ИКОНОМИЧЕСКИ ИЗСЛЕДВАНИЯ
ECONOMIC STUDIES

Volume 30 (7), 2021

Uzma Tabassum[1]
Munazah Nazeer[2]

DETERMINANTS OF ECONOMIC CITY SIZE[3]

The attraction and economic contribution of cities differ across cities in Pakistan. The purpose of this study is to look into what are the drivers of this difference. 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. In-migration is a major factor in determining the economic as well as the physical size of a city. It not only increases the mass but also alters production by increasing labour supply and demand for production. Amenities also significantly influenced city size. Positive amenities of a city tend to increase city size while the negative ones decrease it. Further, the 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 boosts the size of the city. Finally, the effect of being a port city is also significant and positively relates to the economic size of the city.
Keywords: City; Economic Size; Fixed effect; Specialization; Amenities
JEL: C1; R0; O1; O4

1. Introduction

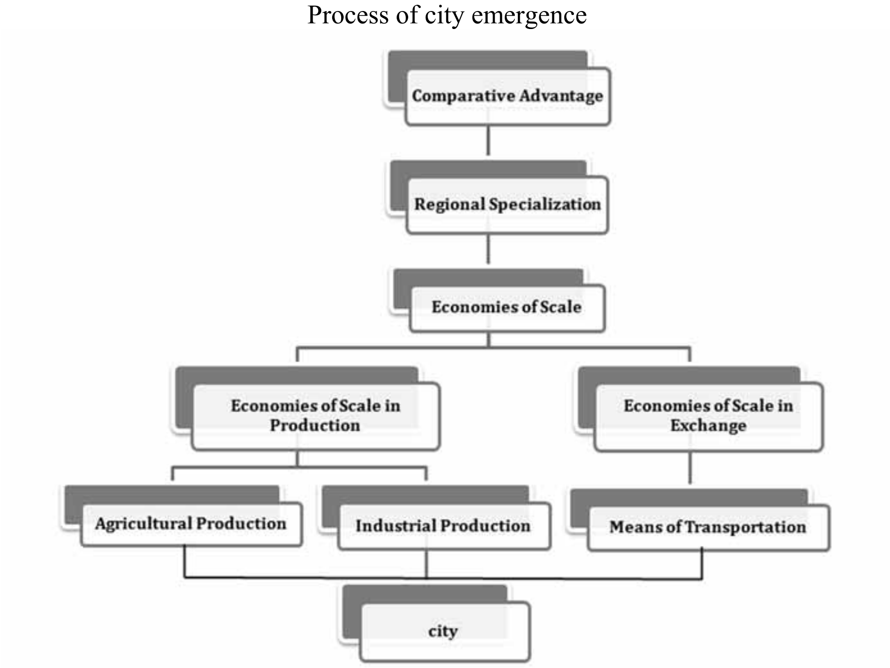
Specialization is a process of effective allocation of abundant resources towards some specific task intending to minimize per-unit cost. Different regions are blessed with different resource allocations and when these regions make effective use of the resources, they become more competitive in relation 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

[1] *Uzma Tabassum, Research Assistant, Ph.D Scholar, Applied Economics Research Center (AERC), Corresponding Author, e-mail: uzma_tbsm@hotmail.com.*
[2] *Munazah Nazeer, Ph.D Scholar, Applied Economics Research Center (AERC), e-mail: Munza_83@hotmail.com.*
[3] This paper should be cited as: *Tabassum, U., Nazeer, M. (2021). Determinants of Economic City Size. – Economic Studies (Ikonomicheski Izsledvania), 30 (7), pp. 107-128.*

and exchange through factor specialization[4] and divisibility of indivisible input cost.[5] 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 of land that causes people to economize on land by occupying small residential units, the result is an increase in population density in a relatively smaller geographical area, an urban area or a city. Now these rural and urban regions prosper by the exchange of what they produce, i.e. agricultural production by the 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 (Figure 1).

Figure 1

Process of city emergence



Source: Author's visualization for city emergence.

The size of a city is the expansion or development of an urban area either in a geographical or economic sense. Physical city size is an expansion of a city geographically, i.e. covering

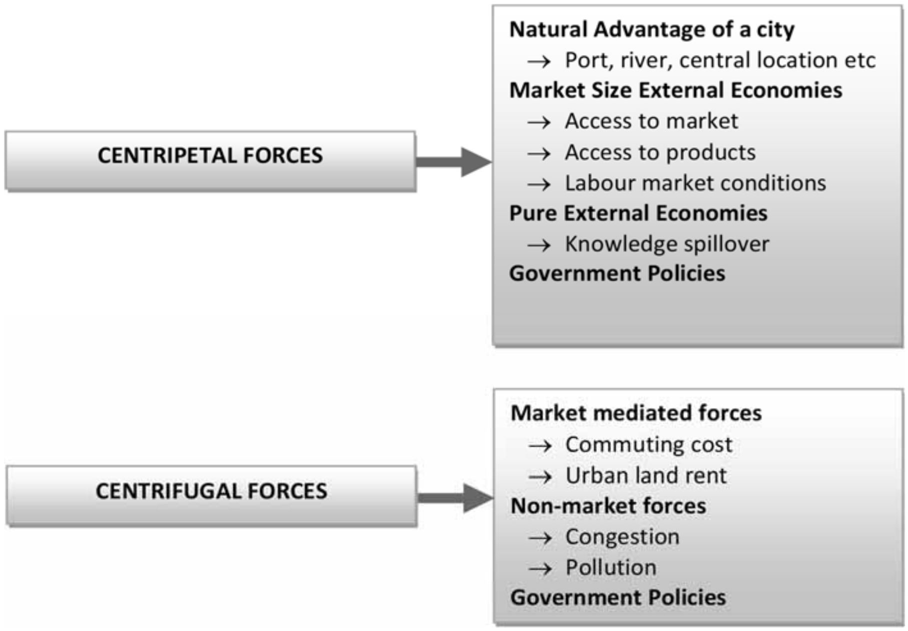
[4] Factor specialization is a process by which worker's skill and efficiency increases with repetition and spend less time switching between tasks

[5] 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.

greater land area. Economic city size refers to degree of participation of a city in the economic development of a country. The city population has a dual role to play towards city size determination. It affects physical city size by increasing the number of persons, who require a greater land area for their accommodation. On the other hand, it contributes towards greater aggregate demand via increased labour supply and a number of consumers. The 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 the 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 on Figure 2.

Figure 2

Factors for Uneven Regional Development



Source: Author's presentation.

Centripetal forces tend to expand the city, while centrifugal forces tend to shrink it. It is the war between these two forces that affect 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 look into what are the drivers of this difference, using regression analysis.

The rest of the paper is organized as follows. The next section provides a review of the 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 the next section (section five), determinants of the economic size of cities are empirically investigated. Finally, section six discusses the main conclusions of this research along with contribution and policy recommendations.

2. Review of Literature

This section provides a 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 grow over time. Population growth on its own is economically important for city growth because more population means more investment is required in housing and infrastructure (roads, hospitals, 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 places are attracted by such area amenities (Rosen, 1979; Roback, 1982). Economic size of the 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 the trade-off between agglomeration economies and urban cost for the existence of the city, along with impressive implications for its population growth. Henderson developed a general equilibrium model of city size on the basis of optimization behaviour of labour, firms and capital owner. Henderson defines the optimum size of the city and the equilibrium size of the city on the basis of social and economic considerations. The optimum size of the city is defined as that which maximizes the participant's potential welfare in the economy and the Equilibrium size of the 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 the 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 the 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 of 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 that come to play because of the agglomeration of localization[6] and urbanization economies,[7] while for the emergence of centrifugal forces, commuting costs and land rents within the city play their part (Krugman, 1994). But by no means it's necessary that these market forces do result in the emergence of an optimal city. Random urban system theory steps forward to present the rationale for this and state that the city's size distribution is actually their type distribution where the type distribution depends on the 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, the ratio of manufacturing to services employment and primary school population coverage have a positive significant relationship with net district domestic product per capita. Bere et al. (2014) looked at the drivers of economic growth for seven Romanian cities using data from 1996-2010. The study found research & development expenditure and migration as a significant positive factor for the growth of these cities while unemployment and population growth influence the growth process negatively in Romania. Da Meta et al. (2005) 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 the labour force in agriculture and in industry, trade openness, as measured by export to GDP ratio, the proportion of foreign assistance to GDP and regional dummies significantly explained the size of the population. Without the inclusion of regional dummies, literacy rates were also a significant determinant of the city population, but after the 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 an increase in city GDP using time series data from 1990 to 2001 for fourteen Turkish cities. Using the fixed effect model, they have shown that population, gross fixed

[6] Localization economies refer to intra-industry benefits enjoyed by firms in a specific industry by locating near to one another.

[7] Urbanization economies refer to inter-industry benefits enjoyed by industries clustering near to one another.

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 States Mills & Lubuele (1995) regressed the MSA's population on the 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 expand 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 2014-15 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.

$$ESC_{jt} = \alpha + \beta_0 HC_{jt} + \beta_1 IFS_{jt} + \beta_2 FDI_{jt} + \beta_3 EX_{jt} + \beta_4 IM_{jt} + \beta_5 U^{+}_{jt} + \beta_6 U^{-}_{jt} + \beta_7 DL + \beta_8 \text{In\_Mig}_{jt} + \mu \quad (1)$$

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

The tabulation on Table 1 briefly explains the symbols of the models and the sign they are expected to take with respect to the dependent variable.

A balanced panel data set that has an 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 are taken from various sources for the years 2005-06 to 2012-13.[8] 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.

[8] 2011-2012 data is not available for LFS based variables.

Table 1

Variable Description and Expected sign with respect to regressand

Variable Symbol	Variable Description	Expected Sign w.r.t dependent variable
ECS	Economic city size	Dependent variable
In_Mig	Migration inflows	Positive
FDI	Foreign direct investment	Positive
IFS	Informal sector	Positive
EX	Exports	Positive
IM	Imports	Negative
HC	Human capital measured by average years of schooling	Positive
DL	Dummy for location	_
UI^-	Index for Negative Urban Amenities.	Negative
UI^+	Index for Positive Urban Amenities.	Positive

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 reported crimes
Congestion and Transport	Number of vehicles

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 1. Despite the simplicity of the model, the pooled OLS might disfigure the real depiction of the relationship between the regressand and the regressors across the cross-sections.

Fixed Effect (FE) Model

The fixed Effect (FE) model investigate the relationship among predictor and predictand variables within 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 be 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 the cross section's error term and predictor variables. The FE model eliminates 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 give better results; this is the main reason for applying the Hausman test.

Random Effect (RE) Model

The basis for applying the random-effects model is that in contrast to 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. The 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^2change.

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

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 significant individual impact.

Pool Vs Random effect Model

To choose between the 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 the FE model is more appropriate or the 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 a significant P-value, the FE model is appropriate; otherwise, it is safe to use the 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 from the published data from the State bank of Pakistan (SBP), Education Statistics of Pakistan, Pakistan Statistical Yearbook, Com Trade (United Nations), Pakistan Telecommunication Authority, Banking Statistics of

Pakistan etc. The construction of variables is a bit complex for the majority of the variables. Hence each had to be discussed one by one along with their relationships.

Economic City Size (ECS)

The city's economic size is best reflected by its contribution to the Real Gross Domestic Product (RGDP) of a nation. In Pakistan, the city-level GDP data is not readily available from secondary sources; thus, using a top-down approach, the national level GDP disaggregate at the city level. The methodology used for generating city level RGDP is stated below.

Estimation of Gross Domestic Product (GDP) at City Level

The city-level real GDP is calculated using a top-down approach, a statistical technique, for disaggregating the annual aggregate value of sector-wise real GDP using a suitable base for this disaggregation. These sectors include agriculture, manufacturing and services. For obtaining City-wise, the real GDP production of these three sectors is added up at the City level as per the production method for GDP measurement.

Deciding Base for Disaggregation

The base of disaggregation, industry-wise employment, is suggested by the very basic production equation regarded as a cornerstone in the foundation of production theory. Consider the Cobb-Douglas production function.

$$Y = AL^{\alpha}K^{1-\alpha} \qquad (3)$$

Here A is the factor productivity, α is the labour share, $(1-\alpha)$ is the share of capital and $K\ and\ L$ are the labour and capital, respectively. As capital is fixed in the short run, labour became the base for disaggregation, which tends to be considered even a stronger base for disaggregation when it is applied to abundant labour countries like Pakistan.

Production of industry belongs to the sector mentioned above is also dependent upon the same production function as

$$\boldsymbol{RGDP_s = \sum_{j=1}^{n} K_j L_j} \qquad (4)$$

Here,

$RGDP_s$ = Real GDP of sector s

K_j = Capital in industry j

L_j = labour employed in industry j

Estimation of GDP

After identifying the base for disaggregation, estimation of district-wise real GDP was conducted as per the formulation below

$$\boldsymbol{RGDP_{ct} = \sum_{s=1}^{3} \frac{RGDP_{st}}{L_{st}} * L_{stc}} \quad (5)$$

$$\boldsymbol{Subject}^{9}\ \boldsymbol{to\ \sum_{c=1}^{n} RGDP_{stc} = RGDP_{st}} \quad (6)$$

Here s stands for the sector, c for city and t is for the year.

Trade Openness

City-wise trade openness is calculated using the following formulation

$$\mathbf{TO_{jt} = \frac{IM_{jt} + EX_{jt}}{RGDP_{jt}}} \quad (7)$$

Where

TO_{jt} = degree of trade openness in city j at time t.

EX_{jt} = Total export of city j at time t.

IM_{jt} = Total import of city j at time t.

$RGDP_{jt}$ = Real GDP of city j at time t.

The above formulation for trade openness demands city-wise data for imports and export, which is not available, but industry-wise import and export data is available at a country level. Hence it had to be generated using industry-wise establishments engaged in production in city j. Imports/exports for cities are calculated by summing their industry-wise share in total import/exports by individual industries on the basis of the share of establishments belonging to all industries in city j out of the total establishments belonging to all industries in the country.

The following equations are used to generated exports and imports by cities under consideration.

For export

$$\mathbf{EX_J = \sum_{i=1}^{n} \frac{s_{ij}}{s_i} (EX_i)} \quad (8)$$

Where

EX_j = Total export of city (j).

[9] Subjective function is based on the assumption that LFS covers all existing regions in the country.

$S_{ij} =$ Total no of the industrial establishment (i) in city j.

$S_i =$ Total no of the industrial establishment (i) in all cities.

$EX_i =$ Total export of establishment i.

A greater number of industrial establishments in a city would result in more production by the city for local consumption and export purposes.

For import

$$\boldsymbol{IM_J} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \frac{\boldsymbol{s_{ij}}}{\boldsymbol{s_i}} (\boldsymbol{IM_i}) \tag{9}$$

Where:

$IM_j =$ Total import of city (j).

$S_{ij} =$ Total no of the industrial establishment (i) in city j.

$S_i =$ Total no of the industrial establishment (i) in all cities.

$IM_i =$ Total import of establishment i.

Production needs inputs that are locally available as well as those that have to import from abroad. Thus, with more industrial establishments locating in the region, more materials from abroad are expected to be imported.

Foreign Direct Investment (FDI)

Foreign direct investment in Pakistan is inclined more towards the services sector and within the services sector, financial businesses and telecommunications are the major heads receiving such investment (Nazeer et al.; 2017). A weighted index for foreign direct investment is calculated using the number of foreign banks and the number of foreign telecommunication franchises for resembling FDI inflow in cities. Symbolically, the formula for the index is

$$\mathbf{FDI\ Index_{jt} = w_1 FB_{jt} + w_2 FTF_{jt}} \tag{10}$$

Where, $w_1 = 1/3$, and $w_2 = 2/3$ are the weights given to foreign bank branches (FF) and foreign telecommunication franchises (FTF) in city j (j = 1, 2...., 14) at time t respectively depending on their degree of consumption, ease of access and spread spatially. Telecommunication is given more weight than foreign banks because it can be observed easily that individuals demanding foreign telecommunication company's services stand far above those demanding services from foreign banks. Interestingly, one may not have a bank account these days, even in a local bank, but they usually have a cell phone which again may be used for multiple SIMs. FDI is positively related to economic city size as more investment means more employment opportunities, more migration, increased aggregate demand and more production. Thus economic activities in the region further accelerate.

Positive urban amenities index

Index for positive city-wise amenities is a weighted average of three-dimension indices, education health and financial institutions.

$$\boldsymbol{UI_{jt}^{+} = w_1 EI_{jt} + w_2 HI_{jt} + w_3 FI_{jt}} \quad (11)$$

Where, w represents fractional weights (w_1=w_2= 2/5 and w_3 =1/5) allotted to the education index (EI_{jt}), health index (HI_{jt}) and financial institutions (FI_{jt}), j represents city and t is for a time following the methodology of Human Development Index (HDI) constructed in the United Nation Development Report (UNDR) 1990. In the construction of dimension index for education, the methodology of UNDR (2010) is adopted as indicators within a dimension are non-mutually exclusive and in this case, arithmetic mean is not appropriate; therefore, the geometric mean is a more suitable measure.

Dimension Index for Education

$$\boldsymbol{EI_{jt} = \sqrt[7]{PS_{jt} \times MS_{jt} \times SS_{jt} \times DC_{jt} \times TC_{jt} \times IC_{jt} \times VI_{jt}}} \quad (12)$$

Where,

Education index (EI_{jt}): Geometric index of a number of educational institutions including Primary School (PS_{jt}), Middle School (MS_{jt}), Secondary School (SS_{jt}), Degree College, intermediate college (IC_{jt}), and vocational institute (VI_{jt}).

Dimension Index for Health

The dimension index of health is calculated using the formula

$$\boldsymbol{HI_{jt} = \frac{NB_{it}}{NH_{it}} \times 1000} \quad (13)$$

Where

HI_{jt} = Beds per hospital for 1000 persons.

NB_{jt} = Total no of beds in city i.

NH_{jt} = Total no of hospitals in city i.

Dimension index for a financial institution

FI_{jt} = number of local financial intermediaries (banks) in a city.

An increase in positive urban amenities index tends to increase a city's size economically as they accelerate economic activities such as production, employment, migration, etc.

Negative urban amenities index

The formula for urban amenity indicator index for negative amenities (UI_{jt}^{-}) is

$$\boldsymbol{UI_{jt}^{-}} = \frac{\boldsymbol{CTI_{jt}+CI_{jt}}}{\boldsymbol{n}} \tag{14}$$

Where

UI_{jt}^{-} = Negative urban amenities index

CTI_{jt} = Dimension index for Congestion and Transportation.

CI_{jt} = Dimension index for crime.

Dimension index for Congestion and Transportation (CTI_{jt}): Index calculated from the number of vehicles in a city giving more weight to heavy traffic than the lighter one. This data is gathered for cars, bicycles, buses, taxies, rickshaws constituting light traffic while trucks, tractors, pickup/delivery vans, etc., for heavy traffic representation.

$$\mathbf{CTI_{jt}} = \boldsymbol{W_1}(\boldsymbol{LT}) + \boldsymbol{W_2}(\boldsymbol{HT}) \tag{15}$$

Dimension index for Crime (CI_{jt}): Index for crime is estimated using cases of murders (M), attempt to murders (AM), kidnapping (K), dacoity (D), robbery (R) and vehicle theft (VT) and snatching (S) with more weight given to murders.

$$\mathbf{CIjt} = \boldsymbol{W_1}(\boldsymbol{M}) + \boldsymbol{W_2}(\boldsymbol{AM}) + \boldsymbol{W_3}(\boldsymbol{K}) + \boldsymbol{W_4}(\boldsymbol{D}) + \boldsymbol{W_5}(\boldsymbol{R}) + \boldsymbol{W_6}(\boldsymbol{VT}) + \boldsymbol{W_7}(\boldsymbol{S}) \tag{16}$$

Informal Sector (IFS)

Data for employment in the informal sector was generated as per its definition in LFS. LFS define the informal sector at the household level. It includes employment in all own-account enterprises whatever their size is, secondly enterprises with ten or less employed persons who may be the owner(s) himself/themselves, the contributing family workers, the employees, whether employed on an occasional or a continuous basis, or as an apprentice and lastly it excludes all enterprises engaged in agricultural activities or wholly engaged in non-market production, Symbolically

$$\boldsymbol{IFS_{jt}} = \boldsymbol{OAE_{jt}} + \lim_{\boldsymbol{L\leq 10}} \boldsymbol{SE_{jt}} - \boldsymbol{AE_{jt}} \tag{17}$$

IFS_{jt} = employment in informal sector area in j at time t.

OAE_{jt} = Own account enterprise employment in j at t.

$\lim_{L\leq 10} SE_{jt}$ = Small enterprise employment limited individually to 10 or less labour in j at t

AE_{jt} = Agriculture enterprise employment in j at t.

Human Capital (HC)

For an individual city's human capital attainment, an average year of schooling is calculated using data from LFS. Economic size of a city is expected to be raised with rising human capital attainment as more human capital attainment means more productive labour force and more production in economic terms.

5. Empirical Results

Descriptive statistics and graphical representation of the variables used in this analysis are reported in the appendix at the end, along with the correlation matrix (A1 to A3). Prior to the 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 the pool and fixed effect model is conducted. The results of this test are reported in Table A4 in the appendix, supporting that the fixed effect model is more appropriate. Similarly, results of the test performed to select among the pool and random effect model are also reported in Table A5 favouring the random effect model. In both tests, the 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, a 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, the 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 Davidsons and Mckinnon test for log or linear transformation of the model. This test supports the linear model rather than its log transformation. To account for correcting the problems of heteroscedasticity, 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 behaviour of the number of panels. Further, in finite samples, the size of the cross-sectional dimension does 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 Table 2.

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 the 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, the 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.

Table 2

Regression Results with Driscoll-Kraay standard errors

Regression Results with Driscoll-Kraay standard errors				
Method: Fixed-effects regression				
Number of groups = 14			Maximum lag:	2
F(8, 13) = 3105.19			Number of obs = 98	
Prob > F = 0.0000			within R-squared = 0.9137	
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

Source: Authors' estimation.

Finally, the effect of being a port city is also significant and positively relates to the economic size of the city.[10] Being a port city accelerated the trade activities and industries also tend to be located near the port city so as to minimize their unit cost, thus leading to the greater physical and economic size of the port city.

6. Conclusion

Cities are the centre 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, the 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 the fixed effect model. Similarly, tests performed to select among the pool and random effect model, favour the random effect model. In both tests, the 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 the Hausman test significantly rejects the null

[10] 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.

hypothesis against the alternative one. Thus fixed effect model is selected for estimating the regression model.

The results mentioned that expansion of the informal sector and migration inflows cause the economic size of the 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. The volume of trade (import and exports) has a significant positive 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 the conclusionp drawn from the analysis few policies are suggested for accelerating the city's economic growth leading to the growth of the national economy.

- Foreign trade plays an 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 policymaker should take into consideration this piece of information while formulating policies about growth. Government should facilitate those industries which are export-oriented, like agro-based industries, to increase foreign exchange earnings. These foreign exchange earnings can be used to established new industries, that require foreign inputs and also discourages monopolies.
- Karachi is the only port city so far, developing Gawadar as the second port city will facilitate not only 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 the education system in the cities and at large, to increase the growth rate of the nation. The positive amenity index that captures the provision of education also appears to be statistically significant, endorsing the importance of the education system. Considering these facts government should formulate policies for targeting both improvements in the provision, by allocating supplementary budget on education, and attainment via providing awareness regarding the importance of education. Similarly, the role of the health sector in promoting economic size is also imperative. This research has established significant linkages among health services and the economic size or growth of cities. Unfortunately, Pakistan's budget allocation on health as a percentage of GDP is lowest in the South Asian region[11]. The health sector requires serious attention of policymakers as Pakistan need far-reaching reforms of the health sector.
- When cities grow to a certain level, they start to produce negative amenities such as congestion, pollution and an increase in crime rate. These negative amenities have a significant impact on contracting the economic city size. The role of policymakers is to

[11] Antonia Settle (2010) Post Budget Orientation Series, Federal Budget Health Sector.

minimize these negative amenities of cities by improving transportation and the judicial system as per the city requirement.

References

Bere, R. C., Adrian, O., Bucerzan, I. (2014). Determinants of economic growth in cities acting as growth poles in regions from Romania. – Procedia Economics and Finance, 10, pp. 357-365.
Black, D., Henderson, J. V. (1999). A theory of urban growth. – Journal of Political Economy, 107, pp. 252-284.
Cohen, B. (2004). Urban Growth in Developing Countries: A Review of Current Trends and Caution Regarding Existing Forecast. – World Development, 32(1), pp. 23-51.
Da Mata, D., Deichmann, U., Henderson, J. V., Lall, S. V., Wang, H. G. (2005). Determinants of City Growth in Brazil. – National Bureau of Economic Research (NBER) Working Paper 11585.
Driscoll, J. C., Kraay, A. C (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. – Review of Economics and Statistics, 80, pp. 549-560.
Duranton, G., Puga, P. (2004). Micro-foundations of urban agglomeration economies. – Handbook of Regional and Urban Economics: Cites and Geography, 4(5), pp. 2063-2117.
Duranton, G., Puga, D. (2002). Diversity and Specialization in Cities: Why, where, and when does it matter?. – Urban Studies 37(3), pp. 533-555.
Erdem, E., Tugcu, C. T. (2011). Investigating the Macroeconomic and Qualitative Dynamics of Urban Economic Growth: Evidence from the Most Productive Turkish Cities. – International Journal of Business and Social Science, 2 (13), Special Issue.
Fujita, M., Thisse, J. (2002). Economics of Agglomeration. Cambridge University Press
Glaeser, E. (1994). Cities, Information, and Economic Growth. – A Journal of Policy Development and Research, 1(1), pp. 9-47.
Hausman, J. A. (1978). Specification Tests in Econometrics. – Econometrica, 46(6) pp. 1251-1271.
Heckscher, E. (1919). The effect of foreign trade on distribution of income. – Economisk Tidskrift, pp. 497-512.
Helsley, R. W., Strange, W. C. (1990). Matching and Agglomeration Economies in a System of Cities. – Regional Science and Urban Economics, 20(2) pp. 189-212.
Henderson, J. V., Wang, H. G. (2007). Urbanization and City Growth: the role of institutions. – Regional Science and Urban Economics 37, pp. 283-313.
Henderson, J. V., Wang, H. G. (2005). Aspects of the rural-urban transformation of countries. – Journal of Economic Geography, 5, pp. 23-42.
Henderson, J. V. (1974). The Sizes and Types of Cities. – American Economic Review, 64, pp. 640-657.
Huff, G., Angeles, L. (2011). Globalization, Industrialization and Urbanization in Pre-World War II Southeast Asia. – Explorations in Economic History, 48(1) pp. 20-36.
Krugman, P. (1994). Urban Concentration: The Role of Increasing Returns and Transport Costs. Conference Paper, World Bank Annual Conference on Development Economic.
Mills, E. S., Becker C. M. (1986). Studies in Indian Urban Development. Oxford University Press for the World Bank.
Mills, E. S., Lubuele, L. S. (1995). Projecting Growth of Metropolitan Areas. – Journal of Urban Economics, 37, pp. 344-360.
Moomaw, R. L., Shatter, A. M. (1996). Urbanization and economic development: A bias toward large cities?. – Journal of Urban Economics, 40(1), pp. 13-37.
Nazeer, M., Tabassum, U., Alam, S. (2017). Banking and Telecommunication influencing migration in major cities of Pakistan. – Pakistan Journal of Applied Economics, 27(1), pp.101-120.
Ohlin, B. (1933). Interregional and international trade Cambridge. MA: Harvard University Press.
O'Sullivan, A. (1993). Urban Economics. Homewood, 2nd Edition.
Ricardo, D. (1817). On the Principles of Political Economy and Taxation. Batoche Books, Canada.
Roback, J. (1982). Wages, Rents, and the Quality of Life. – Journal of Political Economy, 90, pp. 1257-1278.
Rosen, K. T., Resnick, M. (1980). The size distribution of cities: An examination of the Pareto law of primacy. – Journal of Urban Economics, 8(2), pp. 165-186.
Rosen, S. (1979). Wage-based indexes of urban quality of life. – In: Mieskowski, P., Straszheim, M. (eds.). Current issues in urban economics. Johns Hopkins University Press, Baltimore, pp. 74-104.
Sridhar, K. S. (2010). Determinants of City Growth and Output in India. – Review of Urban and Regional Development Studies, 22(1), pp. 22-38.

APPENDIX

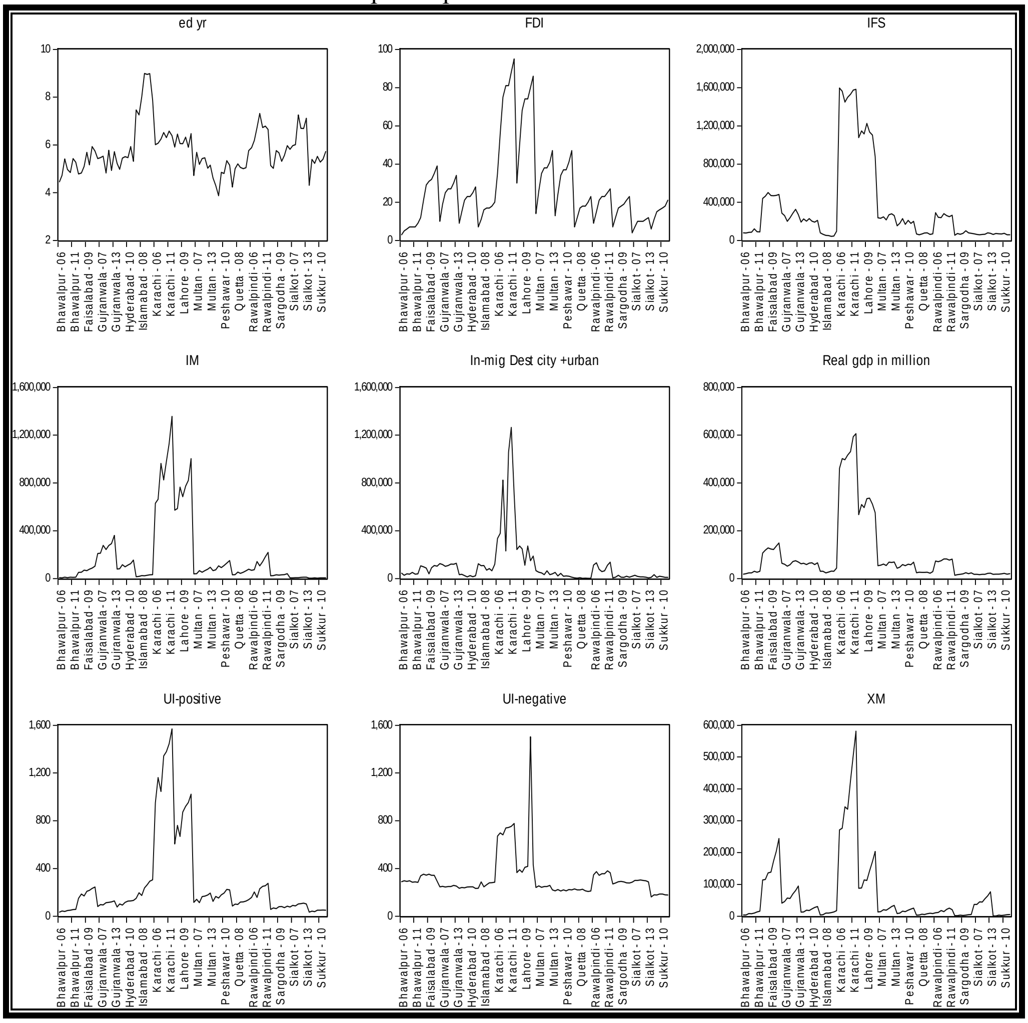
A1: Descriptive summary of the variables

Descriptive Summary	ED_YR (Years)	FDI (units)	IFS (No.)	IM (Rs.ML)	IN_MIG_ DEST_CITY__U RBAN (No.)	REAL_GD P_IN_ (Rs. ML)	UI_ POSITIVE (No.)	UI_ NEGATIVE (No.)	XM (Rs.ML)	TO
Mean	5.763013	27.07143	336588.1	182098.9	105660.1	101142.4	262.6837	318.6748	64772.38	246871.3
Median	5.523397	21	199955	65755.88	40226	56382.76	128.5	280.417	18729.79	87029.93
Maximum	8.993373	95	1595665	1357262	1264857	606593.3	1570	1503.745	580982.4	1938245
Minimum	3.870027	3	43786	2589.206	254	13278.73	35	161.2182	1320.03	4058.361
Std. Dev.	0.980423	21.27544	428385.2	290823.2	197550.2	140419.3	346.1547	177.389	108182.9	386719.6
Skewness	1.230167	1.567222	1.938967	2.188231	4.053929	2.290143	2.309001	3.893935	2.748149	2.390035
Kurtosis	5.008825	4.822002	5.467351	6.946606	20.73622	7.23577	7.307846	22.8663	10.93993	8.323851
Jarque-Bera	41.19522	53.67307	86.26526	141.8106	1552.936	158.9264	162.8576	1859.227	380.7778	209.0359
Probability	0	0	0	0	0	0	0	0	0	0
Sum	564.7753	2653	32985633	17845691	10354690	9911959	25743	31230.13	6347693	24193384
Sum Sq. Dev.	93.23928	43906.5	1.78E+13	8.20E+12	3.79E+12	1.91E+12	11622837	3052285	1.14E+12	1.45E+13
Observations	98	98	98	98	98	98	98	98	98	98

A2: Correlation matrix of the variables used

Correlation Matrix	ED_YR	FDI	IM	IFS	UI_POSITIVE	UI_NEGATIVE	XM
ED_YR	1.000000						
FDI	0.131208	1.000000					
IM	0.175977	0.696227	1.000000				
IFS	0.135645	0.486097	0.595186	1.000000			
UI_POSITIVE	0.290174	0.750571	0.647468	0.646330	1.000000		
UI_NEGATIVE	0.225746	0.668115	0.744020	0.763746	0.749897	1.000000	
XM	0.164023	0.777890	0.746580	0.687993	0.787938	0.716715	1.000000

A3: Graphical presentation of variables



A4: Choice of model between pool OLS and fixed effect model

Choice of Model
Pool vs FE/LSDV Model
Ho: Pool model is better i.e
(1) 2007.Years = 0
(2) 2008.Years = 0
(3) 2009.Years = 0
(4) 2010.Years = 0
(5) 2011.Years = 0
(6) 2013.Years = 0
(7) 2.cities = 0
(8) 3.cities = 0
(9) 4.cities = 0
(10) 5.cities = 0
(11) 6.cities = 0
(12) 7.cities = 0
(13) 8.cities = 0
(14) 9.cities = 0
(15) 10.cities = 0
(16) 11.cities = 0
(17) 12.cities = 0
(18) 13.cities = 0
(19) 14.cities = 0
F(19, 70) = 14.38
Prob > F = 0.0000

A5: Choice of model between pool OLS and random effect model

Choice of model		
Pool vs RE Model		
Breusch and Pagan Lagrangian multiplier test for random effects		
ECS[cities,t] = Xb + u[cities] + e[cities,t]		
Ho: variances across entities is zero		
Estimated results:	Var	sd = SQRT(Var)
ECS~n	1.97E+10	140419.3
E	2.73E+07	5220.311
U	4.43E+07	6658.969
Test: Var(u) =0		
chibar2(01)	=	56.58
Prob > chibar2	=	0

A6: Choice of model between random and fixed effect model

Choice of Model				
Huaseman Test Results				
FE vs RE Model				
Coefficients	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	xtreg_fe	xtreg_re	Difference	S.E.
Inmig~n	0.02351	0.027816	-0.0043065	.
UIpositive	90.9131	100.6254	-9.712341	4.811898
UInegative	-13.616	-14.1443	0.5278915	.
edyr	688.255	-296.905	985.1598	766.4804
IFS	0.23485	0.205194	0.0296536	0.0164018
FDI	85.5836	78.39586	7.187772	49.41431
im	0.0302	-0.00407	0.0342702	0.0147467
xm	0.10878	0.163318	-0.0545381	0.0314189
b = consistent under Ho and Ha; obtained from xtreg				
B = inconsistent under Ha, efficient under Ho; obtained from xtreg				
Test: Ho: difference in coefficients not systematic i.e RE model is more efficient and consistent				
chi2(4)	=	(b-B)'[(V_b-V_B)^(-1)](b-B)		
	=	-38.84		
Prob>chi2	=	0		

A7: Wooldridge test for autocorrelation in panel data

Wooldridge test for autocorrelation in panel data		
H0: no first-order autocorrelation		
F(1, 13)	=	25.873
Prob > F	=	0.0002

A8: Modified Wald test for group-wise heteroskedasticity in FE regression model

Modified Wald test for group wise heteroskedasticity in FE regression model		
Ho: Heteroskedasticity exists		
chi2 (14)	=	3268.63
Prob>chi2	=	0

A9: Pesaran's test of cross sectional independence

Pesaran's test of cross sectional independence		
Ho: No cross-sectional dependence	Coeff.	Prob.
Pesaran's test of cross sectional independence	3.209	0.0013

A10: Davidson and MacKinnon Test for log or linear model transformation

Davidson and MacKinnon Test		
Model	Decisive variable	Probability of coefficient
Linear	Ho: Log model is better Fitted_log	 0.6776
Logarithm	Ho: Linear model is better Fitted_linear	 0.0271

A11: Panel regression results using LSDV approach

Panel regression results using LSDV approach				
No. of obs		=	98	
F (27, 70)		=	3047.4	
Prob> F		=	0	
Dep. Var.: ECS	Coef.	Std. Err.	T	P>|z|
In_Mig	0.022058	0.006479	3.4	0.001
UIpositive	85.58398	16.11579	5.31	0
UInegative	-12.8301	5.218983	-2.46	0.014
edyr	4140.011	1706.095	2.43	0.015
IFS	0.247207	0.016865	14.66	0
FDI	468.1515	167.3241	2.8	0.005
im	0.01794	0.01977	0.91	0.364
xm	0.12782	0.040189	3.18	0.001
dport	68928.83	29294.36	2.35	0.019
_cons	-16914.3	8653.41	-1.95	0.051
Cities				
2	-38509.7	7784.46	-4.95	0
3	-33500.5	6104.303	-5.49	0
4	-8598.98	4341.29	-1.98	0.048
5	-22005.6	6514.786	-3.38	0.001
6	0	(omitted)		
7	-97799.5	23971.46	-4.08	0
8	-28956.2	5512.479	-5.25	0
9	-19432.6	4774.132	-4.07	0
10	-4917.52	2990.616	-1.64	0.1
11	-21489.6	6403.877	-3.36	0.001
12	-9740.58	3178.432	-3.06	0.002
13	-17324.4	3841.903	-4.51	0
14	-4541.6	3106.729	-1.46	0.144
Years				
2007	-2099.48	2276.86	-0.92	0.356
2008	-5397.34	3107.326	-1.74	0.082
2009	-8486.53	3458.757	-2.45	0.014
2010	-10198.8	3266.3	-3.12	0.002
2011	-11624.9	3553.834	-3.27	0.001
2013	-11985.8	3853.288	-3.11	0.002