URBANIZATION AND ECONOMIC GROWTH: A CASE OF BRAZILIAN CITIES

MOIN MARILYN CHEPTEKEI

102239

Submitted in partial fulfilment of the requirements for the degree of Bachelor of Business Science in Financial Economics at Strathmore University

Strathmore Institute of Mathematical Sciences

Strathmore University

Nairobi, Kenya

February, 2021

DECLARATION

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the Research Project contains no material previously published or written by another person except where due reference is made in the Research Project itself.

MARILYN C. MOIN ..... [*Name of Candidate*]

..... [*Signature*]

12TH FEBRUARY 2021 ..... [*Date*]

This Research Project has been submitted for examination with my approval as the Supervisor.

GILLIAN KIMUNDI ..... [*Name of Supervisor*]

..... [*Signature*]

12th Feb 2021 ..... [*Date*]

Strathmore Institute of Mathematical Sciences

Strathmore University

Table of Contents

List of Tables

List of Figures

Abstract

The nexus between urbanization and economic growth is one that has caught the attention of many as global population numbers continue to rise. The two variables, especially in the context of a country, can either enhance or deter growth. The study aims to analyse the relationship between the two variables in Brazilian cities and how their contribution to GDP may vary with urbanization and other factors that contribute to agglomerate economies within the cities. The analysis uses panel data from 5 Brazilian cities ranging from 1999-2015 and makes use of analysis techniques such as Fixed Effects and Random Effects Model, and the Hausman test to verify the appropriate model. A Dynamic Panel Data Estimation is also carried out to incorporate adjustment effects in the dependent variable. Variables such as Gross Domestic Product of the country and the cities, the urbanization levels and agglomeration factors shall be captured to assess the relationship they may have with the economic growth of the country.

CHAPTER ONE: INTRODUCTION

1.1 Background of the Study

There has been a rapid increase in urbanization globally over the past 70 years. With 55% of the world's population residing in urban areas (4.2 billion in urban areas versus 3.4 billion in rural) compared to 30% in the 1950s as shown in *Figure 1* and *Figure 2*. This is expected to grow to 68% by 2050, driven by an increase in overall population and increased number of people residing in urban areas (United Nations, 2019). However, the increase in population is not uniform as most of the increase may be attributed to a small number of countries, of which Africa is expected to be the major contributor, followed second by Asia (United Nations, 2017). The World Urbanization Prospects (2018) highlights countries that host half of the world's urban population. with China being the largest (837 million). followed by India (461 million). While another five countries consist of the United States of America, Brazil, Indonesia, Japan and the Russian federation.

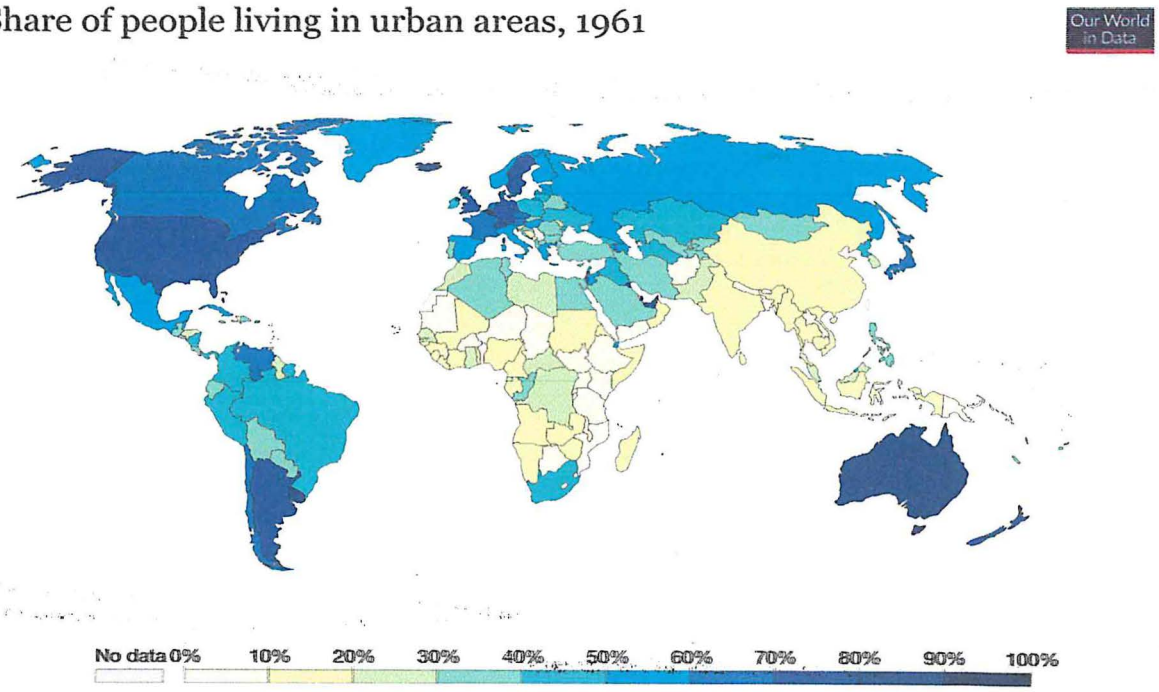

Figure 1)Percentage of population living in urban area in the 1960s. *Showing geographical coverage from global to regional from United Nations (2018). Source: Ritchie (2018)- "Urbanization".* Published online at OurWorldInData.org. *Retrieved from: https://ourworldindata.org/urbanization.*

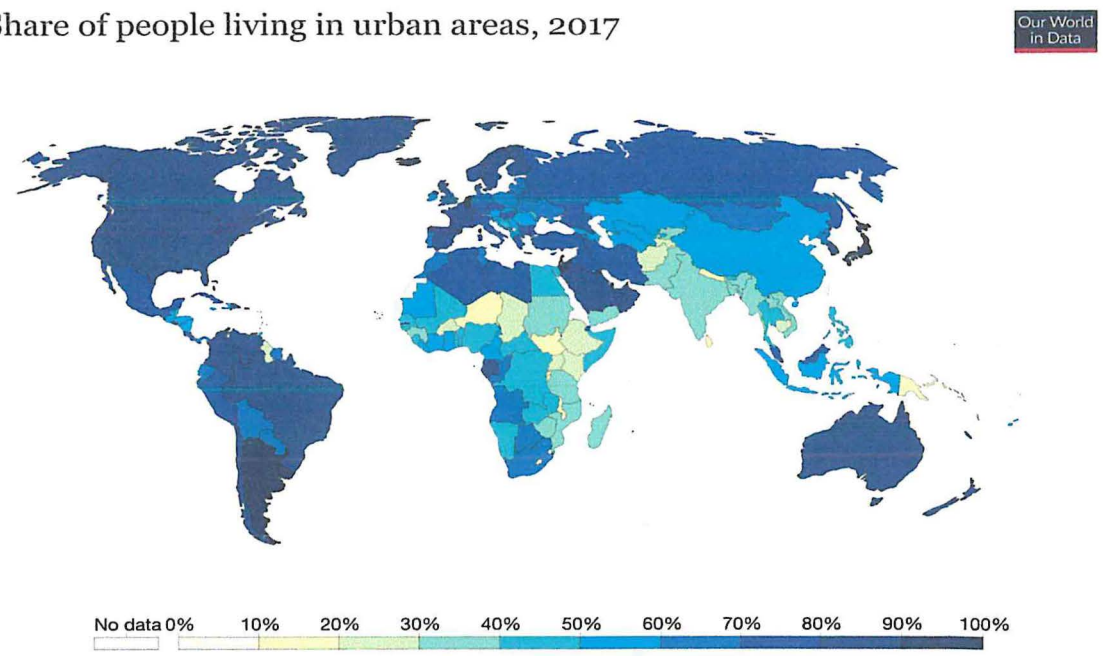

Figure 2) ***Percentage of population living in urban area in 2017.*** Showing geographical coverage from global to regional from United Nations *(2018)*. Source: Ritchie *(2018)*- "Urbanization". Published online at OurWorldInData.org. Retrieved from: *https://ourworldindata.org/urbanization*.

The prospects further highlight least developed countries as having a higher urban proportion annual average rate of change in 2015-2020 with 3.9%, followed by less developed regions with 2.9% and more developed regions with 0.5%. It also indicates the geographic proportion of the population in cities. Large cities with inhabitants of 5 to 10 million has doubled from 21 in the 1990s to 48 in 2018 and is estimated to grow to 66 by 2030 and house an approximate of 9% of the global urban population. While megacities have tripled from 10 in 1990 to 33 in 2018 and account for 13% of the world's urban population, with more than two-thirds being located in the developing regions.

It is implied that urbanization process is associated with the economic growth when looking at the global pattern (Chen M, 2014). This may be attributed to the mega-cities that have higher incomes and productivity levels that also attract populations due to jobs and wealth-creation. Moreover, mega-cities take advantage of the economies of scale and agglomeration that allow a conducive environment for businesses to grow (Daniels, 2004). However, there is a disparity between mega-cities socially and economically. The 'rich' mega-cities profit from the inclusion in global socio-economic and political networks that allows production and service centres in the global market. While 'poor' megacities are an absorption pool to those migrating from rural areas with a majority of the population living below the poverty line (Kraas, 2007). In many cities, the proportion of "urban poor" is between 30% and 60%, while in others it is

higher: in Addis Ababa (79%), in Luanda (70%) and Calcutta (67%) (Daniels, 2004). The "urban poor" characteristic tends to be common in countries in the Southern hemisphere or developing countries. According to Daniels (2004), people move to cities in search of wealth and job opportunities but in several cases labour supply in developing economies is usually greater than that demanded. This results in an increase in the informal sector as people search for jobs and is evident as the informal employment sector in Africa ranges from 54% to 3% or less in high income countries. Additionally, rapid growth of cities in the developing countries and the southern hemisphere are financially constrained to improve urban infrastructure to accommodate the growing urban areas (UN Habitat, 2015).

In terms of economic growth, large cities have been seen to contribute more to the country's GDP than others. In Latin America, the top ten cities in the region generate 30% of GDP, in China the top ten cities contribute 20% of GDP, while in India it is estimated that large cities will contribute 50% of GDP by 2025 (Cadena, Dobb, & Remes, 2012). According to McKinsey Global Institute (2011) 600 urban centres generate 60 per cent of global GDP and estimate 600 cities to generate 60 per cent of GDP by 2025, but not the same cities. Of the 600, the 380 developed regions accounted for 50 per cent of global GDP in 2007 with more than 20 per cent coming from 190 North American cities alone. The 220 largest cities in developing regions contributed another 10 percent which include the cities from China generating 4 percent and Latin America's largest cities contributing 4 percent. They expect 136 new countries to enter the top 600 all of the from developing regions such as China, India and Latin America. The projections for the contribution by the top 600 cities are as illustrated below in Figure 3.

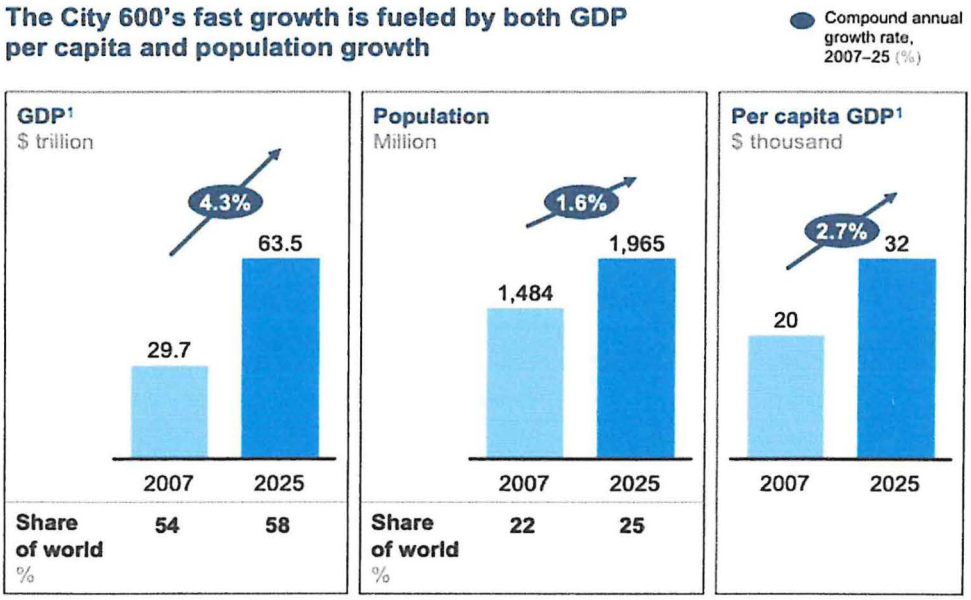

Figure 3) ***Projections of combined GDP from top 600 cities.*** *(McKinsey Global Institute, 2011).*

Taking into consideration that there is no "one size fits all" in terms of urbanization and the growth of cities, depending on where they are positioned to take advantage of the economic environment, some cities will continue to grow as others decline. Cadena (2012) highlights how in the United States, cities well positioned for industrialization led to the rise of the Iron Belt cities but turned to what is referred to as the Rust Belt after transitioning to a service economy. However, others have mitigated this by diversifying their industries and strategies to cater to a wider range of the population and services. Compared to developing regions who are still in different stages of development but can learn to manage the rapid urbanization from developed countries.

One of the characteristics of urbanization in developing countries is urban primacy. Urban primacy as defined by Mark Jefferson is a city which has accommodated a population more than twice of that of the second largest city of the country (Faraji, Qingping, Valinoori, & Komijani, 2016). It indicates to what extent economic development or activities are concentrated in in a few cities. However, urban primacy is not only in developing countries, but was present in developed European cities like London and Paris in the 19th century. Moreover, urban primacy is not present in all developing nations like China and India (Faraji, Qingping, Valinoori, & Komijani, 2016). Studies show that countries in Latin America have had higher rates of urban primacy compared to other developing regions as measured by the index of primacy. This index is calculated by the division of the population of the largest national city to that of the second largest city (Short & Kim, 2010). Examples of this include Lima, the capital of Costa Rica, Panama, Guatemala and the Dominican Republic. Lima was in 1940 8 times, and in 2000 10 times greater than the second largest city in Peru's Arkupya. While in 2012 Lima had 25% of the country's population of 30 million (Faraji, Qingping, Valinoori, & Komijani, 2016). Latin America's rapid population growth was among the highest in the middle of the 20th century with an annual growth of 2.8% in the 1960s. This slowed down in the 1970s as a result of a decline in fertility but the population continued to expand (Population Reference Bureau , 2003). However urban growth continues to expand as a result of rural to urban migration.

Even though countries in Latin America have high urban primacy, countries like Brazil and Columbia seem to have the lowest urban primacy index amongst the Latin America countries (Faraji, Qingping, Valinoori, & Komijani, 2016) as exhibited in Figure 4. This would seem reasonable as some countries may not need more than one large city due to the land mass area.

However, Brazil's uniqueness in this situation may follow the world-wide pattern where countries with large areas exhibit little to no primacy (Browning, 1989).

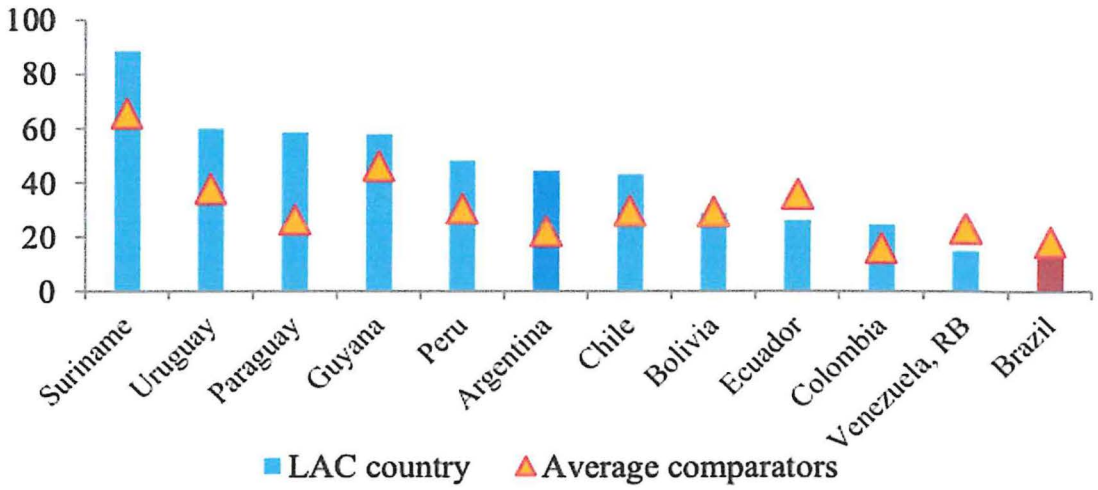

*Figure 4)**Rates of Urban Primacy in the Latin American Region.** Source: Reprinted from Raising the bar for productive cities in Latin America and the Caribbean (World Bank Group, 2018).*

Taking a closer look at Brazil, the most populous country in Latin America, with a population of 212 million (approximately 32% of the Latin America population) and largest land area of 8.5 million square kilometres, making it the fifth largest country in the world and sixth most populous in the world. Brazil's declining population growth is a result of decrease in births per woman from an average of 6 births per woman in the 1950s to 1.75 births per woman in 2018 (World Population Review , 2020). The country has two megacities, Sao Paulo and Rio de Janeiro, with an estimated population of 22 million and 13 million respectively (World Population Review , 2020). The country follows the high consistent rates of urbanization in Latin America with approximately 80% of the population living in urban areas by 2016 and was more urbanized than the United States of America, China and India (Ritchie & Roser , Urbanisation, 2020). Brazilian cities have among the highest population densities, 92.75%, above the global median. High densities come with negative effects of congestion that can deter productivity but may also allow for positive agglomeration economies as suggested by evidence (Ferreira & Roberts, 2018). According to Ferreira (2018) based on night-time lights, out of 10 cities with the highest economic outputs in the Latin America and Caribbean regions, 5 are Brazilian cities. However, the country's productivity still lags behind the global frontier and has high productivity dispersion across urban areas and high skill inequalities that lead to large disparities in income which may be as a result of high-skilled individuals residing in large cities.

This may lead to widening productivity gaps amongst the cities and resulting in lower contribution to the national productivity of the country. Moreover, governing bodies may be biased to aid in developing larger cities due to the potential for higher productivity compared to smaller or medium sized cities. Considering the complexity in the multicity agglomerations in Brazil, focusing on developing the different cities in their different specializations may benefit the country in light of their lag behind the global productivity frontier (Ferreira & Roberts, 2018). In light of the low urban primacy amongst other Latin American countries and being a developing economy, Brazil has experienced rapid urbanization growth compared to other large developed economies (Figure 5). Other large developed economies such as the United States had gradual urbanization accompanied by a relatively high gross domestic product to foster smooth growth and achieve a more spread-out economic activity (Henderson V. , Urbanisation in Developing Countries, 2002).

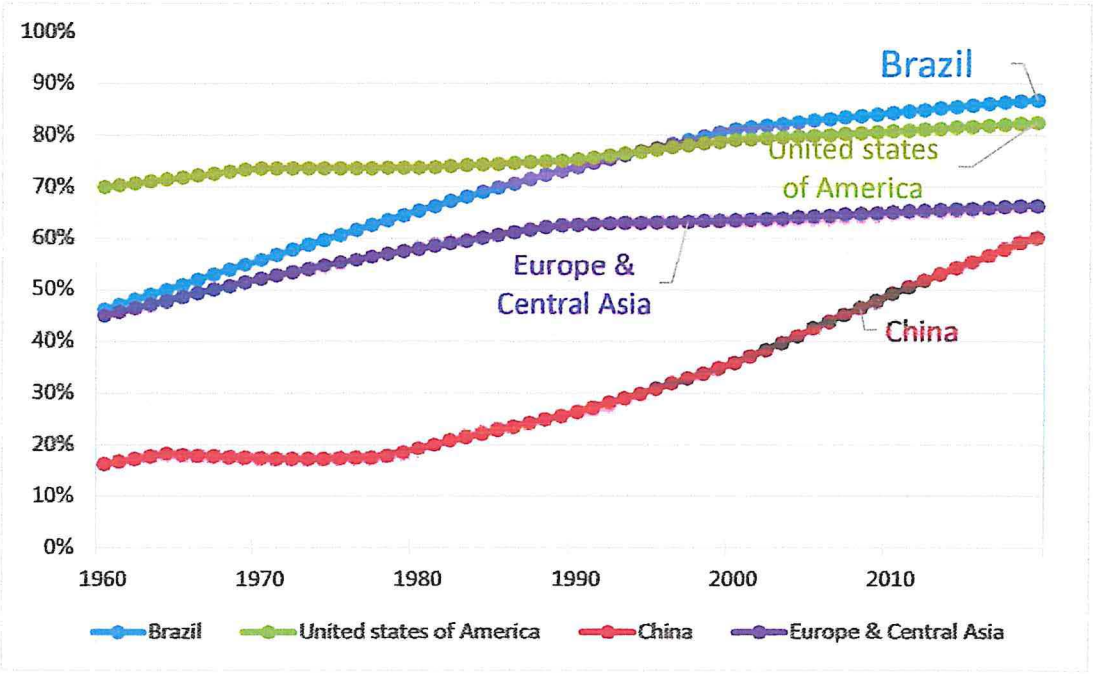

*Figure 5)**Percentage of urban population in the United States, Brazil, China and European Union.** Source: The World Bank (United Nations, 2017)*

1.2 Problem Statement

There are high global urbanization projections, 68% by 2050 (United Nations, 2019), with the highest urbanization rates being in developing regions. This could be as a result of pull factors such as employment, better institutions or facilities or more opportunities for individuals in cities. As firms cluster in these cities due to factors such as accessibility to the demand pool and complimentary facilities, they are able to enjoy economies of scale. These agglomerate

economies in cities has been seen to have a positive effect on economic growth (Nguyen & Nguyen , 2018), but up to a certain extent.

Problems facing developing economies involves rapid urbanisation that does not allow time for planning or financing the expansion of cities. This results in a strain on resources such as housing and proper distribution of clean water and sewage facilities, congestion and air pollution amongst others.

Chen et.al. (2014) found a positive correlation between urbanisation and economic growth at a global level but also show significant evidence of no correlation between urbanisation speed and economic growth. Brülhart & Sbergami (2009) agree with this, but on the aspect of agglomeration economies boosting economic growth but up to a certain extent when looking at the link between spatial concentration, urbanization and economic growth across countries. Delving deeper, Frick & Rodriguez-Pose (2017) looked at the link between city size and economic growth and their results show that growth is context dependent as different levels of urbanisation enhance growth in different sized countries. However, Henderson (2003) points out that growth was negative while rapid urbanization took place but rather the form in which urbanisation manifests itself is what would drive productivity growth.

However, not a lot of studies have been done to analyse how the cities and their sectoral composition affect growth. Therefore, this paper attempts to examine the existence of sectoral development of cities in Brazil by analysing the relationship between urbanization rates of cities and their contribution to GDP. Additionally, to see what factors may contribute to the disparity in the different cities.

1.3 Research Objectives

To analyse the impact of urbanization on economic growth, the following shall be carried out for the different cities in the country:

1. Analyse the relationship between urbanization and economic growth.
2. Assess whether there are significant moderating effects from Gross Added Values of different Sectors, literacy levels, minimum wage levels, years of schooling, people belonging to an urban labour union and economically active persons on the relationship between urbanization and economic growth

1.4 Significance of the Research

The relevance of the research is to study the impact that urbanization has on economic growth, especially on a city level. This is because cities have been known to be the drivers or engines of economic growth. In Brazil, a city like Sao Paulo contributes 10.6% of the total GDP, seven municipalities have contributed 25 % of the national GDP compared to 1,324 municipalities that contributed as low as 1% to the country's GDP as per 2017 (Souza, 2019).

However, this may lead to higher percentage of the population migrating to cities like Sao Paulo, looking for better opportunities, but may later result in adverse aftermaths such as congestion and inadequate infrastructure to support the rapid increase in population.

Therefore, in identifying the impact of urbanization on economic growth and identifying what attributes of agglomeration may drive growth, cities may be able to plan adequately for growth in terms of urban infrastructure and networks that can improve productivity and the quality of life of its inhabitants. Moreover, in analysing the growth rates and sectoral composition of cities, we may also identify the inequalities between the cities and the origin of the inequalities.

This may inform policy makers on areas that require more funding or investment, identify areas for further research and possibly highlight unique drivers of growth in the different cities.

CHAPTER TWO: LITERATURE REVIEW

2.1 Theoretical Literature Review

2.1.1 Urbanization Theories

Urbanization is no novel revelation, but what is new is the rapid rates of urbanization globally, especially in the developing countries. Urbanisation theories have evolved over time and tend to intertwine concepts regarding industrialization, globalisation and cities. In the pre-industrial era, a movement from the hunter and gatherers lifestyle in the Neolithic population, to utilization of inventions and tools to better farming practices resulted in aggregation of people in towns who did not grow their own food (Davis K. , 1995). This attests to the theory of self-generated urbanisation that requires two separate conditions: i) the generation of surplus products that sustain people in non-agricultural activities (Childe, 1950) and ii) the achievement of a social level development that allows large communities to be capable of successfully working alone (Lampard, 1961). However, Davis (1995) emphasises that growth of cities in the ancient world were limited by inefficiencies in their tools, but the integration of the two conditions were strongly felt during the industrialization era in the late 18th century, leading to the urban revolution in Great Britain, North America and Europe (Peng, Chen, & Cheng).

The theory supports the rural to urban shift as a source of urbanisation, highlighting industrialization as a driving force. As town or centres become more productive, they are able to trade surplus foods and develop enterprises that support trade. Hence, movement away from rural areas, where agriculture takes place to urban areas. Overlapping this is the Modernization theory that was predominant from the 1950s to 1970s, explaining the development of Western Europe and North America industrially. The theory highlights that urbanization takes place at the initial onset of modernization and that neither of the terms can be divorced, the importance of technological application in comparison to social organization and cultural diffusion that was higher with collectivist cultures like Japan than individualistic cultures such as the U.K and U.S.A[(Bodo, 2019); (Bond & Smith, 1996)]. Therefore, once modernization takes place, processes such as urbanization, bureaucracy, democracy and mass consumption follow.

Additionally, the modernization theory suggests that societies are more productive as a result of efficiency associated with use of technology, there is better education for children and more welfare for the poor (Reyes, 2001), which usually results in rural to urban migration in search of a better life or better living standards. These can be termed as push and pull factors that

have been brought forth by Ravenstein (1889) in *"The Laws of Migration"* that highlights the process of migration as a result of push-pull factors with pull factors emanating from favourable external opportunities. A reformulation of this theory by Everett Lee (1966) that points out the individual characteristics of a person and surrounding obstacles that may hinder migration. These could include, social ties or dependents, age, gender, political barriers, distance and social class amongst others. Moreover, urban bias theory argues that the disparities between urban and rural regions in terms of provision of resources is as a result of favouritism policies towards urban regions. Hence, resulting in rural to urban migration (Tetty, 2005).

In conjunction with the modernization development theories are theories like Rostow's five stages of development (Rostow, 1959). In this theory, Rostow (1959) highlights 5 stages that an economy must go through to develop: traditional society; pre-conditions for take-off; take-off; drive to maturity and high mass consumption. This theory may be used to study urbanization from the modernization. However, one problem with modernization theories is the homogeneity of its application to countries globally, considering it originated from Western countries like the USA. Moreover, it drives the idea of convergence between developed and developing countries even though this may not necessarily happen. Also in regards to urbanization, modernization theories have been criticised to fail to account for external factors that may limit urban growth or an explanation to causes and consequences in developing nations. This allowed for the dependency theory to arise.

Since the theory was proposed by an Argentine economist, Raul Prebisch, in the late 1950s, the dependency theory has been tackled and revised by many authors. However, it still stands to highlight the existence of a core and periphery economy (rich core and poor periphery), where the core controls or manipulates the growth of the periphery economy through investments or set policies. This stemmed from countries that were usually colonised and functioned as providers of raw materials to the now developed nations and would purchase the finished goods at higher prices. This theory would help comprehend underdevelopment and inequality in developing countries. Despite several contributions, four main theoretical propositions based on Angotti (1981) and Bath (1976)that highlighted: i) underdevelopment in the Third World countries is closely linked to expansion or advancements in the capitalist countries; ii) the division of the world into two parts consisting of the industrialized countries and periphery of underdeveloped countries and can also be found in internally; iii) international

trade between the core and periphery is an unequal exchange between, high-wage, high-profit and low-wage, low-profit; and iv) dependency conditions are more complex and deeper in terms of influence on policy making processes, culture elements and social structure (Namkoong, 1999).

The developed or core countries can influence urbanisation in the developing countries by exploiting or extracting natural resources or raw materials through foreign investments in those sectors. This would in turn either displace those settled on agrarian land and leave them to search for jobs in urban areas if they cannot find one. Or the setting up of factories to produce goods may leave residents to assume that there may be higher paying jobs in those locations, hence a rural to urban shift of labourers in search for higher incomes (Bradshaw, 1987). This may alternatively result in a rise in the informal sector when a mismatch of skills arises. The rise in industrialization as a result of external capital has been evident in cities such as Sao Paulo, Lagos and Buenos Aires.

Kasarda & Crenshaw (1991) highlight the global urban network, where "world cities" of developed countries use primate capitals of the developing countries to act as transmitters of wealth and in turn hinder equality of development that would allow them to profit, as well as maintain the unequal social structures that would allow them to prosper. Davis (1991) highlights the growth of Mexico City's import substitution industrialization through dependent development that resulted in rapid urbanization to the city. Moreover, the unequal trading relationship was evident as Mexico was offered low-interest rates to undertake major urban infrastructural changes that would benefit the developed lending countries.

Switching from the economic perspective to a more political perspective, the urban bias theory proposed by Lipton (1977) advanced the theory that poverty in rural areas was as a result of uneven distribution of resources between urban and rural areas. Assumptions of Olson's (1982) distributive coalition's theory that can be related to Lipton's (1977) urban bias theory include: i) Political institutions are more powerful in fostering development compared to other institutions; ii) these theories suggests that the alliance formed around self-interest are for the purpose of exerting pressure on the state or to dominate it and they usually involve diversion of social surplus to cater for special interests; and iii) the theories assume that coalition-building is not the soul property of any particular economic regime (Kasarda & Crenshaw, 1991).

These assumptions imply that political influence has the power to drive policies that will favour either the elite or those who are in coalition with the elite. This means that those whose self-interest lie in developing businesses in urban areas or would like to enhance activity in their business area, policies set will be in their favour. Sadly, this will be at the expense of the rural areas as resources are shifted from rural to urban even though a higher population of society may be residing in rural areas. Moreover, the urban areas enjoy produce that they charge minimal rates from the rural areas, leaving those in urban areas with higher profit margins and lower costs while those in rural areas carry the burden. As a result, there is a higher chance of rural to urban migration as people seek for better facilities and wages due to the disparity that exists between rural and urban areas.

2.1.2 Theoretical Linkages between Urbanization and Economic Growth

Taking a general look at the theories surrounding economic growth, several authors have improved upon the concept or considered a different perspective on the same topic. Starting off with Schumpeter (1934) who suggested that economic growth and development was driven by innovation and creativity of entrepreneurs based on the assumptions of efficient and existing markets that would support these new inventions (Pietak, 2014) as compared to the accumulation of capital as a driver of growth. However, Schumpeter's theory only considered countries that were democratic and developed economically. Arthur Lewis stepped in to explain the case for poor countries by highlighting the existence of a modern and traditional sector with surplus labour moving from the traditional to modern sector (Lewis, 1954). Assuming that labour wager rates remain low, the modern sector is able to save its surplus profits and re-invest to invest in capital to catapult economic growth.

Simon Kuznet supported this theory with "Kuznet's curve" that highlights how development from agricultural to modern sector results in increased income per capita that results in economic inequality but up to a certain extent (Kuznets, 1971). After which the economic equality starts to decline. Additionally, Kuznets recognised there was a positive association of economic growth and increased share of urban population (Pietak, 2014). Other theories and models of growth included Rostow's 5 stages of economic growth (Rostow, 1959) and models that focused on savings and investments such as the Harrod-Domar growth model (Harrod, 1939) & (Domar, 1946).

Considering the role that urbanization plays in the growth trajectory of cities in the coming years, it is an important aspect to look at. Black & Henderson (1999) examine how urbanization affects the efficiency of growth process and how growth influences urbanization patterns for countries experiencing endogenous economic growth and exogenous population growth. They found that the growth of urban centres is as a result of local human capitalization and knowledge spill overs which "promotes agglomeration and human capital accumulation fosters endogenous growth." This was supported by Henderson, Shalizi & Venables (2001), Krugman and Quigley (2009) who were in agreement that firms could benefit from locating in areas with highly skilled labour that would eventually result in attracting greater population and firms to urban centres.

Henderson's extensive work in the field has led to literature that explains the spatial clustering of economic activity in an area as well as the relationship between urbanization and economic development, factors that contribute to agglomeration, linkages within the same industries and movement of the factors of production and an advancement in the explanation of growth in relation to national policies and urban concentration through data and tests (Singh, Swami, & Prasad, 2014). This has been able to explain further how different aspects of urbanization can affect economic growth down to the level of industry participants and city centres when it comes to agglomeration.

These cost advantages presented to consumers and producers largely from the formation of cities are known as agglomeration economies. They come in two forms: urbanization economies that are effects that arise as a result of general growth of a concentrated geographic region, while localization economies are effects that are captured by a specific sector in the economy (Todaro & Smith , 2014). Quigley (2009) highlights the external effects of urbanization and the roles that cities play when it comes to growth. These external effects of agglomeration include: specialization whose gains are enhanced and contribute to lower transaction costs and better complementarities in the industry as well as the concentration of skilled workers in an urban area fostering externalities of productivity. Moreover, the law of large numbers allows for a large amount of cost savings.

Puga (2009) highlights the theoretical literature by Duranton & Puga (2004) to explain how urban agglomerate economies come to exist. The three classifications put across are: i) a larger market allows for sharing of facilities and infrastructure and skilled labour and workers that would result in greater efficiency; ii) better matching is achieved between buyers and suppliers, employees and employers amongst others in larger markets; iii) larger markets facilitate learning through adoption of technologies and new innovations. However, it has been noted that moving from the theoretical to modelling these classifications is not as direct. However, it has been found that agglomeration and inequality have an influence on long-run economic growth. Castells-Quintana & Royuela (2014) show that high-inequality levels limit economic growth in the long-run even though high-levels of urban concentration have a relationship with economic development which may be as a result of high urban economic activity associated with agglomerate economies. They also point out that growth in low-income countries can be as a result of increasing agglomerate economies be it from increasing urbanisation or increased urban concentration. Furthermore, low-inequality countries are more likely to benefit from

increasing agglomeration economies whether they are low-income or high-income countries (Castells-Quintana & Royuela, 2014).

Agglomeration economies are known to come from the formation of cities and many studies have shown the positive relationship between urbanization rates and economic growth. Historically, it has been evident that the presence of large cities resulted in high economic growth and development as in Rome or Alexandria and in the 1950s for high-income countries, as low-income countries did not have any large cities apart from a few exceptions like China, Mexico and Brazil. However, Frick & Rodriguez-Pose (2017) take on a different angle by focusing on whether the city size does contributes to economic growth and whether additional contextual factors such as governance effectiveness, sectoral composition and sanitation affect growth as well. The study found a non-linear relationship between city size and growth dependent on the country's size. Additionally, there was a threshold on urban population living within a city that would contribute to growth, past which it would be detrimental due to the diseconomies that may arise from congestion and large informal sectors. Therefore, in some countries, smaller cities may be better drivers for growth compared to large ones. Although, they also found that contextual factors such as government effectiveness, urban infrastructure and the share of industries that benefit form agglomerations also has an impact the growth drivers for the city, emphasising that context is an important factor during analysis.

Still in the scope of cities and growth, Duranton (2014) digs deeper to examine the productivity of workers in agglomerate economies and cities. Their results show that workers are productive in larger cities in the long-run and the short-run and link this to the possibility of higher wages and the presence of innovation and entrepreneurship. Moreover, the productivity of workers is also related to the functioning specialization of firms. He highlights how agglomerate economies focus on doing different things differently, bringing about the aspect of sector composition or specialization. Implying that larger cities should not aim to do everything as it may not add to diversification workers are looking for to be productive and it may also hold back small cities. However, there is limited literature on sectoral composition of cities and its impact on growth.

2.2 Empirical Literature review

To study the effect of urbanization on economic growth, researchers have been able to measure and study urbanization and its different facets on a county or city level, and the effect it has on

economic growth. A vast array of data in past studies has ranged from 80 to 100 countries as well as different continental regions. However, the studies seem to arrive at a similar conclusion that urbanization and economic growth do have a positive relationship.

Brülhart & Sbergami (2009) examined the impact of spatial concentration of economic activities or agglomeration growth within a country on economic growth of the country. They used cross-section, panel data from 84-105 countries from continents such as Africa, Asia, Europe, Australia and South America and a European sample that includes sectoral information spanning between 1960-2000 (divided into 8 five-year intervals). They move on to carry out their analysis on both the cross-sectional and panel data to arrive at a conclusion. They begin by testing two hypotheses: 1) The "Williamson hypothesis" that GDP is boosted by agglomeration growth but only up to a certain extent. 2) The "openness hypothesis" that examines how an increase in trade openness, how open the country is to international trade, weakens growth-promoting effects brough about by intra-country agglomeration.

First, to analyse the cross-sectional data, a cross-section OLS regression is carried out for estimation with per capita GDP growth rate as the dependent variable, with the explanatory variables including an agglomeration variable and control variables based on Sala-i-Martin et.al. (2004). While the panel regressions make use of "system GMM" or SYS-GMM to examine growth effects on the agglomeration proxies on the world sample. For both datasets, the number of countries used varies depending on the variables that are available. The results show consistency with the "Williamson hypothesis" implying that agglomeration benefits do accrue to economic growth but up to a certain level, approximately 10,000 US dollars in 2006 prices which is close to the per capita income of Brazil or Bulgaria. However, the estimations did not consistently support the openness hypothesis.

Another study by Chen et.al. (2014) look at urbanisation and economic growth on a global perspective and includes cross-section and panel data from 226 countries and regions over a span of 30 years (1980-2011) from countries in South America, South East Asia, Africa, North America and Australasia. They also examine whether urbanization speed does have an effect on economic growth rate. Gross domestic product (GDP) per capita and level of urbanization are used to calculate the speed of urbanization and economic growth rate in order to investigate the difference in correlation between the level and speed of urbanization and economic growth.

Chen et.al. (2014) run a correlation analysis in the above mentioned using both the cross-sectional and panel data, they use a pooled model, a random effects model and a fixed effects model to analyse the hypothesis. In addition to this, they use a geographic information system to visualize global spatial data on urbanization levels from 1980-2011. The study finds that there is a correlation between urbanization and economic growth but they suggest that factors such as urban concentration and agglomeration may have aided in inducing the growth. However, they do not find any significance in the correlation between urbanization speed and economic growth.

Narrowing down to the city level, Frick & Rodriguez-Pose (2017) study the relationship between the size of a city and economic growth on a national level, of which city size is considered as the percentage of urban population living in the city. The paper seeks to find if there are particular city sizes that enhance growth and whether other factors such as agglomeration influence the city size and growth relationship. Panel data of 113 countries between 1980-2010 is used to test if there is a relationship.

An extended econometric Solow growth model is employed based on a GDP per capita growth equation and the growth rate of the country is estimated in five-year periods as a function of the size of a country's cities as well as other control variables usually used in growth regressions. (Frick & Rodriguez-Pose, Big or Small Cities? On city size and economic growth, 2017). They model the data using the Hausman Taylor estimator and include an analysis of country fixed effects as well as stress tests to test for robustness. They also include data on sectoral composition and contextual factors to test their effect on growth. Findings from the study show that there is a non-linear relationship between city size and economic growth. However, this growth is dependent on the city size as a larger urban population in small cities with less than 500,000 people enhances growth in small and medium sized countries. While in medium sized cities with a population between 500,000 and 3 million in small countries, growth is inhibited. Moreover, large populations in megacities (a population beyond 10 million) only enhances growth in large cities. However, these results are also context dependent on factors such as urban infrastructure, government effectiveness and an environment that allows for benefits of agglomeration to be enjoyed.

Lastly, based entirely on South East Asian Nations, the relationship between urbanization and economic growth is tested from 7 ASEAN countries: Thailand, Vietnam, Cambodia, Brunei,

Malaysia and Indonesia (Nguyen & Nguyen , 2018). Panel data from the selected countries from 1993-2014 includes data such as the per capita income and the proportion of urban population to total population as well as information regarding age structure and population density.

Nyugen et.al. (2018) use a static and dynamic model to analyse the impact of urbanization on economic growth. For the static panel data regressions, Fixed Effects, Random Effects, Driscoll and Kraay are used. For the dynamic panel data regressions D-GMM and PMG are used. Additionally, Granger causality testing is also done. Nyugen et.al. (2018) find that there is a non-linear relationship between urbanization and economic growth for the ASEAN countries as well as causality between the two factors. For 3 out of the 7 countries there was a two-way causal relationship, for 2 out of 7 there was a one-way causal relationship (urbanization impacting growth) and for another 2 countries the one-way causal relationship was the opposite (growth impacting urbanization). However, they also highlight a threshold past which growth is hindered (69.9%-static model; 67.9%-dynamic model).

The studies highlighted above from a global to a city perspective arrive at the conclusion that urbanization does have an effect on growth, be it from agglomeration benefits or urban concentration. Nevertheless, these studies also highlight a threshold or extent to which urbanization is beneficial to the country, past which it becomes detrimental.

CHAPTER THREE: RESEARCH METHODOLOGY

3.1 Research design

This exploratory study assesses the impact of urbanization on economic growth. It explores the relationship between similar variables, in addition to urbanization, from a city level on the national output of the country over the years using secondary data. The study shall also explore how similar variables from the different cities vary over time and whether it has any effect on the contribution to the country's economic growth.

3.2 Population and Sampling

The study will analyse data from the top 5 Brazilian cities that contributed the most to the GDP of the country as per 2017 (Souza, 2019). These cities include São Paulo, Rio de Janeiro, Brasília, Belo Horizonte and Curitiba. The annual data collected from the five cities covers 16 years and spans from 1999-2015.

3.3 Data collection

To analyse the impact of urbanization on economic growth using panel data, Gross Domestic Product (GDP) per capita is used as the dependent variable with the independent variable as urbanization. Additionally, agglomeration variables are included to assess the sectoral composition of cities and other variables such as literacy are included. All the data was collected from the Brazilian Institute of Geography and Statistics, originally known as the Instituto Brasileiro de Geografia e Estatistica (IBGE).

The variables are summarized in the table below:

Variable	Measurement	Source
Gross Domestic Product per city	The sum of all final services and goods produced by a city in a year	IBGE- National Accounts
Gross Domestic Product per capita	The sum of all final services and goods produced by a city in a year and accounts for the population in that city	IBGE- National Accounts

Population	Number of inhabitants in the country/city	IBGE- Population and Population Census
Urban population	The percentage of the total population residing in urban areas	IBGE- Population and Population Census
Urbanisation rate	Calculated by dividing the population residing in the urban areas over the total population in the same area	
Gross Added Value of Agriculture	The value of the services and goods produced in the Agricultural sector	IBGE- National Accounts
Gross Added Value of Industry	The value of the services and goods produced in the Industrial sector	IBGE- National Accounts
Gross Added Value of Services	The value of the services and goods produced in the Services sector	IBGE- National Accounts
Literacy	The ability of a person to read and write	IBGE- Population and Population Census
Years of Schooling	The number of years an individual has attended formal schooling (more than 8 years)	IBGE- Population and Population Census
Economically active *(work)*	Persons aged 10 and over and economically active in the week of reference	IBGE- Population and Population Census
Wage	Value of average monthly income of persons aged 10 and over	

Table 1 Variables to be considered in the models

3.4 Data Analysis

To analyse the data, we apply the following panel data analysis techniques mentioned below after testing for stationarity with the following unit root test:

3.4.1 Unit root tests

For this study, we shall use the Fisher's type test to test for stationarity in our panel data. Previous tests have had stringent conditions that would have the alternative hypothesis assume that all individual series in the panel would be stationary if the null hypothesis does not hold. This may not be the case as some of the individual series may be stationary while others non-stationary.

Therefore, Maddala & Wu (1999) and Choi (2001) propose an approach that combines the p-values of N-cross-sectional unit root tests. This results in the following null and alternative hypothesis:

$$H_0: \delta_i = 0 \rightarrow for\ all\ i = 1,2, \dots, N$$

$$H_1: \delta_i < 0 \rightarrow for\ all\ i = 1,2, \dots, N_1$$
$$\delta_i = 0 \rightarrow i = N_1 + 1, N_1 + 2, \dots, N$$

where the null hypothesis shows presence of unit root for all i cross-sections for each series, while the alternative allows for some (but not all) of the individual series to have unit roots.

Furthermore, to overcome the restrictive and inflexible nature of previous tests, Choi (2001) proposes a test based on combining p-values. These p-values match the test statistics from each unit root tests applied to each cross-section. The proposed Fisher type test is as follows:

$$P = -2 \sum_{i=1}^{N} \ln p_i$$

where p_i is the p-value of a unit root test for cross-section i and P is asymptotically chi-distributed. This test can handle unbalanced panels and is also the best because it can distinguish the null and alternative hypothesis at the highest power.

3.4.2 Fixed Effects Model

The fixed effects model employs use of Ordinary Least Squares (OLS) in the cross-section regressions, where the GDP per capita of the city is estimated as a function of the urbanization of the city i at time t as shown below:

$$Y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \qquad (3.1)$$

Where the independent variable X_{it} captures the urbanization rates of city i at time t. In this model the error term ε_{it} is split into two:

$$\varepsilon_{it} = \mu_i + v_{it} \qquad (3.2)$$

Where μ_i captures variables that would affect the variable Y_{it} cross-sectionally but not varying with time, a variable such as culture of the city. These variables can be encompassed through dummy variables. While v_{it} captures the unobserved factors varying across time and cities. Therefore, rewriting equation 3.1 as:

$$Y_{it} = \alpha + \beta X_{it} + \mu_i + v_{it} \qquad (3.3)$$

The fixed effects model aims to take care of the unobserved heterogeneity captured by μ_i to ensure that the estimates produced are the best linear unbiased estimator (BLUE). This can be done through two main methods: within transformation and first difference estimation. Within transformation works to eliminate μ_i by subtracting the time-mean of each cross-section from the variable. This eliminates μ_i and we are able to estimate β, which is unbiased under the assumptions that the model is linear, independent, strictly exogenous and no there is no perfect collinearity in small samples.

The first difference estimation is able to exclude μ_i by subtracting the first period's observation of an individual from the second period's observation of the same individual, allowing us to estimate β. Once both methods have eliminated the unobserved heterogeneity, β is estimated through OLS. However, both methods may remove explanatory variables from the equation if they do not vary with time. This can be corrected by interacting such terms with time-varying variables.

3.4.3 Random Effects Models

This model uses Generalised Least Squares (GLS) and is also referred to as the error components model. It also captures the different cross-sectional components of the city i at time t as the independent variable, with the city's GDP per capita as the dependent variable, Y_{it}. The Random Effects model can be written as:

$$Y_{it} = \alpha + \beta X_{it} + \varphi_{it} \qquad (3.4)$$

where

$$\varphi_{it} = \epsilon_i + v_{it} \qquad (3.5)$$

In this case, the ϵ_i term incorporates heterogeneity of the variables cross-sectionally. However, this model also requires the assumption that the expectation of the error term, ϵ_i, be independent of the explanatory variable X_{it} and have a mean of zero holds.

3.4.4 The Hausman Test

In order to identify which model to use between the Random Effects model and the Fixed Effects model, the Hausman test is carried out. The null hypothesis for the test follows the Random Effects model assumption where the error terms, ϵ_i and v_{it}, are both supposed to be independent of the explanatory variable X_{it}. While the alternative hypothesis follows the less stringent assumption of the Fixed Effects model, where the expectation of the μ_i term and the independent variable is not equal to zero as summarised below.

$$H_0 = \mathrm{E}[\mu_i, X_{it}] = 0$$

$$H_1 = \mathrm{E}[\mu_i, X_{it}] \neq 0 \qquad (3.6)$$

This highlights one of the disadvantages of the Random Effects model, as its strict assumptions require φ_{it} to have no correlation with the dependent variables, which may not always be the case.

3.4.5 Dynamic Panel VAR

Due to the possibility of high interdependencies or interactions among the variables, we consider a dynamic panel vector auto-regressive model. This can be represented by the model below if we assume the number of variables used are two:

$$Y_{1it} = \beta_{10} + \beta_{11}Y_{1it-1} + \alpha_{21}Y_{2it-1} + \mu_{1i} + \varepsilon_{1it}$$

$$Y_{2it} = \beta_{20} + \beta_{21}Y_{2it-1} + \alpha_{11}Y_{1it-1} + \mu_{2i} + \varepsilon_{2it}$$

The following variables may be determined by Ordinary Least Squares (OLS) or fixed effects but will be biased due to the presence of lagged variables. Therefore, Generalized Method of Moments (GMM) estimation is used as it has been proposed to produce more consistent estimates. In this case, we shall consider the following models: Arellano & Bond (1991) Arellano & Bover (1995) & Blundell & Bond (1998). The Arellano & Bond model, also referred to as the difference GMM, first eliminates the individual fixed effects by first differencing then uses GMM to get rid of the present endogeneity. After the endogeniety is corrected, lagged variables of the dependent variable can be used as instrument variables from

the second and third lags; the lagged levels. However, by using the first difference, the Arellano & Bond model magnifies the gaps in unbalanced panel data. Therefore, the Arellano-Bover (1995)/ Blundell-Bond (1998) ,also known as system GMM, improves this by using forward orthogonal deviations (FOD) transformation (Arellano & Bover , 1995) to eliminate the individual fixed effects.

FOD transformation works by subtracting the average value of all available future observations from the current value, while dropping the last observation for each individual. This also allows the model to use both lagged levels and lagged differences as instruments, implying more instruments can be used for the model that can improve efficiency.

Furthermore, GMM estimation can either be one-step or two-system GMM. Where one-step assumes homoscedastic errors while two-step GMM corrects for patterns of heteroskedasticity and cross-correlation between error terms. These models will allow us to solve for problems associated with heteroskedasticity, serial correlation and endogeneity. However, the weaknesses associated with system GMM concern the number of instruments used, as more instruments may lead to misspecification of the model.

The diagnostic tests used to test the validity of the dynamic panel data model are the Test for Autocorrelation of the residuals in the differenced equation and the Sargan/Hansen test of overidentifying restrictions that looks at whether instruments are exogenous. In both cases, a higher p-value is preferred compared to the usual condition of a p-value that is less than 0.05, where serial correlation is only allowed in the first order for the test of autocorrelation.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Descriptive Statistics

A general descriptive analysis of the variables to be used are reported in both *Table 2* and *Table 3*. *Table 2* shows that all cities have the following commonalities: high urbanisation rates ranging from 84% to 96%; high literacy rates ranging from 93% to 94%; low gross added value from the agriculture industry and a moderately low contribution from the industry sector as well compared to the high contribution from the service sector that contributes 62% to 71.1% of the city's GDP. Furthermore, the percentage of economically active people *(work)* in all the five cities falls within a small range of 57% to 65%. This also applies to the years of schooling that falls within a range of 52% to 62%, with the exception of Sao Paulo at 95%.

state	ln_GDP~a	Urbani~n	AgroR	IndR	ServR	School~g	Lit	lnwage	Urbanemp	Work
Belo Horizonte	9.646866	.84066	.0080527	.2001374	.6345251	.5267134	.9301668	3.533689	.8429305	.655122
Brasilia (DF)	10.49627	.948897	.0037311	.0614495	.7111755	.623608	.9396068	3.274033	.8634418	.6452507
Curitiba	9.898047	.8460323	.0002751	.235184	.6029933	.5718648	.9433479	3.229856	.7627202	.6544196
Rio de Janeiro	10.00575	.9671532	.0002898	.1559887	.6259088	.5892883	.8902065	3.218059	.7191955	.5770279
Sao Paulo	10.19831	.9497136	.0000738	.2062163	.6323088	.951365	.9404531	3.537847	.7919604	.6366847
Total	10.04905	.9104912	.0024845	.1717792	.6413807	.6525663	.9287562	3.358539	.7960499	.633701

Table 2 Mean of all variables by state

On a more general level, according to the mean presented in *Table 3,* there is high urbanisation, high literacy and a high percentage of workers belonging to an urban employment union. Moreover, the service sector is the highest contributor to the cities' GDP, followed distantly by the industrial sector and the agricultural industry contributing the least. Additionally, we see high standard deviations and variance for total population, log of GDP per capita, years of schooling and log wage. For kurtosis and skewness, we have variables like literacy and years of schooling being skewed.

stats	ln_GDP~a	Urbani~n	AgroR	IndR	ServR	School~g	Lit	lnwage	Urbanemp	Work
mean	10.04905	.9104912	.0024845	.1717792	.6413807	.6525663	.9287562	3.358539	.7960499	.633701
p50	10.08269	.9424588	.0003103	.1523262	.6468439	.5881367	.942795	3.637586	.8093094	.6341379
sd	.6394638	.0576376	.0045745	.0947741	.1177769	.6193314	.0990415	.9136316	.0753926	.0486413
variance	.408914	.0033221	.0000209	.0089821	.0138714	.3835714	.0098092	.8347227	.005684	.0016517
skewness	-.1838343	-.4969055	2.644312	.9394636	.6715119	8.255691	-8.113088	-1.47242	-.4425277	-.0418507
kurtosis	1.979289	1.675134	9.394164	3.316228	4.622191	79.48428	68.89051	5.961539	2.227405	2.552486
N	85	85	85	85	85	75	75	75	79	75
sum	854.169	77.39175	.2111829	14.60123	54.51736	48.94247	69.65672	251.8904	55.72349	47.52758
range	2.436297	.2031931	.0202824	.3912893	.5485833	5.619355	.9684907	4.317488	.319427	.1617874
min	8.775131	.7784268	.0000319	.046365	.377952	.2787508	.0949634	0	.5987682	.5553876
max	11.21143	.9736199	.0203142	.4376543	1.026455	5.898105	.9634542	4.317488	.9181952	.717095

Table 3 Summary statistics of all variables

Figure 6 below presents the trends of the variables over time with respect to the different cities. Log of GDP per capita for all the cities has been on a gradual increase while urbanisation has

been quite stable past the year 2000. However, the gap shows that Curitiba and Belo Horizonte do not match the high urbanisation of the other three cities even though it is still highly urbanised. As for the gross added value of the service industry, its proportion to GDP has been steady for the cities until 2011 when there was a small decrease with the exception of Brasilia, that shows a dramatic decline over the years and a sudden drop on 2010. Furthermore, all cities have had a small and gradual increase in the percentage of economically active persons with a slight sharp increase in 2011. However, Rio de Janeiro seems to lag behind in this variable as can be seen in the graph even though it has the highest urbanisation rate. These factors would be interesting to carry out further investigation.

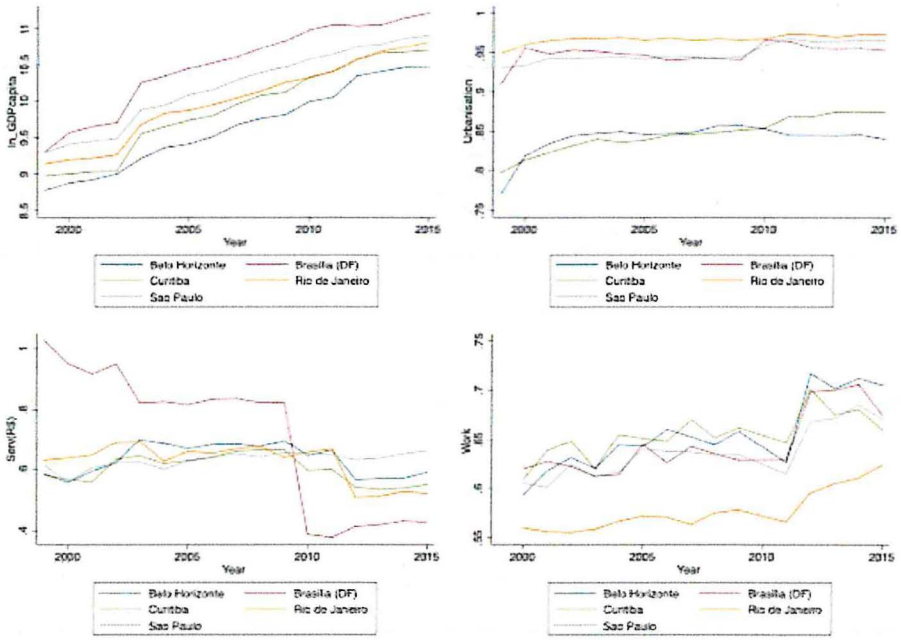

Figure 7 Variables trend over time

4.2 Regression results

This section highlights the results from the regressions using both the Random Effects model and the Dynamic Panel VAR. In both models, log of GDP per capita is used as the independent variable with close attention to urbanisation as the main independent variable.

Prior to running the models, all the variables are tested for presence of unit root (stationarity) by using the Fisher-type test of Dicky Fuller and Phillips- Perron as suggested by Choi (2001)and Maddala & Wu (1999). All the variables are deemed to be stationary except for log of GDP per capita, Gross Added Value from agriculture, industry and services and work. After

these variables are differenced the Fisher type test show them to be stationary and allow for estimations to carried out.

4.2.1 Static model estimation and results

The Hausman test is used to aid the selection of a suitable estimation technique (either Fixed Effects or Random Effects) for the data being used. The null hypothesis of the test favors the selection of the Random Effects model and concurs with the strict assumption that the error terms are to be independent of the explanatory variable while the alternative hypothesis favors the Fixed Effects model. The test results give a Chi2=1.99 (Prob>Chi2=0.9916), thus we fail to reject the null hypothesis at a 5% level of significance and select the Random Effects model as our model of estimation.

The results of the regression as highlighted in *Table 4* show the total number of observations ranging from 65 to 70 depending on the variables selected, with urbanisation representing the urbanisation rates in the city while agriculture, industry and service represent the agglomerate effects; the remaining variables act as control variables. Column (1) presents our base model, showing all variable estimates in the model, with agglomerate variables being statistically significant at a level of 5% but urbanisation showing significance at a 10% level of significance. However, the agglomerate variables exhibit a negative relationship with the growth rate of GDP per capita where agriculture exhibits the largest negative effects. This suggests that an increase in the ratio of the city's GDP presented by agriculture, results in a decrease of the growth rate of GDP per capita implying that agriculture may not be as value adding as compared to the industrial or service sector.

Dependent variable: GDP per capita growth rate	(1)	(2)	(3)
Urbanisation	0.0772*	-30.83***	-44.62***
	(0.0429)	(4.563)	(13.82)
Agriculture	-28.70***	24.47	
	(5.031)	(58.43)	
Industry	-2.026***	2.342	
	(0.161)	(2.265)	
Service	-1.561***	-2.642	
	(0.390)	(4.910)	
Work	-0.773***	-0.0569	
	(0.234)	(1.853)	
Schooling years	-0.00728***	2.769**	4.467**
	(0.00121)	(1.315)	(2.106)
Literacy	-0.00703	-30.87***	-42.25***
	(0.0180)	(4.402)	(12.87)
lnwage	0.00710	0.0145	
	(0.00537)	(0.0876)	
Urban labour union	0.161***	-2.905**	-7.429**
	(0.0453)	(1.421)	(3.083)
Urb_Industry		-4.758*	
		(2.450)	
Urb_Agriculture		-59.50	
		(69.12)	
Urb_Service		1.138	
		(5.386)	
Urb_Work		-0.841	
		(1.991)	
Urb_Schooling years		-2.934**	-4.734**
		(1.391)	(2.226)
Urb_Literacy		31.86***	43.64***
		(4.550)	(13.29)
Urb_Urban labour union		3.166**	7.784**
		(1.508)	(3.366)
Urb_wage		-0.0144	
		(0.101)	
Constant	-0.130***	29.87***	43.29***
	(0.0280)	(4.387)	(13.41)
Observations	65	65	70
Number of states	5	5	5

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4 Regression results from the Random Effects model

Furthermore, the control variables found to be statistically significant (*work, urban labour union and schooling years*) have a relatively low economic impact on GDP per capita growth rate even though an increase in the percentage of economically active persons and persons who

have attained more than eight years of schooling results in a reduction of the GDP per capita growth rate.

The second model, Column (2), shows an increase in the R-squared from 0.7854 to 0.8157, implying better explanation of the sample variance in the growth rate with the introduction of the interaction terms. Moreover, urbanisation is now statistically significant at a 5% level but now displays a negative relationship with GDP per capita growth rate. Column (2) shows that an increase in urbanisation rate results in a decrease in the GDP per capita growth rate by 31%. Theoretically, the expected relationship may be positive, but from the empirical studies of Brülhart & Sbergami (2009) and Nguyen et.al. (2018) they highlight a threshold beyond which urbanisation or agglomeration no longer boost GDP, also referred to as the Williamson hypothesis. This may also be due to the population growing at a much faster rate than the GDP of the city resulting in lower GDP per capita.

Additionally, all agglomeration variables are no longer statistically significant while the control variables (*schooling years, literacy and urban labour union*) are significant. An interesting observation is the interaction term between literacy and urbanisation. The effect of literacy given a higher urbanisation rate would result in a positive impact on the GDP per capita growth rate, implying that literacy may play a role in the productivity of an individual or chances of earning an income in the economy. Additionally, given high urbanisation, years of schooling and the industrial agglomerate variable would impact GDP per capita growth rate negatively but would have a relatively low economic effect even though statistically significant. Similarly, the ratio of employees belonging to an urban labour union given high urbanisation would also have a low but negative economic impact on the dependent variable.

When the statistically insignificant variables are dropped to form the third model, column (3), the resulting R-squared significantly drops to 0.1209 even though the economic significance of the statistically significant variables increase. This implies there may be high correlation of the dependent variables with the regressors that resulted in higher a higher R-squared previously.

4.2.2 Dynamic model estimation and results

To further analyse the relationship in question, the Differenced-General Method of Moments (D-GMM) and the System-General Method of Moments (SYS-GMM) estimation techniques

brought forth by Arellano & Bond (1991) and Arellano & Bover (1995)/Blundell & Bond (1998) are utilised. These methods help to correct for serial correlation, endogeneity and heteroskedasticity in the dataset that may bias the estimates from the regressions.

The panel still consists of the five states with data from 1999 to 2015, with an introduction of the lagged dependent variable as an explanatory variable. This applies across all models presented in *Table 5*, all consisting of lagged variables of the dependent variables from time t-2 to t-5 as instruments together with all other independent variables. The disparity in the number of observations arises from the D-GMM using the difference method to eliminate the unobserved heterogeneity, resulting in increased gaps in the data, while SYS-GMM avoids this issue by using forward orthogonal deviations. This allows the SYS-GMM to have more instrumental variables from the lagged level and lagged differenced variables compared to D-GMM.

Table 5 reports the estimates of all original variables in Column (1) and (3) using D-GMM and SYS-GMM respectively, whereas Column (2) and (4) reports the additional terms interacted with urbanisation rates to better explain urbanization impact. All the models produce valid instruments in regards to the Hansen test with a Prob>chi2=1.0 after running robust models.

Dependent variable: GDP per capita growth rate	(1)	(2)	(3)	(4)
	D-GMM		SYS-GMM	
L.dln_GDPcapita	0.0773	0.344*	-0.587*	0.445
	(0.170)	(0.201)	(0.317)	(0.380)
Urbanisation	0.584	21.60	-0.182	-10.22
	(1.226)	(19.56)	(0.136)	(21.60)
Agriculture	-20.12*	-598.1***	-23.66***	-592.2***
	(10.49)	(77.48)	(7.172)	(117.3)
Industry	-0.752***	-6.056	-0.0838	-11.19
	(0.174)	(3.914)	(0.246)	(7.262)
Service	-0.732***	6.905	-0.624**	8.219
	(0.132)	(5.719)	(0.258)	(6.856)
Work	0.543**	10.62	-0.394	12.69
	(0.262)	(10.19)	(0.327)	(10.96)
Schooling	-0.0116***	-29.79***	-0.00306	-13.14
	(0.00241)	(7.991)	(0.00405)	(8.891)
Lit	-0.0317***	46.76**	0.0483**	4.186
	(0.0115)	(23.10)	(0.0239)	(27.54)
lnwage	-0.00208	0.167***	-0.00800	0.107**
	(0.00682)	(0.0480)	(0.00560)	(0.0463)
Urbanemp	-0.441*	-10.03***	0.0220	-7.763***
	(0.248)	(2.423)	(0.107)	(1.271)
Urb_Industry		5.088		10.91
		(4.329)		(7.718)
Urb_Agriculture		669.7***		665.8***
		(84.49)		(135.4)
Urb_Service		-7.761		-9.168
		(6.058)		(7.235)
Urb_ Schooling years		31.48***		13.88
		(8.449)		(9.402)
Urb_Work		-10.87		-13.19
		(11.01)		(11.69)
Urb_Literacy		-48.30**		-4.384
		(23.84)		(28.42)
Urb_Urban labour union		10.80***		8.353***
		(2.659)		(1.344)
Urb_wage		-0.186***		-0.124**
		(0.0532)		(0.0531)
Constant			0.275*	9.598
			(0.164)	(21.02)
Observations	35	35	45	45
Number of states	5	5	5	5

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 Dynamic Panel Data Regressions

However, only the model in Column (2) passes for validity under the Sargan test while we reject the null hypothesis ($H_0 = valid\ overidentifying\ restrictions$) for the remaining models. Nonetheless, all the models report an AR (1) and AR (2) p-value (Pr>z) greater than

0.05, implying absence of serial correlation.

The introduction of the lagged dependent variable turns out to be significant, only column (2), with interaction terms, and column (3) that uses SYS-GMM but at a significance level of 10% and with low economic impact and a negative relationship for column (3). When looking at the impact of urbanization, the variable is found to be insignificant across all models even with the introduction of interaction terms which differs from what was observed in the static model. With the display of high but steady urbanisation rates in *Figure 6,* this may correspond to the idea that high urbanisation may not necessarily drive economic growth especially when the population is growing faster than GDP (Chen M, 2014).

The variable on agriculture is consistent in terms of statistical significance and it's negative impact on the GDP per capita growth rate across all models. Although, with the introduction of the interaction terms, it is evident that agricullture has a positive impact on GDP per capita growth rate given high urbanisation rates in both the column (2) and (4). Other variables of interest exhibiting statistically significant coeffecients are the control variables of years of schooling, literacy, belonging to an urban labour union and wage when interacted with urbanisation.

To begin with, the effect of years of schooling given high urbanisation rates has a positive impact on GDP per capita growth, while the interaction term indicates that the effect of literacy given high urbanisation rates has a negative impact on GDP per capita growth. This may imply that the more educated one is, the more they can contribute to the economy. On the other hand, a highly literate urbanised population may not necessarily contribute to GDP growth as there may be disparities between the skills needed and those possessed as pointed out by Ferreira & Roberts (2018). However, both these interaction terms bear no statistical significance in the SYS-GMM.

There is some consistency between both the D-GMM and SYS-GMM when it comes to the interaction terms of workers belonging to an urban labour union and wage as both are found to be significant. Belonging to an urban union given high urbanisation has a positive impact on dependent variable while the the average wage of an individual given high urbanisation has a negative impact on GDP per capita growth.

CHAPTER 5: CONCLUSION & RECOMMENDATIONS

5.1 Conclusion

The empirical analysis investigates the impact of urbanization (measured as the ratio of urban population to total population) on economic growth on a city level with economic growth captured through the GDP per capita growth rate. We also include variables that capture the sectoral composition of the cities and their contribution to the city's GDP. The methods of estimation used include static (Random Effects) and dynamic model estimation (i.e., D-GMM and SYS-GMM). In regards to our main objective, the impact of urbanization on economic growth, we find inconsistencies across the two estimation techniques with the static model reporting statistical significance for some of the variables while the dynamic model reports coefficients that are not statistically significant for the same variables. However, even with the coefficients being statistically significant, we find low economic impact and a negative economic impact of urbanization on the growth of GDP per capita even with the introduction of interaction terms.

This concurs with study by Brülhart et.al. (2009) with respect to the Williamson hypothesis even though we are not able to identify whether the threshold also exists for these cities and the study by Chen (2014) that highlights high rural to urban migration may be driven by high GDP growth but urbanisation may not necessarily be a strong driver for economic growth. Furthermore, the agglomeration variables of agriculture, industrial and the service sector are only statistically significant in the original model without the interaction terms. Additionally, they all bear negative coefficients– with agriculture having the highest negative impact. This suggests that agriculture may not be as value adding to the economic growth of a city and resources may be better used in another sector.

Furthermore, the interaction terms of urbanisation with the years of schooling, literacy and ratio of employees belonging to an urban labour union have statistically significant coefficients. However, with the random effects model, the effect of literacy given high urbanisation yields a positive and economically significant coefficient. Unlike in the D-GMM model where it yields a negative but also economically significant coefficient. While years of schooling given high urbanisation has a negative but low impact in the random effects and a positive but large economic impact. Nonetheless, this implies that an educated population may either negatively or positively impact economic growth in the face of urbanisation.

5.2 Limitations

As with any study that relies on census data, there are the known limitations of using such data that should be noted. With data spanning from 1999 to 2015, there may have been changes in methods of measurement and collection resulting in inconsistencies with the data used. Additionally, apart from 2000 and 2010 when the actual census was carried out, the other years are based on household sample surveys.

5.3 Policy recommendations

In line with the results of the study, it is evident that urbanisation has either a minimal or a negative impact on economic growth. This implies that high urbanisation cannot be viewed as a main driver of economic growth. In this case, the government can try to shift the focus from rural to urban migration as the marginal benefit of further urbanisation may have already been extinguished. Therefore, in highly urbanised areas they can focus on aspects such as literacy, education and the quality or type of jobs available to such cities to ensure individuals can still be productive to contribute to the growth of GDP. Policies on these areas may have spill over effects that aid in productivity. Moreover, they can also look at urbanising through decentralisation rather than migration to ensure that cities continue to stay productive by keeping the economic growth rates higher than the population growth rates.

5.4 Further research

Even though the results point to urbanisation having minimal or negative effect on economic growth in highly urbanised cities, there are still some areas that require further investigation. These include the aspects of urbanisation such as the population density and primacy index to find out if a there is a threshold past which cities do become less productive. Moreover, further investigation into variables such as years of schooling and literacy could be examined by using alternative proxies such as the human development index that incorporates the education index from sources such as the United Nations Development Programme that could better the robustness of the models. Additionally, other factors such as numbers of teachers allocated per class or per student could be included to find the specific aspects of these factors that impact growth for better policy formulation. Lastly, this same investigation could be done on other cities with lower contribution to national GDP to see if they return similar results.

Bibliography

Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*, 48-63.

Angotti, T. (1981). The Political implications of Dependency Theory. *Latin America Perspectives, 8*, 126-127.

Arellano , M., & Bond , S. R. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277-297.

Arellano, M., & Bover , O. (1995). Another look at the instrumental variables estimation of error-components models. *Journal of Econometrics, 68*, 29-51.

Bath, R., & James, D. (1976). Dependency Analysis of Latin America: Some Criticism, Some Suggestion. *Latin American Research Review, 11*(3), 5.

Black , D., & Henderson , V. (1999, April). A Theory of Urban Growth. *Journal of Political Economy , 107*(2), 252-284.

Blundell, R., & Bond , S. (1998). nitial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115-143.

Bodo, T. (2019, November). Rapid Urbanisation: Theories, Causes, Consequences and Coping Strategies. *Annals of Geographical Studies, 2*(3), 32-45.

Bond, R., & Smith, P. B. (1996). Culture and Conformity: A Meta-Analysis ofStudies Using Asch's (1952b, 1956)Line Judgment Task. *Psychological Bulletin , 119*(1), 111-137.

Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*, 48-63.

Bradshaw, Y. W. (1987, April). Urbanization and Underdevelopment: A Global Study of Modernization, Urban Bias, and Economic Dependency. *American Sociological Review, 52*(2), 224-239.

Browning, C. E. (1989). Urban Primacy in Latin America. *Yearbook. Conference of Latin Americanist Geographers. 15*, pp. 71-78. University of Texas Press.

Cadena, A., Dobb, R., & Remes, J. (2012). The Growing Economic Power of Cities. *Journal of International Affairs, 65*(2), 1-17.

Castells-Quintana, D., & Royuela, V. (2014, January). Agglomeration, inequality and economic growth. *The Annals of Regional Science, 52*, 343-366.

Chen M, Z. H. (2014). The Global Pattern of Urbanization an Economic Grwoth : Evidence from the Last Three Decades. *PLOS One*.

Childe, V. G. (1950, April). The Urban Revolution. *The Town Planning Review, 21*(1), 3-17.

Choi , I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Finance, 20*, 249-272.

Daniels, P. (2004, December). Urban challenges:the formal and informal economies in megacities. *Cities, 21*(6), 501-511.

Davis, D. E. (1991). URBAN FISCAL CRISIS AND POLITICAL CHANGE IN MEXICO CITY: From Global Origins to Local Effects. *Journal of Urban Affairs, 13*(2), 175-199.

Davis, K. (1995, March). The Origin and Growth of Urbanization in the World. *American Journal of Sociology, 60*(5), 429-437 .

Domar, E. (1946, April). Capital Expansion, Rate of Growth, and Employment. *Econometrica, 14.*

Duranton, G. (2014). *Growing through Cities in Developing Countries.* The World Bank Research Observer . Oxford University Press.

Faraji, S. J., Qingping, Z., Valinoori, S., & Komijani, M. (2016, April). Urban Primacy in Urban System of Developing Countries; Its Causes and Consequences. *Human: Journal for Interdisciplinary Studies , 6*(1), 34-45.

Ferreira, & Roberts. (2018). *Worldbank Flagship Report 2018: Raising the Bar for Productive Cities in Latin America and the Caribbean.*

Frick , S. A., & Rodriguez-Pose, A. (2017). Big or Small Cities? On city size and economic growth. *Growth and change.*

Frick, S. A., & Rodriguez-Pose, A. (2017). Big or Small Cities? On city size and economic growth. *Growth and Change.*

Harrod, R. F. (1939, June). An Essay in Dynamic Theory. *Economic Journal, 49.*

Henderson , J. V., Shalizi, & Venables, A. J. (2001). Geography and Development . *Journal of Economic Geography* , 81-105.

Henderson, V. (2002). Urbanisation in Developing Countries. *The World Bank Research Observer, 17*(1), 89-112.

Henderson, V. (2003, March). The Urbanization Process and Economic Growth: The So-What Question. *Journal of Economc Growth, 8*, 47-71.

Kasarda, J. D., & Crenshaw, E. M. (1991). THIRD WORLD URBANIZATION: Dimensions, Theories, and Determinants. *Annual Review of Sociology* , 467-501.

Kraas, F. (2007, January). Megacities and global change: Key priorities. *The Geographical Journal.*

Kuznets, S. (1971). *Economic Growth of Nations. Total Output and Production Structure.* Belknap Press.

Lampard, E. E. (1961, October). American Historians and the Study of Urbanization. *The American Historical Review, 67*(1), 49-61.

Lee, E. S. (1966). A Theory of Migration. *Demography, 3*(1), 47-57.
Lewis, W. A. (1954). Economic Development with Unlim.ited Supplies of Labour. 139-191.

Lipton, M. (1977). *Why poor people stay poor : a study of urban bias in world development.*

Maddala, G. S., & Wu, S. (1999). Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics*(Special Issue), 631-652.

McKinsey Global Institute. (2011). *Urban world: Mapping the economic power of cities.* McKinsey & Company .

Namkoong, Y. (1999, March). Dependency Theory: Concepts, Classifications, and Criticisms. *International Area Review, 2*(1).

Nguyen , M. H., & Nguyen , D. L. (2018). The relationship between urbanization and economic growth: An empirical study on ASEAN countries. *International Journal of Social Economics, 45*(2), 316-339.

Olson, M. (1982). *The Rise and Decline of Nations: Economic Growth, Stagflation and Social Rigidities.* Yale University Press.

Parnreiter, C. (2019, April). Primacy. *The Wiley Blackwell Encyclopedia of Urban and Regional Studies.*

Peng, X., Chen, X., & Cheng, Y. (n.d.). Urbanization and its Consequences. *Demography, 2.*

Pietak, L. (2014). Review Of Theories And Models Of Economic Growth. *Comparative Economic Research, 17*(1).

Population Reference Bureau . (2003, March). Retrieved from Population Reference Bureau : https://www.prb.org/populationdynamicsinlatinamericapdf318kb/
Puga, D. (2009). The magnitude and causes of agglomeration economies. *Journal of Regional Science.*

Quigley, J. M. (2009). *Urbanisation, Agglomeration and Economic Development." In Urbanisation and Growth,*. The World Bank.

Quigley, J. M. (2009). Urbanization, Agglomeration, and Economic Development. *Urbanization and Growth: Commission on Growth and Development* . The World Bank .

Ravenstein, E. G. (1889, June). The Laws of Migration. *Journal of the Royal Statistical Society, 52*(2), 241-305.

Reyes, G. E. (2001). Four Main Theories of Development: Modernization, Dependency, World System and Globalization. *Revista Crítica de Ciencias Sociales y Jurídicas.*

Ritchie, H. (2018). *Urbanization.* Retrieved from OurWorldInData.org: https://ourworldindata.org/urbanization

Ritchie, H., & Roser , M. (2020). *Urbanisation.* Retrieved from Our World in Data: https://ourworldindata.org/urbanization#citation

Rostow, W. W. (1959). The Stages of Economic Growth. *The Economic History Review, 12*(1), 1-16.

Sala-i-Martin , X., Doppelhofer , G., & Miller , R. I. (2004). Determinants of long-term growth: A Bayesian Averaging of Classical Estimates (BACE) approach. *American Economic Review, 94*(4), 813-835.

Schumpeter, J. A. (1934). *The Theory of Economic Development.* Harvard Economic Studies.

Singh, R., Swami, N. S., & Prasad, B. C. (2014, March). Urbanization and Economic Growth: An Empirical Study of Pacific Island Economies. *School of Economics Working Paper Series.*

Souza, D. P. (2019, December). *City of São Paulo has GDP of 4.3 thousand Brazilian municipalities together.* Retrieved from Agência IBGE Notícias: https://agenciadenoticias.ibge.gov.br/en/agencia-news/2184-news-agency/news/26407-pib-da-cidade-de-sao-paulo-equivale-ao-da-soma-de-4-3-mil-municipios-brasileiros-2

Tetty, C. (2005). *URBANIZATION IN AFRICA IN RELATION TO SOCIO-ECONOMIC DEVELOPMENT: A MULTIFACETED QUANTITATIVE ANALYSIS.*

Todaro, M. P., & Smith , S. C. (2014). *Economic Development .*

UN Habitat. (2015, March). *FINANCING SUSTAINABLE URBAN DEVELOPMENT: CHALLENGES AND OPPORTUNITIES.* Retrieved from https://www.un.org/esa/ffd/ffd3/wp-content/uploads/sites/2/2015/03/Financing-Urban-Development_UN-Habitat.pdf

United Nations. (2017). *World Population Prospects: The 2017 Revision.* Department of Economic and Social Affairs, Population Division.

United Nations. (2018). *United Nations (UN) Population Division's World Urbanization Prospects.*

United Nations. (2019). *World Urbanization Prospects 2018: Highlights.* Department of Economic and Social Affairs, Population Division.

United Nations University. (1996). Mexico City: No longer a leviathan? In U. N. Univerisity, *The Mega-city in Latin America.*

World Bank. (2020). *The World Bank| Data.* Retrieved from The World Bank: data.worldbank.org

World Bank Group. (2018). *Raising the bar for productive cities in Latin America and the Caribbean.*

World Population Review . (2020). Retrieved from https://worldpopulationreview.com/countries/brazil-population/

Worldometer. (2020). *Worldometer*. Retrieved from Wolrdometer: worldometers.info