SPATIAL DETERMINANTS OF IMPORT TRAFFIC DISTRIBUTION AT PORT HARCOURT (NIGERIA)

by

Soddy I. Ionyang

ABSTRACT

This paper highlights the result of a study carried out to examine the geographic determinants of import traffic distribution at the Port Harcourt Port. A substantial aspect of the study involved building a regression model to estimate import distribution from the Port. The import function was specified in log-linear form. The adequacy of the model was then tested; this involved statistical experiments to obtain the $R^{2}$, as well as t and f values. Further tests on the adequacy of the model were conducted through diagnostic exercises designed to check for multi-collinearity and heteroscedasticity in the data used. Results obtained show that road distance and manufacturing industries are significant geographic determinants of import cargo distribution at the study Port. It is therefore important to take the two variables into account in any policy or planning exercise at Port Harcourt Port.

INTRODUCTION

In the past, several studies, have interested in the task of studying Nigerian Ports (Ogunwane 1966, Nucleo 1970, Bigninski 1979 and Shenwerson et al. 1979). Inspite of such attempts very few studies have conducted to examine port traffic from a spatial perspective; that is, few studies have been engaged with examining the role of the hinterland on port traffic. This is inspite of the fact that hinterland correlates (Socio-economic variables) largely determine the volume and direction of flow of imports. Furthermore, effective planning demands that such spatial variables be identified and their significance scientifically tested; it is only within such a framework, that quantitative estimates, can be made, for purposes of policy consideration.

To the best of our knowledge, the pioneering work of Onakomaiya and Smith (1972) remains the only work which has scientifically examined import traffic distribution in Nigeria. The study utilized data for 1964 and concerned the distribution of imported Cargo by rail from two Nigerian ports (Apapa and Port

* Soddy I. Ionyang is a lecturer at Department of Geography, Nigerian Defence Academy, Kaduna, Nigeria.
Hance, this study indicated that income from agricultural exports and distance from the two ports were the two most significant variables. However, a number of points are worthy of consideration concerning this study. Firstly, the study was conducted at a time when the Nigerian economy was predominantly agricultural. Consequently, it would be expected that agricultural earnings would be significant in explaining import distribution. Besides, at the period, the level of industrialization and level of industrialization was found not to be significant. Thirdly, the use of railway distance as an importance variable in accordance with the importance of the railway at that period. However, given the immense industrial, urban and infrastructural development that has taken place in Nigeria (between the 1950's and 1990) it becomes necessary to further examine the presence of an import distribution function for Port Harcourt.

The specific objective of this paper is to examine import traffic distribution from Port Harcourt (using more recent data); particular attention is therefore focused on the statistical significance of spatial variables such as population of towns, level of industrial development as well as road distance from the study Port. Furthermore, a major distinguishing feature of this paper is the variety of distance measure to which the estimated distribution function is subjected to.

**CONCEPTUAL AND THEORETICAL ASPECTS**

The Gravity model provides, a conceptual and theoretical base for the model to be used. The gravity model has been traditionally applied to migration studies (Gardiner 1983). However, variations of the model have been applied to traffic studies (Britton 1960; Black 1971; O'Sullivan and Rashian 1974 and Openshaw 1976).

In its simplest form, the Gravity model states that interaction between any two points is a function of some attractive factor (or mass term) as well as an impedance factor. The attractive variable is usually population, while the impedance factor is distance. Consequently, the larger the mass (suitably defined) the higher the traffic; on the other hand, the larger the operating distance the lower the traffic (all things being equal).

The definition of attractive and impedance variables may however take many forms; the mass term may be defined in terms of population, income, retail sales, number of commercial activities as well as industrial development. The fractional effect of distance may be defined in terms of two or time involved in over-casting space.

However, several problems are associated with the gravity model. A major problem concerns its predictive ability: the model has been criticized for failure to replicate observed trip patterns (Wills 1986; Sikdar and Machonin 1981). Another problem has to do with the fact that the model is basically misspecified (Fotheringham 1981, 1983 and Baxter 1980); that is, it fails to take into account the spatial structure of origins and destinations.

However, many factors have increased the validity and adoption of the gravity model. For example, Wilson (1967) has proved that the doubly constrained gravity model is equivalent to entropy maximization; Foslin and Tong (1960) showed that the gravity-entropy model is a generalized form of the linear
programming transportation problem. This work therefore takes the gravity model as given, the model therefore provides a strong theoretical base for our model specification.

The conceptual model is therefore as follows:

\[ Y = f(x_{1}, x_{2}, x_{3}) \]

Where

- \( Y \) = Quantity of imports flowing from the port to inland locations
- \( X_{1} \) = Population of inland towns
- \( X_{2} \) = Number of manufacturing industries
- \( X_{3} \) = Road distance between the port and inland locations
- \( \alpha, \beta_{1}, \beta_{2} \) are coefficients to be estimated

In this work, the use of income of towns as an additional variable was dropped because of the following reason:

(i) The difficulty of obtaining adequate data on income at the urban level (that is, income generated by towns from taxation of individuals, industries and commercial activities).

(ii) Since the study is interested in a function the meets the needs of policy, the variables to be used must be obtainable on a continuous basis for the past, present and future.

(iii) The strong possibility of a high correlation between population as well as the number of manufacturing industries on the one hand and income generated on the other. This is because it is assumed the income generated at the urban level is a fraction of the number of taxable adults as well as the number of industrial and commercial enterprises. A high correlation may thus be expected thus raising the problem of multicollinearity (see Johnson (1984) pp 246-268).

We were further discouraged from using income generated from agricultural exports because of the insignificance of such exports in revenue generation.

With regard to the model earlier specified, we should expect (on theoretical grounds) the following findings:

(a) The coefficient of population \( b_{1} \) should have a positive sign given the gravity model which postulates a direct relationship between volume of cargo flow and population.

(b) The coefficient of number of manufacturing industries is also expected to have a positive sign.

(c) The parameter \( b_{3} \) (coefficient of distance) is expected to have a negative sign, this is because theoretically an inverse relationship is expected between quantity of flow and distance.
METHODOLOGY

Methods

Spatial distribution models are usually classified into two groups; these are the multiplicative (or logarithmic) and linear models. The multiplicative models are non-linear in nature and the parameters of such models are usually estimated using maximum likelihood methods (Batty and Mackie 1972; Wilson 1970 and Mackett 1979).

On the other hand linear models are usually estimated using least squares method (Ceseriter 1974, 1975). In practice the two methods are found to be sufficient and yield similar results (Gopenshaw 1976). This study therefore uses the least squares method in estimating an import distribution function for Port Harcourt.

Multiple regression is therefore utilized in this study; the specific functional form used is the log-linear function (This function is used to take into account possible non-linearities in the data).

The log-linear model is specified as follows:

\[ \ln y = b_0 + b_1 \ln x_1 + b_2 \ln x_2 + b_3 \ln x_3 + U \]

Where the b’s are coefficients to be estimated

\( U = \) error term

(other variable in the function have been earlier defined)

The reliability and adequacy of the regression results are tested at various stages. The first stage relates to the theoretical postulates about the sign of the parameters. The second level makes use of first order statistical test; such test involve the use of the R square, f and t values to assess the reliability of the regression results. Further diagnostic analysis involved testing for multicollinearity and heteroscedasticity in the data used in this study.

Multicollinearity in regression analysis usually occurs when new linear dependences occur between the independent variables (Koutsoyiannisa 1977 p.233). This problem leads to inflated sampling variances, imprecise parameter estimates and is a measure of the accuracy of significance test (Hanushek and Jackson 1977 pp.86-96; Johnston 1964 p.240). The seriousness of multicollinearity in the data used was tested for using three methods. The first is the intercorrelation between the explanatory variables; a high correlation index normally indicates the presence of multicollinearity. The second method used was the variance inflation factors (VIF). The factor is defined as

\[ I = \frac{1}{(1-R^2)} \]

where \( R^2 \) is the squared multiple correlation coefficient when an explanatory variable is regressed on the remaining variables (Thornton 1984 p.247). The VIF index quantifies the magnification of the sampling variance due to collinearity relative to the absence of collinearity (VIF = 1) (Barba 1991); on the other hand VIF’s in excess of 10 are taken as indicating severe multicollinearity (Maraquda 1990). The third method used to test for the severity of multicollinearity is a method due to Koutsoyiannisa, (see Koutsoyiannisa 1977 pp.238-241); the method is a revised version
of Frisch's 1934 conference Analysis. The method involves a basic step. The first step is to regress the dependent variable on each one of the explanatory variables separately. The results of the elementary regressions are then checked on the basis of a priori and statistical criteria. The second step involves adding additional variables and checking their effect on individual coefficients, standard errors and the R squared (Koutsoyannis 1977 p.239). The presence of multicollinearity is indicated by changes in the signs of coefficients; and a drop in standard errors as more variables are added.

A major assumption in least squares concerns the pattern of the errors; that is the error is expected to have a constant variance; or the errors should not be correlated with the explanatory variables - that is absence of heteroscedasticity. The presence of heteroscedasticity results in inefficient parameter estimates as well as invalid standard errors. There are various test for heteroscedasticity; such test include those of Breusch - Pagan (1979); Goldfeld - Quandt(1965); Gleister (1969) and the rank correlation test. In this study, the rank correlation test is used; this is because it is the simplest; besides Johnson (see Johnston 1972 p.22) seems to accept that the rank correlation test as well as the Goldfield and Quandt test are preferable to other test.

Data

The port traffic data used in this study was for the year 1980; they were obtained from the Nigerian Ports Authority. Although, more recent data are available from the Port Authority, such data are at an aggregated level. That is, current data reflect movement of goods from port to state, such Zonal data are associated with the problems especially when applied in regression analysis (Douglas and Lewis 1970). The main problem is that Zonal regression can account for only the variation between zones and not for within zone variations (White and Senior 1973 p.155). As a result, relationships between variables representing aggregated data may be inaccurate (Dalvi and Martin 1977). It is therefore better for analysis to be undertaken at the most spatially disaggregated level possible (Dalvi and Martin 1977).

The preceding discussion therefore justifies the use of 1980 data for estimating port traffic distribution. (The data used in this study are disaggregated and involve flow of goods from Port Harcourt to twenty Nigerian towns. (See appendix D). Information on population of Nigerian towns were obtained from the National Population Commission for the year 1963. The data were then projected to obtain 1980 estimates using the formulae Pn = P0e^t

Where

Pn = Future population
P0 = Present population
e = exponential
t = rate
t = time

The growth rate of Lagos for 1980 is given as 4.7%; other Nigerian towns have a growth rate 3.2% Information on industries were obtained from the Federal Ministry of Industries while road distances between Port Harcourt and inland locations were obtained from the Federal Ministry of Work and Housing, (Road

1 Using

square efficient squares external account

various of the results.

by de-
p,233, pns are .86-96; 
tested

natality t

tive to

cow of

third

due to
distances are used because it is assumed that a large proportion of Port-inland traffic movement now use the highway network rather than the railway.

THE STUDY PORT

Eastern Nigerian rail line and as an outlet for the export of coal from Enugu to Lagos and certain West African Countries. The Port is located 65.6 kilometres up the Benin River and was named after Louis Harcourt the then British Secretary of State.

Port Harcourt has witnessed extensive development since its established. The first was between 1956 and 1964, the second between 1962 and 1966. The last development took place between 1970 and 1974. The port is capable of berthing eight vessels and has seven stacking areas. The main quay covers an area of 47 hectares and is 1360 metres long. The port remains the most important port in the Eastern part of Nigeria (see Figure 1 for a map of the Port).

EMPIRICAL RESULTS

Table 1 shows the regression results; equation number 1 in the table relates to the regression results for all parameters. The R squared for the equation is 65.3% and the regression is statistically significant as indicated by the F ratio. Furthermore, all the regression parameters have the correct sign. For example, the population parameter comes out with a positive sign; however, the coefficient of the population parameter is not statistically significant. The industrial parameter also has the correct sign but is not statistically significant; while road distance is statistically significant and also has the expected sign.

Given this the R squared obtained, was far above average and that only one variable was significant, we decided to examine the importance of individual parameters. In order to do this, we estimated simple functions using each variable one at a time. The result of this exercise is shown on Table 1 from equation Number 2 to 4. Equation Number 2 relates to the population parameter. it shows that population alone explains 12% of the variation in the dependent variable; this variable was found to be insignificant using an F test at the 5% level with 1 and 18 degrees of freedom). Equation Number 2 relates to manufacturing industries; this variable explains 28% of the variation in the import data and was found to be statistically significant. Furthermore, road distance was found to be significant; this variable recorded an R² of 36%.

The preceding paragraph therefore shows that distance and industries are two important variables that explain the distribution of imported goods from Port Harcourt. However, given that in the complete equation (equation number 3) the industrial parameter was found to be insignificant we decided to further check the adequacy of the result obtained by carrying out diagnostic tests. Such test is important since it is not possible to obtain meaningless results despite having high R-squared and significant 1 and 6 ratios. Besides a violation of major assumptions governing the application of least squares could lead to wrong conclusions concerning the variables. For example, we could reject a variable as being insignificant while it is indeed very significant.
TABLE 1 Regression Result

\[ \ln Y = \ln b_0 + b_1 \ln x_1 + b_2 \ln x_2 + b_3 \ln x_3 + \ln U \]

<table>
<thead>
<tr>
<th>Equation number</th>
<th>Constant</th>
<th>Population</th>
<th>Industries</th>
<th>Real estate</th>
<th>R Squares</th>
<th>Adjusted R Squares</th>
<th>R value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.708</td>
<td>1.544</td>
<td>-0.361</td>
<td>-1.4552</td>
<td>(2.1)</td>
<td>(1.88)</td>
<td>(2.55)</td>
</tr>
<tr>
<td>VF</td>
<td>2.1031</td>
<td>2.1185</td>
<td>1.5267</td>
<td>-</td>
<td>65.3%</td>
<td>58.8%</td>
<td>10.046</td>
</tr>
</tbody>
</table>

**DIAGNOSTIC TEST**

Table 2 shows the correlation matrix between the explanatory variables. The table shows a high collinearity between population and manufacturing industries. Thus, the presence of multicollinearity in the data is confirmed. However, other tests were carried out to examine the severity of the collinearity in the data. The variance inflation factor (VIF) was used to assess the severity of the multicollinearity. Results in Table 1 indicate that the VIF for all the parameters were quite low. For instance no variable had a VIF of up to 10; consequently multicollinearity is not a serious problem. However, we note that all the variables have some collinearity since all the VIF’s are above 1.

To further examine the severity of multicollinearity in the data we adopted a method suggested by Koutsoyannakis. The result of the exercise is presented in table 3. We used distance as the first elementary regression (Y = b0) since it recorded the highest R square. The introduction of manufacturing industries significantly improves the R square; the sign of the parameters are correct and both variables are highly significant.
The regression with all three variables shows the effect of multicollinearity. That is the inclusion of population results in the industrial parameter becoming insignificant. Thus population has a detrimental effect on the value of the industrial parameter. However, it should be noted that the inclusion of population the R-square is increased and the signs of the parameters are still correct. Consequently, this method also indicates that multicollinearity is not severe, this is because the addition of collinear data did not result in a change of the signs of the parameters.

<table>
<thead>
<tr>
<th>X1</th>
<th>X2</th>
<th>X3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.773</td>
<td>1</td>
</tr>
<tr>
<td>.725</td>
<td>0.20</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Regression Results

<table>
<thead>
<tr>
<th>Equation number</th>
<th>b0</th>
<th>b1</th>
<th>b2</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.6251</td>
<td>-1.2916</td>
<td>-1.3075</td>
<td>56%</td>
</tr>
<tr>
<td>2</td>
<td>11.3137</td>
<td>-1.3010</td>
<td>-1.800</td>
<td>57.6%</td>
</tr>
<tr>
<td>3</td>
<td>1.708</td>
<td>-1.4550</td>
<td>-1.915</td>
<td>55.2%</td>
</tr>
</tbody>
</table>

Note: An unlined coefficient is significant at the 5% significance level (using a two-tailed test).
From the preceding discussion we can therefore state that road distance and manufacturing industries are important variables that explain the distribution of imports from Port Harcourt. The non-significance of the industrial parameter in the complete equation is due to the collinearity between population and manufacturing industries. The significance of manufacturing industries is not surprising given the emphasis placed on industrial development. Furthermore, the nature of the industries found in Nigeria sheds more light as to the significance of the industrial parameter; most industries in Nigerian assemble components which to large extent are imported; consequently it is expected that there should be a positive correlation between imports and manufacturing industries.

The spearman rank correlation test was used to test for heteroscedasticity in the data. The rank correlation coefficient between the absolute values of the errors and each of the explanatory variables in turn was then computed. A t test was then used to examine the significance of the rank correlation coefficient. The test used is in the form:

$$ t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}} $$

where

- $r$ = rank correlation coefficient
- $n$ = sample size
- $n-2$ = refers to the degrees of freedom used in the test

The results obtained show that the rank correlation coefficient for population against the errors was 7%. The t value obtained was 279.6. This value was compared to the tabular value at the 5% significance level with 18 degrees of freedom; the calculated t value was found to be insignificant. We therefore conclude that heteroscedasticity did not exist in the population data.

In the case of industries the rank correlation coefficient obtained was 5.1%; the computed t value was found to be insignificant when compared to the tabular value. Consequently heteroscedasticity did not constitute a problem in the industrial data. The test carried out on the data concerning road distance also indicated a lack of heteroscedasticity in the data; (The rank correlation coefficient obtained was 4.7% at t test indicated that this value was not significant).

**CONCLUDING REMARKS**

The objectives of this paper have been to provide evidence on the spatial determinants of import traffic distribution from Port Harcourt. We had therefore set out to examine the statistical significance of variables such as population, manufacturing industries and road distance in explaining the distribution of imports. The main conclusion of the paper are therefore as follows:

In the first place road distance was found to be very significant in explaining the distribution of imports. This finding is plausible given the extent of road development in Nigeria; most of the major highways have been extensively reconstructed or expanded into Expressways. This fact has helped to develop the
road haulage industry to the extent that importers prefer to pay more to have their goods delivered by road rather than by rail. Consequently, any policy on import distribution in Nigeria must take into account road distance.

Secondly, a number of manufacturing industries significantly influence the volume of imports that flow into an inland location. The importance of this variable is illustrated by the distribution of industrial land use in Nigeria. It is also worth noting that most of such industries depend almost entirely on imported raw materials. Consequently, it is necessary for policy makers to take into account the impact of industrialization on the economy when examining import distribution. Population was found to be insignificant in explaining the variation in the import data. The variable is insignificant since its interaction with the other variables is significant. The results obtained in this study, to a large extent, meet theoretical expectations; however, they appear to be reasonably reliable (given the results of the diagnostic exercises). We may therefore conclude that the methodology adopted as well as the results obtained may prove useful for planning a Port Harcourt. (See Figure 1)

REFERENCES


Cesario E.J. (1975) 'Linear and non-linear regression model of spatial interaction' Economic Geography Vol.51 No.1 pp.69-77


DaVi M.Q. and Martin K.M. 'The effects of spatial aggregation on predicting trip rate: some empirical results', working paper No.72 submitted for Transport seminar, University of Leeds, Leeds.


ext on land
of
bees,

usual of
action
squares

ing
trip
airport

49

Linear
impervious

den in
in the

anomalous

1

urban

1, New

s of

eory of

systems,

al of the

utistical

s

'nominal

linen

impervious

ods in

in the

models'
## DATA USED IN THE REGRESSION ANALYSIS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Port Harcourt</td>
<td>774,924</td>
<td>2,144,06</td>
<td>338</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Abe</td>
<td>31,632</td>
<td>2,302</td>
<td>180</td>
<td>64</td>
</tr>
<tr>
<td>3</td>
<td>Umuahia</td>
<td>972</td>
<td>1,507</td>
<td>36</td>
<td>116</td>
</tr>
<tr>
<td>4</td>
<td>Ogbota</td>
<td>244,720</td>
<td>2,000</td>
<td>60</td>
<td>206</td>
</tr>
<tr>
<td>5</td>
<td>Igboogho</td>
<td>106,857</td>
<td>2,497</td>
<td>105</td>
<td>253</td>
</tr>
<tr>
<td>6</td>
<td>Umaram</td>
<td>315,756</td>
<td>1,555</td>
<td>824</td>
<td>499</td>
</tr>
<tr>
<td>7</td>
<td>Calabar</td>
<td>61,077</td>
<td>1,553</td>
<td>47</td>
<td>198</td>
</tr>
<tr>
<td>8</td>
<td>Uyo</td>
<td>186,235</td>
<td>2,998</td>
<td>40</td>
<td>122</td>
</tr>
<tr>
<td>9</td>
<td>Boki</td>
<td>241</td>
<td>1,537</td>
<td>12</td>
<td>1029</td>
</tr>
<tr>
<td>10</td>
<td>Akwa</td>
<td>326</td>
<td>2,645</td>
<td>56</td>
<td>1307</td>
</tr>
<tr>
<td>11</td>
<td>Zark</td>
<td>62</td>
<td>2,058</td>
<td>19</td>
<td>1902</td>
</tr>
<tr>
<td>12</td>
<td>Kere</td>
<td>2361</td>
<td>5,251</td>
<td>215</td>
<td>1274</td>
</tr>
<tr>
<td>13</td>
<td>Bera</td>
<td>1603</td>
<td>1,751</td>
<td>150</td>
<td>355</td>
</tr>
<tr>
<td>14</td>
<td>Wani</td>
<td>1727</td>
<td>982</td>
<td>51</td>
<td>383</td>
</tr>
<tr>
<td>15</td>
<td>Akwa</td>
<td>155</td>
<td>462</td>
<td>6</td>
<td>217</td>
</tr>
<tr>
<td>16</td>
<td>Abe</td>
<td>6307</td>
<td>16,280</td>
<td>163</td>
<td>846</td>
</tr>
<tr>
<td>17</td>
<td>Akwa</td>
<td>1,502</td>
<td>960</td>
<td>12</td>
<td>907</td>
</tr>
<tr>
<td>18</td>
<td>Gbas</td>
<td>4059</td>
<td>11,166</td>
<td>125</td>
<td>658</td>
</tr>
<tr>
<td>19</td>
<td>Bokita</td>
<td>1,685</td>
<td>3,710</td>
<td>15</td>
<td>846</td>
</tr>
<tr>
<td>20</td>
<td>Calabar</td>
<td>5,005</td>
<td>18,710</td>
<td>3</td>
<td>463</td>
</tr>
</tbody>
</table>

Sources: 1. Nigerian Census Authority
2. National Population Commission
3. Federal Ministry of Industry
4. Federal Ministry of Works and Housing