Determining Economic Size of Major Cities in Pakistan: 
An Empirical Analysis

Uzma Tabassum¹

Abstract

Cities are center of economic growth, creativity and modernization. The economic structure of cities is of immense importance not only from the point of view of city development and growth but also for the national development and growth. To choose the suitable estimation technique prior to estimation of coefficients of variables determining the size of a city, several pre-estimation tests are conducted. The results of pooled vs. fixed effect test supported fixed effect model. Similarly, test performed to select among pool and random effect model favored random effect model. In both tests, pool model is found inappropriate for estimation. Now to choose between the two suggested models, fixed and random effect, Hausman (1978) test is applied with the null hypothesis that fixed effect model and random effect model estimators do not differ substantially and in such a case random effect model is preferred, otherwise fixed effect model is more appropriate. The result of Hausman test significantly rejects the null hypothesis against the alternative one. Thus fixed effect model is selected for estimating the regression model. The results mentioned that expansion of informal sector and migration inflows causes economic size of city to be larger. Positive amenities as reflected by the provision of education, health and banking services also found to have a significant impact in expending city size, on the other side negative amenities like congestion and crime rate of a city contract city's economic size. Volume of trade (import and exports) has a positive significant impact in enhancing city growth economically. Finally, the effect of being a port city is also significant and positive. A port is more prone to increased concentration of trade activities and industries generating substantial employment opportunities which in turn enhance consumption and production.

1. Introduction

Specialization is a process of effective allocation of abundant resources towards some specific task intending to minimize per unit cost. Different regions are blessed with different resource allocations and when these regions make effective use of the resources they become more competitive in relation to other regions. This process is referred to as Regional Specialization. As per the neo-classical theory of trade, the concept of comparative advantage is what explains the specialization patterns of a region in terms of relative production cost (Ricardo, 1817) and relative factor endowments (Heckscher, 1919; Ohlin, 1933). The comparative advantage leading to regional specialization frames the basis of city emergence via scale economies (O Sullivan, 1993). Economies of scale can be achieved in production and exchange through factor specialization and divisibility of indivisible input cost. The presence of specialization accelerates the process of urbanization. To fully exploit economies of scale the trading firms locate at places that can efficiently collect and distribute large volumes of output. The agglomeration or concentration of trading workers bids up the price

¹ Research Assistant/PhD Scholar, Applied Economics Research Center (AERC), Karachi, Email: uzma.tabassum@aerc.edu.pk and uzma_tbsm@hotmail.com

² Factor specialization is a process by which worker’s skill and efficiency increases with repetition and spend less time switching between tasks

³ Indivisible input cost is the fixed cost of capital that must be bore for production which then spreads over the entire production. The more one produce the less will be the unit fixed cost.
of land that cause people to economize on land by occupying small residential units the result is increase in population density in relatively smaller geographical area, an urban area or city. Now these rural and urban regions prospers by the exchange of what they produce i.e agricultural production by rural sector and manufactured goods & services by urban sectors. The pace of this prosperous growth of both sectors will be dependent on the means of transportation between them. The more efficient the means of transportation are, the faster will be the growth of these regions. Cities differ in their sizes depending on the type of agglomeration economies. The pace, number and variety of firms clustering in an area defines its size along with the technology a firm adopted. As shown from figure-1

**Figure 1: Process of city emergence**

![Process of city emergence](image)

**Source:** Author’s visualization for city emergence

The size of a city is the expansion or development of an urban area either in geographical or economic sense. Physical city size is expansion of a city geographically i.e covering greater land area. Economic city size refers to degree of participation of a city in the economic development of a country. City population has dual role to play towards city size determination. It affects physical city size by increasing number of person which required greater land area for their accommodation. On the other hand, it contributes towards greater aggregate demand via increased labour supply and number of consumers. Population’s physical contribution may or may not dominate its economic contribution. It is the economic city size rather than physical that matters more because with population growth, a city’s
output might grow further, may remain stagnant or can even decline (Cohen, 2004) and (Sridhar, 2010). The economic size of city is different for different cities depending on its centripetal and centrifugal forces as reflected by various urban growth models. Centripetal and centrifugal forces are shown in (Figure -2)

Figure 2: Factors for Uneven Regional Development

Centripetal forces tend to expand city while centrifugal forces tends to shrink it. It is the war between these two forces that affects the city size. The economic size of the city will increase if centripetal forces out-weight centrifugal forces and will shrink if centripetal forces short-length centrifugal forces. The attraction and economic contribution of cities differ across cities in Pakistan as well. The objective of this study is to looks into what are the drivers of this difference, using regression analysis.

The rest of the paper is organized as follows. The next section provides review of important theoretical and empirical literature. Section three describes the econometric methodology of research. Data sources and construction of variables is discussed in section four. In next section (section five) determinants of economic size of cities are empirically investigated. Finally, section six discusses the main conclusions of this research along with contribution and policy recommendation.

2. Review of literature

This section provides the review of previous literature that supports the research design both by theory and empirics. Theoretical literature provides linkages between the key variables while the empirical literature is equipped with evidences and techniques of estimation regarding these linkages.
2.1 Theoretical review of literature

Cities won’t flourish at the same pace though they usually grew over time. Population growth on its own is economically important for city growth because more population means more investment required in housing and infrastructure (roads, hospital, schools, sanitation etc) for facilitating their accommodation and commuting. The easiness in travelling, accessibility of housing, and the level of income determine the population size of cities as individuals from other regions or place are attracted by such area amenities (Rosen, 1979 & Roback, 1982) Economic size of city in terms of its earning and productivity itself is linked with a city’s population size. (Fujita, 1988; Helsley & Strange; 1990, Glaeser, 1994; Duranton & Puga, 2002) explicitly acknowledge agglomeration benefits or city’s production advantages. Henderson (1974) provided his seminal contribution to city size literature which focused on trade off between agglomeration economies and urban cost for existence of city along with impressive implications for its population growth. Henderson developed general equilibrium model of city size on the basis of optimization behaviour of labour, firms and capital owner. Henderson defines optimum size of city and equilibrium size of city on the basis of social and economic consideration. Optimum size of city is defined as that which maximizes the participant's potential welfare in the economy and the Equilibrium size of city is determined by the decisions about investment and perceived location of labourers and capital owners, every one attempting to attain their own welfare level.

Equilibrium city size is economic size of the city as it is based on the rational behaviour of economic agents. The market behaviour of factor owners is depicted by labourers moving between cities to maximize welfare and capital owners investing to maximize rent of capital. It is the behaviour of firms that determines city size. The size of a city varies depending on the type of production specialization of different goods and services traded domestically or internationally. Different degrees of scale economies in production across cities have different levels of commuting and congestion costs which in turn defines cities with different sizes. The above discussion supports neo-classical urban system theory which states that it is the tension between centripetal and centrifugal forces that determines optimal city size. Centripetal forces are the forces which come to play because of agglomeration of localization4 and urbanization economies5 while for emergence of centrifugal forces, commuting costs and land rents within city play their part (Krugman, 1994). But by no means it’s necessary that these market forces do result in emergence of an optimal city. Random urban system theory step forward to present rationale for this and state that the city’s size distribution is actually their type distribution where the type distribution depends on city’s individual characteristics which then determine a city’s economic size.

2.2 Empirical review of literature

The literature on the size of cities predicts that a country’s urban population growth, induced by industrialization or technological change, will be contained by growth in both city population sizes and the number of cities in a country (Black & Henderson, 1999; Henderson & Wang 2005 & 2007). Mills & Becker (1986) founds that a city’s population grows faster with faster industrial employment growth and national population growth in cities of India (Sridhar, 2010) also estimate the determinants of city growth in India using District level data from 1999-2006. Their main findings were that increase in Literacy rate, ratio of manufacturing to services employment and primary school population coverage have positive significant relationship with net district domestic product per capita. Da Meta, et al (2005)

4 Localization economies refer to intra-industry benefits enjoyed by firms in a specific industry by locating near to one another.
5 Urbanization economies refer to inter-industry benefits enjoyed by industries clustering near to one another.
analyzed the factors that influence the growth of Brazilian cities. They came to the conclusion that improvements in transportation facilities, increase in rural population supply, and labour force educational attainment inclination have sturdy impacts on the pace of growth of Brazilian cities. They also found that crime rate measured by homicide rate limiting city growth. Moomaw & Shatter (1996) estimate city growth, as measured by percent urban population, by using 1960, 1970 & 1980 data of 90 countries. They found that GDP per capita, percentage of labour force in agriculture and in industry, trade openness, as measured by export to GDP ratio, proportion of foreign assistance to GDP and regional dummies significantly explained size of population. Without inclusion of regional dummies literacy rate were also significant determinant of city population but after inclusion of dummies it became insignificant.

Huff & Angeles (2011) took 32 cities of six South East Asian countries as a unit of analysis. They established a conclusion that Globalization measured by Industrial production, main city dummy and government expenditures have a positive and significant impact on city population growth. Erdem & Tugcu (2011) also empirically investigate the city growth reflected from increase in city GDP using time series data from 1990 to 2001 for fourteen Turkish cities. Using fixed effect model they have shown that population, gross fixed capital formation, call deposited bank loan and exports notably explained city level growth rate of GDP. On the other hand imports have no noteworthy relationship with city growth.

Using 1970, 1980 and 1990 data sets of Metropolitan Statistical Areas (MSA’s) of United State Mills & Lubuele (1995) regressed the MSA’s population on lag of population, square lagged of population, wage, employment and regional dummies. Results indicated that MSA’s population was influenced strongly by wages, employment and lagged population. Black and Henderson (1999) explored the determinants of city population growth of 318 MSAs in the 48 States of USA on the basis of time series data from 1940 to 1990. They set up strong evidence that it is human capital growth that becomes the basis of city growth. Employment moves parallel to investment in human capital. Manufacturing employment was also found significant. Increased education reflecting higher human capital relatively benefit larger cities more than the smaller ones as concluded by Henderson & Wang (2007) backed by a data set comprising major city from 142 countries and a time span of 40 years (1960-2000). They further identified that openness is more likely to expend port city’s growth.

3. Econometric Methodology for Estimation:

The study provides empirical evidence on the determinants of economic city size which will be performed using data from 2005-06 to 2012-13 from various secondary sources. The model used for finding the impact of various variables influencing the size of a city by time is expressed symbolically in equation-1. Panel analysis with fixed effects accounting for individual city characteristics by time is applied for regressing this model.

\[
ES_{jt} = \alpha + \beta_0 HC_{jt} + \beta_1 FS_{jt} + \beta_2 FDI_{jt} + \beta_3 EX_{jt} + \beta_4 IM_{jt} + \beta_5 U^+_{jt} + \beta_6 U^-_{jt} + \beta_7 DL + \beta_8 In_Mig_{jt} + \mu \\
\]

Where

\( j \) represent the cross sectional unit, that is, city (j=1,...,14), \( t \) shows time (t =2006-2013) and \( \mu \) represent error term.

The tabulation (table-1) below briefly explains the symbols of the models and the sign they are expected to take with respect to dependent variable.
Table 1: Variable Description and Expected sign w.r.t Regressand.

<table>
<thead>
<tr>
<th>Variable Symbol</th>
<th>Variable Description</th>
<th>Expected Sign w.r.t dependent variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECS</td>
<td>Economic city size</td>
<td>Dependent variable</td>
</tr>
<tr>
<td>In_Mig</td>
<td>Migration inflows</td>
<td>Positive</td>
</tr>
<tr>
<td>FDI</td>
<td>Foreign direct investment</td>
<td>Positive</td>
</tr>
<tr>
<td>IFS</td>
<td>Informal sector</td>
<td>Positive</td>
</tr>
<tr>
<td>EX</td>
<td>Exports</td>
<td>Positive</td>
</tr>
<tr>
<td>IM</td>
<td>Imports</td>
<td>Negative</td>
</tr>
<tr>
<td>HC</td>
<td>Human capital measured by average years of schooling</td>
<td>Positive</td>
</tr>
<tr>
<td>DL</td>
<td>Dummy for location</td>
<td>–</td>
</tr>
<tr>
<td>UI⁻</td>
<td>Index for Negative Urban Amenities.</td>
<td>Negative</td>
</tr>
<tr>
<td>UI⁺</td>
<td>Index for Positive Urban Amenities.</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Positive Amenities include

Education | Number of educational institutions
Health    | Beds per hospital
Financial Institutions | Number of local financial intermediaries (Banks)

Negative Amenities include

Crime | No of reported crime
Congestion and Transport | Number of vehicles

A balanced panel data set which has equal number of observations for fourteen cross sectional units (cities) is used for analyzing determinants of economic city size. Data on the above mentioned variables is taken from various sources for the years 2005-06 to 2012-13⁶. Considering the heterogeneity of the dataset, different types of techniques are applied to estimate model-1 for comparative purposes and then the most appropriate one is finalized for estimation. These include the pooled OLS, Fixed effect, i.e. Least Square Dummy Variable (LSDV) and random effect model.

**Pooled OLS:**

In pooled OLS it is assumed that all coefficients are constant across time and cross sectional units so there is neither significant temporal nor cross sectional effects. In pool OLS, all the data is pooled as one and ordinary least Square regression is performed on model

⁶ 2011-12 data is not available for LFS based variables
Despite of the simplicity of the model, the pooled OLS might disfigure the real depiction of the relationship between regresand and the regressors across the cross-sections.

**Fixed Effect (FE) Model**

Fixed Effect (FE) model investigate the relationship among predictor and predictant variables with in a cross section (country, cities, etc). Each cross section has its own individuality that may or may not influence the predictor variables. The FE model assumes that something within the individual may influenced or biases the predictor or outcome variables and that need to be controlled. This is the rationale behind the assumption of the correlation between cross section’s error term and predictor variables. FE model eliminate the effect of those time-invariant characteristics and gives the net effect of the predictors on the outcome variable. Additionally FE model assumed that those time-invariant characteristics are unique to the individual and should not be correlated with other individual characteristics. Each cross section is different therefore the cross section’s error term and the constant (which captures individual characteristics) should not be correlated with the others. In the case of correlation between error terms the FE model is not suitable since inferences may not be correct and Random Effect (RE) model may gives better result this is the main reason for applying Hausman test.

**Random Effect (RE) Model**

The base for applying random effects model is that, in contrast the FE model, the variation across cross sections is assumed to be random and uncorrelated with the predictor or independent variables included in the model and allows for time-invariant variables to play a role as explanatory variables.

**Model Specification Test**

To check which model is better, a formal test for the two models is used. Pooled regression model is used as the baseline for our comparison. We can perform this significance test with an F test resembling the structure of the F test for $R^2$ change.

\[
F = \frac{(R^2_{FE} - R^2_{POLS})/(N - 1)}{(1 - R^2_{LSDV})/(NT - N - k)}
\]

Where

- $T$ denotes time,
- $N$ is the no. of cross sectional units and
- $k$ is the no. of regressors in the model.

The significant probability of F statistics indicates that each cross sectional unit is not statistically zero and does have its individual significant impact.

**Pool Vs Random effect Model**

To choose between pool and random effect model, Lagrange Multiplier (LM) Test proposed by Breusch–Pagan is conducted under the null hypothesis that pool OLS is better against the random effect estimation of the model.

**Random Vs Fixed effect Model**

To decide, whether FE model is more appropriate or RE model, Hausman (1978) test is commonly used which tests the null hypothesis that the coefficients estimated by the RE model are the same as the ones estimated by the FE model. With significant P-value, FE model is appropriate otherwise it is safe to use RE model.
4. Data Sources and Variable Construction

This research covers a micro-panel dataset of seven years (2005-06, 2006-07, 2007—08, 2008-09, 2009-10, 2010-11 and 2012-13) and fourteen major cities defined by LFS. These fourteen cities are Karachi, Hyderabad, Sukkur, Lahore, Faisalabad, Rawalpindi, Islamabad, Bahawalpur, Sargodha, Sialkot, Gujranwala, Multan, Peshawar and Quetta. The data for the variables used for this research is principally obtained from Census of manufacturing industries (CMI), Labour force survey (LFS) and Federal Bureau of Statistics for fourteen cities of Pakistan. The research also gets benefited by the published data from the State bank of Pakistan (SBP), Education Statistics of Pakistan, Pakistan Statistical Yearbook, Com Trade (United Nation), Pakistan Telecommunication Authority, Banking Statistics of Pakistan etc.

5. Empirical Results

Descriptive statistics and graphical representation of the variables used in this analysis is reported in the appendix at the end along with the correlation matrix (A1 to A3). Prior to estimation of coefficients of variables determining the size of a city, a number of pre-estimation tests are conducted so as to choose the correct type of model and technique to be used. As per the correlation matrix (A2) multi-collinearity is not found to be an issue though autocorrelation is (A7).

For choosing the correct type of model, three tests are performed. Considering the heterogeneous nature of the dataset, first a test to choose between pool and fixed effect model is conducted. The results of this test are reported in table A4 in the appendix supporting that fixed effect model is more appropriate. Similarly, results of test performed to select among pool and random effect model is also reported in table A5 favoring random effect model. In both tests, pool model is found to be inappropriate for estimation in this case though the alternate in both is accepted. Now to choose between the two suggested models, fixed and random effect models, a third test proposed by Hausman (1978) is applied with the null hypothesis that the difference between the two models is inconsistent and in such a case random effect model is preferred, otherwise fixed effect model is more appropriate. Hausman test results in table A6 is significant rejecting the null hypothesis against the alternative one. Thus fixed effect model is selected for estimating the regression model presented in equation 1.

Further, Pasaran test for observing cross-sectional dependence is also presented in the appendix in table A9 which is found significant indicating the dependence across cross sections. Heteroskedasticity is also encountered in the model (A8). Table A10 reports the Davidsions and Mckinon test for log or linear transformation of the model. This test support linear model rather its log transformation. To account for correcting the problems of heteroskedascity, autocorrelation and cross sectional dependence this research followed Driscoll-Kraay’s (1998) procedure to deal with these problems. Standard errors produced by this procedure are robust to general forms of cross-sectional (spatial) and temporal dependence as this non-parametric procedure of estimating standard errors imposes no restrictions on the limiting behavior of the number of panels. Further in finite samples, size of the cross-sectional dimension do not constitute a constraint on feasibility even if the number of panels is much larger than T. The results of the FE model with Driscoll-Kraay standard errors are reported in the Table-2.
Table 2: Regression Results with Driscoll-Kraay standard errors

| Dep var: ECS | Coef. | Std. Err. | T     | P>|t| |
|--------------|-------|-----------|-------|-----|
| In_Mig       | 0.02351 | 0.01268   | 1.85  | 0.087 |
| UI           | 90.9131 | 21.5209   | 4.22  | 0.001 |
| UI           | -13.616 | 3.01136   | -4.52 | 0.001 |
| HC           | 688.255 | 1268.61   | 0.54  | 0.597 |
| IFS          | 0.23485 | 0.03173   | 7.4   | 0    |
| FDI          | 85.5836 | 87.1743   | 0.98  | 0.344 |
| IM           | 0.0302  | 0.01569   | 1.93  | 0.076 |
| EX           | 0.10878 | 0.04442   | 2.45  | 0.029 |
| Constant     | -18759  | 15139.8   | -1.24 | 0.237 |

As per the results reported above, except for FDI and years of schooling all other variables are found to be significant with the correct sign relationship with the dependent variable. In-migration is a major factor in determining the economic as well as physical size of a city. It not only increases the mass but also alters production by increasing labour supply and demanders for production. Economic size of a city is significantly influenced by the amenities it holds. Positive amenities of a city tend to increase city size while the negative ones decrease it by attracting / repelling migrants and enhancing/ turning down the productive efficiency of the city respectively. Further, greater the size of the informal sector in a city, the greater it contributes to national growth and GDP. Imports and exports both tend to raise production and consumption in the city which eventually boost the size of the city.

Finally, the effect of being a port city is also significant and positively relates with the economic size of the city. Being a port city accelerated the trade activities and industries also tend to located near the port city so as to minimize their unit cost thus leading to greater physical and economic size of the port city.

6. Conclusion

Cities are center of economic growth, creativity and modernization. The economic structure of cities is of immense importance not only from the point of view of city development and growth but also for the national development and growth. It is better to understand city dynamics for understanding national growth and development. Despite of the crucial importance of cities, unfortunately in Pakistan city level analysis is rarely cited. This research makes an attempt to fill this gap in the existing literature in the context of Pakistan.

To choose the suitable estimation technique prior to estimation of coefficients of variables determining the size of a city, several pre-estimation tests are conducted. The results of pooled vs. fixed effect test supported fixed effect model. Similarly, test performed to select among pool and random effect model favored random effect model. In both tests, pool model is found inappropriate for estimation. Now to choose between the two suggested

---

7 See table A11 in the appendix for FE regression results incorporating dummy variable. This regression does not follow Driscoll-Kraay’s procedure. FE with Driscoll-Kraay command does not allow for manual regression.
models, fixed and random effect, Hausman (1978) test is applied with the null hypothesis that fixed effect model and random effect model estimators do not differ substantially and in such a case random effect model is preferred, otherwise fixed effect model is more appropriate. The result of Hausman test significantly rejects the null hypothesis against the alternative one. Thus fixed effect model is selected for estimating the regression model.

The results mentioned that expansion of informal sector and migration inflows causes economic size of city to be larger. Positive amenities as reflected by the provision of education, health and banking services also found to have a significant impact in expending city size, on the other side negative amenities like congestion and crime rate of a city contract city's economic size. Volume of trade (import and exports) has a positive significant impact in enhancing city growth economically. Finally the effect of being a port city is also significant and positive. A port is more prone to increased concentration of trade activities and industries generating substantial employment opportunities which in turn enhance consumption and production.

On the basis of conclusion drawn from the analysis few policies are suggested for accelerating city's economic growth leading to growth of the national economy.

- Foreign trade plays essential role in the process of growth and development of a region. This fact is also apparent from this research as both regression and causality results demonstrate that export and import have a significant impact in expanding the size of major cities of Pakistan. The policy maker should take into consideration this piece of information while formulating polices about growth. Government should facilitate those industries which are export oriented like agro based industries to increase foreign exchange earnings. These foreign exchange earning than can be used to established new industries that required foreign inputs and also discourages monopolies.

- Karachi is the only port city so far, developing Gawadar as the second port city will not only facilitate trade but also the economic growth of Gawadar, which has the potential to be in the major cities of Pakistan. Human capital as measured by the average years of schooling has a significant impact on cities economic participation (GDP). This shows the importance of education system in the cities and at large, to increase the growth rate of nation. The positive amenity index that captures provision of education also appears to be statistically significant endorsing the importance of education system. Considering these facts government should formulate policies for targeting both improvements in provision, by allocating supplementary budget on education, and attainment via providing awareness regarding importance of education. Similarly the role of health sector in promoting economic size is also imperative. This research has established significant linkages among health services and economic size or growth of cities. Unfortunately Pakistan's budget allocation on health as a percentage of GDP is lowest in South Asian region\(^8\). Health sector required serious attention of policy makers as Pakistan need far reaching reforms of health sector.

- When cities grow to certain level, they start to produce negative amenities such as congestion, pollution, and increase in crime rate. These negative amenities have significant impact in contracting the economic city size. The role of policy makers is to minimize these negative amenities of cities by improving transportation and judicial system as per the city requirement.

References


---

• Ohlin, B. (1933), Interregional and international trade Cambridge, MA: Harvard University Press.