

THE ECONOMIC POTENTIAL VIA THE GRAVITY MODEL OF TRADE

by

N. L. Beronilla¹, P. J. H. Esguerra², and J. Ocampo³

ABSTRACT

Trade, as a gauge of economic activity, can be used as an indicator to identify the relative economic potential among the areas. At the policy level this may aid in the allocation of resources and in designing incentives. For example, if the goal is to create economic opportunities away from the traditional growth centers, trade estimates could identify towns with high relative economic potential among the poor towns in which government resources could be poured. On the other hand, if the goal is to shift investments to less developed areas, the tax incentives could be offered to areas with low trade estimates.

Given the potential of its uses, trade estimate is available only at regional disaggregation, even though planning mostly happens down to the municipal level. This research fills this gap by estimating trade at town or city level via the Gravity Model of Trade. Normally it is estimated using Ordinary Least Regression but this research uses Poisson Regression which is better at handling zero trade values without transformation. The research results have been applied by couple of government agencies in their respective programs.

Key Words: *Gravity Model of Trade, Poisson Regression, economic potential for growth*

¹ Consultant of the National Anti-Poverty Commission
Email Address: nlberonilla@gmail.com

² Undersecretary of the National Anti-Poverty Commission;
Email Address: usec.napc@gmail.com

³ Former Staff of the National Anti-Poverty Poverty Commission

Introduction:

The research effort to identify towns with high economic potential arose from the policy question posed by the President's in early 2012.⁴ Given that poor towns are already identified using Small Area Statistics estimates, the next logical question is which among these towns have the potential for economic growth. Determining the economic potential is important as this may aid in allocating resources and policy incentives. For example, if the goal is to create economic opportunities away from the traditional growth centers, resources could be poured to poor towns with high relative economic potential. On the other hand, if the goal is to shift investments to less developed areas, the tax incentives could be offered to areas with low economic potential.

Despite its importance, the research on economic potential for the Philippine towns is patchy. One study is the "Behind the Veil of Conflict", which covers the towns of Mindanao (Judd and Esguerra, 2010). The paper estimated a proximity indicator of a particular town which is a function of output of its trading town (e.g., GDP) and the distance between them. The indicator in the paper is similar in functional form used by Battersby (2006) and Yoshida and Deichmann (2009), which were originally intended to estimate the flow of labor (labor migration).

The proximity indicator is directly proportional to town's GDP of the trading partner town (the growth center) and inversely proportional to distance between the town and the growth center. The distance is assumed to have an exponential decay which is assumed by Judd and Esguerra (2010) to be 5, equivalent to the number of the identified growth centers. Assuming ₱ 1,320 mn as an average town GDP, the simulation shows that at exponential decay of 5, the proximity indicator would be close to zero beyond 4 kilometers.

Another method to estimate the economic potential is the gravity model of trade (GMT) which was originated by Tinbergen in 1962, as cited by several authors (e.g., Linders and de Groot, 2006). Recently, it was used to simulate the level of development around the world from Africa (Buys et al, 2006; Coulibaly and Fotagné, 2004)), and Latin America (Carillo and Li (2002)).

GMT uses the variables found in the proximity indicator plus other variable like infrastructure level. In addition, the exponential decay factor is not assumed but estimated from the data set. In GMT, trade (which is a proxy of economic potential) is assumed to be governed similarly to Newton's universal law of gravitation. With strong theoretical foundation, flexibility to accommodate other variables, and no ad hoc assumption of exponential decay, the GMT is adapted in this paper to estimate economic potential. The succeeding sections are arranged as follows: details of the gravity model of trade, result, application, summary and limitations of the study.

Details of the Gravity Model of Trade

Trade between towns is a good indicator of potential for growth, but unfortunately, estimate of trade flow is available only at the regional level. Fortunately, the gravity model of trade could estimate the closed-form formula using the regional data and applied it at the municipal level. This is possible because the predictor variables used in the model are both found at the regional and at the municipal levels.

As mentioned before, the gravity model of trade bears a strong similarity to Newton's

⁴ Personal communication with Usec P. J. Esguerra

formula of gravitation. In this model, the two trading areas could be viewed as celestial objects and the value of trade could be viewed as the gravitational pull. The trade (~gravitational pull) is dependent on the GDPs (~mass) of the two trading areas, and their physical distance. The bigger the GDP (~mass) between the two trading areas (~celestial objects) the greater is the trade (~gravitational pull). The trade between the two areas decays exponentially as distance increases (~ decrease in gravitational pull by the square of distance). The similarities end there as GMT can take other variables like infrastructure (~sources of friction). Formally, the gravity model of trade is specified below:

$$(1) \quad T_{ij} = K \frac{M_j^a X_i^b R_j^c R_i^d E_j^e E_i^f U_j^g U_i^h W_j^k W_i^l}{D_{ij}^m}$$

T_{ij} = trade value, imported by municipality/town/city⁵ j from i (in mn ₪)

K = constant (~ gravitational constant, a parameter/coefficient)

M = GDP of municipality j (importer) (in mn ₪)

X = GDP of municipality i (exporter) (in mn ₪)

R = Road quality

E = % Household w/ electricity

U = Unemployment rate

W = % Household w/ access to water

D = Distance in km

a, b, c, d, e, f, g, h, k, l, m = parameters/coefficients

j = subscript for the importing municipality

i = subscript for the exporting municipality

Another difference is that the decay factor of distance for Newton's formula is 2, while in GMT the decay factor of distance “ m ” could vary depending on the data. The parameter “ m ” and other parameters could be interpreted as elasticities of trade with respect to the predictor variables. For example the parameter “ m ” is the distance elasticity which measures the percentage change in trade value caused by a percent change in distance.

There are many regression methods to estimate the parameters and constant of the GMT model. The simplest method is the Ordinary Least Squares (OLS) which can be applied after taking the natural logarithm of equation 1 (Buys et al, 2006). In most trade data sets, however, some trading areas have zero trades which could become undefined if log transformation is applied. This problem could be remedied if zero values are replaced by 0.1 to make log transformation possible. However, Santos Silva and Tenreyro (2006) found that replacing zeroes with 0.1 makes the OLS estimates of parameters biased. The same authors, suggested to use another method called Poisson regression.

There are two estimators of Poisson Regression, namely, Poisson pseudo maximum likelihood (PPML) and other is Quasi-Poisson. According to Root (2011), the later is the most appropriate if the zero values are frequent (say more than half of the trading areas), although both estimators yield the same parameter estimates. Quasi-Poisson gives a larger standard error to correct the bias emanating from frequent zero values which makes it likely to reject predictor variables that are significant under PPML.

Another set of authors, Linders and de Groot (2006) and Gómez Herrera (2010) suggested using Heckman sample selection. Gómez Herrera (2010) went further by trying all the estimation methodologies and found that the Heckman sample selection method is the best, although, Linders and de Groot (2006) conceded that it is difficult to find an identification restriction, a pre-requisite in the Heckman sample selection.

⁵ Used interchangeably in the remaining of this paper.

From the survey of regression methods for the GMT, the most appropriate for the purpose of estimating economic potential is Poisson regression. Given that zero trades in the regional data set are less than half, the estimator to be used is PPML. In Poisson regression, the zero values are handled by the model without the need for transformation, and the parameters of the model are estimated in multiplicative form which is shown below:

$$(2) \quad T_{ij} = \exp(K + aM_j + bX_i + cR_j + dR_i + eE_j + fE_i + gU_j + hU_i + kW_j + lW_i - mD_{ij})$$

Estimating the GMT Parameters

With appropriate regression method and the estimator selected, the next step is to estimate the parameters of the GMT model. The data set needed is at regional level with three outlier data points removed, namely trade within NCR, trade within Region 6 and trade between Caraga to Region 7. Their summary statistics are shown below.

Table 1. Summary Statistics of Regional Data Set

	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
Trade (₱ mn)	0.0	218.3	2214.7	6300.4	6192.6	33158.2
Distance (km)	0.0	186.0	566.0	691.9	1176.0	1582.0
Regional GDP (₱ bn)	65.7	162.0	215.1	451.7	518.3	2813.8
% Paved Road	47%	66%	78%	77%	88%	100%
Unemployment	3%	5%	6%	6%	8%	11%
% HH w/ Water	8%	24%	33%	32%	37%	62%
% HH w/ Electricity	56%	78%	82%	82%	87%	99%

The data sets above are collated from different sources. For example, inter-regional trade is available from the National Statistics Office website, while the regional GDPs are obtained from the National Statistics and Coordination Board (NSCB). The raw data on road quality is coming from the Department of Public Works and Highways. Unemployment is taken from Family Income and Expenditure Survey 2009. The number of households with electricity and access to water are taken from the National Household Targeting System (NHTS). The road distance between regions is measured between regional centers and obtained using Google Earth.⁶

Using the regional data sets, the GMT model is estimated using an open-source called R base package, with parameters shown below.

Table 2. Variables and Parameter Estimate

	Estimate	Std. Error	z value	Pr(> z)	Significance
Constant (~Gravitational Constant)	4.9590 (K)	0.0138	358.5300	0.0000	***
Exporter Area GDP [X]	0.0006 (a)	0.0000	316.0000	0.0000	***
Importer Area GDP [M]	0.0005 (b)	0.0000	178.9200	0.0000	***
Distance (km) [D]	-0.0003 (m)	0.0000	-92.4800	0.0000	***
% share of paved road [R] (importer)	0.6493 (c)	0.0182	35.7800	0.0000	***
Households w/ electricity [E] (exporter)	0.0000 (f)	0.0000	143.8200	0.0000	***
Unemployment [U] (exporter)	6.0830 (h)	0.1298	46.8700	0.0000	***

⁶ Extraction was automated using a script.

The variable that is not significant like % household with access to water is remove in the estimation of the final model. All the signs of parameters are positive, except for distance “c” which is negative to indicate that trade between the municipality and growth center decreases with distance.

Estimating Economic Potential

With the parameters of the gravity model of trade estimated, the next step is to estimate the economic potential. This is done by plugging the municipal level variables that are identified as statistically significant as shown in Table 2. The final form of the GMT model is in exponential form specified below,

$$(3) \quad T_{ij} = \exp^{K + aX + bM - mD + cR_j + fE_i + hU_i}$$

In the regional data set, a region may trade to all other regions. This could be mirrored at the municipal level. However, this may lead to complexity as extraction of distance among 1,600 municipalities could be very tedious. To manage the computational complexity, a particular municipality is assumed to trade only to three closest regional centers or other growth centers.⁷ Mathematically, this could be summarized below,

$$(4) \quad \sum_{j=1}^3 T_{ij}$$

Both the import and the export values of municipal j from i are computed using the equation 4 and its average is the economic potential indicator.

Some of the municipal datasets to be plugged in equation 3 are readily available as primary data like proportion of household with electricity, and access to water. However, some datasets like municipal GDP and unemployment are not available; hence, they must be estimated from secondary data sources. To address this constraint, the method adopted is the Structure Preserving Estimation (SPREE) technique pioneered by Purcell and Kish (1979) which is also used by Judd and Esguerra (2010).

The SPREE uses ratio and proportions to estimate the missing variables at lower level. For example, to get the municipal GDP, the regional GDP is first divided into provincial outputs using the number of establishments as weights, and finally, the provincial output is divided into municipal outputs based on share of municipal population. To get the unemployment at the municipal level, the share of rural and urban unemployment at the provincial level are multiplied to urban and rural population to get the count. The estimated count of the unemployed at the municipal level is then expressed as share of unemployment. The description of the data sets and their origin are summarized in the table below.

Table 3. Municipal Level Variables and Their Sources

Variables	Level	Sources	Indirect	Method	Allocation Variables
GDP	Municipal	NSCB*,NSO**	Yes	SPREE ****	# of Establishment, Population

⁷ These growth centers are 60 towns or cities that both appear in the NEDA list of 101 highly urbanized areas based on 125 thousand population cut-off; and on the NAPC list of 99 high growth towns based on estimated municipal GDPs.

Distance	Municipal	Google Maps	No	NA	
% share of road	Municipal	DPWH***	No	NA	
HH w/ electricity	Municipal	NHTS****	No	NA	
Unemployment	Municipal	NSCB,NSO	Yes	SPREE	Rural & Urban Population

* NSCB - National Statistics Coordination Board

** NSO – National Statistics Office

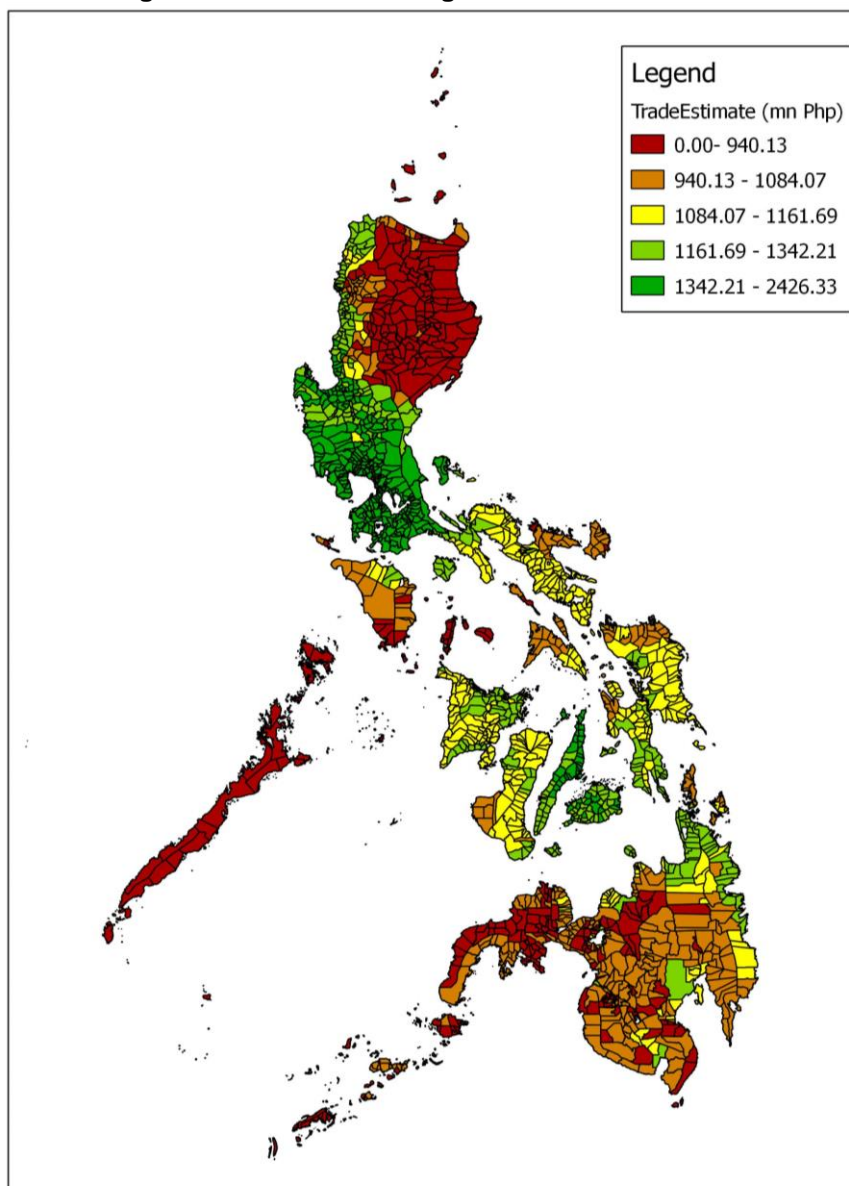
*** DPWH – Department of Public Works and Highways

**** National Household Targeting System

***** SPREE - Structure Preserving Estimation: the allocation variables are used to distribute the values at the lower levels of disaggregation. For example, regional GDP is distributed to province based on the number of establishment, and to municipalities based on population. For detailed information on SPREE, see Purcell, N.J. and L. Kish (1979)

On road quality, the lowest disaggregation is at the engineering district level. With this constraint, it is assumed that municipal level road quality is the same at engineering district level. On the other hand, the road distance from the municipal to the growth center is taken from Google earth, but in cases where road is not existent, straight line is used. Given that there is no road, the road quality used is the minimum across the country.

Figure 1. Estimated Average or Economic Potential



Results and Application

The result of economic potential is mapped using heat map color scheme in Figure 1 with QGIS software. The high economic potential areas are shown in green while low potentials are shown red. The result confirms the common knowledge that the National Capital Region (together with its neighboring provinces) and Cebu are areas with high economic potential. The map also shows the areas with low trades which are mostly located at the peripheries like Cagayan Valley, and central to western portion of Mindanao.

The economic potential indicators were used in the two policy applications, and in both instances in conjunction with other variables. For example, in the 2013 Bottom Up Budget planning in which the goal is to shift the growth to poor areas, economic potential is used in conjunction with poverty count. This was done by selecting the top 650 towns with high economic potential (or high trade) on one side, another top 650 poor towns with high poor count on the other side. The towns and cities that appear in both interim listings are considered poor town or city with high economic potential. To ensure actual economic potential, the resulting list was further trimmed down to town/city with in-migration above 1,000 as in-migration below 1,000 indicates minimal economic activity.

In another policy application, in which the goal of the Board of Investments is to grant tax incentives to new businesses that operate in less developed areas, the economic potential was used in conjunction with Municipal Income Class classification.⁸ Similar to the method above, two interim lists were created, one is the list of less developed areas as proxied by low economic potential (i.e., less developed areas), and the other is the list of low income class towns (4th to 6th class). The towns that both appear in these lists were the ones granted with tax incentives.

Summary and Conclusion

This research is able to fill the gap of the policy question, which among the towns/municipalities have the potential for economic growth. It uses the gravity model of trade which is estimated using Poisson regression and calibrated using regional data sets. The economic potential estimates were computed using municipal level data sets that are either collected from the ground or generated using secondary sources of data. Aside from filling the gap, the more important thing this research has achieved is to aid and clarify the selection the municipalities based differing policy objectives, but in conjunction with other variables, like poverty count, income classification, and in-migration.

Lastly, the model could still be improved: one by using more accurate datasets especially the distance of towns without roads in the Google Maps, and the estimates of road quality; two by experimenting with other general linear models like Negative Binomial and Zero Inflated Poisson regressions.

⁸ Income Class is an indicator of revenue in a town, hence towns with lower income class have lower revenue and smaller tax base. A business that will locate in lower income class town will help broaden the town's tax base.

REFERENCES

- Battersby, B. (2006).** "Does Distance Matter? The Effect Of Geographic Isolation On Productivity Levels", OECD Economic Studies No. 42, 2006/1.
- Buys, P., Deichmann, U., and Wheeler, D. (2006).** "Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa", Development Research Group World Bank, February 2006.
- Carrillo, C. and Li, C. (2002).** "Trade Blocks and the Gravity Model: Evidence from Latin American Countries," University of Essex, Discussion Paper.
- Coulibaly, S. and Fontagné, L. 2004,** "South – South Trade: Geography Matters," CEPII Working Paper No. 2004-08, CEPII Research Center.
- Herrera E.G., (2010)** "Comparing alternative methods to estimate gravity models of bilateral trade". Department of Economic Theory, University of Granada Unpublished version
- Judd M and Esguerra J (2010).** "Behind the Veil of Conflict: Moving Towards Economic Integration for Sustained Development and Peace in Mindanao Philippines", The International Bank for Reconstruction and Development/The World Bank Washington, DC 20433, USA
- Linders, G.M. and de Groot, H. L.F. (2006)** "Estimation of the Gravity Equation in the Presence of Zero Flows" TI 2006-072/3 Tinbergen Institute Discussion Paper
- Purcell, N.J. and L. Kish (1979)** Estimation of Small Domains. Biometrics, Vol. 35
- QGIS Development Team, 2012.** QGIS Geographic Information System. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>
- Redding, S. and Venables, A.J. (2002)** "The Economics of Isolation and Distance" Nordic Journal of Political Economy, Volume 28 Pages 93-108.
- Root, E (2011)** "Poisson Regression Using count data", University of Colorado, Unpublished lecture notes on Introduction to Quantitative Methods,
- R Core Team (2013).** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Santos Silva, J.M.C. and Tenreyro, S., (2006).** "The Log of Gravity" The Review of Economics and Statistics 88, 641–658.
- Yoshida, N., and Deichmann, U. (2009)** "Measurement of Accessibility and Its Applications" Journal of Infrastructure Development
- Zeileis, A., Kleiber, C and Jackman S., (2010).** "Regression Models for Count Data in R" <http://www.cran.r-project.org/web/packages/pscl/vignettes/countreg.pdf>.

APPENDIX A. R SOFTWARE SYNTAX

```
data1 <- read.table("C:\\the location of the file\\data1.csv", header=TRUE, sep=",")
```

```
summary(run1 <- glm(trade ~ exporter + importer + distance +  
other_variables_acting_as_friction , poisson, data1))
```