

# Explaining the Phenomenon of Third World Urban Giants: The Effects of Trade Costs

Rasha Gustavsson  
*Lund University*

## Abstract

*Traditional push-and-pull factors offered partial explanations to the size of large urban areas in the third world. Moreover, the growing literature in economic geography identifies an additional factor exacerbating the phenomenon, namely trade costs. The present study tests econometrically the proposed hypothesis, whether higher trade barriers intensify the concentration forces in third world urban centers. That is, whether trade costs interact with the traditional push-and-pull effects and exacerbate the concentration effects. In particular, controlling for the traditional push-and-pull effects we incorporate two additional variables in order to capture trade costs within and across borders, respectively. The results suggest mixed support for the traditional push-and-pull effects. Higher trade costs within borders seem to reduce the concentration forces in urban centers. In addition, we find no evidence that trade barriers across boarders encourage concentration. Finally, the evidence from sensitivity analysis suggests that the results need to be treated with caution. (JEL-Classifications: F12, F13, F15, R12) <Key Words: urban giants, metropolises, trade barriers, concentration, agglomeration.>*

---

\* Correspondence Address: Department of Economics, Lund University, Box 7082, S-220 07 Lund, Sweden; (Email) Rasha.Torstensson@NEK.LU.SE.

## I. Introduction

An important issue that has received renewed attention in the economic debate is the explanation of the existence of third world *metropolises*<sup>1</sup>. Conventional development theories have focused on push-and-pull factors that certainly give partial explanations to the phenomenon. These factors include, among other things, wage differentials as pull factors to the metropolis, and population density in the countryside as a push factor to urban centers (this variable is meant to capture declining living conditions and working possibilities in rural areas). Although development economics can explain urbanization, a problem has been the lack of an overall theoretical model explaining both urbanization and the location of major economic centers. However, the recent development of economic geography seems to offer a framework that helps to add further insights into these issues. In short, research in this relatively new area tries to explain the location of production in geographical space and analyzes the question of why industries tend to cluster within regions, assuming that firms and people are relatively mobile across regions. As a consequence of external economies more firms are drawn to the same region. On the demand side, firms can offer higher wages in the *now* larger market and thereby attracting more workers, a process that further increases the demand for final products, thus leading to more concentration.<sup>2</sup>

The literature on spatial economics has recently identified additional variables that interact with push-and-pull factors and exacerbate the degree of concentration. These variables include scale economies, monopolistic competition and trade costs. In particular, it is suggested that transportation costs and import substitution industrialization policies (ISI), or generally protectionistic trade policies, have significantly increased the concentration effect, perhaps particularly in third world countries.

These issues are interesting for the analysis of third world metropolises

---

1. *Metropolises* are defined as very large concentrations of population, often capitals.

2. It should be noted that both the traditional literature on urbanization and the new trade and geography literature rely on the same mechanism, *i.e.* it is in both theories essentially the wage differentials that drive urbanization and agglomeration within a country. See Williamson [1988] and Krugman [1991, 1993].

for two reasons. First, trade costs in many third world nations are high by virtue of a generally less developed infrastructure, and second because of the extensive implementation of ISI-policies in those economies. In addition, it is a fact that a number of the largest cities are located in developing countries and the concentration of workers and industries is often higher in the developing compared to the developed world.<sup>3</sup> The theory of agglomeration is general in nature, and may thus be relevant to industrialized as well as third world countries. However, it may be argued that the interaction between the traditional push-and-pull effects giving rise to urbanization, together with trade costs giving rise to more agglomeration, may be particularly interesting and strong for third world cities. At any rate, in this empirical study we have chosen to concentrate on third world countries.<sup>4</sup>

Thus, based on the new theoretical insights that recent models offer in the area of economic geography, the present study considers the forces behind urbanization and agglomeration in a developing country context. Since there are very few empirical studies analyzing the relationship between the development of trade costs in general and the rise of large urban areas, more empirical investigation is necessary. Furthermore, earlier work typically focuses on the situation in Mexico, particularly because Mexico City is the world's largest urban center. Nonetheless, generalizations drawn solely from Mexico would be rather limited due to its special geographical location (sharing a common border with the US). Moreover, in another empirical study conducted by Ades and Glaeser [1994] agglomeration effects are analyzed for a sample of developed as well as developing countries, but they do not control for urbanization attributed to traditional push-and-pull effects.

Therefore, this study offers a cross-sectional analysis of both urbanization and agglomeration, using data from a number of third world countries covering the time period 1980-1990. Thus, we explore the determinants of agglomeration, considering traditional push-and-pull factors, together with

---

3. For example, more than half of the manufacturing value-added in Mexico were located in Mexico City in 1980 (Krugman and Elizondo [1996]).

4. As for the econometric calculations, it may be argued that we then work with a somewhat more homogenous group than if we had considered the entire world.

trade costs as explanatory variables.

The study attempts to extend earlier empirical work in three basic ways. To begin with, we seek to control for push-and-pull factors, which is essential in order to separate the trade policy effect on agglomeration. To our knowledge this has not been done in earlier studies treating this particular phenomenon. Secondly, contrary to Ades and Glaeser (1994), only third world countries are included in our sample. Lastly, due to the lack of knowledge in this particular area emphasis is placed on specification testing.

The results do not clearly suggest a positive and significant relationship between agglomeration and the variables capturing the push-and-pull effects. The evidence is mixed on this point.

When we included trade costs across and within borders as additional explanatory variables, the evidence revealed that lower trade costs within borders seem to encourage agglomeration. In addition, the results do not support the hypothesis that trade barriers across borders reduce agglomeration. These findings contradict some of the ideas of recent theoretical literature.

The division of the study is as follows. Section II offers a brief theoretical background and a summary of earlier empirical findings. In section III we go into the econometric methodology and data description. In section IV our results are presented and analyzed. Section V contains some sensitivity tests of the derived results, and finally section VI offers a few concluding remarks.

## II. Theoretical Background and Previous Studies

While conventional explanations of city concentration in developing countries stress push-and-pull factors to the urban areas, the economic geography literature uses the terms centripetal and centrifugal forces (Krugman, 1991). Centrifugal forces, *i.e.* forces causing individuals to move toward the center, are synonymous with both the push and pull factors toward urban centers. However, centripetal forces, *i.e.* forces causing individuals to move away from centers, refer to the formation of urban satellites (sub-urbs) an issue that is beyond the focus of this study.

A first attempt to provide a general theoretical model of urbanization with

the basic features of the new economic geography was provided by Krugman [1993]. In particular, he adds the spatial dimension to a monopolistic competitive general equilibrium model in order to explain the size and location of cities. In addition to agglomeration effects due to forward and backward linkages between firms as explanations of why cities grow, the spatial dimension might enable us to explain also the actual geographical location of cities. In Krugman's setting, the formation of cities is driven by economies of scale at the firm level in accordance with cumulative causation<sup>5</sup>, whereas the location of cities follows a central place pattern.

For the purpose of our study, it should be noted that the degree of openness to trade affects city formation in these types of models. Thus metropolises, defined as large concentrations of economic activities, can partly be explained endogenously by agglomeration forces driven by the interaction of transportation costs and scale economies. Furthermore, Krugman and Elizondo [1996] consider the interaction between trade policy and forward-backward linkages between firms and suggest that import-substitution industrialization policies tend to contribute to the expansion of third world cities.

A general observation pertaining to the trade and geography models is the fact that they generally fall into two basic categories, those that implicitly model intermediate production (or so called input-output linkages), *e.g.* Krugman [1991] and Krugman and Elizondo [1996], and those that model the intermediate production explicitly, *e.g.* Krugman and Venebles [1996]. However, the mechanism relating trade costs to agglomeration is basically the same. Thus, this feature does not have a bearing on this particular study.<sup>6</sup> In addition, it ought to be noted that these models are by no means general, but build on quite specific assumptions.

The general intuition in Krugman and Elizondo [1996] goes as follows. In an economy that is relatively closed, firms must rely on the domestic market for inputs. This leads to strong agglomeration forces where firms want to locate close to each other. The tendency will be further exaggerated if

---

5. In other words, at the same time as production tends to locate close to large markets, markets become larger where production locates.

6. Bear in mind that the present study is entirely empirical.

transport costs are high between regions.<sup>7</sup> As a consequence, the concentration of industries will attract workers since higher wages can be offered. With trade liberalization, however, firms will be able to engage in international trade: importing inputs and exporting final goods. Thus, the linkages between domestic firms become weaker and the advantage of geographical closeness becomes less important. In other words, closed markets may tend to huge metropolises, whereas open markets may discourage agglomeration. Nevertheless, it ought to be noted that the relationship between so-called “iceberg” trade barriers (assumed in all the models described above) and agglomeration is not monotonic in all models. We will return to this point later on when we discuss our results, where it seems to be of considerable relevance.

These new theoretical ideas seem, according to some studies, to find empirical support. Hanson [1997] examines the structure of relative regional wages in Mexico before and after trade liberalization and finds that while industry was heavily concentrated in Mexico City before trade liberalization took place, it reallocated to some extent close to the US border afterwards. Moreover, in areas far from the industrial center and the US border, nominal wages tended to be lower.<sup>8</sup>

A related empirical study is provided by Ales and Glaeser [1994]. Using cross-sectional analysis, including both developed and developing countries, they identify a few factors that influence the concentration of population in certain urban centers. In particular, high tariffs, or more generally high trade costs, are found positively related to the size of the metropolis. Conversely, high openness measured as share of trade in GDP had a negative effect on average agglomeration. In addition, they found that the degree of concentration is declining in the size of the domestic population and (rather naturally) in the share of population employed in agriculture.

---

7. This story is especially plausible for developing countries where there is insufficient infrastructure.

8. The case of Mexico city is, however, special due to the fact that Mexico shares a common border with the US.

### III. The Econometric Analysis

The basic hypothesis we will be testing is that higher levels of trade costs, *ceteris paribus*, lead to increased concentration in third world metropolises. These trade costs fall into two categories, within country trade costs here expressed by transportation costs, and across country trade costs, which may depend on the degree of protectionistic trade policy.

This section offers an empirical study of urban agglomeration, as measured by the relative size of the metropolis, employing a cross-sectional analysis including only third world countries. Thus contrary to Ades and Glaeser [1994], we do not include developed countries, but concentrate on what we might consider to be a third world phenomenon. However, the theoretical arguments are, as mentioned, not necessarily restricted to third world countries, but may be valid also for developed countries.<sup>9</sup> We study the effects trade costs may have on concentration in the largest city. Since we adopt a general framework, we need, as suggested, to control for a few factors that are identified in the literature both in earlier studies of urbanization and migration and in newer economic geography models.

#### A. Methodology and Data Description

Our base function is  $CONC = f(AGR DENSITY, INFRA, TRADE)$ . However, since  $CONC$  lies in the range between 0 and 1, we assume that the logarithm of  $CONC = CONC^*$  and not  $CONC$  itself is a linear function of the explanatory variables:

$$CONC^* = \alpha + \beta_1 AGR DENSITY + \beta_2 INFRA + \beta_{3i} TRADE + \varepsilon_j \quad (1)$$

$i = 1, 2, 3$ , *i.e.* we use three measures for trade policy defined below.

$j = 1, 2, 3, 4$ , *i.e.* we have four residuals resulting from running four regressions.

$CONC$  is defined as average urban agglomeration over the period 1980-

---

9. It can, though, be argued that looking at the 1980's, the traditional push-and-pull effects may be more relevant for third world countries.

1990. It is measured as the ratio between population in the country's largest city and total population.<sup>10</sup> (Source: Prospects of World Urbanization [1998] and Summers & Heston [1994]).<sup>11</sup> *AGR DENSITY*, agricultural density, expresses the number of people per squared kilometer in agricultural areas, which we take as a proxy for the living conditions in the countryside (Source: Social Indicators of Development [1993]).<sup>12</sup> *INFRA* is a proxy for how well the infrastructure is developed it measures road density in kilometers per million persons (Source: World Development Report [1995]). Finally, *TRADE* is a proxy for trade policy. We employ three different measures for trade policy: *TARIFF*, *NTB* (non-tariff-barriers), (Source: Pritchett [1991]), and *OPEN* (openness) which measures trade volumes as a share of GDP, a higher figure is supposed, *ceteris paribus*, to imply lower trade barriers<sup>13</sup> (Source: and Summers & Heston [1994]). Figures on *CONC* and *AGR DENSITY* are averaged over the period 1980-1990, figures on *INFRA* are from 1992, *TARIFF* and *NTB* are from 1988 and *OPEN* is averaged over 1980-1990. Finally  $\varepsilon$  is an independent stochastic error term assumed to be normally distributed with zero expected mean.

### ***B. Expected Signs***

*AGR DENSITY* is population density in the countryside. It can be interpreted as a proxy for the living conditions in rural areas and is expected to

---

10. An empirical problem with this variable is that in some cases the variable statistics take account of suburbs, yet in other cases they do not.

11. In (1), the definition of the dependent variable is:

$$CONC^* = -\ln\left(\frac{1}{CONC} - 1\right), \text{ i.e. } -\infty < CONC^* < +\infty. \text{ For simplicity, we take the negative of the natural logarithm so that } CONC^* \text{ rises with } CONC \text{ and decreases with it.}$$

12. A difficulty with this measure is connected with the degree of extensive agriculture. In case the agricultural land is mainly fit for cattle, the area may seem large in relation to the number of farm people, still there may be shortage of land and excess supply of labor.

13. Also it reflects to some extent the relative size of the country, as very large countries have typically small values of *OPEN*.

have a positive impact on agglomeration. Higher population density implies deteriorating living conditions, and thus more rural urban migration. This very same variable implicitly proxies a pull effect. Higher rural population density implies excess labor supply in the countryside that pushes rural wages downwards. This in turn widens the urban rural wage gap creating an incentive to move to the city. This latter effect (pull) works in the same direction as the former effect (push), according to the theory. The expected sign is, therefore, definitely positive.

*INFRA* this variable measures road density in kilometers per millions of persons. A higher figure thus implies a more developed infrastructure, which means less internal trade cost, and less concentration, *i.e.* according to some newer economic geography models, the expected sign of *INFRA* would be negative. Some other models or theories do not give a definite answer to this question. One might, intuitively, argue that a more developed infrastructure could tend to increase concentration, as it becomes easier for people to move to the large city. Moreover, once an economy is open to international trade, the internal trade costs might be of less importance, since firms can import inputs and export finished products. The expected sign, and for that matter the significance of *INFRA*, become, ambiguous.

*TRADE* is supposed to express protectionism. Three alternatives are attempted, *TARIFF*, *NTB* and openness *OPEN*. The tested hypothesis suggests that higher trade barriers lead to increased concentration. Higher values of *TARIFF* and *NTB* imply more protection and, therefore, stronger agglomeration. According to this theoretical idea, the expected sign is, thus, positive for those two variables. A higher value of *OPEN* implies less trade costs and less agglomeration the expected sign is therefore negative.

Before going through the results, two points ought to be noted. To begin with, strictly we expect changes in our trade policy variables to affect concentration with a fairly long distributed lag. However, due to data availability we work with a static model, and with levels. The idea is that a positive shift in e.g. *AGR DENSITY* tends to higher concentration, and after a while an equilibrium may develop where, *ceteris paribus*, *AGR DENSITY* is somewhat lower but still at a high level and *CONC* is somewhat higher than before the shift in *AGR DENSITY*. In the same manner, in countries where traditionally trade cost is at a high level, *CONC* is, *ceteris paribus*, also at a relatively high

**Table 1**  
**Regression Results**  
**(Dependent Variable: *CONC\**)**

Variables	(i)	(ii)	(iii)	(iv)
Constant	-2.2867 (-12.39)*** [0.00]	-1.7057 (-5.71)*** [0.00]	-1.6687 (-6.02)*** [0.00]	-3.2775 (-7.05) [0.00]
<i>AGR DENSITY</i>	0.0001 (25.67)*** [0.00]	0.0001 (15.16)*** [0.00]	0.0001 (19.53)*** [0.00]	0.00003 (1.65)* [0.09]
<i>INFRA</i>	0.0001 (1.95)** [0.05]	0.0001 (1.82)* [0.07]	0.0001 (1.92)** [0.05]	0.0001 (1.75)* [0.08]
<i>TARIFF</i>		-0.0156 (-3.35)*** [0.00]		
<i>NTB</i>			-0.0118 (-2.32)** [0.02]	
<i>OPEN</i>				2.7126 (2.42)** [0.02]
Obs.	34	34	34	32
$\bar{R}^2$	0.05	0.14	0.23	0.15

Note: parentheses ( ) give heteroscedasticity consistent t-statistics, and [ ] give p-values.

\*; significant at the 10% level \*\*; significant at the 5% level, and \*\*\*; significant at the 1% level.

level according to the hypothesis.

In addition, it would have been more appropriate to control also for the differences in urban and rural wages. However, this data only existed for seven developing countries; and besides, measuring wages in non-agricultural activities and wages in agriculture is not quite the same as measuring urban and rural wages. In other words, even those variables were imperfect measures for urban-rural wage differences. This may explain why earlier

empirical studies did not attempt to control for the pull effect.<sup>14</sup> In the calculations below we control, as mentioned above, for the push effect directly and the pull effect indirectly by including the variable *AGR DENSITY*.<sup>15</sup>

#### IV. Results

In Table 1 the employed sample excludes Hong Kong and Singapore, despite data availability for those two countries.<sup>16</sup> We have, nevertheless run the very same regressions including these two countries and the results were very close to those reported in Table 1, with one essential exception the coefficient of determination was exceptionally high ranging between 0.89-0.90. The main reason for excluding these two countries is that both are characterized by exceptionally high urban concentration and agricultural densities.

In Table 1 we present four regressions. In column (i), besides *AGR DENSITY*, only the *INFRA* is included. We then separately introduce the three trade policy variables *TARIFF*, *NTB* and *OPEN* in columns (ii), (iii) and (iv), respectively.<sup>17</sup> The results suggest strong support for the traditional push-and-pull hypothesis, the effect of *AGR DENSITY* is positive and significant at least at the 10 % level in all four estimations. The relationship between trade costs within countries and concentration, however, does not conform to the hypothesis derived from some of the newer theories. *INFRA* is positive and significant at least at the 10% level in all four estimations. This result implies that a more developed infrastructure may encourage concentration.

Last but not least, *TARIFF*, *NTB* and *OPEN* are significant yet carry unexpected signs. The results suggest that higher trade costs within and across borders discourage agglomeration. This is a rather remarkable

---

14. We experimented with pooled data for seven countries. Although the results are tentative, they nevertheless largely support the results presented below.

15. Needless to say, the variables employed in the present study are also imperfect proxies for the variables they are supposed to represent. However, they are to our knowledge the best available alternatives.

16. See Appendix I, The Data.

17. This is done to avoid possible multicollinearity problems. See the Appendix I and II reporting the data and simple correlation coefficients.

**Table 2**  
**Regression Results**  
**(Dependent Variable: *CONC*\*)**

Variables	(i)	(ii)	(iii)	(iv)
Constant	-2.3618 (-9.41) <sup>***</sup> [0.00]	-1.8855 (-6.08) <sup>***</sup> [0.00]	-1.6349 (-4.96) <sup>***</sup> [0.00]	-3.3094 (-7.59) <sup>***</sup> [0.00]
<i>AGR</i> DENSITY	-0.0007 (-1.43) [0.15]	-0.0007 (-1.31) [0.19]	-0.0008 (-1.59) [0.11]	-0.0008 (-1.30) [0.19]
<i>INFRA</i>	0.00008 (1.95) <sup>**</sup> [0.05]	0.00007 (1.85) <sup>*</sup> [0.06]	0.00008 (1.86) <sup>*</sup> [0.06]	0.00006 (1.74) <sup>*</sup> [0.08]
<i>TARIFF</i>		-0.0138 (-3.10) <sup>***</sup> [0.00]		
<i>NTB</i>			-0.0113 (-2.45) <sup>***</sup> [0.01]	
<i>OPEN</i>				2.5568 (2.89) <sup>***</sup> [0.00]
<i>GDPCAP</i>	0.00014 (1.46) [0.14]	0.00014 (1.84) <sup>*</sup> [0.07]	0.00009 (1.18) [0.24]	0.0001 (1.69) <sup>*</sup> [0.09]
Obs.	34	34	34	32
$\bar{R}^2$	0.13	0.24	0.25	0.24

Note: parentheses ( ) give heteroscedasticity consistent t-statistics, and [ ] give p-values.  
<sup>\*</sup>; significant at the 10% level <sup>\*\*</sup>; significant at the 5% level, and <sup>\*\*\*</sup>; significant at the 1% level.

result since it contradicts the expectations according to the new economic geography theories. We will return to the interpretation of these results in the following sub-section. At this stage a preliminary and tentative conclusion is that higher trade barriers across countries, contrary to a priori theoretical expectation, seem to discourage agglomeration.

One interpretation of our results may be that the trade policy variables may to some extent capture a level of development effect, *e.g.*, the variable *OPEN*, as measured here, tends to be higher the more developed a country is. We, therefore, control for GDP per capita to examine whether this might change the results.<sup>18</sup> The *GDPCAP* variable is averaged over the period 1980-1990 (Source: Summers and Heston, [1994]) and its expected sign is positive. The results are reported in Table 2 below.

The results reported in Table 2 suggest no support for our push-and-pull hypothesis. *AGR DENSITY* is now insignificant in all estimations. The GDP per capita variable is itself, as would be expected, positive and significant at least at the 10% level in two cases, suggesting a positive relationship between concentration and the level of development. An interpretation of this result is that higher GDP per capita is associated with, in addition to a higher level of development, a higher degree of industrialization, which according to the theoretical literature is expected to have a positive impact on concentration. The *INFRA* variable is, as in Table 1, positive and significant.

Interestingly, the trade policy variables: *TARIFF*, *NTB* and *OPEN* are still significant and once again carry unexpected signs, *i.e.* even when controlling for GDP per capita. The results, thus, to a large degree support the conclusions drawn from Table 1.<sup>19</sup>

### A. Further Evidence

To take our empirical study a step further, we use an alternative measure for our dependent variable, namely urban as a percentage of total population (*URB\**)<sup>20</sup>. This is done in order to examine the robustness of the results

18. Note that the introduction of this variable on the right hand side may give rise to a reverse causality problem.

19. For a discussion on how income may affect population growth in cities, see Glaeser *et al*, [1995].

20. This variable ranges between  $0 < URB < 100$  since it measures urban population as a percentage of total population. That means that it has to be transformed in order to lie in the range  $-\infty < URB < +\infty$ . Thus we make the following transformation:

As in the case of *CONC\** the negative sign before the logarithm  $URB^* = -\ln\left(\frac{100}{URB} - 1\right)$  secures that *URB\** increases and decreases with *URB*.

**Table 3**  
**Regression Results**  
**(Dependent Variable : the Degree of Urbanization)**

Variables	(i)	(ii)	(iii)	(iv)
Constant	-0.5199 (-2.57) *** [0.01]	-0.3154 (-1.12) [0.26]	0.2623 (0.79) [0.43]	-0.7956 (-1.69) * [0.09]
<i>AGR DENSITY</i>	-0.00008 (-0.26) [0.79]	-0.00005 (-0.19) [0.85]	-0.0002 (-0.83) [0.41]	-0.0001 (-0.49) [0.62]
<i>INFRA</i>	0.00025 (4.69) *** [0.00]	0.00024 (4.69) *** [0.00]	0.00025 (4.80) *** [0.00]	0.0002 (5.22) *** [0.00]
<i>TARIFF</i>		-0.0057 (-1.13) [0.26]		
<i>NTB</i>			-0.0143 (-3.62) ** [0.00]	
<i>OPEN</i>				1.00667 (0.98) [0.33]
Obs.	34	34	34	32
$\bar{R}^2$	0.17	0.17	0.38	0.16

Note: parentheses ( ) give heteroscedasticity consistent t-statistics, and [ ] give p-values.

\*; significant at the 10% level \*\*; significant at the 5% level, and \*\*\*; significant at the 1% level.

with respect to the choice of measure of the dependent variable. The results are reported in Tables 3 and 4. In Table 3 we report the results, excluding the GDP per capita variable.

The *AGR DENSITY* variable is insignificant in all four estimations. The *INFRA* variable is positive and highly significant suggesting that a more developed infrastructure encourages urbanization. As for the trade policy, the variables *TARIFF* and *OPEN* are now insignificant *i.e.* the relationship

between those two variables and urbanization is inconclusive while *NTB* is, as in Tables 1 and 2, negative and significant.

Before drawing any conclusions, it would (as in Table 2) be instructive to include GDP per capita (*GDPCAP*) in our set of explanatory variables to examine the robustness of our results. We do this in Table 4.

As in Tables 2 and 3, *AGR DENSITY* is insignificant and *INFRA* is positive

**Table 4**  
**Regression Results**  
**(Dependent Variable : the Degree of Urbanization)**

Variables	(i)	(ii)	(iii)	(iv)
Constant	-0.91796 (-2.79) <sup>***</sup> [0.01]	-0.6901 (-1.62) <sup>*</sup> [0.10]	0.1353 (-0.28) [0.78]	-1.2791 (-3.46) <sup>***</sup> [0.00]
<i>AGR DENSITY</i>	-0.00027 (-0.78) [0.43]	-0.00024 (-0.76) [0.45]	-0.00032 (-1.15) [0.25]	-0.00029 (-0.93) [0.35]
<i>INFRA</i>	0.00024 (5.51) <sup>***</sup> [0.00]	0.00024 (5.69) <sup>***</sup> [0.00]	0.00025 (6.01) <sup>***</sup> [0.00]	0.00022 (6.35) <sup>***</sup> [0.00]
<i>TARIFF</i>		-0.0066 (-1.47) <sup>***</sup> [0.14]		
<i>NTB</i>			-0.0121 (-2.89) <sup>***</sup> [0.00]	
<i>OPEN</i>				1.2462 (1.36) [0.17]
<i>GDPCAP</i>	0.00017 (1.22) [0.22]	0.00017 (1.33) [0.18]	0.00012 (0.98) [0.33]	0.00015 (1.22) [0.22]
Obs.	34	34	34	32
$\bar{R}^2$	0.27	0.29	0.41	0.26

Note: parentheses ( ) give heteroscedasticity consistent t-statistics, and [ ] give p-values.  
\*; significant at the 10% level \*\*; significant at the 5% level, and \*\*\*; significant at the 1% level.

and highly significant. Once again, *NTB* is negative and significant; however *TARIFF* and *OPEN* are insignificant. Interestingly, the GDP per capita variable *GDPCAP* shows no significance when the degree of urbanization is employed as our dependent variable. The basic result derived from Tables 3 & 4 is, therefore, that lower trade costs within borders seem to encourage urbanization, and that the relationship between trade policy and urbanization is inconclusive. The relationships between *AGR DENSITY* and *GDPCAP* on the one hand and urbanization on the other are also inconclusive.

One basic difficulty in this study has been data availability, a problem exacerbated by the fact that we work with a third world country sample. For example, one interesting potential variable that might affect people's expectations of the possibility of getting a job in the city is the level of unemployment (including disguised unemployment) in the country. Although official data did exist for a sub-sample of 18 countries, the figures themselves looked extremely dubious. In fact, the official figures of unemployment were in many countries below 3 percent, indicating overfull employment, while it was well-known that there was considerable unemployment. The idea to include this variable had thus to be abandoned.

## ***B. Discussion***

In this section we will attempt to offer some interpretations of the empirical evidence reached so far. The results reported in Table 1 suggest that the push-and-pull hypotheses as emphasized in the development literature receive empirical support. The proxy variable measuring the push-and-pull effects, *AGR DENSITY*, carries a positive sign and is significant. We may thus far conclude that our results suggest some support for the traditional push-and-pull hypotheses. This may be explained by the fact that individuals simply respond to economic incentives. Significant differences in urban and rural conditions attract people to the city, and declining living conditions in rural areas as a result of high rural population density push people out from the countryside.

The variable meant to capture trade costs within borders, *INFRA*, gives by far the most robust result. It is highly significant in all estimations and it carries a positive sign. There is more than one explanation for this result. First, economic theory is not really clear on this particular relationship.

Some considerations suggest that a more developed infrastructure would discourage agglomeration, yet other theoretical conclusions suggest that the more open an economy is, the less important becomes the effect of internal trade costs. Secondly, *INFRA* may not be an adequate measure for developed infrastructure in all cases. We may, however, tentatively conclude that the results suggest that trade costs within a country seem to encourage concentration.

Moreover, the variables supposed to capture trade policy, *i.e.* *TARIFF*, *NTB* and *OPEN*, are in the first two tables significant yet carry unexpected signs according to newer economic geographers. When trade policy is measured by protectionism, the coefficients are negative implying that higher trade barriers, contrary to the reported theoretical conclusions, seem to discourage agglomeration. When it is measured by openness, it carries a positive sign suggesting that more liberal trade, other things being equal, has a positive and significant effect on the degree of concentration. This, as well, does not conform to the expectations derived from the new trade and geography theory. The question, therefore, is what explanations can we offer for our empirical findings and what possible conclusions can be derived?

There could be several explanations to the results derived. To begin with, the specific nature of the trade and geography models does not make them quite appropriate to derive inferences about third world agglomerations. In particular, these models consider monopolistic competition, scale economies and intermediate production. Although they add over and above the so-called, trade costs, the whole process hinges on the fact that firms are dependent on intermediate inputs which creates forward and backward linkages between firms. This supposedly causes firms who produce intermediate goods, as well as those producing finished goods, to cluster within the same region. In a closed market framework this agglomeration process becomes stronger due to the heavier reliance on domestic intermediate production.

When we consider the rather specific nature of these models, it becomes more apparent why this scenario might probably be more devised to suit industrialized countries and less appropriate when dealing with third world countries. Third world economies rely more on production of homogenous goods, which in turn does not involve heavy reliance on intermediate inputs.

Furthermore, there may be a difference between industrial agglomeration and population concentration in general. If we think instead of trade costs within a specific country, the picture might be simpler. If trade costs are restrictively high within a country it is intuitively conceivable that customers and firms will tend to cluster into relatively large concentrations where firms can offer higher wages.

Comparing this to earlier empirical studies, Ades' and Glaeser's [1994] is to our knowledge the only cross-sectional study that supports the hypothesis that more restrictive trade costs within and across borders have a positive effect on agglomeration. Considering the sample they used, it consists of a cross section of 85 countries, both developed and developing, *i.e.* they do not in principle treat the phenomenon of third world urban giants. Bearing in mind the earlier discussion on the nature of trade and geography models, which seem more applicable in an industrialized country setting, we may argue that Ades and Glaeser's results may be due to their choice of country sample.

When the described hypothesis was tested for Mexico City in Hanson [1997], it received strong empirical support. Our evidence suggests that the effects of trade costs in this case are typical for most other third world countries.

Another explanation may be that more developed/industrialized countries are generally more open to trade. Openness in its turn encourages national and international investment, which in turn exacerbate industrial agglomeration. For example, some case studies on London and Buenos Aires both trade cities that have comparative advantage in commerce, suggest that these cities seem to grow with trade volume, thus, implying that openness to trade encourages concentration.

At this point we might suggest that there may be a difference between industrial agglomeration and the general concentration of economically active populations, indicating that economic development, which is associated with more industrialization, leads to further concentration *per se*.

However, it is more difficult to offer explanations to the finding that the relationship between more restrictive trade policies and agglomeration is negative. One explanation may be that the trade policy variables capture effects of some omitted variables. That openness has a positive effect on

agglomeration may be explained by the fact that a more open market may be more attractive for foreign investors. It is probably more attractive for multinational firms to locate in relatively open markets, a fact which in turn would have a positive effect on agglomeration.

In order to examine whether the trade policy variables capture some omitted variables' effects we controlled for GDP per capita. In consequence the trade policy variables, with the exception of *NTB*, became insignificant suggesting that the effects of GDP per capita may to some degree have been captured by our trade policy variables.

## V. Sensitivity Analysis

Due to the scarce empirical evidence analyzing the relationship between trade costs and agglomeration it seems warranted to attempt a sensitivity analysis in order to gain reasonable confidence in the inferences derived. Some basic objections to our results may arise due to omitted variables, measurement errors, and the normality of the error terms. In the following we perform some sensitivity tests on our preferred regressions, those reported in Table 2 column (iii) and Table 4 column (iii).

To begin with, we use the Ramsey [1969] test for omitted variables. The null hypothesis of no contemporaneous correlation could not be rejected at the 5% level. Thus, we may tentatively conclude at this point that omitted variables need not be an important problem.

A second, and perhaps more important, objection could be that measurement errors in our independent variables can affect the true coefficient for our variables of interest. It has been shown by Klepper and Leamer [1984] that the bounds of the true maximum likelihood estimates can be obtained by performing reverse regressions on all variables that are expected to be measured with error. Tables 5A, and 5B, present the results of these reverse regressions, where we have performed reverse regressions for estimations (iii) in Tables 2 and 4, that is, we test whether *NTB* (the most robust variable) is measured with error when employing *CONC\** and *URB\** as dependent variables. Reverse regressions are performed using all the independent variables expected to be measured with error as dependent variables. We have then solved each equation for the implied coefficients of

**Table 5A**  
**Reverse Regression Results**

Direction of Minimization					
	<i>CONC</i> *	<i>AGRDENSITY</i>	<i>INFRA</i>	<i>NTB</i>	<i>GDPCAP</i>
<i>AGRDENSITY</i>	-	-	-	+	+
<i>INFRA</i>	+	+	+	-	-
<i>NTB</i>	-	+	+	-	-
<i>GDPCAP</i>	+	-	-	-	+

**Table 5B**  
**Reverse Regression Results**

Direction of Minimization					
	<i>URB</i> *	<i>AGRDENSITY</i>	<i>INFRA</i>	<i>NTB</i>	<i>GDPCAP</i>
<i>AGRDENSITY</i>	-	-	+	+	+
<i>INFRA</i>	+	-	+	-	-
<i>NTB</i>	-	+	+	-	-
<i>GDPCAP</i>	+	-	-	+	+

the independent variables.<sup>21</sup> If there are no changes in sign when estimating the reverse regressions, this suggests that the estimates are robust to measurement errors.

The results suggest that the problem of measurement errors cannot be overlooked. All the variables change sign in at least one estimation suggesting that the coefficients are not bounded. Thus, extra caution is called for when drawing conclusions based on these two regressions.

We have also tested for normality of the error terms by performing a joint test for skewness and kurtosis as suggested by Shapiro and Wilk [1965]. The hypothesis of normality could not be rejected at the 5% level. This, thus,

21. Consider the regression:

$$CONC = \alpha + \beta_1 AGRDENSITY + \beta_2 INFRA + \beta_3 TRADE, \text{ where we assume that } INFRA \text{ is measured with error. Then, we estimate:}$$

$$INFRA = \gamma + \eta_1 AGRDENSITY + \eta_2 TRADE + \eta_3 CONC$$

Then, we solve for the implied coefficients equal to:

$$CONC = \frac{\gamma}{\eta_3} - \frac{\eta_1}{\eta_3} AGRDENSITY - \frac{\eta_2}{\eta_3} TRADE + \frac{1}{\eta_3} INFRA$$

implies that our OLS estimates are efficient.

## VI. Concluding Remarks

This study takes some of the recently growing literature referred to as trade and location as a theoretical point of departure to empirically analyze the phenomenon of third world urban giants. This literature offered new insights explaining the increased concentrations in urban centers. In addition to the traditional push-and-pull factors identified by earlier studies of urbanization and migration, we also have trade costs variables as additional factors. This is probably of particular relevance to third world countries due to the heavy reliance on ISI policies and to the underdeveloped infrastructure in these economies.

The present study, therefore, empirically analyzed the effect of trade costs on the size of third world metropolises. In particular, we employed two types of trade costs variables: within country trade costs (or transportation costs) and across country trade costs (or trade policies), both of which are in theory expected to have a positive effect on the degree of agglomeration.<sup>22</sup>

The results offer some mixed support for the push-and-pull hypotheses. Also, the variables capturing trade costs show mixed results. Lower trade costs within countries tend to encourage agglomeration; but, contrary to what would be expected based on the theoretical literature, trade costs across borders do not have the expected signs, or the relationship between trade costs across borders and concentration is inconclusive. Thus, we deduce that the empirical evidence particularly in the case of across countries trade costs is inconclusive. Furthermore, performing reverse regressions, we found that the problem of measurement errors could not be ignored.

We argue that the explanation to the somewhat mixed results is twofold: First, there are difficulties pertaining to model specification and omitted variables. It is not at all clear at this point which explanatory variables should theoretically be included when analyzing agglomeration. Secondly,

---

22. Conversely, we also tested whether openness discourages agglomeration.

there are, as we have discussed, problems relating to the very special nature of the economic geography models.

In light of earlier empirical studies combined with the results derived from the present study, a few basic conclusions may be derived. The impact of trade costs on agglomeration may better be treated on an individual country basis. On a general level there is need for further empirical and theoretical analysis of the phenomenon of third world urban giants. Particularly as far as empirical work is concerned we are still at a very preliminary stage. As for the theoretical analysis, new models that are particularly suited to a third world economic environment may aid our understanding of the phenomenon. In addition, more models that distinguish between industrial agglomerations as opposed to general population concentrations would certainly add to our understanding.

### References

- Ades, A.F. and E. L. Glaeser [1994] "Trade and Circuses: Explaining Urban Giants," *NEBR Working Paper* No. 4715.
- Glaeser, E. L., J. A. Scheinkman and A. Shleifer [1995], "Economic Growth in a Cross-section of Cities," *Journal of Monetary Economics* 36; 117-143.
- Hanson, G. [1997], "Increasing Returns, Trade and the Regional Structure of Wages," *The Economic Journal* 107; 113-133.
- Klepper, S. and E.E. Leamer [1984], "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica* 52(1); 163-183.
- Krugman, P.R. [1991], "Increasing Returns and Economic Geography," *Journal of Political Economy* 99; 483-499.
- Krugman, P. [1993], "Number and Location of Cities," *European Economic Review* 37; 293-298.
- Krugman, P. and R. L. Elizondo [1996], "Trade Policy and the Third World Metropolis," *Journal of Development Economics* 49; 137-150.
- Krugman P. and A. J. Venables [1995], "Globalization and the Inequality of Nations," *Quarterly Journal of Economics* 110; 857-880.
- Krugman, P. and A. J. Venables [1996], "Integration, Specialization and Adjustment," *European Economic Review* 40; 959-967.
- Pritchett, L. [1991], "Measuring Outward Orientation in Developing Coun-

- tries; Can it be Done?" *The World Bank Working Paper* No. 566.
- Ramsey, J.B. [1969], "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," *Journal of The Royal Statistics Society Series B* 31; 350-371.
- Shapiro, S. S. and M. B. Wilk [1965]. "An Analysis of Variance Test for Normality," *Biometrika* 52(4); 591-611.
- Summers, R. and A. Heston [1991]. "The Penn World Table (Mark 5.6, PWT5.6): An Expanded Set of International Comparisons, 1950-1988," *Quarterly Journal of Economics* 106(2); 327-368.
- White, H. [1980], "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica* 48(4); 817-838.
- Williamson, J. G. [1988], "Migration and Urbanization," in H. Chenery and T. N. Srinivasan (eds), *Handbook of Development Economics* I, Amsterdam, North Holland.
- World Bank [1993], *Social Indicators of Development*, Washington DC: World Bank.
- World Bank [1995], *World Development Report*, Washington DC: World Bank.
- United Nations [1988], *Prospects of World Urbanization*.

**Appendix I: The Data**

	<i>URB</i>	<i>URB*</i>	<i>CONC</i>	<i>CONC*</i>	<i>GDPCAP</i>	<i>AGRDEN</i>	<i>INFRA</i>	<i>TARIFF</i>	<i>NTB</i>	<i>OPEN</i>
Algeria	47.53333	-0.09875	0.113792	-2.05258	2841	63	2403	18.2	68.4	0.467656
Cameroon	36.4	-0.55804	0.073127	-2.53962	1302.333	75	299	37	15.3	0.283742
Congo	38.1	-0.48531	0.314547	-0.77895	2279.667	22	509	28.7	100	0.722994
Egypt	44.8	-0.20875	0.165887	-1.61506	1836.667	1973	633	41.4	38.6	0.311368
Ghana	31.73333	-0.76605	0.093929	-2.26658	890	187	474	31	38.4	
Ivory Co	38.6	-0.46416	0.168459	-1.59659	1516	69	357	22.9	20.6	0.543138
Kenya	19.8	-1.39884	0.05683	-2.80919	872	58	324	73	36.9	0.355825
Morocco	44.53333	-0.21954	0.109259	-2.09834	2016	81	179	34.6	39.7	0.428572
Mozambique	19.76667	-1.40094	0.080484	-2.43579	810.6667	32	343	24.5	100	0.146532
Nigeria	31.1	-0.79545	0.070716	-2.57575	1165	157	376	20.5	18.4	0.422428
Senegal	36.56667	-0.55085	0.197728	-1.40056	1147.333	66	542	29.9	14.9	0.484001
Tanzania	24.56667	-1.12186	0.032681	-3.38775	317.6667	59	142	35	100	0.34807
Tunisia	53.16667	0.126836	0.195691	-1.41345	2731.667	103	2080	27.5	77.6	0.624388
Uganda	9.5	-2.25406	0.041755	-3.13328	542.6667	185	118	18.1	100	
Zambia	45.23333	-0.19125	0.107688	-2.11457	822.6667	22	795	25.8	100	0.413987
Zimbabwe	24.7	-1.11468	0.086634	-2.35545	1201.333	124	1406	20.8	100	0.274607
Guatemala	38.3	-0.47683	0.099472	-2.2031	2263.667	274	320	17.3	13.1	0.330723
Mexico	69.53333	0.825173	0.195106	-1.41717	5834	85	1019	13.4	24.1	0.210097
Argentina	84.63333	1.706128	0.339265	-0.66657	5512	18	1856	38.6	21.2	0.515405
Brazil	70.7	0.880858	0.099478	-2.20304	4120.667	59	929	75.2	44.1	0.184162
Chile	83.63333	1.631195	0.341577	-0.65627	3899	72	808	20.2	16.1	0.438042
Ecuador	51.53333	0.061353	0.137961	-1.83233	2968.667	130	476	39.1	51	0.40313
Peru	67.36667	0.724816	0.277857	-0.95511	4554	69	347	56.1	55.6	0.244353
Uruguay	84.63333	1.706128	0.41582	-0.33996	6560.33	21	2106	27.6	20.6	0.328061
Venezuela	83.06667	1.590359	0.153123	-1.71031	1230.333	89	10269	31.4	46.1	0.43282
Bangladesh	13.8	-1.832	0.045573	-3.04179	11389	1056	59	67.1	55.1	0.219807
India	25.03333	-1.09684	0.012899	-4.33763	1635	460	893	140	87.4	0.129308
Indonesia	26.3	-1.03043	0.012982	-4.33111	3288	530	160	18.4	92.5	0.397151
Jordan	60	0.405465	0.293515	-0.87838	4661	261	1767	27.1	16.8	0.541079
Korea S	64.56667	0.600045	0.034359	-3.33593	4661	1912	1090	22.7	14.2	0.620966
Pakistan	29.96667	-0.84889	0.065865	-2.65201	1255.333	423	826	68.5	85.4	0.310923
Philippines	40	-0.40547	0.126253	-1.9345	1728	653	242	29.8	63.6	0.418454
Thailand	19.9	-1.39256	0.102634	-2.16829	2740.333	242	841	36.9	20.2	0.542848
Turkey	53.13333	0.125498	0.107481	-2.11674	3230.667	151	5514	44.8	90.6	0.291608

## Appendix II

### Simple Correlation Matrix

	<i>INFRA</i>	<i>TARIFF</i>	<i>NTB</i>	<i>OPEN</i>	<i>GDPCAP</i>
<i>AGR DENSITY</i>	0.195	0.484	0.391	0.497	0.537
<i>INFRA</i>		0.429	0.482	0.535	0.417
<i>TARIFF</i>			0.769	0.700	0.684
<i>NTB</i>				0.761	0.586
<i>OPEN</i>					0.726

### Covariance / Correlation Matrix

	<i>AGR DEN</i>	<i>INFRA</i>	<i>TARIFF</i>	<i>NTB</i>
<i>AGR DENSITY</i>	316155.90625	0.1948370972	0.4835721384	0.3905503384
<i>INFRA</i>	250733.37500	5238171.25000	0.4289353380	0.4822878574
<i>TARIFF</i>	12290.00000	44373.29062	2043.05438	0.7694413858
<i>NTB</i>	13316.03438	66933.48125	2108.93281	3677.00719
<i>OPEN</i>	115.32416	504.86343	13.04331	19.02183
<i>GDPCAP</i>	1098988.87281	3468803.84312	112469.54921	129205.86348

	<i>OPEN</i>	<i>GDPCAP</i>
<i>AGR DENSITY</i>	0.4978611862	0.5372974457
<i>INFRA</i>	0.5354546456	0.4166410201
<i>TARIFF</i>	0.7004652322	0.6840173572
<i>NTB</i>	0.7614539357	0.5857430090
<i>OPEN</i>	0.16972	0.7261940733
<i>GDPCAP</i>	1088.28325	13232921.58813

### Covariance / Correlation Matrix for the Transformed Dependent Variables

	<i>CONC*</i>	<i>URB*</i>
<i>CONC*</i>	5.302475344202	0.4831116332
<i>URB*</i>	1.147123288829	1.063277307848

### Covariance / Correlation Matrix for the Dependent Variables before Transformation

	<i>CONC</i>	<i>URB</i>
<i>CONC</i>	0.030065763	0.9088403126
<i>URB</i>	7.762490535	2426.357385621

**Summary Statistics**

Series	Obs.	Mean	Std Error	Minimum	Maximum
<i>URB</i>	34	44.476471	21.489132	9.500000	84.633333
<i>URB*</i>	34	-0.244932	1.016705	-2.254058	1.706128
<i>CONC</i>	34	0.140307	0.103414	0.012899	0.415820
<i>CONC*</i>	34	-2.098640	0.961980	-4.337633	-0.339958
<i>GDPCAP</i>	34	2759.519510	2240.34526	317.666667	11389.000000
<i>AGRDEN</i>	34	287.67647	472.659126	18.00000	1973.000000
<i>INFRA</i>	34	1191.235294	1903.845137	59.000000	10269.000000
<i>TARIF</i>	34	37.150000	24.459109	13.400000	140.000000
<i>NTB</i>	34	52.544118	32.691730	13.100000	100.000000
<i>OPEN</i>	32	0.387070	0.143298	0.129308	0.722994