BEYOND UNIVARIATE MEASUREMENT OF SPATIAL AUTOCORRELATION:
DISAGGREGATED SPILLOVER EFFECTS FOR INDONESIA

Jennifer Day
Lecturer in Urban Planning
Faculty of Architecture, Building, and Planning
The University of Melbourne
Parkville, VIC 3010
Australia
+61 451 054 878
jday@unimelb.edu.au

Blane Lewis
The World Bank
Poverty Reduction and Economic Management Unit
The World Bank Office, Jakarta
Indonesia Stock Exchange, Tower 2, 12th floor
Jl. Jenderal Sudirman Kav 52-53, Jakarta, 12190 – Indonesia
blewis@worldbank.org

Keywords
Indonesia
Spatial statistics
Spillovers
Economic growth
Durbin representation

Introduction
The purpose of this paper is twofold. First, we wish to demonstrate empirically that accounting for spatial effects in analyses of economic growth and productivity can improve model estimations, and provide deeper insight into the determinants of growth in Indonesian districts. Second, we hope to test the hypothesis that the differential effects of various important predictor variables – and specifically their spatial effects – are discernible for Indonesia.

A rich theoretical and empirical literature exists that links economic growth to initial levels of income, private capital and investment, human capital, public infrastructure, institutional capacity, and many other variables. The literature includes both cross-country and individual country sub-national studies. More recently, economists have begun to incorporate spatial effects into models of economic growth, both at national and sub-national levels. In this paper, we innovate by providing tests of spillover effects in the sub-national context for Indonesia.
Spillover effects are externalities resulting from spatial proximity; that is, benefits or costs that accrue to a place, firm, or other entity, as a result of the condition of that place, firm, or other entity’s neighbors. This study focuses on place-based spillovers, i.e., those that accrue to spatial areas such as cities, counties, and administrative districts.

Spillovers can be positive, if benefits accrue to a region, or negative, if a place suffers costs related to its neighbors’ economic growth, changes in public infrastructure, demographics, or other characteristics. Easy intuitive connections exist that allow us to imagine both scenarios. For example, if a region with a largely service-based economy is located near other regions with high-polluting industries, then that region could face negative spillover effects in the form of pollution. If the spatial units are fine enough so that regions share labor, that same place could in turn benefit from the demographic makeup of its neighbors – for instance, if its neighboring districts had relatively-high proportions of working-age or educated labor, which it could access for its own growth.

In area-based growth models where geographic spillovers are not accounted for, the models are generally mis-specified due to spatial autocorrelation in the error terms. This spatial dependence between regions leads to biased estimation results in OLS. Baumont et al (2001) show that improved estimations can be achieved when spatial dependence between observations is accounted for in growth models.

Our findings support both of the above hypotheses. Spillover effects turn out to be significant; in some cases they appear to be more important than own-district effects. Further, the inclusion of spillover effects improves model fit. Also, the disaggregated effects of different types of spillovers (Gross Regional Domestic Product [i.e., GRDP] levels and growth, human capital, capital, infrastructure, and others) are detectable for Indonesia, and our method of including spillovers allows for some of these effects to be observed. Such effects are typically masked by non-spatial models. Contradictory findings between our model and more-traditional spillovers measures, i.e., the Moran’s I, also point to a need for more exploration with models of this type, which account for disaggregated spillover effects at the sub-national level. We discuss these contradictory findings in Section 4.

**Background**

Virtually any characteristic of one place can spill over and affect other places – for example, war and conflict (Murdoch 2002; Vothknecht and Sumarto 2011), technology (Paci and Pigliaru 2001), and innovation (Bottazzi and Peri 2003; Ciccone 1996), among others. For our study, the two most relevant areas of related work are in the areas of economic productivity and growth, and infrastructure. The purpose of this section is to review the empirical literature on these two types of spatial spillovers. We largely ignore other area-based spillover effects here, for the sake of brevity.1 We also disregard studies of non-area-based spillovers, such as the innovation, knowledge, and growth flows from

---

1 See Ye and Wei (2012) for an introduction to and review of other spatial and spillover effects not covered here including those related to foreign direct investment, trade, and innovation, among others.
foreign to local companies in developing countries (Hermes and Lensink 2003; Lipsey and Sjöholm 2004; Sjöholm 1998; Todo and Miyamoto 2006; Takii 2005). These are important types of spillovers, but they are only tangentially relevant to this study. Section 3 below reviews the literature specifically related to our model selection.

Recent papers in spatial econometrics have significantly advanced the state of the practice in this branch of econometrics. However, the methods are complicated, the software routines are still being developed, and thus, the empirical literature applying these techniques to regional science is still emerging. Recent work (Lee and Yu 2010) shows that using fixed effects models (as some authors noted below do) to estimate spatial panel models with spatial fixed effects will result in an inconsistent variance parameter ($\sigma^2$) estimate if the sample size (N) is large and the number of panel waves (T) is small, and will also yield inconsistent estimates of all parameters of a model with spatial and time-period fixed effects if both N and T are large. They propose a bias correction procedures for these scenarios, which we use to produce our fixed effects models here.

Even when biases are corrected for, though, the model coefficients cannot be directly used to interpret spatial effects. Elhorst (2011) and LeSage and Pace (2009) point out that using the point coefficient estimates from spatial models to test for spillover effects can lead to erroneous conclusions. LeSage and Pace present a partial derivative interpretation as a more valid basis for testing the hypotheses that spillover effects are present. A number of Stata routines have been developed that use these new techniques. In this paper, we use some of these routines to estimate our models.

Before these new advances in computational methods, area-based economic spillovers had been examined at the country, region, and sub-national levels. The remainder of this section examines the existing empirical literature on growth spillovers, and situates our work within that literature. Barro and Sala I Martin (1994) measure country-level spatial spillover effects by incorporating a weighted average of neighboring countries’ GDPs per capita into their growth models, and find a significant and positive, but small, effect. Ramon and Trehan (1997) demonstrate that a country’s growth rate is positively influenced by growth rates in neighboring countries. More recently, Ramirez and Loboguerrero (2002) examine 98 economies between 1965 and 1995, testing the hypothesis that interaction among countries is a significant factor in growth. They estimate a growth model in which a country’s economic growth is a function of its neighbors’ growth, but fail to find effects when they control for regional (continent) effects. Numerous studies have studied spatial effects at the national and regional levels in the European Union, and found positive results (Ertur, Gallo, and Baumont 2006; Fingleton 1999; Esther et al. 2004; López-Bazo et al. 1999; Baumont, Ertur, and Gallo 2001).

Sub-national studies provide a finer look at the differences between areas in the same country. Rey and Montouri (1999) examine states (from the United States) in a panel dataset spanning the seven decades between 1924 and 1994 using an unconditional convergence model (i.e. in which growth is not dependent on initial income levels). They find strong evidence of spatial

---

2 These include: spregfext to estimate fixed effects models, spglsxt to estimate models using Generalized Least squares estimators, and spgmmxt to estimate models using Generalized Method of Moments estimators.
autocorrelation, and further conclude that the level of spatial effect is correlated with regional incomes in the country. Niebuhr (2001), Arbia (2003), and Bernat (1996) focus on state-based spillovers in West German, Italy, and the United States, in panel studies. However, all of these studies focus only on the impact of economic output of neighbors, and neglect other important spillover effects such as public capital, human capital, and infrastructure. In this paper, we consider these effects.

Fewer studies examine outcomes for developing countries at the sub-national level – very likely due to limitations related to data availability. Magalhaes et al. (2000) examine growth in Brazilian states in a panel dataset spanning 1975 to 1995, and tests for the presence of spatial growth effects using an unconditional convergence model. They find strong evidence of spatial correlation in growth rates between the states in Brazil. Ying (2003) uses a spatial model in a dataset covering 1978 to 1998, in attempt to understand GDP-based growth spillovers among Chinese regions. Ying’s highly-significant coefficient for GRDP levels in neighboring provinces suggests a polarizing process that encourages growth in developed areas to the disadvantage of peripheral regions, according to the author. Ying also conjectures that there are “a variety of spillover effects underlying the polarizing process,” and that accounting for these spillover effects through one variable (GRDP) improves the estimates of coefficients for labor and capital. In this paper, we attempt to improve upon the methods used in the aforementioned studies by going beyond conjecture about differential effects of spillovers, to actually testing these differentiated effects.

A handful of studies have incorporated spatial effects into sub-national growth models specifically for Indonesia – namely, (McCulloch and Sjahrir 2008; Day and Ellis 2012a, 2012c, 2012b; Day and Ellis 2012d; Wardaya and Landiyanto 2005). However, like the above studies in developed countries, all of these analyses focus on the economic characteristics (particularly GRDP per capita) of neighbors, neglecting other spillover effects. McCulloch and Sjahrir (2008) conclude that nearness to a growing region has enhanced economic growth and convergence. In a study of regional GDP (GRDP) in the East Java province of Indonesia between 1983 and 2002, Wardaya and Landiyanto (2005) grouped administrative into rich and poor “clubs,” and considered spillover effects. The major finding from this research (that districts in poor clubs diverged faster than rich ones) was not spillover-related. However, they found spillover effects to be significant. In a number of sector-specific and general economic panel growth models for data spanning from 1993 to 2007, Day and Ellis (2012a, 2012b, 2012c) conclude that spatial effects are among the most significant in predicting economic growth. In this paper, we improve upon this Indonesia-specific work by testing for the presence and magnitude of differentiated spatial effects.

The process by which infrastructure contributes to economic growth can work either directly (by enhancing productivity of existing firms), or indirectly (by making the region more attractive to investment and firms looking to relocate). Chandra and Thompson (2000) find both effects to hold in a study of the United States’ Interstate Highway System. The conclude that some industries grow as a result of reduced transportation costs, while others shrink as complementary industry resettles elsewhere.
In addition to direct and indirect spillover effects, some authors differentiate between infrastructure’s contributions to the stock of local public capital (e.g., power, water), and impacts on connectivity (e.g., roads and rail). This is useful in growth models to differentiate between infrastructure factors that would be likely to have different effects on a given district. For instance, in a panel study of multiple countries using Penn World Data from 1950 to 1992, Canning and Pedroni (2004) examine the long-run effects of telephone, road, and electricity infrastructure, and conclude that a causal link exists between infrastructure provision and long-run growth. Hulten et al (2006) differentiate between direct and indirect infrastructure effects using panel data in India. Neither of these studies account for spatial effects in their models. In a study of Spanish regions, Moreno and Lopez-Bazo (2003) do include spatial effects, and conclude that when inputs are mobile, transport infrastructure in a region’s neighbors can draw production away from that region.

In developed countries, there is no empirical consensus on whether infrastructure-based spillovers exist and in which direction (i.e. positive or negative) they act. In a study of the United States’ Interstate Highway System (begun in the 1950s and completed over the next four decades), Chandra and Thompson (2000) find that highways have differential impacts in space. In counties where highways pass through, growth occurs, but the freeways tend to draw activity away from adjacent counties not receiving highways. However, Holtz-Eakin and Schwartz (1995) include state effects and conclude that the same system has had no effect on economic productivity. We recognize the importance of infrastructure spillover effects, and include them in this study.

Part of the reason for the lack of consensus could be the different approaches, including different spatial units (county versus state), and different measurement units (per capita economic growth versus worker productivity). A number of authors have also argued that early studies examining spillover effects of infrastructure inadvertently picked up the effects of spurious growth factors because they did not include controls for regional effects (this concurs with findings mentioned previously in this literature review, by Ramirez and Loboguerrero, on spillover effects of GDP per capita). This resulted in erroneously significant, large, and positive effects for infrastructure because the wrong estimation methods were used (Garcia-Mila, McGuire, and Porter 1996; Holtz-Eakin and Lovely 1996; Perret 2011). In an attempt to overcome such problems, our study focuses on regional effects.

A problem with the literature that focuses on the effects of infrastructure in developed, industrialized countries is that these places, with already-rich infrastructure stocks, can only benefit marginally from infrastructure improvements. Indeed, there is some evidence that thresholds exist after which infrastructure improvements have little marginal benefit for growth (Moreno and López-Bazo 2003; Moreno, Lopez-Bazo, and Surinach 1997). However, there is relatively little empirical evidence from developing countries on area-based spillovers (as opposed to firm-to-firm spillovers) – perhaps because reliable sub-national data are more difficult to obtain in these countries. Notable exceptions include: Lall (2007), who finds significant spatial spillover effects from roads and communications infrastructure among Indian states; and Perret (2011), who considers both local infrastructure such as power and water, and connectivity infrastructure on value added growth in the Russian Federation. Perret finds
significant but very modest effects for local private capital (elasticities of 0.06 percent) and roads infrastructure (elasticities between 0.03 and 0.05 percent). Our work adds to this emerging body of literature.

**Method and data**
This paper uses two techniques to explore spillover effects in Indonesia. First, we use nonparametric, univariate tests. Then, we model spillovers using a spatial Durbin panel model. This section describes these techniques and their applicability to our problem.

**Tests for spatial autocorrelation**
We provide two standard tests for spatial autocorrelation among districts. They are: Moran’s I and the Geary C. For both of these tests, the null hypothesis is that the test statistic is the same as the expected value of the statistic if spatial distribution of values were random. Thus, rejection of the null hypothesis indicates the presence of spatial autocorrelation (Anselin 1988; Getis and Ord 1992; Getis and Ord 1996).

Moran’s I can also be computed locally, for individual districts. In this context, the Moran’s I statistic provides a measure of association between some attribute of spatial units that are in proximity to each other, i.e., the effect on one place of being nearby another place (Anselin 1988; Getis and Ord 1992; Getis and Ord 1996). In our context, these spatial associations represent spillover effects between districts. We follow other notable authors in this approach, for example, (Perret 2011; López-Bazo et al. 1999 ). The formula for the local Moran’s I is:

\[
I_i = \frac{N}{\sum (x_i - \bar{x})^2} \sum_{i,j} w_{ij} (x_i - \bar{x})(x_j - \bar{x})
\]

(1)

where \( x \) is a variable of interest, e.g., GRDP per capita, \( i \) is a district, \( j \) refers to all other districts that could influence \( i \), and \( w \) is an analyst-defined weight matrix.

A negative Moran’s I implies that a district’s performance on measure \( x \) is inversely associated with its neighbors’ performance – i.e., negative spillover effects are present. A positive Moran’s I implies that a district’s performance on a given metric is directly associated with its neighbors’ performance – i.e., positive spillover effects. Our weight matrix in computing the Moran’s I is the same as the weight matrix used in the econometric modeling. We specify the weight matrix below.

**Econometric model background**
In general, three econometric model types comprise the bulk of empirical studies of spatial dependence. They are the spatial autoregressive (or spatial lag) model, the spatial cross-regressive model, and the spatial error model (Baumont, Ertur, and Gallo 2001; Anselin 1988; Fingleton and López-Bazo 2006; Anselin, Gallo, and Jayet 2008). A fourth model, the spatial Durbin model, is less common in empirical work. This section discusses the appropriateness and limitations of
each model to this study, and summarizes how we arrived at our final model specifications.

The spatial lag panel model generally takes the form (LeSage and Pace 2009; Elhorst 2010):

\[ y_d = \delta \sum_{j=1}^{N} w_{ij} y_{ji} + x_{it} \beta + \mu_i + \epsilon_u \]  

(2)

For a growth model such as ours, \( y_{it} \) is annual GRDP per capita growth, \( x_{it} \) is a matrix of exogenous variables, \( w \) is a weight matrix relating district \( i \) to all other districts \( j=1 \) to \( N \) (discussed in more detail below), and \( t \) denotes a time wave. Also, \( \delta \) is a coefficient to be estimated and \( \beta \) is a vector of coefficients to be estimated, \( \mu \) is a spatial time-invariant fixed effect, and \( \epsilon \) is the standard (well-behaved) error term.

The spatial lag model, notably, does not introduce exogenous spatially-lagged regressors. The spatial cross-regressive model includes such spatially-lagged exogenous regressors, and takes the general form in panel models:

\[ y_d = \delta \sum_{j=1}^{N} w_{ij} y_{j,i-1} + x_{it} \beta + \mu_i + \epsilon_u + \epsilon_{it} \]  

(4)

Thus, the spatial cross-regressive model extends the spatial lag model by incorporating spatial lags of both initial GRDP per capita levels \( y_{i,1} \) and exogenous control variables – instead of collecting all of the effects of \( X \) into a single variable \( g_y \), as the spatial lag model does. This allows for differentiated effects to be discerned among the factors that contribute to growth; for example, infrastructure, capital, human capital, land, and labor.

In the spatial error model, since spillover effects are accounted for in the error term, growth in a region is treated as a function of random shocks in neighboring regions. The spatial error model takes the general panel form (Elhorst 2010):

\[ y_d = x_{it} \beta + \mu_i + \phi_u \]  

(5)

\[ \phi_u = \rho \sum_{j=1}^{N} w_{ij} \phi_{j,1} + \epsilon_{it} \]  

(6)

where \( \epsilon \) is a well-behaved error term and \( \rho \) is a scalar parameter that indicates the degree of spatial correlation between residuals.

The spatial error model can also be expressed in the Durbin representation, which nests the spatial lag and spatial error. The spatial Durbin model contains both a spatially lagged dependent variable and spatially lagged independent variables (Baumont, Ertur, and Gallo 2001; Fingleton and López-Bazo 2006; LeSage and Pace 2009; Elhorst 2010). In unconstrained form, the model is as follows:
The spatial Durbin representation includes all elements of models previously discussed: initial GRPD per capita \((y_{it-1})\) and the influence of other districts' initial levels, spatially-lagged growth, and exogenous variables from both the home district \(i\), and neighboring districts \(j=1\) to \(N\). Thus, it incorporates three types of spatial effects: initial level of GRP per capita, growth, and exogenous factors in neighboring regions. If \(\gamma=0\), then the Durbin simplifies to the spatial lag model; if \(\gamma+\delta\beta=0\), then it simplifies to the spatial error model. These conditions can be tested using the common factor test (Baumont, Ertur, and Gallo 2001; Burridge 1981). Because we are interested in disaggregated spatial effects, we begin with the full Durbin specification. We eliminate some estimated relationships due to identification problems (discussed in the next paragraph). We then measure empirically whether the model simplifies to the lower models. These empirical results are presented in the Findings section.

One potential problem with the spatial Durbin model in the form above is that its parameters are not identified (Anselin, Gallo, and Jayet 2008). Elhorst (Elhorst 2012) summarizes possible adjustments to the Durbin specification that address this identification problem: 1) exclude the exogenous interaction effects \(WX_t\) from the model; 2) exclude the endogenous interaction effects \(WY_t\); or 3) exclude lagged endogenous interaction effects \(WY_{t-1}\), among others. For our purposes, because our core inquiry is in the exogenous interaction effects, eliminating the \(WX_t\) term is not appropriate. Here, we choose to eliminate the lagged endogenous interaction effects \(WY_{t-1}\), rationalizing that they are less important than current endogenous interaction effects. Thus, our final specification is as follows:

\[
y_a = \nu y_{i,t-1} + \delta \sum_{j=1}^{N} w_{ij} y_{j,t-1} + \lambda \sum_{j=1}^{N} w_{ij} y_{j,t} + x_a \beta + \sum_{j=1}^{N} w_{ij} y_{j,t} + \mu_i + \epsilon_i \tag{8}
\]

Depending on test results as outlined by (Baumont, Ertur, and Gallo 2001), the Durbin representation could collapse to the spatial error model – though other studies on spillover effects in Indonesia suggest that collapse to the simpler model is unlikely because spatial growth effects appear to be present in Indonesia (Day and Ellis 2012a, 2012c, 2012b; Day and Ellis 2012d). This specification follows the advice of Fingleton and Lopez-Bazo (2006), who review studies incorporating spatial effects on long-run growth. These authors warn that spatial dependence in growth models should be of the substantive type (those that include spatial lags of dependent variables), rather than the nuisance error dependence type (i.e., the spatial error model). In reality, most studies that incorporate spatial effects use the spatial error model. Thus, our work improves upon much of the existing literature by using a more-appropriate model. As it turns out, our specification does not collapse to the spatial error model, and thus is in line with best practices.

Another good reason to use the Durbin representation is pointed out by Elhorst (2011). Elhorst points out that, when faced with an identity problem, the best option is to exclude the spatial autocorrelation in the error term and instead...
consider endogenous and exogenous interaction effects, i.e., the Durbin model. According to LeSage and Pace (2009, pp. 155-158), conclude that ignoring spatial dependence in the dependent variable or the independent variables is more serious than ignoring spatial dependence in the error terms. They rationalize that if relevant explanatory variables are omitted from the model, the estimator will be biased and inconsistent for the variable coefficients in the model. Ignoring spatial dependence in the disturbances, on the other hand, will only result in a loss of efficiency.

Two important issues are worth noting in regard to estimating spatial autoregressive models. First, they must be estimated using Maximum Likelihood (ML), Generalized Method of Moments (GMM), or Instrumental Variables (IV) methods, rather than Ordinary Least Squares (OLS), since the error term will always be correlated with \( w \), even if the residuals are identically and independently distributed (Anselin 1988; Baumont, Ertur, and Gallo 2001). For the spatial error model, the specification assumes that the spatial dependence is represented in the variables omitted from the analysis, rather than in the control variables contained in \( x \). In this formulation, the spatial error model does not suffer from systemic correlation between the error term \( v \) and \( w \). However, OLS is still not a viable estimation method since its use with non-spherical error terms yields inefficient (though unbiased) estimators (Baumont, Ertur, and Gallo 2001). The second issue noted by Baumont et al (2001) is that the models assume that a random shock in region \( i \), affects growth in all other regions \( j \). If there are theoretical reasons why this assumption may not hold, this formulation should be interpreted with care.

We take two approaches to estimating Equation 8. First, we estimate a fixed effects model (Elhorst 2003). Although the previous discussion indicates that fixed effects will be biased, recent work offers a process by which the data can be transformed to avoid this bias (Lee and Yu 2010). In choosing fixed effects models, we follow other notable authors that control for area-specific factors but do not consider possible endogenous relationships in estimating spillover effects (Garcia-Mila, McGuire, and Porter 1996; Perret 2011; Holtz-Eakin 1994).

One serious problem with fixed effects models is that, when the lagged dependent variable is included as a regressor, the estimation can be biased – in some cases by up to 20 percent (Kennedy 2003, 308). Thus, our fixed effects models exclude the lagged dependent variable. This is problematic because, in theory, lagged GRDP level is a significant predictor of current GRDP levels. Our second set of models allows for the inclusion of the lagged dependent variable. Using Stata’s \textit{spregdpd} routine, we estimate the Durbin model using a maximum likelihood estimator (MLE) – in particular, a Spatial Panel Arellano-Bond Linear Dynamic Regression (Arellano and Bond 1991). The \textit{spregdpd} routine allows for automatic testing of several hypotheses regarding the Durbin specification, including those that collapse it to the spatial error and spatial lag models as described above. All of our variables are either log-transformed or presented as percentages, so the effects can be interpreted as elasticities.

As per Equation 8, our formulation of spillover effects considers neighbors’ endowments and characteristics in conjunction with a district’s own effects. This is in contrast to Moran’s I, a more simple measure. Moran’s I is based on an average of neighbors’ effects on a given district (Equation 1).
**A note about interpreting coefficients**

Elhorst (2011) notes that it is incorrect to directly interpret the coefficient estimates, variance-covariance matrix, standard errors and t-values, that arise from spatial models. LeSage and Pace (2009) define a partial derivative approach for producing interpretable effects from the model coefficients estimates. Specifically, they demonstrate that the effects of individual variables in a model is comprised of a partial derivative of a combination of all model coefficients and the weight matrix. That process is too detailed to be presented here, but it is the process we follow in this paper. The effects presented in this paper, then, are not direct coefficient estimates, but rather, are the partial-derivative effects that LeSage and Pace identify. Fortunately, the authors of Stata’s `spregdpd` routine provide these estimates as part of the routine output. They also provide appropriate t-values, computed based on the partial derivative framework. Elhorst (2011) notes of this partial derivative method that “only these effects estimates should be used to draw inferences regarding the relationships we are modeling (pp. 17).”

**The weight matrix**

As we will show in our results sections, the formulation of the spatial weights matrix is important in the model specification, as variations in this matrix can significantly affect the estimation results. A growing literature that warns analysts of the susceptibility of econometric models to the specification of the weight matrix, for example (Anselin 1988; Anselin 1999; Tiefelsdorf, Griffith, and Boots 1999). The weight matrix, \( w \), in general reflects the influence of a given district \( j \), on the home district \( i \). The diagonal of \( w \) contains zeroes, as a district’s own effect on itself is accounted for elsewhere in the model.

The literature is divided between papers that use only one weight matrix formulation, and those that check the robustness of their estimations using multiple weight matrices. Some studies have begun to suggest processes to identify weight matrices empirically (Bhattacharjee and Jensen-Butler 2006). We use two weight matrices in order to provide some test of robustness.

It is possible to formulate the weights in \( w \) in a large number of ways, for example: adjacency (a district receives a value of 1 if adjacent to district \( i \), and a zero otherwise), distance (as-the-crow-flies or road distance separating districts \( i \) and \( j \); generally an inverse distance formulation is used), time (travel time between districts \( i \) and \( j \)), economic attractiveness, among others. We provide model estimates using an inverse-distance weight matrix (where all off-diagonal values in the matrix consist of 1 divided by the distance). This ensures that higher values reflect higher levels of closeness. We then innovate on the literature by using a gravity-based weight matrix. The gravity approach generally involves weighting districts by their attractiveness, tempered by some measure of separation:

\[
W_{ij} = \frac{\text{ATTRACTION}_j}{\text{SEPARATION}_{ij}}
\]  

[9]
where *ATTRACTION* refers to an attribute of a neighboring district \( j \), in this case GRDP *per capita*, and *SEPARATION* is the road travel time between the centroids of the two districts \( i \) and \( j \).

We choose the gravity method for development of \( w \), reasoning that spillover effects for a given district in Indonesia would be influenced by not only the distance separating it from its neighbors, but also by the magnitude of influence that those neighbors exert on district \( i \). The formulation of our weight matrix assumes that richer neighbors exert stronger influence than poorer neighbors, i.e., that neighbors with higher GRDPs *per capita* exert stronger influence over their neighbors than those with lower GRDPs *per capita*. As a measure of separation, we use road travel time, generated from road network data and estimated travel speeds in ArcGIS. The gravity approach is superior to a distance or adjacency-based approach in that it does not assume reciprocal spillover effects between two regions, but rather, allows for asymmetric effects between districts.

Asymmetric spatial weights are a common feature of spatial models, and are particularly useful when there is reason to believe that there are asymmetric effects between spatial units of analysis (Anselin 1988, 23-24). Common applications of asymmetrical spatial weights matrices include core-periphery models and models of network flows (Anselin 1999; Bhattacharjee and Jensen-Butler 2006). Because Indonesia’s districts vary widely in terms of their distance to economic agglomerations of various sizes, and because this analysis includes remote rural districts as well as large urban agglomerations, it would not be appropriate to treat the attractions between a given district pair as symmetric in all cases. A rural area near an urban agglomeration will experience different effects from its urban neighbor than that urban neighbor receives from it. Our final weight matrices are row normalized, i.e., all terms in a row sum to one.

Elhorst (2003) points out that the asymptotic properties of the maximum likelihood (ML) estimator depend on the features of the weight matrix. Elhorst particularly notes that when the spatial weight matrix is an inverse-distance formulation (as ours is), do not undermine the consistency of the ML estimator, particularly when panel data are used in the estimation. Both of the weight matrices we use in the final specifications are row normalized.

**Data, variables, and spatial units**

In deciding which variables to include as regressors, we follow a neoclassical Solow-style growth model (Barro and Sala-i-Martin 1994), focusing on standard factors of production (land, labor, and capital). We include infrastructure specifically in the specification equations, rather than just as a subset of capital, because there is a need in our growth models to differentiate between infrastructure factors that would be likely to have positive versus negative spillover effects on a given district. Differentiated infrastructure effects are commonly accounted for in the measurement of spillover effects; for instance, in Moreno and Lopez-Bazo (2003) and Perret (2011).

All data used in the analyses here come from various surveys and censuses conducted by Indonesia’s Central Bureau of Statistics (BPS) and the Ministry of Finance. GRDP *per capita* and other monetary values are given in real 2000 Indonesian Rupiah (IDR). In 1999, Indonesia began its “Big Bang”
decentralization process, transferring some spending and government functions to local (i.e., district) governments. There were 298 administrative local-government districts before the process began, including six in the Jakarta metropolitan area. By 2007, there were 434 districts, and the number continues to grow. In order to compare the same spatial units over time, we collapse district data to pre-Big Bang configuration (298 districts).

The analysis includes all districts for which data are available. As specified in the equations above, we use indicators of land, labor, capital, and infrastructure as determinants of growth. In particular, we are interested in testing whether spillovers related to these variables are responsible for explaining some portion of the economic growth in Indonesian districts.

To reflect the availability of land available for economic development, we employ as an indicator the reported unused land area for villages in the district. These data come from the Village Potential Survey (Potensial Desa, or Podes), which was conducted in years 2003, 2005, and 2008 of our study period. Intermediate years were interpolated linearly.

As broad indicators of the size of the labor force and the capacity of that labor force to generate economic activity, we use population, the percentage of the population that is of working age (i.e. 15-64 years old), and mean years of schooling of district residents. These data come from the annual National Socioeconomic Survey (Susenas).

As indicators of capital availability and effectiveness, we include capital owned by medium and large manufacturing firms, as a percent of value-added, and EXPY, a measure of the level of “upscaleness” of the economic outputs of the district – that is, the capacity of an economy to move up the production food chain to producing higher-end, higher value-added goods. The data for this indicator also come from the Statistik Industri. We build upon Hausmann et al’s (2007) concept of EXPY to build an indicator of upscaleness. EXPY is an indicator of the sophistication of a country’s export basket – i.e., whether a country is producing goods that are on the “upscale” end of the product spectrum. The core idea behind the EXPY computation is that, if all other factors are held constant, “an economy is better off producing goods that richer countries import.” Hausmann’s results show that high EXPY values – that is, countries whose exports are more upscale – are associated with high levels of economic growth. EXPY is always positive, with a minimum of zero. The higher the number, the more sophisticated the export basket of an area. There is no upper bound on the range.

We apply Hausmann et al (2007) framework to sub-national data, to develop an indicator of the competitiveness of Indonesia’s sub-national (district) economies. Specifically, we looked at manufacturing sophistication, using manufacturing value-added data (from SI). Manufacturing data were used because, in the context of a developing country in the middle-income category (like Indonesia, among others), manufacturing industries are significant drivers of growth (Shen 2000; UNIDO 2009; and Yusuf and Nabeshima 2010). Medium and large firm data were further considered to be representative of the substantial portion of the manufacturing industry, since by nature manufacturing requires economies of scale to be effective.

Following theory that distinguishes the attractive powers of some infrastructure (power, water, etc.) and centrifugal features of other
infrastructure (roads), we develop two types of indicators of a district's infrastructure. The first of these reflects the total amount of electricity used (in kilowatt hours) as a proportion of value added; another variable measures the percent of electricity use that is self-generated (i.e. as opposed to supplied by the National Electricity Company—PLN). These data come from the annual Large and Medium-Scale Manufacturing Survey, or Statistik Industri (SI). The second variable indicates the proportion of villages within a district that is served by asphalt or gravel roads (from Susenas).

We acknowledge that electricity and capital data gathered from manufacturing firms (i.e. from SI) are non-ideal for measuring whole-economy growth. However, we view them as suitable proxies in the absence of other data on these attributes of districts, in light of evidence pointing to industry as the largest sectoral beneficiary of public infrastructure provision compared with other economic sectors (Holtz-Eakin and Lovely 1996; Moreno, Lopez-Bazo, and Surinach 1997). Table A.1 in Appendix A gives summary statistics for all variables included in the final model specifications, for the year 2008.

**Findings**
The findings appear to confirm the two hypotheses framing this study: first, that estimations of growth equations can be improved by accounting for spatial spillover effects in Indonesia; and second, that these effects are discernible and important at the level of individual predictor variables. A third and unexpected set of findings has arisen from the comparison between the local Moran’s I (Equation 1) and the econometric model results, which appear to contradict each other. All of these are discussed in this section.

Tables 1 and 2 summarize the findings from the hypothesis tests for the presence of spatial autocorrelation, i.e. Moran’s I and the Geary C statistic, respectively. The Moran’s I test confirms the presence of spatial autocorrelation in all of our model’s predictor variables, with high degrees of significance on all variables. The Geary C test provides similar results, but fails to confirm spatial autocorrelation in two of the variables (GRDP per capita growth and working age population), at the 5 percent significance level.

[TABLES 1 AND 2 ABOUT HERE]

Figures 1 and 2 show schematic mapping of the local Moran’s I, computed using 2008 GRDP per capita, for Western and Eastern Indonesia, respectively. The maps are shaded to indicate positive (hatched in the print copy; blue in the online version) and negative (shades of grey in the print copy; red in the online version) spillover effects. Different hatch patterns or darker hues of grey indicate more-intense spillover effects (in the print version). In the online version, darker hues of red and blue indicate more-intense spillover effects. Indonesia’s administrative districts are divided into kota (city) and kabupaten (regency). Kota are urban districts and kabupaten are non-urban, and the two are distinguishable by geographic size (kota are generally much smaller), demographics, and economic structure. All of the country’s kota (cities) are highlighted in black in Figures 1 and 2, and some select kota are labeled in the figures. Island names are indicated in boldface type.
Perhaps the most striking feature of this schematic representation is the stark contrast between urban regions and their rural hinterlands on Java. The six kota comprising Jakarta indicate highly positive economic spillover effects for the national capital, and are surrounded by districts with highly negative spillovers. Although there are some kabupaten that show positive spillover effects near Jakarta, the large majority of kabupaten on Java are have negative spillover effects. In addition, the smaller positive-spillover kotas (Bandung, Sukabumi, etc.) scattered around Java – all surrounded by negative-spillover hinterlands – also suggest that cities are benefitting economically from their rural hinterlands, but that the reciprocal is not occurring.

The same pattern holds true – though it is not so extreme – in many other areas of the country. On Sumatra, Medan for instance, shows positive spillover effects, while its hinterlands have negative spillover effects. Kota Banda Aceh shares the benefits of urbanization with its immediate hinterlands, but not with those further away. On Kalimantan, Kota Pontianak is an island of positive spillovers surrounded by negative spillover effects. All kotas on Sulawesi, Kota Ambon on Maluku, and Kota Jayapura on Papua, all share this patterning. This patterning does not hold universally, however. Four kabupaten near Jakarta show positive economic spillovers, as do a large number of kabupaten on the eastern coast of Kalimantan, and several near Samarinda and Balikpapan. This implies that at least some kabupaten have been able to leverage proximity to urban areas or other higher-income kabupaten, for their own economic prosperity.

These trends in the Moran’s I for GRDP per capita, while illustrative, inform us only about one dimension (GRDP per capita) of a multi-dimensional problem. In the remainder of this section, we interpret the results of the econometric model output, which provide a more-disaggregated examination of the spatial drivers of growth.

Tables 3 and 4 provide the model estimation results for the fixed effects and MLE models, respectively. In Table 3, Model 1 shows the estimations for the annual model with no spillover effects included, where the outcome variable is GRDP per capita levels. Models 2 and 3 contain spillover effects, first with an inverse-distance $w$, and then with the gravity-based $w$. In Table 4, results are presented for a based model estimated using random effects with MLE (Model 4), and then two MLE models with spillover effects, again with inverse-distance and gravity-based $w$, respectively.

We note, in preface, that there are apparent contradictions between the Moran’s I statistics reported in Figures 1 and 2, and the model estimations in Table 3. We explore the implications of this mismatch at the end of this section.
cannot comment on the improvement in this measure of model fit in Table 4, as
the random effects model does not provide an R-squared-type measure of model
fit. However, for the fixed effects model, the improvement in fit is large, from
0.429 to above 0.99 in both Models 3 and 4.

Second, the fixed effects and MLE models generally agree in their
estimates of the signs of the coefficients, with one major exception: the
coefficient on population, which is positive in Models 2 and 3 but negative in
Models 5 and 6. The reason for this is the inclusion of lagged y as a regressor in
the MLE models, but not in the fixed effects models. Because population is highly
correlated with GRDP per capita (larger places are wealthier), failure to account
for initial GRDP per capita in the fixed effects models results in the coefficient for
population absorbing the effects of initial GRDP per capita. This is evident in
Models 7 and 8, which show very strong effects of the lagged GRDP per capita
levels.

Third, all of the spatial models produce more significant coefficients, and
more highly-significant p-values, than the non-spatial models. In several cases,
the spatial models have significant coefficients, whereas the non-spatial models
do not, e.g., working age population and capital (Table 3) and population and
years of schooling (Table 4). In several cases, also, the inclusion of the spatial
variables in the model significantly alters the magnitude and levels of
significance of some of the estimated own-district effects, e.g., average years of
schooling, population, and roads. In some cases, only the spatial variable is
significant, e.g., capital and both electricity variables (in Table 4). These
differences between the base models and the spatial models indicate that
controlling for spatial effects reduces misspecification bias that masks the
significance of some variables in the non-spatial models.

Fourth, the models using the gravity-based weight matrix tend to produce
larger coefficients. This is logical, considering the weighting scheme is designed
to weight the effects of wealthy neighbors more heavily than poorer neighbors.
And fifth, despite high levels of significance, many of the effects shown in the
tables are very small. There are a few notable exceptions, which we discuss
below.

Because of the similarity in results between the fixed effects and MLE
models (excluding the lagged dependent variable effect, discussed above), and
also because of the better specification of the MLE models (resulting from the
inclusion of the lagged dependent variable), the focus of the remainder of this
section is on the maximum likelihood estimations. For the MLE models
presented in Table 4, three sets of hypothesis tests were run to test whether the
model would collapse to a simpler model, as discussed in the model section
above. All of these hypothesis tests have a null hypothesis of insignificant spatial
effects for some term in the model. In only one case did we fail to reject the null
hypothesis. This was in the tests for nonzero effects in the spatially lagged
dependent variable, shown in the table as the LM Lag Anselin Test and the LM
Lag Test (robust). Thus, our Durbin representation collapses to a model that
does not include spatially-lagged dependent variable effects. In order to test the
hypothesis that there are differential effects for urban and rural areas, we
further ran the models with an interaction variable computed by interacting
spatially lagged GRDP per capita with an urban dummy variable. The test for this
effect being nonzero was also insignificant. These are the final models presented in Table 4.

Consistent with theory and previous empirical work, economic growth in Models 5 and 6 is positively correlated with initial *per capita* GRDP levels. However, when spillover effects of GRDP *per capita* levels are accounted for, the effect of initial GRDP *per capita* increases in absolute value from 98.7 percent to nearly 159 percent, depending on the weight matrix used. This suggests that simply using GRDP *per capita* levels in a model, without considering the effects of neighboring districts, under-estimates the magnitude of the effect significantly. Table 4 further shows that in fact most explanatory variables included in the analysis of economic growth (i.e. in addition to GRDP *per capita*, GRDP *per capita* growth, and the urban interaction variable on growth spillovers) have noteworthy spillover effects. In particular, the output demonstrates that a region’s economic growth apparently depends to a significant extent on its neighbors’ unused land, population, working-age population, residents’ education level, EXPY, private capital investment, firm electricity use and reliability, and roads, all else being equal. All of the spillover effects modeled here are significant in at least one of Models 5 and 6.

An interesting example of the bias reduction in the spatial models is in the effect of population. Model 4 shows no effect of population on GRDP *per capita* levels. However, in Models 5 and 6, the effects are large and pronounced, with the spillover effects having opposite signs to the own-district effect. This sign mismatch masked the effect of population on economic levels in the non-spatial model.

All of the effects shown in the spatial models in Table 4 have effects that read as expected. Higher concentrations of unused area is correlated with lower GRDP *per capita* in Model 5, reflecting lower levels of firm investment. Higher own-district population is negatively correlated with GRDP *per capita*, which is reasonable because we have controlled for initial GRDP *per capita* levels (places with similar initial GRDPs *per capita* but higher population will see those GRDPs *per capita* diluted by the extra people. However, Model 6 indicates that having higher-population neighbors is good for productivity. Working-age population is positively correlated with GRDP *per capita* in own district and neighbors, as is years of schooling. The latter implies that there are returns to education within district and also for neighbors.

The EXPY variable indicates that own-district upscaleness is good for GRDP *per capita*, but competitive neighbors detracts from a district’s wealth. Capital investment and electricity variables are only significant in the spillover effects. For capital, this implies that neighbors’ capital investments positively affect a district’s GRDP *per capita*. The spillover effects of electricity reliability is also consistent with expectations: as neighbors’ electricity reliability weakens, the associated dis-benefits also tend to flow back to the home district (Model 6). This suggests that non-reliable energy sources can impede growth even in neighboring economies. The spillover effects of road accessibility are also positive in both Models 5 and 6, indicating that having neighbors with more roads is associated with own-district GRDP *per capita*.

There are two variables in Table 4 whose effects are not consistent between Models 5 and 6. They are: the spillover effect of electricity per value added and the own-district effect of roads access. These cases illustrate that the
composition of the weight matrix is important in producing interpretable findings.

As indicated above the findings in Table 3 appear to contradict the Moran’s I statistics displayed in Figures 1 and 2. This contradiction lies in the observation that the Moran’s statistics indicate a pattern of urban areas benefitting from spillover effects (positive spillovers) while rural areas fail to benefit (negative spillovers). This is in direct contrast with the spatially lagged GRDP per capita having no effects for both the overall case and for urban areas (described above). It is in the lack of effect on lagged GRDP per capita that we can resolve the mismatch between the Moran’s I and the econometrics. That the spatially lagged dependent variable is insignificant when other spatial effects are included, implies that the other model variables are responsible to the effects that the Moran’s I attributes to GRDP per capita.

Relying on a simple univariate measure like Moran’s I, which aggregates all effects (land, labor, capital, etc.), fails to account for different types of spillovers, each having different underlying processes. In contrast, by disaggregating spillover effects into constituent land, labor, capital, etc., effects, the regression model unmasks sometimes-large effects that the Moran’s formulation does not. This is illustrated in the coefficients in the econometric models. Because these effects – including the large negative effect of own-district previous GRDP per capita, which increases with the estimation including spillovers – are separated from the effects of general growth spillovers by the econometric models, they cannot distort the measure as they can when the more-simple Moran’s I is used. Thus, our model output highlights the improvement of our model formulation on more simple measures like the Moran’s I.

The contradiction between the Moran’s I and the model outputs raises many questions as to the causes of the mismatch, which we cannot test in the space allowed here. For instance, it is possible that the intense spillover effects in and around urban areas (indicated by the greys and hatched patterns in Figures 1 and 2) overshadow smaller effects in rural areas, thus causing the model to register the overall effect as positive when the real effect might actually be much more nuanced. This could happen because the magnitude of economic activity in urban areas is much larger in Indonesia, and thus could cause the coefficients of GRDP per capita spillovers to lean toward the positive on average, even though many particularly rural areas experiences negative spillover effects.

Such averaging is a drawback of econometric models. However, while the Moran’s I examines localized effects, it is guilty of the same problem as the model: that it cannot discern the directionality or nuance of spillover effects. In the formulation provided here, it cannot tell us, for example, whether rural areas’ negative spillovers (where they occur) arise from interactions with urban neighbors or other rural neighbors.

The apparent contradictory findings between the Moran’s I and the econometrics, then, also point to a need for more exploration with models of this type, which account for disaggregated spillover effects, particularly at sub-national levels. The model contained in this paper arrives at overall average positive spillover effects for all types of districts. That the study covers such a wide variety of districts – from rural districts that are adjacent to urban ones, to
distant and isolated rural districts; from isolated urban districts to urban districts in massive metropolitan agglomerations; from large local economies to tiny ones. Future studies should use this framework to examine more-nuanced subsets of areas within national boundaries. Examples could include urban-to-rural spillovers – i.e., spillovers that accrue from urban centers to rural hinterlands, and vice versa, where urban places and their hinterlands are defined in economic terms instead by administrative boundaries (Wang et al 2005). The power of our model is that it provides a framework to think not only about the directionality of influence, but also the disaggregated land, labor, capital, human capital, and infrastructure factors underlying that influence.

Conclusions
Perhaps the most significant conclusion from this paper is that various types of spillovers affect a place in different ways and magnitudes, and accounting for these effects in growth models may improve the efficacy of those models. Our findings suggest that, for Indonesian districts, the influence of neighbors includes demographics, capital, human capital, and infrastructure components. Significant relationships were found in land availability, working-age population, schooling, capital, and electricity provision. In our study, few variables had an effect that remained localized (maintained entirely within that district) when spillover effects were accounted for.

These disaggregated effects may pose challenges for Indonesia, a country working hard to decentralize financial and administrative functions. The findings related to working-age population suggest that districts should recruit labor from surrounding districts in order to increase their own productivity and also to diminish growth competition from neighbors. This has implications for national migration policy, and raises concerns about the ability of districts to handle large influxes of population in terms of water, sanitation, schools, and other needs. Somewhat antithetically to the ability of local governments to respond with appropriate policy, our findings also suggest that own-district education policy is not as important as being situated close to districts with good education outcomes – and similarly with capital and roads. Future work should also continue to disaggregate the effects of spillovers on economic growth. This disaggregation should occur among factors (land, labor, capital, etc.), but also among geographies tested – for instance, urban-to-rural spillovers. Our model provides a framework within which future works could discern which types of districts (small and emerging, large and developed, urban, rural, etc.) benefit from different types of spillover effects.
### Table 1. Moran's I Tests for Spatial Autocorrelation in Predictor Variables, 2008

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Moran's I</th>
<th>Expected I</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of GRDP per capita</td>
<td>1.052</td>
<td>-0.003</td>
<td>0.055</td>
<td>19.142</td>
<td>0.000</td>
</tr>
<tr>
<td>GRDP per capita growth (g_y)</td>
<td>3.410</td>
<td>-0.003</td>
<td>0.137</td>
<td>24.831</td>
<td>0.000</td>
</tr>
<tr>
<td>Unused area (percent of total area)</td>
<td>0.107</td>
<td>-0.003</td>
<td>0.010</td>
<td>11.098</td>
<td>0.000</td>
</tr>
<tr>
<td>Working-age population (aged 15-64 years)</td>
<td>2.365</td>
<td>-0.003</td>
<td>0.065</td>
<td>36.621</td>
<td>0.000</td>
</tr>
<tr>
<td>Average years of schooling</td>
<td>1.476</td>
<td>-0.003</td>
<td>0.062</td>
<td>23.823</td>
<td>0.000</td>
</tr>
<tr>
<td>EXPY</td>
<td>0.821</td>
<td>-0.003</td>
<td>0.054</td>
<td>15.255</td>
<td>0.000</td>
</tr>
<tr>
<td>Capital (as a percent of manufacturing value-added)</td>
<td>15.178</td>
<td>-0.003</td>
<td>0.135</td>
<td>112.251</td>
<td>0.000</td>
</tr>
<tr>
<td>Electricity (percent firm self-generated)</td>
<td>0.453</td>
<td>-0.003</td>
<td>0.153</td>
<td>2.976</td>
<td>0.001</td>
</tr>
<tr>
<td>Roads (percent of villages with access to an asphalt or stone road)</td>
<td>0.159</td>
<td>-0.003</td>
<td>0.015</td>
<td>10.568</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Table 2. Geary's Tests for Spatial Autocorrelation in Predictor Variables, 2008

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Geary C</th>
<th>Expected C</th>
<th>Standard Error</th>
<th>Z-score</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of GRDP per capita</td>
<td>0.431</td>
<td>1.000</td>
<td>0.076</td>
<td>-7.444</td>
<td>0.000</td>
</tr>
<tr>
<td>GRDP per capita growth (g_y)</td>
<td>2.297</td>
<td>1.000</td>
<td>0.896</td>
<td>1.447</td>
<td>0.074</td>
</tr>
<tr>
<td>Unused area (percent of total area)</td>
<td>1.790</td>
<td>1.000</td>
<td>0.118</td>
<td>6.679</td>
<td>0.000</td>
</tr>
<tr>
<td>Working-age population (aged 15-64 years)</td>
<td>0.909</td>
<td>1.000</td>
<td>0.202</td>
<td>-0.449</td>
<td>0.327</td>
</tr>
<tr>
<td>Average years of schooling</td>
<td>0.317</td>
<td>1.000</td>
<td>0.086</td>
<td>-7.902</td>
<td>0.000</td>
</tr>
<tr>
<td>EXPY</td>
<td>0.391</td>
<td>1.000</td>
<td>0.076</td>
<td>-8.001</td>
<td>0.000</td>
</tr>
<tr>
<td>Capital (as a percent of manufacturing value-added)</td>
<td>4.313</td>
<td>1.000</td>
<td>1.181</td>
<td>2.805</td>
<td>0.003</td>
</tr>
<tr>
<td>Electricity (percent firm self-generated)</td>
<td>0.206</td>
<td>1.000</td>
<td>0.477</td>
<td>-1.666</td>
<td>0.048</td>
</tr>
<tr>
<td>Roads (percent of villages with access to an asphalt or stone road)</td>
<td>0.413</td>
<td>1.000</td>
<td>0.083</td>
<td>-7.039</td>
<td>0.000</td>
</tr>
</tbody>
</table>
Table 3. Elasticities from Fixed Effects Regression Models for Annual GRDP per capita Levels in Indonesian Districts, 2003-2008

<table>
<thead>
<tr>
<th>Model Number</th>
<th>Fixed Effects</th>
<th>Spatial Panel Fixed-Effects Durbin Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight Matrix</td>
<td>Lagged GRDP per capita</td>
<td>N/A</td>
</tr>
<tr>
<td>Unused area (percent of total area)</td>
<td>-0.001*** (0.014)</td>
<td>0.0002 (0.459)</td>
</tr>
<tr>
<td>Spillover effects of unused area</td>
<td>-0.000 (0.462)</td>
<td>0.004*** (0.000)</td>
</tr>
<tr>
<td>Population</td>
<td>1.137*** (0.000)</td>
<td>0.387*** (0.000)</td>
</tr>
<tr>
<td>Spillover effects of population</td>
<td>0.000009 (0.939)</td>
<td>0.084*** (0.000)</td>
</tr>
<tr>
<td>Working-age population as percent of population (aged 15-64 years)</td>
<td>0.000 (0.132)</td>
<td>0.001** (0.031)</td>
</tr>
<tr>
<td>Spillover effects of working-age population</td>
<td>0.0003*** (0.010)</td>
<td>0.090*** (0.024)</td>
</tr>
<tr>
<td>Average years of schooling</td>
<td>0.066*** (0.000)</td>
<td>0.019*** (0.008)</td>
</tr>
<tr>
<td>Spillover effects of average years of schooling</td>
<td>0.0002*** (0.000)</td>
<td>0.014*** (0.000)</td>
</tr>
<tr>
<td>EXPY</td>
<td>0.067*** (0.000)</td>
<td>-0.0004 (0.282)</td>
</tr>
<tr>
<td>Spillover effects of EXPY</td>
<td>-0.00003*** (0.005)</td>
<td>0.000 (0.000)</td>
</tr>
<tr>
<td>Capital (as a percent of manufacturing value-added)</td>
<td>-0.000 (0.758)</td>
<td>-0.000003 (0.780)</td>
</tr>
<tr>
<td>Spillover effects of capital</td>
<td>0.00000005*** (0.041)</td>
<td>-0.000 (0.640)</td>
</tr>
<tr>
<td>Electricity reliability (percent firm self-generated)</td>
<td>-0.000*** (0.023)</td>
<td>0.443 (0.037)</td>
</tr>
<tr>
<td>Spillover effects of electricity reliability</td>
<td>-0.000006*** (0.010)</td>
<td>-0.000 (0.039)</td>
</tr>
<tr>
<td>Electricity used per value added</td>
<td>-0.034*** (0.000)</td>
<td>-0.005 (0.063)</td>
</tr>
<tr>
<td>Spillover effects of electricity per value added</td>
<td>-0.000001 (0.259)</td>
<td>-0.174*** (0.000)</td>
</tr>
<tr>
<td>Roads (percent of villages with access to an asphalt or stone road)</td>
<td>0.002*** (0.008)</td>
<td>-0.0000009 (0.854)</td>
</tr>
<tr>
<td>Spillover effects of Roads</td>
<td>0.000009*** (0.000)</td>
<td>0.028*** (0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>-14.786*** (0.000)</td>
<td>-5.990*** (0.000)</td>
</tr>
</tbody>
</table>

Observations: 1788, 1788, 1788
Cross Sections: 298, 298, 298
R-squared: 0.429, N/A, N/A
Raw Moments R-squared: N/A, 0.9992, 0.9998
Wald Test p-value: N/A, 0.000, 0.000
F Test p-value: N/A, 0.000, 0.000
Hausman Test p-value: N/A, 0.00, 0.00
Fixed/Random Effects: Fixed, Fixed, Fixed

p-values in parentheses (*** p<0.01, ** p<0.05, * p<0.1)

1 All variables (dependent and independent) are log-transformed using natural logarithms, except those given in percent terms. These are: Area (given as percent of total area that is unused), Capital (given as a percent of manufacturing value-added), Electricity (given as the self-generated electricity as a percent of the total consumed by manufacturing firms), and Roads (given as the percent of villages with access to a sealed road).
Table 4. Elasticities from MLE Regression Models for Annual GRDP per capita Levels in Indonesian Districts, 2003-2008

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Model Number</th>
<th>Weight Matrix</th>
<th>Random-effects model using MLE</th>
<th>Spatial Panel Arellano-Bond Linear Dynamic Regression (Durbin Model)¹</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged GRDP per capita</td>
<td>0.987*** 1.5903*** 1.6521***</td>
<td>(0.000) (0.000) (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unused area (percent of total area)</td>
<td>-0.0001 -0.00007***</td>
<td>(0.587) (0.003) (0.134)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of unused area</td>
<td>-0.000002***</td>
<td>(0.000) (0.283)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>-0.0004 -0.4806*** -0.8278***</td>
<td>(0.815) (0.000) (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of population</td>
<td>-0.000004 2.0802***</td>
<td>(0.745) (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working-age population as percent of population (aged 15-64 years)</td>
<td>0.0006*** 0.0008*** 0.002*</td>
<td>(0.000) (0.003) (0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of working-age population</td>
<td>0.0000007***</td>
<td>(0.000) (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average years of schooling</td>
<td>0.003 0.018** 0.011**</td>
<td>(0.910) (0.011) (0.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of average years of schooling</td>
<td>0.00002*** 0.1476**</td>
<td>(0.000) (0.028)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXPY</td>
<td>0.013*** 0.001* 0.0004*</td>
<td>(0.006) (0.014) (0.038)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of EXPY</td>
<td>0.000009 0.0061*</td>
<td>(0.671) (0.061)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital (as a percent of manufacturing value-added)</td>
<td>-0.00000001 0.000004 0.0003</td>
<td>(0.246) (0.866) (0.572)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of capital</td>
<td>0.0008*** 0.001*</td>
<td>(0.000) (0.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electricity reliability (percent firm self-generated)</td>
<td>-0.00004 -0.00007 0.0001</td>
<td>(0.162) (0.219) (0.32)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of electricity reliability</td>
<td>-0.0000003 -0.0011***</td>
<td>(0.021) (0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electricity used per value added</td>
<td>0.009 0.005 -0.0038</td>
<td>(0.846) (0.114) (0.360)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of electricity per value added</td>
<td>0.004*** -0.1682***</td>
<td>(0.000) (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roads (percent of villages with access to an asphalt or stone road)</td>
<td>0.0003** 0.0008*** -0.0038***</td>
<td>(0.015) (0.000) (0.042)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spillover effects of Roads</td>
<td>0.000004*** 0.0096***</td>
<td>(0.000) (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.194*** 4.886***</td>
<td>(0.009) (0.189)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1788</td>
<td>1788</td>
<td>1788</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross Sections</td>
<td>298</td>
<td>298</td>
<td>298</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw Moments R-squared</td>
<td>N/A</td>
<td>0.9994</td>
<td>0.9994</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM Lag Anselin Test</td>
<td>N/A</td>
<td>0.6969</td>
<td>0.6287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LM Lag Test (robust)</td>
<td>N/A</td>
<td>0.7632</td>
<td>0.8004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wald Test p-value</td>
<td>N/A</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F Test p-value</td>
<td>N/A</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>2230.3639</td>
<td>2495.3657</td>
<td>2642.7134</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio Test p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sargan Overidentification Test</td>
<td>N/A</td>
<td>304.55</td>
<td>263.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sargan Overidentification Test p-value</td>
<td>N/A</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹All variables (dependent and independent) are log-transformed using natural logarithms, except those given in percent terms. These are: Area (given as percent of total area that is unused), Capital (given as a percent of manufacturing value-added), Electricity (given as the self-generated electricity as a percent of the total consumed by manufacturing firms), and Roads (given as the)

²Using Stata function spregdpd

p-values in parentheses (** p<0.01, * p<0.05, * p<0.1)
APPENDIX A. Summary Statistics for Model Variables
### Table A.1. Summary Statistics for Model Variables, 2008

<table>
<thead>
<tr>
<th>Variable Description</th>
<th>N</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual GRDP per capita growth (gy)</td>
<td>296</td>
<td>3.19</td>
<td>0.04</td>
<td>2.90</td>
<td>3.49</td>
</tr>
<tr>
<td>GRDP per capita at the beginning of each growth period (t=0)</td>
<td>296</td>
<td>1.77</td>
<td>0.67</td>
<td>0.52</td>
<td>4.66</td>
</tr>
<tr>
<td>Spillover effects of GRDP per capita (t=0)</td>
<td>296</td>
<td>1.08</td>
<td>0.70</td>
<td>0.19</td>
<td>6.57</td>
</tr>
<tr>
<td>Spillover effects of GRDP per capita, for urban districts (interaction)</td>
<td>296</td>
<td>0.47</td>
<td>0.86</td>
<td>0.00</td>
<td>6.57</td>
</tr>
<tr>
<td>Spillover effects of GRDP per capita growth (gy)</td>
<td>296</td>
<td>0.26</td>
<td>0.45</td>
<td>-1.64</td>
<td>5.19</td>
</tr>
<tr>
<td>Unused area (percent of total area)</td>
<td>297</td>
<td>9.11</td>
<td>11.93</td>
<td>0.00</td>
<td>84.86</td>
</tr>
<tr>
<td>Spillover effects of unused area</td>
<td>298</td>
<td>3.37</td>
<td>0.70</td>
<td>1.36</td>
<td>6.16</td>
</tr>
<tr>
<td>Percent of population of working age (aged 15-64 years)</td>
<td>298</td>
<td>64.55</td>
<td>3.70</td>
<td>51.79</td>
<td>74.20</td>
</tr>
<tr>
<td>Spillover effects of working-age population</td>
<td>298</td>
<td>38.99</td>
<td>21.46</td>
<td>7.05</td>
<td>142.23</td>
</tr>
<tr>
<td>Population</td>
<td>296</td>
<td>13.21</td>
<td>0.86</td>
<td>9.84</td>
<td>15.53</td>
</tr>
<tr>
<td>Spillover effects of population</td>
<td>298</td>
<td>14.08</td>
<td>13.77</td>
<td>12.10</td>
<td>15.74</td>
</tr>
<tr>
<td>Average years of schooling</td>
<td>298</td>
<td>2.32</td>
<td>0.46</td>
<td>1.74</td>
<td>3.81</td>
</tr>
<tr>
<td>Spillover effects of average years of schooling</td>
<td>298</td>
<td>1.68</td>
<td>0.44</td>
<td>0.31</td>
<td>3.01</td>
</tr>
<tr>
<td>EXPY</td>
<td>288</td>
<td>16.37</td>
<td>0.28</td>
<td>15.75</td>
<td>17.43</td>
</tr>
<tr>
<td>Spillover effects of EXPY</td>
<td>288</td>
<td>18.17</td>
<td>1.24</td>
<td>12.18</td>
<td>19.65</td>
</tr>
<tr>
<td>Capital (as a percent of manufacturing value-added)</td>
<td>280</td>
<td>117.38</td>
<td>329.41</td>
<td>0.03</td>
<td>3589.14</td>
</tr>
<tr>
<td>Spillover effects of capital</td>
<td>280</td>
<td>4.09</td>
<td>0.56</td>
<td>2.37</td>
<td>5.33</td>
</tr>
<tr>
<td>Electricity reliability (percent firm self-generated)</td>
<td>276</td>
<td>33.36</td>
<td>44.34</td>
<td>0.00</td>
<td>299.74</td>
</tr>
<tr>
<td>Spillover effects of electricity reliability</td>
<td>276</td>
<td>14.43</td>
<td>4.12</td>
<td>3.05</td>
<td>43.86</td>
</tr>
<tr>
<td>Electricity use (kW-hour per value added)</td>
<td>276</td>
<td>0.08</td>
<td>0.11</td>
<td>0.00</td>
<td>1.10</td>
</tr>
<tr>
<td>Spillover effects of electricity use</td>
<td>276</td>
<td>0.04</td>
<td>0.02</td>
<td>0.01</td>
<td>0.13</td>
</tr>
<tr>
<td>Roads (percent of villages with access to an asphalt or stone road)</td>
<td>297</td>
<td>87.29</td>
<td>18.65</td>
<td>4.27</td>
<td>100.00</td>
</tr>
<tr>
<td>Spillover effects of Roads</td>
<td>297</td>
<td>0.15</td>
<td>0.61</td>
<td>0.00</td>
<td>3.97</td>
</tr>
</tbody>
</table>

All variables are log-transformed using natural logarithms, except those given in percent terms or as a ratio. These are: Area (given as percent of total area that is unused), Capital (given as a percent of manufacturing value-added), Electricity reliability (given as the self-generated electricity as a percent of the total consumed by manufacturing firms), Electricity consumed (given as kW-hours per value added in manufacturing) and Roads (given as the percent of villages with access to a sealed road).
References


Moreno, Rosina, and Enrique López-Bazo. 2003. The Impact of Infrastructure on Regional Economic Growth: Some Results on its Spillover Effect. In Grup d’Anàlisi Quantitativa Regional (mimeo).


Figure Captions

Figure 1. Local Moran's I, Computed using 2008 GRDP *per capita*, for Western Indonesia

Figure 2. Local Moran's I, Computed using 2008 GRDP *per capita*, for Eastern Indonesia