, where the annual growth of per capita income over the period 1991-2000 depends on contribution to inequality of each province and a number of conditioning variables. The growth equation is given as:

\[ \log \left( \frac{y_t}{y_{t+k}} \right) = \beta_0 + \beta_1 \log(y_t) + \beta_2 Ineq + \beta_3 Pov + \beta_4 Edu + \beta_5 Urb + \varepsilon, \]  

[11]

where \( y_t \) and \( y_{t+k} \) represent the per capita income at the initial period (1991) and ending period (2000), respectively. \( Ineq \) represents the contribution of each region to inequality, \( Pov \)
represents the poverty incidence of each province, \( Edu \) is a variable capturing the human capital available in the province and \( Urb \) is a dummy taking value 1 when the province is urban and 0 when it is rural.

The estimation procedure starts with an OLS estimation of the model in equation [11]. Spatial diagnostics statistics are used to determine the appropriate spatial specification. To this end, we consider a battery of Lagrange Multiplier tests (see Anselin, 1988; Anselin et al., 1996). For the estimation of the spatial regressions, a distance-based weight matrix is considered. The spatial weight matrix is constructed using the arc distances between the geographical midpoints (centroids) of the 80 provinces. It is a binary weight matrix with elements \( w_{ij} \) taking value 1 when the distance between the midpoint of the provinces \( d_{ij} \) is less than the threshold distance \( T = 126 \) miles, and 0 when the distance is larger than \( T \).\(^2\) The matrix is row-standardized, enforcing row sums to be equal to one. The spatial weight matrix has dimension 80 x 80, with 26.90% of the weights being nonzero. The minimum and maximum number of links between provinces are 1 and 31, respectively, with an average number of links of 21.

**Results**

**Exploratory Data Analysis**

Before presenting the results of econometric estimation, we first provide an insight into the spatial distribution of the main variables. Figure 1 shows the spatial distribution of economic growth over the period 1991-2000, the regional contribution to inequality, poverty incidence and educational attainment. Fast growing provinces are distributed throughout the country in all three

\(^2\) The distance of 126 miles is the minimum cut-off distance needed to ensure that each county has at least one neighbor.
major regions, Luzon, the Visayas and Mindanao. However, most provinces contributing the most to inequality are found on Luzon and cluster around the Metro Manila area and Laguna de Bay with most of the positive values from 0.01 to 0.122 in this locale. Poverty incidence appears to be concentrated on Mindanao, in the south, with some occurrences of high poverty incidence in the Visayas and North-Central Luzon. Human capital follows a pattern that is similar to the contribution to inequality with the most near Metro-Manila and central and northern Luzon, though northern Mindanao shows a high concentration of education level as well.

**Econometric Results**

The estimated models are presented in Table 1. The estimation procedure begins with unconditional growth model where the initial per capita income is the only right-hand side variable (column 1). Subsequently, the inequality variable is entered in the model as well as the other conditioning variables (column 2 and 3). In these estimations, Ordinary Least Square (OLS) is first used. Maximum Likelihood Estimation (MLE) is used for the estimation of spatial process models (column 4 and 5). In the comprehensive model (column 4), the spatial diagnostic tests recommend a spatial process with spatial lag and error parameter (SARAR). Meanwhile, after estimating the SARAR, the spatial lag parameter was insignificant, while the spatial error parameter is. We therefore re-estimated the model and consider only a spatial error process (column 5). The estimated coefficients remain consistent across all models, but their magnitudes vary. For instance, in all estimations the initial per capita income is negatively related to the per capita growth of the period. This denotes the occurrence of beta-convergence. Consequently, poor economies tend to grow faster than rich ones. The annual rate of convergence increases as more conditioning variables are added to the model. The comprehensive model, which is label 5,
has an annual convergence rate of 8.5%.\(^3\) Inequality has a positive and significant effect on per capita income growth. Provinces that contribute more to the national inequality tend to grow faster. However, the poverty incidence has a negative and significant effect on per capita growth. Provinces with high poverty incidence tend to grow less. As expected, human capital has a positive and significant effect on growth. Provinces with high level of human capital tend to have high per capita growth. It is interesting to notice that the magnitude of the effect of poverty incidence as well as human capital remains consistent across models. Finally, as expected, urban provinces tend to grow faster than rural provinces.

In all the estimated models in Tables 1 parameters are considered global, however, it may well be that the estimated relationships are not stable. For robustness check, the growth model in equation [11] is therefore re-estimated with a model that allows parameter variation across space, the geographically Weighted Regression (GWR). The GWR model was estimated for the two weighting functions described previously: the Gaussian and bisquare functions. Given that the estimated parameters are very similar for the two weighting functions, we only presented results for the bisquare function. Figure 2 shows the spatial distribution of the inequality parameters. The map clearly confirms the spatial variation in the inequality-growth relationship. The magnitude of the inequality parameter varies from 0.72 to 3.36. Provinces with larger inequality-growth link concentrated primarily in southern areas of the country. Areas of Mindanao show a high response of growth to inequality, ranging from 1.38 to 3.36 in magnitude. The rest of the Philippines fall under 1.37.

\(^3\) The annual rate of convergence is defined as \(-(1/T)*\log(1+b)\), where b represent the coefficient of the initial per capita income, T the number of years between the growth period under study. The standard error of the convergence rate is approximated as: \((1/T)*SE(b)/(1+b)\), SE(b) represents the standard error of the parameter b.
Conclusion

The primary purpose of this study is to revisit the inequality-growth relationship data at the provincial level in the Philippines over the period of 1991-2000. Based on the findings, income inequality has a positive effect on per capita income growth over the period considered. Provinces with higher computed contribution to inequality were found to have faster growth rate of per capita income. However, the paper shows that the inequality-growth is not spatially stable. Using the Geographically Weighted Regression (GWR), we observe a large variability on the magnitude of the inequality-growth relationship. The poverty incidence has a negative effect on economic growth, and human capital appears to be a significant booster to economic growth. The model suggests that urban provinces are more likely to grow faster than the rural ones.
References


### Table 1. Econometric Estimation of the Inequality-Growth Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unconditional models</th>
<th>Conditional models</th>
<th>Diagnostic tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>OLS (2)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.49**</td>
<td>0.74***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.25)</td>
<td></td>
</tr>
<tr>
<td>initial income 1991</td>
<td>–0.09*</td>
<td>–0.15***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
<td></td>
</tr>
<tr>
<td>contribution to inequality</td>
<td>1.29**</td>
<td>0.87*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.63)</td>
<td>(0.51)</td>
<td></td>
</tr>
<tr>
<td>poverty incidence</td>
<td>–0.002***</td>
<td>–0.002***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>education</td>
<td>0.01***</td>
<td>0.01***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>urban/rural dummy</td>
<td>0.02</td>
<td>0.04*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
</tr>
<tr>
<td>Spatial AR parameter</td>
<td>–0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Error parameter</td>
<td>0.80***</td>
<td>0.70***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td>Diagnostic tests</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moran's I (error)</td>
<td>0.18***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM-error</td>
<td>21.36***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust LM-error</td>
<td>26.37***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM-lag</td>
<td>4.06**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robust LM-lag</td>
<td>9.07***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM-SARMA</td>
<td>30.43***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convergence rate (%)</td>
<td>0.41</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.007)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors of parameters estimates are in parentheses. Significance at the 1, 5 and 10% level is signaled by ***, ** and *, respectively.
Figure 1: Spatial distribution of economic growth, inequality, poverty incidence and educational attainment. Gray indicates provinces with missing data or Laguna de Bay.
Figure 2: Geographically Weighted Regression (GWR): Inequality-Growth Parameters. Gray indicates provinces with missing data or Laguna de Bay.