The London School of Economics and

Political Science

Urban structure, location of economic activity and aggregate growth: empirical evidence and policies

Susanne Andrea Frick

A thesis submitted to the Department of
Geography and Environment of the London School of Economics for
the degree of Doctor of Philosophy in Economic Geography, London,
March 2017

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 56,633 words.

Statement of conjoint work

I confirm that Chapter 1 and Chapter 4 were jointly co-authored with Professor Andrés Rodríguez-Pose and I contributed a minimum of 50% of this work.

Susanne Frick

Abstract

This thesis explores two distinct but related aspects of the relationship between the spatial location of people within a country and national economic growth. The first three essays set out to establish whether a country's urban structure impacts economic growth at the aggregate level. Each paper explores a different aspect of the location of people, including the level of urban concentration and the size of a country's cities. The analyses rely on cross-country regressions and employ different estimation techniques, including fixed effects, system GMM and instrumental variables. Particular attention is paid to whether the relationship differs between developed and developing countries and how contextual factors, such as a country's economic structure and government capacity, shape the relationship. Across the three papers, the results consistently suggest that concentration and larger cities can be growth promoting at the country level; however only in developed countries or if certain contextual factors are met. The findings add to a growing body of empirical literature which questions the universal validity of the benefits of agglomeration for economic growth. Furthermore, they specifically address a gap in the empirical literature which so far had failed to link city size to aggregate growth as opposed to city-level productivity.

The fourth essay moves to the policy level and analyses Special Economic Zones (SEZs) – a policy which is frequently employed to influence the location of economic activity and people. Specifically, it studies the performance drivers of SEZs. A scarcity of data has limited quantitative research on this topic so far. The analysis relies on a novel dataset, which resorts to nightlights as SEZ performance proxy and covers zone and policy characteristics for SEZs in 22 countries. The findings partially confirm, but also refute the dominant knowledge on the viability of SEZs. While larger zones tend to perform better, growth is difficult to sustain over time and particularly hard to achieve for high-technology focused zones. Other factors commonly assumed to matter, such as the nature of the zone operator, the incentive package and programme set-up, seem to be highly context dependent. Furthermore, contextual factors, such as proximity to markets and a pre-existing industrial base, influence zone performance.

Acknowledgements

My first words of gratitude reach beyond this PhD and go without any doubts to my parents and my family. Mama, Papa, not only did you enable me to do this PhD in the first place, but you have always supported my – for you probably sometimes not very obvious – life choices. And you did this with surprisingly little questioning. I am deeply grateful to both of you for this support. A big thanks also to my siblings and their partners. I am grateful to be part of a family which always has each other's backs!

My second, and PhD related biggest, 'thank-you' goes out to my supervisor Andrés Rodríguez-Pose. Andrés, you have guided me during this PhD with your academic, professional and personal advice. Throughout, you have been a mentor and a great friend. I could not imagine having done this PhD without your support.

I am also deeply grateful for my fun and supportive PhD colleagues, who made this PhD so much more enjoyable and rewarding. Many of you have become good friends and, while I cannot say that I will miss the PhD room, I will miss you guys.

I also want to thank everyone else in the department for their support and advice over the years. A special thanks to my review supervisor Riccardo Crescenzi for his valuable feedback during the upgrade. Furthermore, I am very grateful to Olmo Silva for always having the door open to answer questions about regressions, methodology or just to have a quick chat in the hallway.

To the LSE catering crew, in particular the team of the fourth floor restaurant and café: thanks a lot for being smiley, all the laughs and entertaining chats during the coffee breaks!

And finally, my gratitude goes to my husband Stefano. Ste, you have been by my side from the beginning, seen the occasional panic attacks, moments of happiness and general ups and downs. Thank you for your patience, your (sometimes very handy) mathematical knowledge and sharing lots of Lebanese food! Bacio!

Table of contents

List of figures

List of tables

Synopsis

I. Introduction

This PhD thesis explores two distinct but related aspects of the relationship between the spatial distribution of people within a country and national economic growth. The first three essays set out to establish empirically whether a country's urban structure influences economic growth at the aggregate level. Each paper explores a different aspect of the location of people, including the level of concentration of the urban population within a country, the size of a country's cities and contextual factors which may shape the relationship. The fourth essay moves to the policy level and looks at Special Economic Zones – a policy which is frequently employed to influence the location of economic activity and people within a country. Specifically, it analyses the performance drivers of Special Economic Zones in emerging countries, including how a zone's location within the country affects its performance.

The thesis is set in the context of the growing body of literature that examines the urbanization process, its various forms and its influence on a country's economic performance in developing countries (Castells-Quintana, 2016; Castells-Quintana & Royuela, 2014; Duranton, 2008; Glaeser, 2014; Gollin, Jedwab & Vollrath, 2016; Henderson, 2010; Jedwab & Vollrath, 2015). The question of how urbanization – and frequently more specifically cities and agglomeration – interacts with economic development has long attracted the attention from both academics and policy makers (Henderson & Ioannides, 1981; Hoselitz, 1953; Jacobs, 1970, 1985; World Bank, 1979). Limited data for developing countries resulted in most empirical research focusing on the developed country's point of view. In recent years, however, there has been an increasing awareness that more research is needed to understand how these processes may differ between developing and developed countries and that a direct application of empirical evidence, based on developed country data, may be misleading for many countries.

The thesis aims to contribute to closing this gap. Two out of the four substantive chapters employ newly assembled datasets which, on the one side, allow for a more nuanced analysis than previous research (in the case of the chapter 1 on urban concentration); and on the other side, permit to carry out the first large scale quantitative assessment on the topic (in the case of chapter 4 on the drivers of SEZ performance). The other two chapters add a new important perspective to the existing literature, by shifting the emphasis from urban concentration to the size of a country's cities and evaluating potential differences between developing and developed countries as well as underlying contextual factors.

In the following sections, I outline the motivation and relevance of the topic, provide an overview of the related literature as well as summarize the individual chapters and their main contributions. In the final section, I conclude, deduct policy implications and point to further areas of research.

II. Motivation

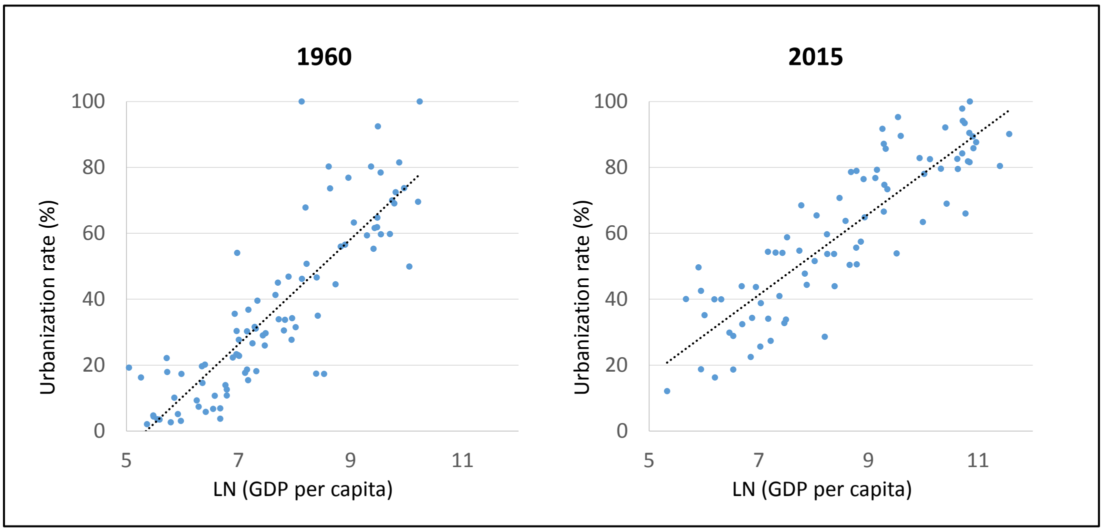
i. Urbanization and city development in a historical perspective

Over the past 60 years, the world has undergone an unprecedented transformation in terms of urbanization patterns and city development. Looking around the globe today, urbanization rates are at unforeseen levels with large cities virtually anywhere in the world – including many African, Asian and Latin American countries (United Nations, 2014). In fact, the majority of the world's largest cities is in developing countries. Until the 1950s, however, the world looked very different. For most of history, cities and urbanization were closely intertwined with economic success and/ or political dominance. Only in the economically and politically most advanced nations, the growth of large cities was possible and in turn, large cities brought about further development. A few examples include Alexandria around 200BC, Rome in 1AD or Chang'an in China in 800AD (Chandler, 1987). The relationship between cities and economic development is believed to be such that social scientists frequently resort to urbanization and cities as a proxy for economic development for periods where no data on GDP is available (Acemoglu, Johnson, & Robinson, 2002; De Long & Shleifer, 1993).

This rule was still very much in evidence in the 1950s. At that time, the majority of the world's biggest cities – 20 out of 30 – were indeed located in high-income countries (United Nations, 2014). The few exceptions to the rule were a handful of large cities in big developing countries, such as China, Mexico and Brazil. The poorest countries in the world at the time lacked, in general, very large cities.

Until today, one can still trace a strong relationship between urbanization and economic development. Figure II-1 plots a country's natural logarithm of the GDP per capita against its urbanization rate in 1960 and 2015: a strong positive relationship between the two can be observed in both points in time. The higher the economic development, the more urbanized a country is. There are, however, also important differences between the two graphs. In 1960, countries with a low level of economic development were largely rural with urbanization rates close to 0%. The graph for 2015, in contrast, shows a different picture. While the slope is still clearly positive, the urbanization level is generally higher with almost all countries being more than 20% urbanized. Furthermore, the slope is marginally less steep, pointing to a less pronounced relationship between the two variables. Jedwab and Vollrath (2015) describe these trends in their historical account and demonstrate that countries today are 25% to 30% more urbanized at any level of economic development than in the past.

Figure II-1: Economic development and urbanization rate[1]



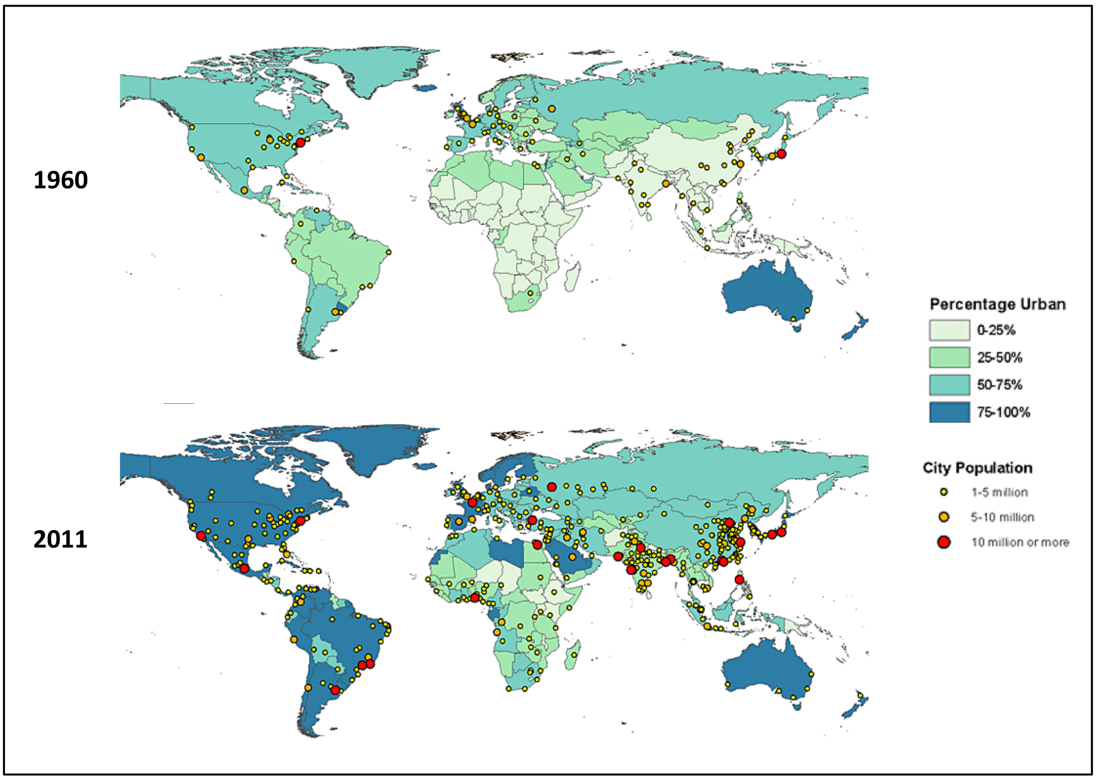
Source: World Development Indicators, GDP per capita at constant 2010US$

This evolution has been driven by galloping urbanization throughout the world over the past decades. Whereas in 1950, less than one billion people lived in cities, the number increased to 3.6 billion in 2010 and will continue to rise to an estimated 5 billion in 2030 (United Nations, 2014). These numbers do not only reflect an increase in the absolute number of people as a result of general population growth, but also a relative shift in the share of people living in urban versus rural areas. The share of the urban population grew from just about 30% in 1950 to 52% in 2010. Since 2006, for the first time in history, more people live in cities than in the rural areas.

The surprising element within this structural shift is the change in the geographical pattern of urbanization. As Jedwab and Vollrath (2015) show, up until 1950, urbanization had mainly taken place in the richer countries of Western Europe and North America. Poorer countries, in contrast, saw stagnating rates. This pattern has, however, reversed since then with the strongest changes occurring in developing countries. The comparative world maps for the years 1960 and 2011 in Figure II-2 illustrate well these changes.

[1] Note: the country sample is kept constant for both years, 1960 and 2015, facilitating the comparison between the two years.

Figure II-2: Urbanization and large cities, 1960 and 2011



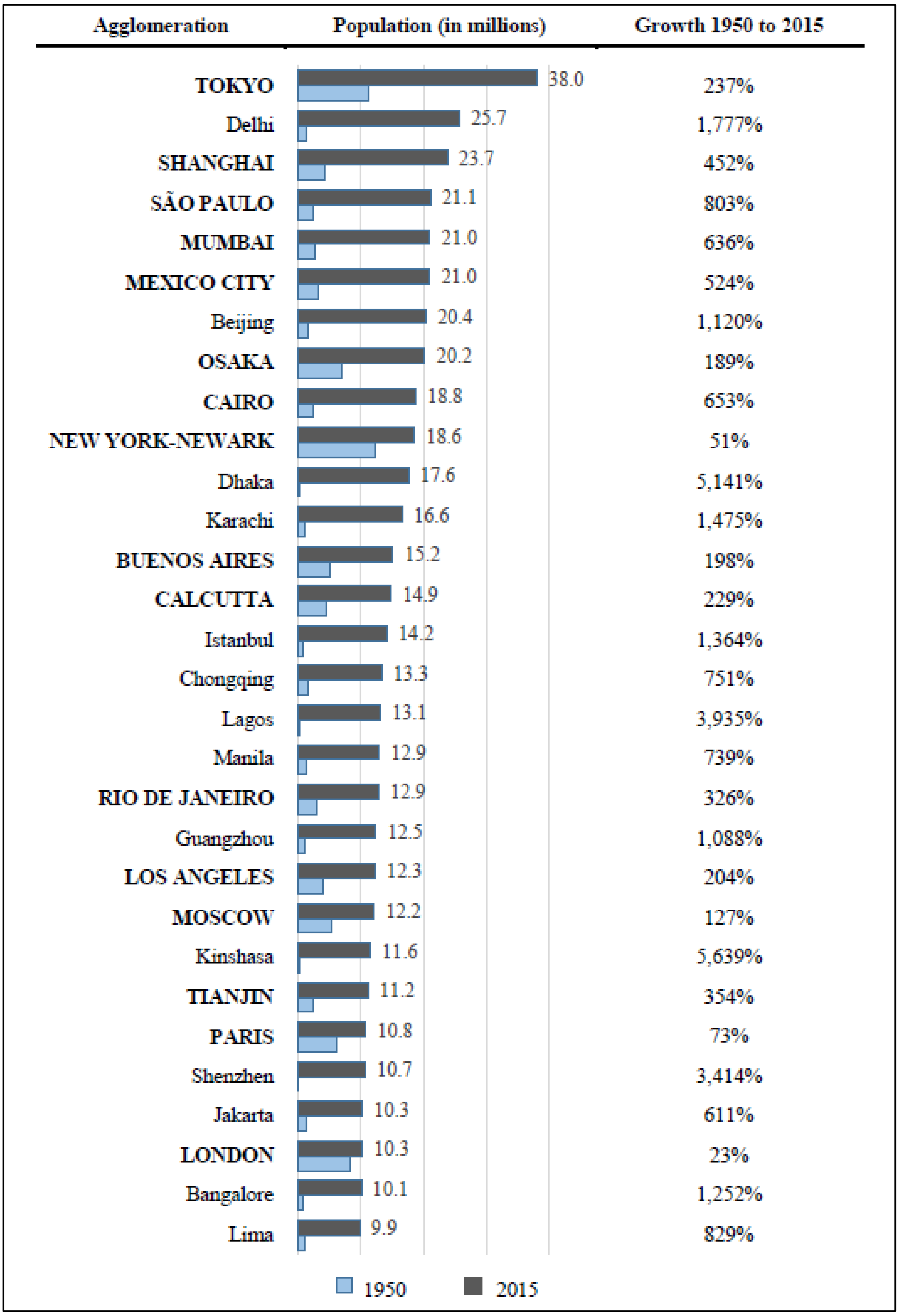
Source: World Urbanization Prospects, 2014

The shading of the countries indicates their level of urbanization with darker colours representing more urbanization. In 1960, (Western) Europe, Oceania and Northern America displayed the highest levels of urbanization together with a handful of countries in Latin America. Most parts of the developing world exhibited comparatively low rates at the time. By 2011, the vast majority of countries generally displays higher rates of urbanization than in 1960 with the strongest changes clearly occurring across the developing world. Low- and middle-income countries have urbanized strongly and frequently display similar levels of urbanization than high-income countries (for example the majority of Latin American countries and Egypt). Estimates suggest that even those countries in Southern Asia and Africa, which are still relatively less urbanized nowadays, are expected to get close to 45% by 2030 (United Nations, 2014).

Figure II-2 reveals another interesting and maybe even more striking development: the emergence of a vast number of mega cities, the majority of them yet again located in the developing world. While there were only 2 cities with more than 10 million inhabitants until the 1970s, their number increased more than ten-fold to 25 by 2010 (indicated by the red dots). The number of cities of 5 to 10 million inhabitants grew in parallel from merely 4 in 1950 to 38 in 2010 (displayed by the orange dots). Again, changes are most visible in the developing world. The large majority of mega cities with more than 10 million inhabitants can be found in the populous countries of Asia. Even among the still relatively less developed and less urbanized countries across the African continent, massive agglomerations of 5 million inhabitants and above have emerged. Out of the 30 largest agglomerations in 2015, 23 are located in the developing world.

In addition to this increase in the number of cities, cities have also grown considerably in size (Figure II-3). The average population of the world's largest thirty agglomerations quadrupled from 1950 to 2015: while the average was around 4 million inhabitants in 1950, it now stands at 16 million (United Nations, 2014). This development is further illustrated by taking the example of the largest city at each point in time. The New York-Newark metropolitan area was the biggest agglomeration in 1950 with roughly 12 million people. Today the largest area is the Tokyo agglomeration with over 38 million people. Other cities, in particular in emerging countries, have seen an even more dramatic increase in their population. Almost half of the thirty largest cities in 2015 were not included in the 1950 list (those cities not capitalized in Figure II-3). The majority of these cities has grown between ten to fifty times over the past decades, with Shenzhen reaching 3,400 times its 1950 size.

Figure II-3: 30 largest agglomerations in 2015[2]



Source: Own elaboration based on World Urbanization Prospects 2014

The growth of the large metropolis in countries with relatively low levels of urbanization (compared to developed countries) has brought about another phenomenon which has received

[2] Bold capitalized cities were among the 30 largest agglomeration in 1950

much attention in the literature: the spatial concentration of people and economic activity in few cities and the associated spatial disparities between primary urban areas and lagging regions. Urban primacy, the share of a country's urban population living in the largest city, is a frequently used indicator to measure this spatial concentration. In the developed countries of Europe and North America, primacy reaches average levels between 10% and 15%. In contrast, many developing countries, in particular low-income countries, display urban primacy rates of over 30%.[3] Similarly, the ratio between the largest and second largest city is on average 2.7 for developed countries, while it raises to 3.9 for the developing countries' average and reaches over 10 in a number of cases, such as Kabul, Addis Ababa, Lima and Buenos Aires.[4]

We are thus living in a world which is – independent of income levels – consistently more urban than 60 years ago, and where urbanization and the formation of large cities seems at least partially disconnected from economic development. In contrast to the mid-20th century, the majority of urban development is happening in the developing world and at an unprecedented scale.

ii. Urbanization, cities and economic development

The aforementioned changes are striking not only because of their sheer scale, but also because they challenge traditional economic theory on urbanization and cities. First, they question the prevailing explanations posited by economists on the mechanisms triggering urbanization. Second, they have sparked a growing debate on whether cities in developing countries have the same catalytic effect on economic development as in developed countries.

On the first point, as described above, urbanization and the existence of large cities in a country were closely connected to the economic development process throughout much of history. In the past, only the most developed countries were highly urbanized and had very large cities. Traditionally, economists explain the urbanization process with a structural shift

[3] Primacy numbers are sourced from the World Development Indicators
[4] Own calculations based on (United Nations, 2014)

from an agricultural to a manufacturing based economy (Bertinelli & Black, 2004; Davis & Henderson, 2003). On the one hand, a sufficiently productive agricultural sector is needed to shift from full absorbance of labour by subsistence farming to available surplus labour in the countryside. On the other hand, it assumes that a growing industrial sector, which is quintessentially urban, attracts former rural workers: higher (expected) wages in the urban industrial sectors stimulate migration from rural areas to cities (Harris & Todaro, 1970; Henderson, 2003; Lewis, 1954). Hence, urbanization becomes a by-product of industrialization. This mechanism, in fact, reflects rather well the development processes at work in today's developed countries: urbanization and city growth were strongly tied to the industrialization of their economies and economic growth (Jedwab & Vollrath, 2015).

This traditional explanation is, however, somewhat at odds with the developments witnessed across the world over the past decades. Many countries with relatively low levels of economic development are far more urbanized today than developed countries were decades ago at similar stages of economic development. Developing countries have, thus, urbanized without the accompanying industrialization which has normally been the mechanism put forward as explanation. Fay and Opal (2000) describe this situation as 'urbanization without growth'. While this phenomenon is not entirely new, there has been renewed interest in it from the economics literature over the past five to ten years, driven by the scale of city growth in many developing countries.[5]

A diverse number of factors has been put forward as further potential mechanisms triggering urbanization in developing countries in addition to the structural transformation argument. So-called push-factors drive people from the rural countryside into cities. Push-factors include conflicts (Glaeser & Shapiro, 2002), as the rural population seeks protection in cities as well as negative agricultural shocks and rural poverty which decrease rural wages and thus incentivize an outward migration to cities (Fay & Opal, 2000; Jedwab & Vollrath, 2015).

[5] Urbanization without industrialization has already been described as early as the 1950s (Hoselitz, 1953)

Urban pull-factors are those drivers which make cities more attractive to people. They include urban amenities such as better urban living conditions through improved access to public services as well as a political urban bias (Ades & Glaeser, 1995; Davis & Henderson, 2003). Glaeser (2014) also proposes that developing countries urbanize earlier due to their ability to import food. In developed countries, by contrast, cities could only increase in population as the agricultural sector became more productive and could provide food for larger non-food producing urban population. Furthermore, Jedwab and co-authors (Gollin et al., 2016; Jedwab & Vollrath, 2015) argue that natural resource exports also drive urbanization in developing countries: 'consumption cities' as opposed to 'production cities' emerge through the income gains generated by the resource exploitation. As a consequence, the share of urban dwellers, working in the non-tradable service sectors, is much higher than in 'production cities' where the urban population works in manufacturing and the tradable service sectors.

The second point, which has been increasingly called into question by the developments over the past 60 years, are the productivity gains arising from cities. Whether we look at the endogenous growth theory, urban economics or (new) economic geography, cities are supposed to bring productivity gains through agglomeration economies. While the explicit mechanisms and emphasis within the different schools of thought differ, the underlying concepts go back to the same ideas already voiced by Marshall (1890) and Jacobs (1970, 1985): bringing people together in cities and agglomerating economic activity causes efficiency gains through the generation of thick labour markets, forward and backward linkages and knowledge spill-overs. Thick labour markets are created in cities as a larger pool of people facilitates firms to find employees with the required skill-sets. Conversely, a large number of companies located in their vicinity eases the job search for employees and reduces the risk of unemployment. Forward and backward linkages arise through the benefits for firms by being located closer to their markets and suppliers (Krugman, 1991). In the presence of increasing returns to scale, closeness to markets brings about efficiency as it allows firms to reduce transport costs. Furthermore, a network of suppliers is formed which can cater to the firm’s specific needs. This is particularly the case for clusters of specific industries. And finally, external economies (Rosenthal & Strange, 2004) create informational spill-overs between different actors. In the words of Marshall (1890, p. 332): “... so great are the advantages which people following the same skilled trade get from nearby neighbourhood to one another. The mysteries of trade become no mysteries; but are as were in the air...”. Proximity, which is created through cities, is key to these sorts of knowledge spill-overs as they have shown to have a strong distance decay effect (Fischer, Scherngell, & Jansenberger, 2009; Jaffe, Trajtenberg, & Henderson, 1993).[6]

Besides these general claims on the advantages of cities, there has been an increasing focus on praising the catalytic effect of large cities in particular. Empirical evidence in the urban economics tradition has stressed the productivity gains from increasing city size: a doubling of city size is accompanied by a 3% to 8% increase in the productivity of the urban worker (Duranton & Puga, 2004; Melo, Graham, & Noland, 2009; Rosenthal & Strange, 2004). Similarly, both theoretical and empirical studies taking a new economic geography point of view – i.e. the level of analysis is not the city but the country – have shown that more concentrated urban structures are growth inducing at the country level since the agglomeration economies generated by the concentration of people increase the overall productivity level (Bertinelli & Strobl, 2007; Brülhart & Sbergami, 2009; Fujita & Thisse, 2003; Henderson, 2003; Martin & Ottaviano, 2001). These findings imply, holding the population constant, that larger cities are indeed strongly beneficial from an economic growth point of view.

It has also been analysed whether this effect hinges on the country’s level of economic development. This idea goes back to the so called ‘Williamson curve’, coined by Jeffrey Williamson in the sixties (Williamson, 1965). He suggested that the relationship between economic growth and spatial disparities within a country follows an inverted U-curve. At a

[6] While the language, used in this short account, is based on the new economic geography literature, other streams describe similar channels – for instance the sharing, matching and learning mechanisms from the urban economics literature. For the sake of simplicity, I use the economic geography language.

low level of economic development, regional disparities will increase when the country experiences growth, at a higher level these disparities will decrease. Williamson stressed the high economic interdependence between regions within a country which would entail labour mobility to equalise any differences in regional income over time. This hypothesis is supported by a number of recent empirical studies, which find that agglomeration is growth inducing at lower levels of economic development while it matters less at higher income levels (Barrios & Strobl, 2009; Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014; Henderson, 2003).

This stream of literature has been influential in policy discussions over the past decade, leading to a set of well-defined policy prescriptions. Agglomeration and thus the concentration of people in few large cities is deemed beneficial for growth and (developing) countries should avoid attempting to reduce spatial disparities between more and less developed regions. Instead, promoting agglomeration is considered the fastest and most direct path for development, since the presence of a large city will multiply growth opportunities. As the World Bank (2009) put it most prominently in its 2009 World Development Report: "Economic growth is seldom balanced. Efforts to spread it prematurely will jeopardize progress" (p. 5-6). Based on this logic, the rationale and usefulness of 'spatially-targeted' policies, which aim to promote the development outside of the main urban areas, have been increasingly called into question. Instead, 'spatially-blind' approaches have been favoured, which focus on investing in people independent of where they live.

The universal validity of such statements, however, has been increasingly called into question for a variety of reasons. First, our improved understanding of the factors driving urbanization in developing countries makes it questionable that processes which are – at least partially – influenced by such different forces should lead to the same outcome. Gollin et al. (2016) for instance suggest that 'consumption cities' in the developing world are less likely to have the same productivity inducing effect as 'production cities', since most people work in non-tradable sectors with less possibilities for on-the-job learning. Second, the sheer magnitude of urbanization and city development in the developing world is unprecedented in history. Developed country's cities, on which most of the empirical evidence is based, are frequently multiple times smaller, so that a one-to-one application of the principle of 'the bigger the better' seems unlikely to hold. In fact, McCann and Acs (2011) show that 38% of the 75 most productive cities in the world have a population of less than 3 million inhabitants; the population of another 42% ranges between 3 and 7 million and only very few cities that we would consider as megacities make it on the list. Similarly, the OECD (2006) shows that the positive effect of city size on income only holds if the largest cities are excluded from the sample. The effect becomes negative when looking at cities with more than 7 million inhabitants.

Finally, researchers also emphasize the role of a multitude of other factors besides size behind changes in city-level productivity (Camagni, Capello, & Caragliu, 2013, 2015). Since these factors differ from city to city, cities in the same size class are not necessarily equally productive (Storper, 2010). Among these factors, industry specific aspects such as industrial composition and specialization (Au & Henderson, 2006; Graham, 2009; Storper, 2010) as well as company size (Faggio, Silva, & Strange, 2014; Rigby & Brown, 2015) figure prominently. A city's sectoral specialization, in particular, has attracted considerable attention. The tradable service sector and high-tech manufacturing industries are assumed to profit more from being located in large cities than other industries (Graham, 2009; Storper, 2010). Countries with a significant presence of such sectors, thus, experience stronger productivity effects from large cities than countries with a bigger and/or more mature manufacturing industry base. Furthermore, city context specific elements seem to be an important determinant for urban productivity levels. A city's government capacity (Ahrend, Farchy, Kaplanis, & Lembcke, 2014; Glaeser, 2014) and urban infrastructure (Castells-Quintana, 2016) as well as network integration (Sassen, 1991; Scott, 2001) and borrowed city size (Alonso, 1973) have been highlighted in these analyses. In the developing countries context, there has been a particular emphasis on the provision of an adequate urban infrastructure. Castells-Quintana (2016), for instance, shows empirically that in countries with an insufficiently developed urban infrastructure (proxied by sanitation), urban concentration can be growth inhibiting. Glaeser (2014) and Ahrend et al. (2014) furthermore highlight the importance of institutional capacity, in order to reduce negative externalities which arise following rapid city growth.

iii. Spatial policies – Special Economic Zones

The increasing size of cities and perceived heightened polarization between primary urban areas and secondary cities have led policy makers to implement a raft of policies. On the one side, these policies aspire to address urgent challenges such as sanitation and transport infrastructure in the emerging cities in order to make them more efficient and liveable. On the other side, they frequently aim to 're-balance' the country's urban structures by promoting economic development in secondary cities and lagging regions, thereby decreasing the gap between these areas.

Among this second group of policies, Special Economic Zones (SEZs) have been particularly popular over the past decades. Their number increased from just 176 in 47 countries in 1986 to 3,500 in 130 countries in 2006 (Singa Boyenge, 2007). Recent estimates suggest that the number reached almost 4,500 in 2014 (The Economist, 2015).

While SEZs are a diverse group of policies and there is no uniform definitions of what constitutes one, some common features can be defined (Asian Development Bank, 2015; OECD, 2009; World Bank, 2011). First, SEZs aim to attract (foreign) investment to specific areas within a country and thereby support the attraction of firms and the generation of employment and exports. Frequently, they are employed with the specific objective to attract economic activity to economically less developed regions. By establishing SEZs, countries also seek to either achieve a first step in their industrialisation process or to increase the value-added of the existing manufacturing sector. Second, the SEZs' attractiveness mainly stems from a favourable regulatory and/ or incentive scheme for firms directing investment into the SEZ rather than to the rest of the country. Reduced corporate tax rates, the exemption from import duties and a reduction of the regulatory burden often form part of the incentive package. Third, SEZs have a clearly demarcated geographic territory. This area can range in size from relatively small industrial park type zones of a few hectares to entire regions of a country. And finally, most SEZs provide infrastructure and other services to resident firms in order to overcome barriers in the local business environment and facilitate the companies' operations. Onsite custom offices, one-stop-shops dealing with obtaining permits for setting up and running companies as well as pre-built factory units and serviced land plots are among the services offered.

Despite these common features, SEZs vary significantly. Different types of SEZs have emerged, ranging from traditional export processing zones and commercial free zones to high-tech parks and large region wide special economic areas. Each type has a different focus, among other things, in terms of target companies (e.g. local versus foreign), industries (e.g. low-tech versus high-tech, services versus manufacturing), incentives offered as well as requirements imposed on firms (e.g. export or foreign ownership requirements). Even within the different SEZ types, zones differ considerably including in size, sectoral focus, location within the country, services offered, incentive package as well as the general context they are placed in.

SEZs have a mixed track record in terms of investment attraction and the generation of employment and exports (World Bank, 2011). In fact, the final verdict on whether SEZs manage to achieve their stated objectives is still unclear. One the one side, the well-known success of Chinese Special Economic Zones has fuelled the hope of many policy makers to achieve a similar growth story in their own county when implementing SEZ policies. On the other side, a vast number of SEZs, even among the Chinese zones, has failed to deliver any meaningful success. Frequently, plans remain on paper and zones never become fully operational (FIAS, 2009).

Taken these facts together – the diversity in the design and set-up of the zones as well as the highly mixed performance – there is surprisingly little quantitative comparative research on the drivers of SEZ success. Case studies, frequently on the more successful cases, abound.

They either remain descriptive or zoom into an individual or a limited group of zones, often within specific countries (Farole & Kweka, 2011; FIAS, 2009; Nel & Rogerson, 2013; OECD, 2009; UNIDO, 2015; World Bank, 2011). The results of these studies have also been highly contrasting. While these studies are informative and contribute to a better understanding of the functioning of SEZs, a larger scale assessment would be desirable to complement the evidence.

However, a lack of comparable cross-country data on the performance of SEZs as well as on the zone-specific regional and/or national policy programs, which are at the very origin of SEZ formation, has represented a fundamental barrier to the development of this type of research. The few studies which attempted to quantitatively assess what contributed and what hindered SEZ success (Aggarwal, 2005; Farole, 2011) rely on a low number of observations, which casts doubt on the generalisability of the results.

III. Summary of the chapters

The main focus of this thesis is on the question of how increases in city size and changes in the urban structure affect economic development at the country level. The aim is to contribute to the growing body of empirical literature which focuses on the developing country experience and to allow for a more nuanced perspective on some of these topics.

The thesis comprises four substantive chapters. Chapter 2 on average city size and economic growth as well as chapter 4 on SEZ performance drivers are co-authored with Professor Andrés Rodríguez-Pose. Chapter 1 on the evolution of urban concentration and growth and chapter 3 on city size, contextual factors and growth are single-authored.

The first three chapters follow in the footsteps of some of the most frequently cited empirical papers covering the relationship between urban structure and economic growth, namely Henderson (2003) and Brülhart and Sbergami (2009). Each chapter examines the topic from a different angle and adds a new, previously not considered perspective to the debate. Chapter 1 provides a detailed description of the evolution of urban concentration from 1985 to 2010, facilitated by a newly constructed data set which allows to overcome data limitations which have plagued research on the topic so far. It then continues to relate the changes in the levels of urban concentration to economic growth. Chapter 2 examines the question from an angle which has received little attention to date by relating economic growth not to urban concentration (as the previous papers on the topic), but to the actual size of a country's cities. This is an important addition to the literature since much of the research examines urban concentration and not city size, but is frequently cited as evidence to support claims on the advantages of larger cities. In the third chapter, I maintain the angle of city size and add the dimension of the aforementioned contextual factors to the analysis. In the final chapter, the focus moves to the policy level and examines the factors driving SEZ performance. To my knowledge, the analysis is the first quantitative assessment of this question. The following sections provide a brief summary of each chapter.

i. The evolution of urban concentration and economic growth

While urban concentration has been a focus of economic analysis for a considerable amount of time, little attention has been paid to describing its levels and evolution in more detail. A general remark on the well-known higher average levels of urban concentration in many developing countries frequently suffices as justification for the subject, in particular in the literature on the relationship between urban concentration and economic growth (see for example Henderson, 2003; Venables, 2005; Brülhart and Sbergami, 2009). Furthermore, the vast majority of the literature uses primacy as indicator for urban concentration. The explanatory power of urban primacy, however, is limited to the upper echelon of the city size distribution and may camouflage important differences underneath.

The aim of the first substantive chapter of the thesis is therefore twofold. First, it describes the evolution of urban concentration from 1985 to 2010 in 68 countries around the world; and second, it analyses the extent to which the degree of urban concentration affected national economic growth in the same period. For this purpose, I built a new urban population dataset which includes all cities with 50,000 inhabitants[7] or more in the 68 countries considered. Equipped with this data, I construct a set of Herfindahl-Hirschman-Indices which capture a country's urban structure in a more nuanced way than primacy.

The analysis of the evolution of urban concentration in the first part of the paper shows that, in line with common perception, urban structures are on average far more concentrated in developing than in developed countries. However, contrary to the general perception of increasing levels, urban concentration has on average decreased or remained stable across the world (depending on the specific indicator used). However, these averages camouflage significant shifts within many countries, which – far from having stable levels of urban concentration – frequently showcase decreasing but also increasing concentration. This is true for both developing and developed countries. While most Latin American countries have lower levels of urban concentration today compared to 1985, many European, Asian and some African countries display higher concentration levels. Furthermore, despite decreasing averages, developing countries remain far more concentrated than developed countries.

The results of the second part of the chapter, covering the econometric analysis of the relationship between urban concentration and economic growth, suggest that there is no uniform association between the two factors. I find that urban concentration is beneficial for economic growth in high-income countries, while this effect does not hold for developing countries. The results are robust to accounting for potential endogeneity issues when reverting to the use of IV analysis. These findings differ from previous analyses that tend to find a particularly important effect of urban concentration at low levels of economic development.

ii. Average city size and economic growth

This chapter is co-authored with Professor Andrés Rodríguez-Pose. It follows a similar approach as chapter 1, but moves away from the question of urban concentration and focuses on the role of the actual size of a country's cities for economic growth. This is an important

[7] In a handful of countries, the city sample is restricted to a higher city size threshold

addition to the existing empirical literature since it provides a so far mainly overlooked perspective. On the one hand, the empirical literature, linking city size to productivity, analyses the relationship from a city perspective, not a country growth perspective. Given that most countries have more than one city, maximising productivity levels of just one city may not be efficient for the country as a whole. On the other hand, the literature, associating urban concentration to economic growth at the national level, does not say much about the effect of city size (even though it is frequently cited in the debate): countries with very different sized cities can have the exact same level of urban concentration as long as the relative size distribution remains the same. In fact, city size is on average negatively correlated with urban concentration since smaller countries tend to be more concentrated while having smaller cities. In the context of agglomeration economies however, city size may be a more important aspect to analyse than the level of urban concentration.

Chapter 2 thus addresses the link between a country's average city size and economic growth in a total of 114 countries for the period between 1960 and 2010. As in the previous chapter, emphasis is placed on understanding potential differences between high-income and developing countries. The analysis – which includes pooled 2SLS, panel data analysis, system GMM, and an instrumental variable (IV) approach – finds the relationship between average city size and economic growth to be far from universally positive. Again, the results vary between high-income and developing countries. In high-income countries, there is consistent evidence of a positive link between city size and economic growth, although the effect decreases as the average city size increases. In contrast, the relationship does not hold for developing countries, for which most of the coefficients display insignificant results or point towards a negative connection between both factors.

iii. Big or small cities? Does city size matter for growth?

The third chapter maintains the city size perspective, but takes a slightly different approach than chapter 2. Instead of examining the role of average city size, I relate the percentage of the urban population living in cities of a certain size to economic growth. Under the assumption that cities of a particular size maximize people's productivity, a larger share of the urban population living in cities of that size should drive productivity and economic growth. The chapter, furthermore, incorporates additional factors, such as a country's industrial composition and contextual factors, into the analysis, in order to understand how these factors interact with city size. I hypothesise that the balance between agglomeration economies and diseconomies which determines the productivity maximising city size, is affected by a country's endowment of these factors.

The econometric section follows a two-step approach. Using a panel of 113 countries between 1980 and 2010, I, first, explore whether there are certain city sizes that are growth enhancing. Secondly, I introduce additional factors into the analysis to test their impact on the relationship. These factors are the sectoral composition of a country's economy (as measured by service-sector and high-tech exports) and contextual factors (government effectiveness and urban infrastructure as proxied by sanitation).

The results suggest a non-linear relationship dependent on the country's size. In contrast to the prevailing view that large cities are growth inducing, I find that for the majority of countries it is beneficial to have relatively small cities of up to 3 million inhabitants. A large share of the urban population in cities with more than 10 million inhabitants is only growth promoting in countries with an urban population of 28.5 million and more. I, furthermore, find that the relationship is highly context dependent: a high share of industries that benefit from agglomeration economies, a well-developed urban infrastructure and an adequate level of governance effectiveness allow countries to take advantage of agglomeration benefits from larger cities.

iv. Special Economic Zones: What makes them truly special?

The final chapter of this thesis moves on to the policy level and looks at SEZs. This is the second co-authored chapter of my thesis and was developed and written in collaboration with Professor Andrés Rodríguez-Pose.

As described above, SEZ type policies are frequently employed across the developing and developed world with the aim of influencing the location of economic activity within a country and thus in the context of the debate of urban concentration. Despite their significant proliferation and the diversity in their design and implementation, there is little quantitative research on the factors that influence their performance due to limited data for SEZ outcomes and characteristics. The aim of this chapter is to fill this gap and contribute to a better understanding of drivers of SEZ performance.

SEZ performance is assumed to be influenced by both internal as well as external factors, including SEZ characteristics (such as size, sectoral focus, private versus public operator, location etc.), specificities of the SEZ policy (incentives offered, requirements, policy set-up etc.) and contextual factors (e.g. a country's level of development, previous industrialization and proximity to large markets). In order to capture the diversity of factors, we rely on an entirely new dataset which covers a wide range of SEZ policy aspects as well as encompasses SEZ characteristics and contextual factors across 345 zones in 22 countries. In order to obtain a comparable proxy for SEZ performance, we resort to night-lights data, which has been shown to be a good indicator for economic activity when other traditional data is not available (Chen & Nordhaus, 2011; Ebener, Murray, Tandon, & Elvidge, 2005; Ghosh et al., 2010; Henderson, Storeygard, & Weil, 2012; Mellander, Lobo, Stolarick, & Matheson, 2015).

The econometric analysis includes three complementary sections. First, the main analysis covers SEZ performance from 2007 to 2012, for which all explanatory variables are available. Second, a reduced dataset is used to analyse the growth performance of each zone in the five years after it became operational. This implies that the period of analysis varies by zone. A reduced number of explanatory variables is available for this section. Lastly, a deep dive into the Vietnamese zones allows to test whether the SEZ specific results hold once all country specific variation is eliminated. Two dependent variables are used throughout: the absolute growth of each zone and the relative growth of the zone compared to country growth.

The analysis reveals that SEZs cannot be seen as panacea for growth. In the majority of cases, SEZs are either less or only equally dynamic as the country they are placed in. This underlines the need for policy makers to carefully consider whether SEZs are the most appropriate tool for their purpose, in particular given their potentially high costs.

The econometric analysis of the drivers of SEZ performance furthermore confirms, but also refutes some of the dominant views on the viability and success of SEZs. First, zone growth is difficult to sustain over time. Most growth happens in the early years of establishment but eventually peters out. Second, trying to upgrade the technological component or value-added of SEZs is challenging. Those SEZs in our sample, which focus on low- and medium-technology sectors, outperformed their high-tech counterparts. Third, zone size matters: larger zones seem to have an advantage in terms of growth potential. And lastly, the country context determines SEZ performance. Firms look for low cost locations, but prefer those in close proximity to large cities. Proximity to big markets as well as pre-existing industrialization seem favourable for SEZ performance. In contrast, the incentive package, nature of the zone operator (public versus private) and other programme specific variables, which are frequently highlighted in the case study literature as performance drivers, seem highly context specific and are found not to be structurally correlated with SEZ performance.

IV. Conclusion

This PhD thesis explored the relationship between the spatial location of people within a country and national economic growth. The first three substantive chapters analysed whether a country's urban structure impacts economic growth at the aggregate level. Each chapter explored a different aspect, namely the level of concentration of the urban population within a country, the size of a country's cities and the role of contextual factors. The fourth essay explored the related policy question: what are the drivers of Special Economic Zone performance – a policy which is frequently employed to influence the location of economic activity and people within a country. The analysis contributes to a growing body of empirical evidence that investigates questions of urbanization, agglomeration economies and economic development in developing countries.

Across the first three chapters, I find consistent evidence to question the emphasis placed on the role of agglomeration and large cities for developing countries in much of the empirical and theoretical literature. The results in this thesis suggest that agglomeration and city size are a loci for economic growth, however only in high-income countries. In developing countries, on average, there seems to be no consistent relationship.

The results of the third paper help to interpret this finding: the city size growth relationship is found to be far from uniform and shaped by a diversity of contextual factors, such as sector specialization, urban infrastructure, institutions and government effectiveness. Only if these factors are favourable, larger cities have a catalytic effect for growth. On average, high-income countries provide this favourable environment, hence the positive relationship between agglomeration, city size and growth. In developing countries, by contrast, economies are frequently dominated by industries which do not benefit from agglomeration economies to the same degree, infrastructure is less developed, government effectiveness tends to be lower, and cities and concentration levels are much larger than in developed countries. This does not suggest, however, that it is impossible for developing countries to benefit from large cities and agglomeration in the same way developed countries do. It is merely a combination of historical circumstances (which led to the emergence of more concentrated city structures and larger cities in the first place) and current factors (such as industrial composition and urban infrastructure), which hinders developing countries to retrieve a similar benefit from agglomeration as their high-income counterpart.

From a policy point of view, these results have important implications. First, the frequently postulated trade-off between efficiency and equity may play less of a role than assumed. Indeed, for many developing countries it may be economically efficient to promote a territorially more balanced development as opposed to promoting more agglomeration. The emphasis in much of the influential ‘spatially-blind policy’ literature on the benefits of large cities and agglomeration or on the assumption that disparities will self-correct over time, seems misguided in this context. Spatial policies, which promote secondary cities and the development in lagging regions, may in fact be both economically efficient and equitable. Second, the results of chapter 3 on the contextual factors emphasise the importance of improving urban infrastructure and government capacities in order to reduce urban diseconomies arising in large cities. As cities are highly persistent over time and massive cities are already in existence, this policy area will require particular priority.

In terms of policy choice and implementation, chapter 4 highlights one of the largest challenges for spatial policies, even if one believes in the economic benefits of promoting growth in lagging areas: the actual feasibility of doing so. The analysis of the drivers of Special Economic Zones, one of the most popular tools assumed to fit this purpose, reveals that SEZs have been mainly successful in vicinity to the largest cities. The promotion of SEZs in more remote areas, which provide a less conducive context, has been proven difficult to substantiate.

The thesis contributes to the growing empirical literature employing developing country data. As with most research, rather than providing a definite answer to the questions at hand, it unveils more questions and the necessity to 'dig deeper'. Besides the need to further improve data quality and the measurement of many of the aspects considered in this thesis, a closer assessment of specific cases would be desirable to complement the more macro level research included here. How do institutions shape urban development in the developing world? Which sort of formal and informal institutions have contributed to the success of certain policies? Which were the characteristics of the underlying institutions in countries where urban development has gone right? How have some secondary cities managed to grow? Both a qualitative case study approach as well as quantitative research based on microdata will help to shed more light on these aspects.

With regards to the research on Special Economic Zones, it will be important to further validate our results as well as include other sorts of SEZ policies which are excluded from the analysis due to methodological reasons. Single country studies with a large number of SEZs will allow to gain a deeper understanding of the importance of SEZ infrastructure, locational choices as well as other services offered. Furthermore, it would be desirable to enlarge the research to include other types of Special Economic Zones, in particular those focusing on tradable services and those encompassing larger territories.

References

Acemoglu, D., Johnson, S., & Robinson, J. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics, 117*(4), 1231-1294. doi:10.1162/003355302320935025

Ades, A. F., & Glaeser, E. L. (1995). Trade and Circuses - Explaining Urban Giants. *Quarterly Journal of Economics, 110*(1), 195-227. doi:10.2307/2118515

Aggarwal, A. (2005). Performance of export processing zones: a comparative analysis of India, Sri Lanka and Bangladesh *Working Paper No.155*. New Delhi: Indian Council for Research on International Economic Relations.

Ahrend, R., Farchy, E., Kaplanis, I., & Lembcke, A. C. (2014). *What Makes Cities More Productive? Evidence on the Role of Urban Governance from Five OECD Countries*. Retrieved from Paris: http://dx.doi.org/10.1787/5jz432cf2d8p-en

Alonso, W. (1973). Urban zero population growth. *Daedalus, 102*(4), 191-206.

Asian Development Bank. (2015). *Asian Economic Integration Report 2015. How can Special Economic Zones catalyze economic development?* Manilla, Philippines: Author.

Au, C. C., & Henderson, J. V. (2006). Are Chinese Cities Too Small? *Review of Economic Studies, 73*(3), 549-576. doi:10.1111/j.1467-937X.2006.00387.x

Barrios, S., & Strobl, E. (2009). The dynamics of regional inequalities. *Regional Science and Urban Economics, 39*(5), 575-591. doi:10.1016/j.regsciurbeco.2009.03.008

Bertinelli, L., & Black, D. (2004). Urbanization and growth. *Journal of Urban Economics, 56*(1), 80-96. doi:10.1016/j.jue.2004.03.003

Bertinelli, L., & Strobl, E. (2007). Urbanisation, urban concentration and economic development. *Urban Studies, 44*(13), 2499-2510. doi:10.1080/00420980701558442

Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*(1), 48-63. doi:10.1016/j.jue.2008.08.003

Camagni, R., Capello, R., & Caragliu, A. (2013). One or infinite optimal city sizes? In search of an equilibrium size for cities. *Annals of Regional Science, 51*(2), 309-341. doi:10.1007/s00168-012-0548-7

Camagni, R., Capello, R., & Caragliu, A. (2015). The Rise of Second-Rank Cities: What Role for Agglomeration Economies? *European Planning Studies, 23*(6), 1069-1089. doi:10.1080/09654313.2014.904999

Castells-Quintana, D. (2016). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*. doi:10.1016/j.jue.2016.02.003

Castells-Quintana, D., & Royuela, V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science, 52*(2), 343-366. doi:10.1007/s00168-014-0589-1

Chandler, T. (1987). *Four thousand years of urban growth: an historical census*. Lewiston, N.Y., U.S.A: St. David's University Press.

Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences of the United States of America, 108*(21), 8589-8594. doi:10.1073/pnas.1017031108

Davis, J. C., & Henderson, J. V. (2003). Evidence on the political economy of the urbanization process. *Journal of Urban Economics, 53*(1), 98-125. doi:10.1016/S0094-1190(02)00504-1

De Long, J. B., & Shleifer, A. (1993). Princes and Merchants: European City Growth before the Industrial Revolution. *The Journal of Law and Economics, 36*(2), 671-702. doi:10.1086/467294

Duranton, G. (2008). Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics, 41*(3), 689-736. doi:10.1111/j.1540-5982.2008.00482.x

Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063-2117). Burlington: Elsevier.

Ebener, S., Murray, C., Tandon, A., & Elvidge, C. C. (2005). From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. *International Journal of Health Geographics, 4*(1). doi:10.1186/1476-072X-4-5

Faggio, G., Silva, O., & Strange, W. (2014). *Heterogenous Agglomeration*. Retrieved from http://eprints.lse.ac.uk/58426/1/sercdp0152.pdf

Farole, T. (2011). *Special Economic Zones in Africa: Comparing Performance and Learning from Global Experience*. Washington D.C.: The World Bank.

Farole, T., & Kweka, J. (2011). *Institutional Best Practices for Special Economic Zones: An Application to Tanzania*. Retrieved from http://siteresources.worldbank.org/INTRANETTRADE/Resources/SEZTanzaniaPolicynoteFinal.pdf

Fay, M., & Opal, C. (2000). Urbanization without Growth: A Not-So-Uncommon Phenomenon. *The World Bank Policy Research Working Paper Series, 2412*. Retrieved from http://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-2412 doi:10.1596/1813-9450-2412

FIAS. (2009). Special Economic Zones. Performance, Lessons Learned, and implications for Zone Development. Washington D.C.: The World Bank Group.

Fischer, M., Scherngell, T., & Jansenberger, E. (2009). Geographic localisation of knowledge spillovers: evidence from high-tech patent citations in Europe. *Annals of Regional Science, 43*(4), 839-858. doi:10.1007/s00168-009-0300-0

Fujita, M., & Thisse, J. F. (2003). Does geographical agglomeration foster economic growth? And who gains and loses from it? *Japanese Economic Review, 54*(2), 121-145. doi:10.1111/1468-5876.00250

Ghosh, T., Powell, R. L., Elvidge, C. D., Baugh, K. E., Sutton, P. C., & Anderson, S. (2010). Shedding Light on the Global Distribution of Economic Activity. *The Open Geography Journal*, *3*, 147-160. doi: 10.2174/1874923201003010147

Glaeser, E. L. (2014). A world of cities: the causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association, 12*(5), 1154-1199. doi:10.1111/jeea.12100

Glaeser, E. L., & Shapiro, J. M. (2002). Cities and Warfare: The Impact of Terrorism on Urban Form. *Journal of Urban Economics, 51*(2), 205-224. doi:10.1006/juec.2001.2262

Gollin, D., Jedwab, R., & Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth, 21*(1), 35-70. doi:10.1007/s10887-015-9121-4

Graham, D. J. (2009). Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science, 88*(1), 63-84. doi:10.1111/j.1435-5957.2008.00166.x

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development - a two sector analysis. *American Economic Review, 60*(1), 126-142.

Henderson, J. V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth, 8*(1), 47-71. doi:10.1023/A:1022860800744

Henderson, J. V. (2010). Cities and Development. *Journal of Regional Science, 50*(1), 515-540. doi:10.1111/j.1467-9787.2009.00636.x

Henderson, J. V., & Ioannides, Y. M. (1981). Aspects of growth in a system of cities. *Journal of Urban Economics, 10*(1), 117-139. doi:10.1016/0094-1190(81)90026-7

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review, 102*(2), 994-1028. doi:10.1257/aer.102.2.994

Hoselitz, B. F. (1953). The Role of Cities in the Economic Growth of Underdeveloped Countries. *Journal of Political Economy, 61*(3), 195-208. doi:10.1086/257371

Jacobs, J. (1970). *The economy of cities*. London: Cape.

Jacobs, J. (1985). *Cities and the wealth of nations : principles of economic life*. New York: Random House Inc.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced. *The Quarterly Journal of Economics, 108*(3), 577. doi:10.2307/2118401

Jedwab, R., & Vollrath, D. (2015). Urbanization without growth in historical perspective. *Explorations in Economic History, 58*, 1-21. doi:10.1016/j.eeh.2015.09.002

Krugman, P. (1991). Increasing Returns and Economic-Geography. *Journal of Political Economy, 99*(3), 483-499. doi:10.1086/261763

Lewis, A. W. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School, 28*(2), 139-191. doi:10.1111/j.1467-9957.1954.tb00021.x

Marshall, A. (1890). *Principles of economics*. London, New York: Macmillan and Co.

Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review, 42*(4), 947-968. doi:10.1111/1468-2354.00141

McCann, P., & Acs, Z. J. (2011). Globalization: Countries, Cities and Multinationals. *Regional Studies, 45*(1), 17-32. doi:10.1080/00343404.2010.505915

Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-Time Light Data: A Good Proxy Measure for Economic Activity? *Plos One*(10), 18. doi:10.1371/journal.pone.0139779

Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics, 39*(3), 332-342. doi:10.1016/j.regsciurbeco.2008.12.002

Nel, E., & Rogerson, C. (2013). Special Economic Zones in South Africa: Reflections from International Debates. *Urban Forum, 24*(2), 205-217. doi:10.1007/s12132-012-9184-7

OECD. (2006). *Competitive Cities in the Global Economy*. Retrieved from Paris: http://www.oecd.org/gov/regional-policy/oecdterritorialreviewscompetitivecitiesintheglobaleconomy.htm

OECD. (2009). Towards Best Practice Guidelines for the Development of Economic Zones. Retrieved from https://www.oecd.org/mena/competitiveness/44866585.pdf

Rigby, D. L., & Brown, W. M. (2015). Who Benefits from Agglomeration? *Regional Studies, 49*(1), 28-43. doi:10.1080/00343404.2012.753141

Rosenthal, S., & Strange, W. (2004). Chapter 49 - Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2119-2171). Burlington: Elsevier.

Sassen, S. (1991). *The Global City*. Princeton, NJ: Princeton University Press.

Scott, A. J. (2001). *Global City-regions Trends, Theory, Policy*. Cambridge: Oxford University Press.

Singa Boyenge, J.-P. (2007). *ILO database on export processing zones (Revised)*. Retrieved from Geneva: http://staging.ilo.org/public/libdoc/ilo/2007/107B09_80_engl.pdf

Storper, M. (2010). Why Does a City Grow? Specialisation, Human Capital or Institutions? *Urban Studies, 47*(10), 2027-2050. doi:10.1177/0042098009359957

The Economist. (2015). Special economic zones. Not so special. Retrieved from http://www.economist.com/news/leaders/21647615-world-awash-free-trade-zones-and-their-offshoots-many-are-not-worth-effort-not

UNIDO. (2015). Economic Zones in the Asean. Industrial Parks, Special Economic Zones, Eco Industrial Parks, Innovation Districts as Strategies for Industrial Competitiveness. Retrieved from: https://www.unido.org/fileadmin/user_media_upgrade /Resources/Publications/UCO_Viet_Nam_Study_FINAL.pdf

United Nations. (2014). World Urbanization Prospects: The 2014 Revision. Retrieved from http://esa.un.org/unpd/wup/

Venables, A. J. (2005). Spatial disparities in developing countries: cities, regions, and international trade. *Journal of Economic Geography, 5*(1), 3-21. doi:10.1093/jnlecg/lbh051

Williamson, J. G. (1965). Regional Inequality and the Process of National Development - A Description of the Patterns. *Economic Development and Cultural Change, 13*(4), 1-84. doi:10.1086/450136

World Bank. (1979). *World Development Report 1979. Structural Change and Development Policy*. Washington, D.C.: Author.

World Bank. (2009). *World Development Report 2009. Reshaping Economic Geography*. Washington, D.C.: Author.

World Bank. (2011). *Special Economic Zones. Progress, Emerging Challenges and Future Directions*. Washington, D.C.: Author.

1 The Evolution of Urban Concentration and Economic Growth

1.1 Introduction

Many countries have experienced sizeable shifts in the geographical distribution of their population over the past decades, in particular developing countries. Population growth paired with galloping urbanization rates have resulted in an increase in the size of some cities, shrinking population in others and birth of new cities. The perceived increasing concentration of people has received much attention in the literature in terms of its impact on different socio-economic aspects such as poverty reduction (Christiaensen, Weerdt, & Todo, 2013; Portes & Roberts, 2005; Sekkat, 2016), CO2 emissions (Makido, Dhakal, & Yamagata, 2012; Mohajeri, Gudmundsson, & French, 2015) and inequality (Castells-Quintana & Royuela, 2015; Oyvat, 2016). The question whether increasing agglomeration is beneficial for productivity and economic growth has been particularly high on the agenda. A growing number of academics and policy-makers have stressed the importance of urban concentration for economic growth: the concentration of people and firms in one place generates agglomeration economies and productivity gains through pooled labour markets, forward and backwards linkages as well as knowledge spill-overs (Fujita & Thisse, 2003; P. Martin & Ottaviano, 2001; Romer, 1986; Rosenthal & Strange, 2004). Many studies even posit a trade-off between a spatially-balanced economic development (equity) and economic efficiency in the form of a greater concentration of population and economic activity in one place which may potentially lead to greater national economic growth (see R. Martin, 2008 for an overview). More concentrated urban structures – in particular at low levels of development – are frequently regarded as growth enhancing (Brülhart & Sbergami, 2009; Henderson, 2003; World Bank, 2009). Policies to stimulate economic development outside of the main cities are, in contrast, often regarded as inefficient and growth limiting.

The literature on the topic, while extensive, has two important limitations. First, there is still limited information on how the levels of urban concentration have evolved across different countries in the world. While a few studies explicitly describe the patterns and evolution of urban concentration, the majority focuses on analysing what drives city size distribution within countries (Anthony, 2014; Moomaw, 2004). When they do analyse the patterns, studies often remain mainly cross-sectional (Short & Pinet-Peralta, 2009) or, when they are longitudinal, they focus on a specific country or region (Aroca & Atienza, 2016; Behrens & Bala, 2013; Cuervo & Cuervo, 2013). Frequently, general references to the well-known higher levels of urban concentration in developing countries and their social consequences also suffice as an introduction to the topic (Castells-Quintana, 2016; Henderson, 2003; Venables, 2005). A more nuanced description is therefore desirable to understand the extent to which perception of increasing concentration matches reality.

The second limitation relates to the measure of urban concentration. Most literature uses either levels of primacy i.e. the concentration of a country's urban population in the largest city (Anthony, 2014; Behrens & Bala, 2013; Cuervo & Cuervo, 2013; Henderson, 2003) or the share of the urban population living in cities above a certain size threshold (Bertinelli & Strobl, 2007; Brülhart & Sbergami, 2009; Sekkat, 2016) as indicators of urban concentration. This is particularly true for studies looking at the link between urban concentration and growth, since both indicators have the advantage of being available for many countries over several periods. They do, however, only imperfectly portray the historical shifts occurring in many countries: the share of the population in cities of a certain size has little to say about the relative distribution of the urban population; and primacy only depicts changes in the largest city and the overall size of the urban population. In the few studies, in which more sophisticated indicators of concentration are used, the samples tend to be restricted to developed countries, mainly in Europe (Gardiner, Martin, & Tyler, 2011). Whether these lessons can be applied to today's developing countries is increasingly called into question given the rapid urban transformations developing countries are undergoing and potential differences in their developing paths.

Our study specifically aims to address these two gaps in our knowledge. By assembling an entire new dataset – which permits the construction of more nuanced indicators of urban concentration for a large number of countries – we first examine how the level of urban concentration has evolved between 1985 and 2010 in different countries. Secondly, we assess how changes in urban concentration have affected economic growth in the same time period. We furthermore specifically test for differences in impact between developed and developing countries. Concerns about the potential endogeneity of the variables are addressed through instrumental variable (IV) analysis.

The paper adopts the following structure. The next section discusses possible indicators to measure the level of urban concentration, introduces the new dataset and describes the evolution of urban concentration in our sample. The following section provides an overview of the relevant literature on the link between agglomeration and growth. Section 4 introduces the model, data as well the estimation strategy. Section 5 examines the impact of a country's urban structure on its economic performance and discusses the results. The final chapter concludes and lays out some further areas for research.

1.2 The evolution of urban concentration

1.2.1 Indices of urban concentration

In order to describe the evolution of the urban structure of different countries around the world, we consider a number of indicators. Most literature concerned with the topic relies either on (i) urban primacy, (ii) the share of the urban population living in cities above a certain size threshold, (iii) Zipf's law or (iv) the Herfindahl-Hirschman-Index (HHI). Among these four, primacy and the share of the urban population in cities above a certain size are the most widely used and have been particularly popular in research that considers the link between urban concentration and growth (Bertinelli & Strobl, 2007; Brülhart & Sbergami, 2009; Castells-Quintana, 2016; Henderson, 2003).

Although there is no universally accepted definition of urban primacy, it is commonly referred to as the percentage of the urban population living in the largest city or the ratio between the population of the largest city over the sum of the population of the two to four next largest cities (Anthony, 2014; Moomaw, 2004; Short & Pinet-Peralta, 2009). Similarly, different thresholds are used for the share of the population living in cities above a certain size, most prominently either 750,000 or 1 million inhabitants (Bertinelli & Strobl, 2007; Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014). The advantage of both indicators and thus their popularity in the literature rely on their availability for many countries and over a relatively long time period. This makes them particularly apt for use in panel regressions.

As discussed in the introduction, there are, however, certain limitations when using these indicators. On the one side, the percentage of the urban population, which lives in cities of a certain size, does not say much about the relative distribution of the urban population across cities: as long as we do not know across how many cities this percentage is split nor how large the remainder of the cities are, this indicator does not reveal much about a country's urban structure. Primacy, on the other side, does address the question of the relative distribution of people across a country's cities. However, the descriptive power of primacy also has limitations as it is only concerned with the very top tier of the city distribution, potentially concealing important changes and differences across countries below the top echelons. Furthermore, both statistics are highly sensitive to how countries define what constitutes the "urban population". For instance in the UK, any settlement with 10,000 and more inhabitants is defined as urban; in the US, urban areas have 50,000 or more inhabitants, while urban clusters have between 2,500 and 50,000. Hence, the base over which both indicators is calculated may differ significantly across countries and bias the measurement.

This leaves us with Zipf's law and the HHI as possible indicators to capture the level of concentration of a country's urban structure. Firstly, Zipf's law (also called rank size rule), describes an empirical phenomenon in which the size of a country's cities follows a pareto distribution (Zipf, 1949). This law suggests that the second largest city within a country is half the size of the largest, the third largest city has a third of the population of the largest city and so forth. We can visualize this relationship by plotting the log of the population of all cities (x-axis) against the log of the cities' rank (y-axis): if the city size distribution follows Zipf's law, a straight downward line emerges with a slope of -1. Theoretically, this coefficient could be used as a measure of urban concentration since a deviation either below or above a coefficient of -1 indicates more or less evenly distributed urban structures. In practice, however, the empirical literature on Zipf's law has mainly aimed to test whether the law holds across different countries and not to describe the status and evolution of a country's urban structure (Giesen & Südekum, 2011; Ioannides & Overman, 2003; Rosen & Resnick, 1980; Soo, 2005). This is due to the fact that in order to calculate the coefficient, information on a large amount of cities and over multiple time periods is needed. Most studies requiring a time-series indicator for urban concentration have, therefore, resorted to the above-mentioned primacy or urban population share.

The final option is the Herfindahl-Hirschman-Index (HHI). The HHI is frequently used in different disciplines to measure concentration and is defined as follows:

$$HHI = \sum_{y=1}^{n_p} \left(\frac{x_{yip}}{x_{ip}} \right)^2$$

where

x_{yip} = population of city y in country i in the beginning of period p

x_{ip} = total urban population in country i in the beginning of period p

n = number of cities in country i in the beginning of period p

The HHI is the sum of the squared shares of each city's contribution to the overall urban population in the beginning of period p. It takes values between 1/n and 1, where 1 indicates complete concentration. A number of authors emphasize the desirability of using the HHI as opposed to other indicators for urban concentration due to its superiority in describing the full city size distribution (as opposed to the upper echelon only) (Castells-Quintana, 2016; Henderson, 2003). Due to the data requirements for its calculation, it has, however, been rarely used (Wheaton & Shishido, 1981).

Based on this discussion, it is clear that Zipfs law and the HHI are more suited than primacy and the share of the urban population living in cities above a certain size to examine a country's urban structure. Between the two, there is no clear advantage of one over the other in terms of their explanatory power. Our choice therefore falls on the HHI for practical purposes: the HHI is simpler to calculate for a large amount of countries and over multiple time periods.

1.2.2 City population dataset

In order to calculate the HHI, we built a new city population dataset from scratch. The dataset covers 68 countries over the period 1985 to 2010. For each country, the dataset contains information about the population of its cities at different points in time. The data was sourced from census data for each country available on citypopulation.de (Brinkhoff, no date) and complemented with information from the 2014 edition of the World Urbanization Prospects (United Nations, 2014). Different world regions are well represented in the dataset with a relatively even distribution between Africa (19%), Americas (21%), Asia (25%), Europe (32%) and Oceania (3%) (Appendix 1 provides the details). As census years vary from country to country, the data is projected to achieve a balanced dataset with information gathered in five-year intervals[8].

A frequent concern with population data for cities is whether the statistic counts the population within the administrative boundaries of a city only or that of the overall agglomeration (which may include several cities from an administrative point of view). For example, the administrative boundaries of Paris suggest a population of 2.2 million and a growth of around 1% since the early eighties. If we consider the agglomeration however, Paris looks very different in terms of its population numbers i.e. with a population of around 10.6 million and a growth of 15% over the past three decades.[9] As we are interested in external economies arising through the concentration of people and these external economies do not stop at

[8] Some countries are not included for the entire period of analysis as only few census years are available.
[9] Numbers sourced from Brinkhoff (no date)

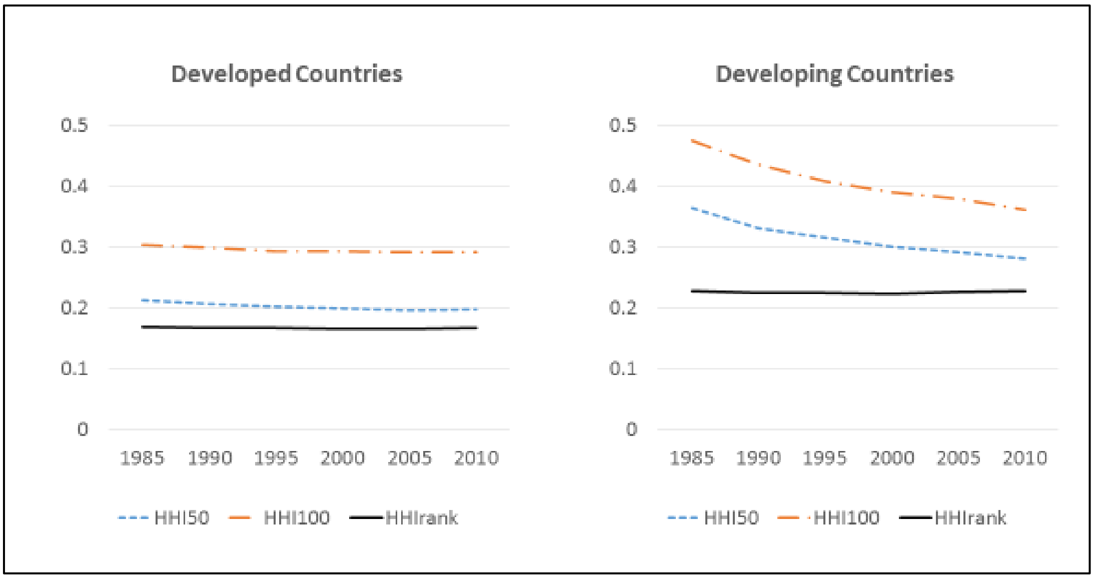
administrative boundaries, we adjust the population data to account for the size of the actual agglomerations. For this purpose, the contiguous boundaries of an agglomeration were defined through the inspection of satellite imagery, and the population of the administrative cities within these boundaries was added up. With this data at hand, the HHI was calculated.

As with other measures related to urbanization and cities, the HHI is sensitive to aggregation bias, i.e. to the number of cities included in its calculation. In order to minimise this problem, we follow the approach suggested by Cheshire (1999) and Rosen and Resnick (1980) and calculate the HHI in multiple ways: 1) based on a size cut-off, including all cities above a certain size threshold and 2) based on a specified number of cities, regardless of their size. Using a size cut-off has the advantage of including all cities considered big enough to generate agglomeration economies. However, using the size cut-off leads to a very different treatment of small and large countries. Small countries often only have a few cities with a population above the defined threshold. Resorting to a specified number of cities solves this problem, but limits the indicator to the upper tail of the city distribution for large countries. This may affect results, for instance in the case of China, where there are more than 100 cities with more than one million inhabitants today. By contrast, in smaller countries with few cities, the indicator includes more cities than if based exclusively on a size cut-off. We experiment with both approaches and calculate the following urban agglomeration measures: a) HHI50 includes all cities of a country with 50,000 or more inhabitants; b) HHI100 all cities with 100,000 inhabitants or more; and c) HHIrank includes the 25 largest cities of a country, independent of their size. This approach also has an important advantage, compared to studies relying on urban primacy, that we employ a uniform definition of what constitutes “urban” across the dataset and do not depend on each country’s definition.

1.2.3 Evolution of urban concentration

Equipped with these HHI indices, we analyse the evolution of urban concentration from 1985 to 2010. Figure 1-1 plots the average of the different HHIs dividing the sample into developed and developing countries.[10]

Figure 1-1: Evolution of urban concentration 1985 – 2010



Three insights emerge from Figure 1-1. First, there are marked differences in the levels of urban concentration. Developed countries have much less concentrated urban structures than developing countries across the three indicators. The difference is particularly marked for HHI50 and HHI100, despite a significant decrease in these indicators for developing countries over the period of analysis.

Second, the average levels of concentration have remained relatively stable for developed countries during the period of analysis, but have changed dramatically in developing countries. While the average level of HHI50 and HHI100 in developing countries was higher than in developed countries, it decreased sharply in the former (from 0.36 to 0.28 and 0.48 to 0.36 respectively), pointing to a decreasing concentration of the urban population. This trend can

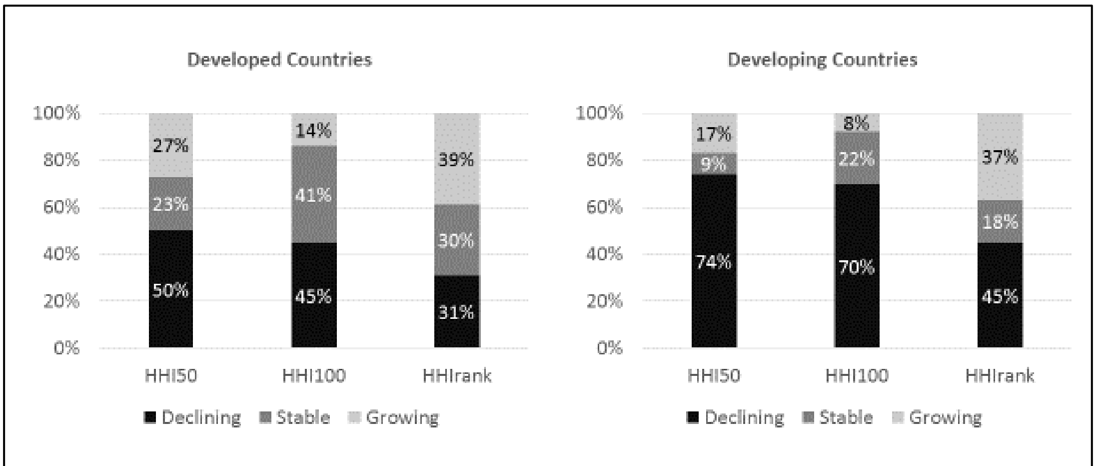
[10] For the purpose of this study, we define developed countries as European countries, US, Australia and New Zealand. Japan and Canada are not included in our dataset as the required data are not available for a sufficiently long time horizon.

be partially attributed to the rapid urbanisation process in the emerging world. Urbanization leads to an increase in the number of cities which pass the threshold for inclusion in the calculation of the indices, thus increasing the base over which the HHI is calculated.

Finally, there are differences between the indicators in terms of their evolution. The average of the HHI50 and HHI100 indicators declined – albeit at a different pace – for both developed and developing countries, suggesting lower concentration levels. In contrast, the average HHIrank indicator remained stable for both groups of countries.

Moving beyond average levels reveals greater heterogeneity between and within the indicators. Figure 1-2 shows the percentage of countries in which the different indicators either displayed a declining, a stable or a growing urban concentration. The “declining” category comprises of countries whose urban concentration levels decreased by more than 5%; “stable” countries hovered around +/- 5% from their initial level; and “growing” includes the countries whose indicator increased by more than 5%. Again, the sample is split into developed and developing countries.

Figure 1-2: Evolution of urban concentration 1985 – 2010, by category



For the developed countries sample, figure 1-2 unveils much greater diversity in urban concentration changes than the relatively stable averages Figure 1-1 suggest. For HHI50 and HHIrank, the stable group represents the smallest of the three groups (23% and 30% respectively), while for the HHI100 – where it is the second group – 41% fall into this category. The large majority of developed countries have experienced significant changes in their urban structure. The picture differs by indicator: HHI50 and HHI100 indicate that more countries had less concentrated urban structures in 2010 compared to 1985. According to the HHIrank however, almost four out of ten developed countries had more concentrated urban structures in 2010 than in 1985. Only three out of ten had less concentration. Countries within the increasing group include Spain, Norway, New Zealand and the UK; Australia, Hungary and the US belong in the decreasing group.

Among developing countries, the picture is more homogenous. Driven by increasing urbanization rates and a general increase in the size of cities, an overwhelming majority of the countries became less concentrated during the period of analysis (as measured by HHI50 and HHI100). When considering the HHIrank, the decreasing group – while not representing the majority of countries – is still the largest group. However, the development is again more diverse for the HHIrank with its fixed number of cities. Almost the same percentage of countries witnessed an increase as a decrease in concentration levels (37% vs 45%). There are clear geographical differences within these groups. The large majority of Latin American countries saw their levels of urban concentration decline regardless of the indicator considered. Many Asian countries in contrast had increasing levels of urban concentration, in particular when considering the HHIrank. One in two Asian countries is more concentrated today than in 1985. The same is true for Africa, where the picture however differs strongly between the indicators. The majority of African countries experienced increased levels of urban concentration for the HHIrank indicator, while the contrary is true for HHI100 and HHI50.

On average, countries have, thus, become less concentrated since 1985 (HHI50 and HHI100). There is limited evidence of significantly more concentrated urban structures (HHIrank). This finding is at odds with the general perception of heightened urban concentration in many countries. Going beyond the averages, however, helps to reconcile this apparent contradiction to a certain degree: individual countries both in the developing and developed world have experienced strong changes in their urban structure. If we consider the HHIrank indicator, with its fixed number of cities, there are as many countries with a more concentrated urban structure as with a less concentrated one. Furthermore, developing countries still remain more concentrated than developed countries which may have contributed to the general perception of increasing concentration.

The differences in the results for the HHI indicators also show that passing judgement on whether we live in a world with more concentration or with less concentration ultimately depends on what we deem to be the most adequate indicator of urban concentration. Each indicator captures a slightly different development and has its advantages as well as disadvantages.

1.3 Urban concentration and economic growth

Turning to our second research question, we analyse how the above described evolution of urban concentration has impacted the countries' economic performance. Whether a country's urban structure and more generally speaking its degree of agglomeration have a bearing on economic development and growth has long attracted the interest of researchers. Already more than a century ago, Alfred Marshall's seminal work on industrial districts (1890) described the productivity gains which companies experience by co-locating with other companies. Six decades later, Williamson (1965) coined the idea of an inverted U–shaped relationship between agglomeration and economic development. In this U-shaped function, levels of concentration rise with economic development and subsequently fall beyond a certain economic development threshold. The interest in the topic has not faded away and an increasing number of researchers has sought to find support for a relationship and to determine the direction of causation. In recent years, New Economic Geography (NEG) and the System of Cities literature have further increased the prominence of the question.

The NEG literature stresses the economic benefits arising from concentrated urban structures. In the basic NEG framework, the balance between centripetal and centrifugal forces determines the location of economic activity and population within a country (Fujita, Krugman, & Venables, 1999; Krugman, 1991). On the one hand, concentration creates economic benefits in the form of centripetal forces, such as forward and backward linkages, thick labour markets, and localized knowledge spill-overs. These generate economic efficiency and incentivize even more demographic and economic concentration. On the other hand, urban concentration also gives raise to centrifugal forces, such as congestion, immobile factors, and increasing land rents. These work against the agglomeration benefits and disperse activity into other regions of the country. Based on this framework, several models have been proposed to examine the effect of a country's level of concentration on economic growth and vice versa. In Fujita and Thisse (2003), localized knowledge spill-overs act as a strong centripetal force. The concentration of economic activity and population is assumed to be welfare enhancing provided that the trickle-down effect is strong enough (Fujita & Thisse, 2003). Martin and Ottaviano (2001) similarly postulate that agglomeration spurs economic growth, as it lowers the costs of innovation through pecuniary externalities. Furthermore, growth fosters agglomeration as new firms tend to locate close to other innovative firms, making economic development and agglomeration mutually reinforcing (P. Martin & Ottaviano, 2001). Puga and Venables (1996) take a slightly different angle by stressing the role of pecuniary externalities as drivers of industrial location. They find that, as economies grow, increases in wages in dynamic areas eventually push some firms to relocate to lower wage regions. Consequently, wage differentials between regions first increase, before decreasing. This finding is in line with Williamson's (1965) U-shaped curve. Overall, the theoretical NEG literature, thus, considers agglomeration beneficial for economic growth. More urban concentration is, hence, beneficial from an economic development point of view.

The urban economics literature takes a more cautious stand on this question. While it emphasizes a positive effect of increased city size and agglomeration on productivity, it also recognized that a highly concentrated urban structure could be growth hindering (Abdel-Rahman & Anas, 2004; Henderson, 2005). Similar to NEG, urban economists consider that agglomerations generate external scale economies through the sharing and matching of inputs, people and ideas (Duranton & Puga, 2004). However, these benefits from increasing city population are at some point outweighed by urban diseconomies, such as congestion and high land rents. A U-shaped trajectory is in evidence again: productivity increases with city size up to the threshold where congestion costs cause productivity to start falling. Workers and firms would thus benefit from a move to another city, creating a more balanced urban structure. Coordination failure, however, may stop people from doing so, as an individual actor is not compensated for the external benefits it creates for others. This in turn may lead to the emergence of strongly concentrated urban structures, with the majority of the urban population in one city. Venables (2005) even maintains that highly concentrated urban structures may confine countries to low economic development. As negative externalities reduce the returns to job creation, they slow down economic development. In the presence of low growth, it becomes more difficult to start a new city, leading to urban systems that are growth inhibiting rather than enhancing (Venables, 2005). Thus, from the theoretical urban economics perspective, whether a country's level of urban concentration is growth enhancing or not depends on whether the concentration of the population in a country's prime city is perceived to be already beyond the tipping point or not.

A growing body of empirical literature aims to test the predictions of the theoretical models on the relationship between agglomeration and economic growth. A first group of studies specifically examines the relationship between urban concentration and economic growth at the national level, using country-level panel data (Bertinelli & Strobl, 2007; Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014; Henderson, 2003). The studies resort to the aforementioned primacy indicator and/or the percentage of the urban population living in cities above a certain size as measures for the level of population concentration. They find a positive relationship between urban concentration and economic growth, although this conclusion is nuanced by a negative sign on the interaction term with GDP per capita. This implies that urban concentration may be particularly important at early stages of economic development.

Some studies point to important regional differences in this relationship. Pholo-Bala (2009) concludes that the relationship between urban concentration and economic growth is far from uniform across different groups of countries. While in the case of Europe the positive relationship holds, a growth trap at medium levels of urban concentration exists in Asia and Latin America. Results for Africa are highly dependent on the measure of urban structure employed (Pholo-Bala, 2009). Similarly Castells-Quintana (2016) identifies a potentially detrimental effect of urban concentration in African countries.

A second group of studies is mainly based on European data and measures agglomeration by the degree of concentration of economic activities, as opposed to urban population concentration. While economic and urban concentration are not equivalent, they tend to go hand in hand. Results emerging from this literature may thus still be informative, in particularly since these studies usually rely on a more diverse and more nuanced set of indicators than the aforementioned literature on urban concentration. Interestingly, the findings in this group are more mixed and depend on the measure employed and scale of analysis. Barrios and Strobl (2009), using the standard deviation of log of regional GDP per capita by country, identify the existence of an inverted u-curve for the relationship between GDP per capita and regional inequalities. Crozet and Koenig (2007) look at whether intra-regional inequality in GDP per capita spurs regional growth in European regions and find a positive relationship, but only in Northern countries. Gardiner et al. (2011) highlight that results depend on the spatial measure and scale used, implying that there is no unambiguous relationship between agglomeration and regional growth. Other research suggests a negative effect of agglomeration on economic growth. Bosker (2007) reports that a higher employment density translates into a lower growth rate, while Sbergami (2002) shows that more equal distribution of economic activities across regions spurs the national growth.

Consequently, the empirical literature on the link between agglomeration and national economic performance is inconclusive. Studies using measures of population concentration generally suggest that a concentrated urban structure is conducive to economic growth up to a certain threshold of economic development, but also point to important regional differences depending on levels of development. The results are also often beleaguered by their reliance on rather crude indicators. Where indicators are more sophisticated, as for the studies on economic agglomeration, studies tend to focus almost exclusively on the European experience. This calls into question the ability to generalise these results, especially to developing countries.

1.4 Model and data

In order to analyse the effect of urban concentration on national economic growth, we follow the existing literature on the topic (Brülhart & Sbergami, 2009; Castells-Quintana, 2016; Henderson, 2003) and specify a simple economic growth model. In this model the dependent variable – GDP per capita growth over five-year intervals – is explained by our different HHI indicators, GDP per capita at the beginning of the period, and a set of control variables. Our model takes the following structure:

$$\Delta GDPpc_{ip} = \alpha + \beta\, HHI_{ip} + \gamma\, GDPpc_{ip} + \delta\, \boldsymbol{X}_{ip} + \vartheta_p + \varepsilon_{ip}$$

where

$\Delta GDPpc_{ip} =$ GDP per capita growth rate of country i in period p

$\mathrm{HHI}_{\mathrm{ip}} =$ urban concentration of country i at the beginning of period p

$\mathrm{GDPpc}_{\mathrm{ip}} =$ln GDP per capita of country i at the beginning of period p

$\mathrm{X}_{\mathrm{ip}} =$ a set of control variables for country i, either measured at the beginning or as an average of period p

$\vartheta_p =$ time fixed effects

$\varepsilon_{\mathrm{ip}}$ is the error term.

Our variable of interest is HHI_{ip}, the measure of the level of urban concentration in each country i at the beginning of period p. We use the three different versions of the HHI as introduced in the section on the evolution of urban concentration. To allow for a possible non-linear relationship between urban concentration and growth, we also add the second order polynomials. The set of control variables comprises those generally included in cross-country growth regressions. These are: a) *GDP per capita* at the beginning of the period, to control for conditional convergence. Countries with a lower GDP per capita are expected to display higher growth rates, thus the coefficient should be negative (Durlauf, Johnson, & Temple, 2005); b) *private investment* due to its contribution to a country's capital stock should be a driver of economic growth and c) *government consumption* as a percentage of GDP given its potential crowding-out effect on private investment. In addition, we control for a country's size and its urbanization rate. We control both for the physical size of a country (land area) and the size of its population. Both variables may affect a country's urban structure. Small countries (both on terms of dimension and population) are likely to have a lower number of cities. This generally implies more urban concentration than larger countries. Furthermore, a country's size also reflects its market potential – larger countries tend to be more attractive to investors – and thus is also believed to influence economic performance (Alesina, Spolaore, & Wacziarg, 2005). Finally, we include a control for a country's political stability and institutional quality. For this purpose we use the state antiquity index by Putterman and Bockstette (2012). The indicator measures for every 50 year period since year 1 A.D. a) if there was a central government, b) what percentage of a country's current land mass has been governed by this government, and c) whether the government was indigenous or foreign. Bockstette, Chanda and Putterman (2002) demonstrate that this indicator is strongly correlated with a country's current institutional ability. Using a historical indicator for institutional quality has the advantage that it is not influenced by today's growth performance, thus we do not need to worry about reverse causality between the indicator and economic growth. Furthermore, alternative indicators such as the World Governance Indicators are only available since the mid-nineties, thus limiting their use in panel regressions. All regressions include time fixed effects and robust standard errors.

GDP per capita growth, initial GDP per capita, private and government investment are sourced from the eighth edition of the Penn World Tables. Population size, urbanization rate and land area are from World Development Indicators. As mentioned, the state antiquity indicator is derived from Putterman and Bockstette's dataset. Appendix 2 and Appendix 3 include a detailed overview of the specific variables used in the analysis, their data sources and descriptive statistics.

We estimate the model using the Hausman-Taylor (HT) estimator to allow for the inclusion of time-invariant variables "land area" and "state antiquity" in a panel setting (Baltagi, Bresson, & Pirotte, 2003). As in the case of fixed effects models, HT uses the within transformation of time-varying variables to estimate consistent coefficients for these variables. It also uses individual means of the time variant regressors as instruments for the time invariant variables. We also report the results of the standard country fixed effects estimator in Appendix 4.

In order to test the robustness of the HT and fixed effect results and address concerns of endogeneity, we resort to an instrumental variable design, which relies on a country's physical geographical characteristics as instruments. The rationale for the instruments and results of the robustness check are presented in a separate section.

1.5 Results

1.5.1 Main results

Table 1-1 shows the results for the HT estimator. Columns 1 – 3 present the results for the world sample, columns 4 – 6 for developed, and columns 7 – 9 for developing countries. The general model works well and the control variables show the expected signs. GDP per capita at the beginning of the period is negative and significant at the 1% level throughout all estimations, pointing towards a conditional catching-up process independent of the data sample considered. Private investment is positive and significant for seven out of nine regressions, with the results being weaker for the developing countries sample. Government consumption retains a negative coefficient throughout all estimates and is significant in the world and the developed countries sample, indicating a possible crowding out of private investments. The indicator for state history is positive and significant in all estimates. This emphasizes the importance of institutions for a country's economic performance. The results for the control variables which are more directly related to a country's urban structure, namely its population, land size and urbanization rate, vary somewhat by sample. Population is negative throughout, but insignificant in the developed countries sample. Land area is positive and significant when considering both samples individually, however not in the world sample. The level of urbanization is weakly correlated with GDP per capita growth for developed countries, but not for developing countries.

Turning to our variables of interest, we find that none of our urban concentration indices (HHIs) or their squared terms displays a significant correlation with economic growth in the world sample (columns 1 to 3). This would – in contrast to some of the previous empirical studies – suggest that a country's urban structure plays a minor role in its growth performance. However, once we divide the sample into developed and developing countries, a clearer picture emerges (columns 4 to 9). For the developed countries group (columns 4 to 6), the coefficients are consistently significant. The main terms of HHI50, HHI100 and HHIrank are positive and significant at the 1% or 5% level. The squared terms of the corresponding HHI are negative, but only significant for the HHI50 and HHIrank indicator. This indicates a positive correlation between urban concentration and economic growth for developed countries, which decreases as urban agglomeration increases.

Table 1-1: Dependent variable: GDP per capita growth in five-year intervals, 1985 – 2010

	World Sample			Developed Countries			Developing Countries		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HHI50	-0.110			4.376***			-0.257		
	(0.577)			(1.268)			(0.585)		
HHI50 squared	-0.136			-5.848***			0.0418		
	(0.434)			(2.021)			(0.458)		
HHI100		0.0106			2.275**			-0.0490	
		(0.586)			(0.908)			(0.558)	
HHI100 squared		0.0788			-2.095			0.0809	
		(0.371)			(1.333)			(0.343)	
HHIrank			-0.644			6.244**			-0.300
			(1.876)			(3.039)			(2.117)
HHIrank squared			-0.127			-14.85**			-0.339
			(3.242)			(6.729)			(3.772)
Initial GDPpc	-0.577***	-0.491***	-0.476***	-0.764***	-0.677***	-0.748***	-0.602***	-0.498***	-0.479***
	(0.0704)	(0.0869)	(0.0838)	(0.0680)	(0.111)	(0.0790)	(0.0750)	(0.0981)	(0.101)
Private investment	0.502**	0.593**	0.519**	0.469**	0.362**	0.540***	0.372	0.513*	0.372
	(0.244)	(0.241)	(0.245)	(0.186)	(0.177)	(0.173)	(0.271)	(0.290)	(0.308)
Government consumption	-0.592*	-0.610*	-0.694**	-1.249***	-1.342***	-1.238***	-0.147	-0.244	-0.327
	(0.310)	(0.317)	(0.299)	(0.185)	(0.272)	(0.236)	(0.344)	(0.332)	(0.307)
State history	1.883***	1.448***	1.473***	0.914**	0.932**	0.978**	1.905*	1.799**	1.614*
	(0.546)	(0.479)	(0.459)	(0.450)	(0.470)	(0.418)	(1.053)	(0.914)	(0.842)
LN (Population)	-0.229***	-0.140**	-0.145**	-0.106	-0.102	-0.166*	-0.644***	-0.580***	-0.547***
	(0.0878)	(0.0703)	(0.0608)	(0.0878)	(0.0724)	(0.0929)	(0.165)	(0.155)	(0.135)
LN (Land area)	0.0274	0.0153	0.0115	0.180**	0.173**	0.159**	0.323**	0.360**	0.381***
	(0.0794)	(0.0651)	(0.0597)	(0.0772)	(0.0682)	(0.0735)	(0.147)	(0.141)	(0.125)
Urbanization rate	-0.00108	0.00234	0.00310	-0.00858**	-0.00687*	-0.00648**	-0.00933*	-0.00481	-0.00514
	(0.00358)	(0.00362)	(0.00352)	(0.00394)	(0.00371)	(0.00309)	(0.00522)	(0.00538)	(0.00554)
Constant	3.875***	3.194***	3.224***	5.011***	4.237***	5.127***	1.114	-0.333	-0.509
	(1.110)	(1.128)	(0.996)	(0.894)	(0.917)	(0.830)	(1.601)	(1.820)	(1.544)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	340	352	364	130	130	136	210	222	228
Number of countries	62	64	66	23	23	24	39	41	42

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

Urban concentration in developing countries has, by contrast, no effect on economic growth. As in the case of the results for the world sample (columns 1 to 3), none of the HHI indicators or their squared terms is significant (Columns 7 to 9). This implies that in the developing and emerging world – and contrary to what is postulated in the World Development Report 2009 (World Bank, 2009) – urban agglomeration cannot be considered a factor in the promotion of greater economic growth.

The country fixed effects results (included in Appendix 4) confirm the findings of Table 1-1. Urban concentration is growth promoting in developed countries, while it does not seem to have a systematic effect in developing countries.

From an empirical perspective, our results contradict the main tenet in the literature (Bertinelli & Strobl, 2007; Brülhart & Sbergami, 2009; Henderson, 2003), that agglomeration economies are particularly important at early stages of economic development. Our results suggest that in fact for lower levels of economic development, urban concentration seems to be wholly disconnected from a country's economic performance. They are, however, in line with the results of Pholo-Bala (2009), who uncovers a positive effect of urban concentration in European countries, which represent the majority of the countries in our developed countries sample.

A potential explanation for the contrasting results may be the differing periods of analysis. Today, rich countries may benefit strongly from agglomeration economies due to a sectoral composition biased towards knowledge-intensive industries. But in less developed countries, continued high levels of urban concentration may lead to a prevalence of urban diseconomies of scale – congestion, pollution, emergence of large slums – as well as other diseconomies, such as inequality, social and political conflict. Urban diseconomies may thus largely undermine any positive effects from the concentration of economic activity in cities.

Comparing our results to the theoretical literature, it seems that NEG models, which predict a growth promoting effect of concentration, are doing a reasonable job in describing current mechanisms in the developed world. However, they do not appear accurate for the situation in many developing countries, where there seems to be no effect of urban concentration on a country's economic performance. The reasons for this discrepancy may be that agglomeration benefits as described by NEG are more prevalent in high-tech and knowledge-intensive industries (Graham, 2009; Henderson, 2010) which only represent a small part of the economy in most developing countries. Similarly, these models may underestimate agglomeration diseconomies in developing countries, which are bound to be stronger than in developed countries, given their lower endowment in urban infrastructure and higher levels of urban concentration. Moreover, the results do not confirm the presence of a low economic development trap in our sample, stipulated as a possibility by the system of cities approach. For developed countries, an efficient system of cities seems to have emerged which promotes economic growth. By contrast, in developing countries, urban structure seems to play second fiddle – if at all – to a raft of other factors which are the real determinants of economic growth.

1.5.2 Robustness check

A recurring concern in the study of the relationship between urbanization, agglomeration and economic development is the question of reverse causality. Does the urban structure of a country drive economic growth or does economic growth drive urban concentration? The reality is likely to be a mix of both. To confirm the robustness of our results, we need to identify the effect flowing from urban concentration to economic growth. In order to do that, we rely on an instrumental variable (IV) design. The aim is to see whether an IV approach confirms the results reported in Table 1-1 and in Appendix 4.

For the purpose of the IV analysis, we need a suitable instrument which is a) relevant, i.e. is driving the urban structure of a country and b) exogenous, i.e. not in any way associated with the country's economic growth performance other than through its impact on the urban structure. As city formation is strongly path dependent, i.e. cities persist and often thrive in the places where they were originally founded (Bleakley & Lin, 2012), the drivers of city formation in the past may provide suitable instruments, capable of predicting a country's urban structure today. Geographical factors have been crucial for the establishment of cities. Cities have been traditionally set up in areas with suitable terrain and easy access to water and fertile land. Proximity to fertile land as a means to secure a regular food supply for the city's population has been the paramount criterion for the establishment of cities (Motamed, Florax, & Masters, 2014). A large supply of fertile land has contributed to the development of larger cities and thus more concentrated urban structures. Limited access to fertile, arable land, by contrast, may have encouraged the formation of smaller cities and therefore provided the foundation for the development of a more balanced urban structure today. Ruggedness also affects the suitability of the land to build cities and in particular to trade between places (Nunn & Puga, 2012). The prevalence of rugged terrains is therefore likely to affect the formation of viable systems of cities. We therefore consider these two geographical properties to determine the suitability of countries to develop large cities and more or less balanced systems of cities.

Concerning the exogeneity of these indicators, it is conceivable that both the soil quality and the ruggedness of a country may impact a country's overall level of economic development and through this today's economic growth performance. A look at our dataset, however, does not reveal any clear correlations. The simple correlation between the share of fertile land in a country and GDPpc growth is -0.03. That between ruggedness and economic growth: -0.02. These very low correlations are an indication that while these factors may have determined the formation of cities and systems of cities in the past, their role in providing urban prosperity today is almost non-existent. The below examples further confirms this notion. Some countries with large percentage of fertile land, such as Benin or the Philippines, are poor, whilst others with similar access to fertile land are highly developed (e.g. France, Germany). Similarly, some countries with high levels of ruggedness are highly developed (e.g. Switzerland, Greece), while others are at the bottom of the pyramid (e.g. Nepal, Rwanda).

For the sake of simplicity, we only include the main term of the HHIs in the IV regression. We estimate the model using a pooled 2SLS estimator with robust standard errors, as our instruments are time-invariant and we can therefore not run the regressions in a panel setting.

All control variables remain the same as in the main regression. Furthermore, we add regional dummies in order to account for some of the heterogeneity between the regions. Table 1-2 displays the second stage results. Our instruments work well with the first stage F-statistics passing comfortably the rule of thumb threshold for strong instruments proposed by Staiger and Stock (1997) and also exceeding the Hausman, Stock and Yogo (2005) threshold values.[11] The first step regressions are included in Appendix 5.

The coefficients of the IV analysis confirm our main results from Table 1-1 and Appendix 4. All HHI indicators are insignificant in the world and developing countries samples, while HHIs are positive and significant in the developed countries sample (Table 1-2). Thus, the positive impact of urban concentration on economic growth is confirmed for developed countries.

[11] Note that in order to pass the standard test for the relevance of the instruments; we use ruggedness as an instrument for the world and the developing countries sample. For the developed countries sub-sample, we rely on both soil quality and ruggedness.

Table 1-2: IV-estimates – Dependent variable: GDP per capita growth, 1985 – 2010

	World Sample			Developed Countries			Developing Countries		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HHI50	0.113			0.379**			0.0781		
	(0.236)			(0.174)			(0.378)		
HHI100		0.119			0.381**			0.134	
		(0.245)			(0.171)			(0.383)	
HHIrank			0.237			0.535**			0.207
			(0.424)			(0.269)			(0.780)
Initial GDPpc	-0.0769***	-0.0852***	-0.0868***	-0.142***	-0.156***	-0.0974**	-0.107***	-0.118***	-0.122***
	(0.0267)	(0.0275)	(0.0249)	(0.0468)	(0.0487)	(0.0410)	(0.0351)	(0.0339)	(0.0367)
Private investment	0.237	0.402***	0.388**	-0.0268	-0.133	-0.206	0.207	0.388**	0.363
	(0.159)	(0.154)	(0.167)	(0.383)	(0.382)	(0.393)	(0.191)	(0.182)	(0.261)
Government consumption	-0.690***	-0.590**	-0.554**	-0.894*	-1.087**	-0.888*	-0.563*	-0.447	-0.415
	(0.252)	(0.267)	(0.225)	(0.472)	(0.491)	(0.467)	(0.297)	(0.316)	(0.276)
State history	-0.0365	0.000810	0.0174	0.00832	0.0572	-0.0581	0.0379	0.0604	0.121
	(0.0635)	(0.0751)	(0.0697)	(0.0970)	(0.0958)	(0.101)	(0.0796)	(0.124)	(0.0849)
LN (Population)	0.0135	0.0282	0.0151	-0.0185	0.00357	-0.0200	0.0211	0.0443	0.0277
	(0.0285)	(0.0389)	(0.0144)	(0.0201)	(0.0245)	(0.0170)	(0.0475)	(0.0609)	(0.0180)
LN (Land area)	0.00195	0.00342	0.00554	0.0213**	0.0146*	0.0166*	-0.00803	-0.00538	-0.00658
	(0.00998)	(0.0100)	(0.0140)	(0.00846)	(0.00874)	(0.00943)	(0.0151)	(0.0148)	(0.0216)
Urbanization rate	0.00190	0.00199	0.00157*	0.00275**	0.00304**	0.00205	0.00313	0.00345	0.00290**
	(0.00133)	(0.00184)	(0.000874)	(0.00132)	(0.00138)	(0.00149)	(0.00193)	(0.00240)	(0.00124)
Constant	0.609***	0.518**	0.537**	1.106**	1.290**	0.916**	0.888***	0.746*	0.835*
	(0.175)	(0.205)	(0.251)	(0.488)	(0.505)	(0.468)	(0.312)	(0.409)	(0.432)
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First stage F-stat	24.47	18.73	18.81	20.49	29.70	19.47	17.39	10.21	13.45
Observations	340	352	364	130	130	136	210	222	228
R-squared	0.302	0.315	0.298	0.213	0.268	0.188	0.365	0.375	0.373

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

1.6 Conclusion

In this paper we have investigated two closely related topics: firstly, the evolution of urban concentration in 68 countries from 1985 to 2010, employing a set of more nuanced indicators than used in the literature thus far. Secondly, the extent to which the degree of urban concentration affected national economic growth in the same period. In order to overcome the data limitations of past empirical literature on the topic, a new city population dataset was built that allows constructing a set of Herfindahl-Hirschman-Indices. We, furthermore, specifically analysed differences between developed and developing countries and are able to account for potential endogeneity through an IV design in our regressions.

The results indicate that contrary to the general perception of heightened urban concentration, levels of urban concentration have decreased on average or remained stable around the world, depending on the indicator employed. These averages, however, conceal strong changes and differences at the country level. Many countries have experienced significant shifts in their urban structure: increased concentration can be found in many developed countries as well as in Asia and some countries of Africa. Levels of concentration have decreased in most parts of Latin America. In general, developing countries still remain much more concentrated than developed countries.

The results of our analysis on the relationship between urban concentration and economic growth suggest that there is no uniform relationship between urban concentration and economic growth. Urban concentration is beneficial for economic growth in high-income countries, but this effect does not hold for developing countries. This contrasts with previous studies that find a particularly important effect for low levels of economic development.

What are the implications of our results for policy makers who face the question of whether to promote further agglomeration or to promote development outside of the primary urban areas? As with most questions, there is no easy answer. On the one hand, the results show that, despite decreasing levels of urban concentration in many developing countries, urban concentration still remains high and many countries saw their levels increase. Urban concentration may, thus, not be self-correcting with economic development as frequently hypothesized. Furthermore, the analysis dispel the prevailing NEG notion that a more concentrated urban structure is best for economic growth, in particular at low levels of economic development. While no uniform relationship can be deducted, the results show that most countries in the developing world are likely to suffer more from congestion generated through increased concentration than benefit from it. Hence, promoting development outside of primary urban areas would be beneficial from an economic point of view.

On the other hand, developed countries with the right urban infrastructure in place and an economy with industries strongly benefiting from agglomeration economies highlight that countries can, in fact, benefit from urban concentration. This implies that sweeping policy recommendations are ill-advised and that more specific, country-based research may be the way forward in order to set up policies that may foster and make best use of the economic potential of cities and urban agglomeration on a case by case basis.

References

Abdel-Rahman, H. M., & Anas, A. (2004). Chapter 52 - Theories of systems of cities. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2293-2339). Burlington: Elsevier.

Alesina, A., Spolaore, E., & Wacziarg, R. (2005). Chapter 23 - Trade, Growth and the Size of Countries. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part B, pp. 1499-1542). Amsterdam: Elsevier.

Anthony, R. M. (2014). Bringing up the past: Political experience and the distribution of urban populations. *Cities, 37*, 33-46. doi:10.1016/j.cities.2013.11.005

Aroca, P., & Atienza, M. (2016). Spatial concentration in Latin America and the role of institutions. *Investigaciones Regionales*, (36), 233-253. Retrieved from http://www.aecr.org/images/ImatgesArticles/2016/12/11_ArocaAtienzareduc.pdf

Baltagi, B. H., Bresson, G., & Pirotte, A. (2003). Fixed effects, random effects or Hausman-Taylor? A pretest estimator. *Economics Letters, 79*(3), 361-369. doi:10.1016/s0165-1765(03)00007-7

Barrios, S., & Strobl, E. (2009). The dynamics of regional inequalities. *Regional Science and Urban Economics, 39*(5), 575-591. doi:10.1016/j.regsciurbeco.2009.03.008

Behrens, K., & Bala, A. P. (2013). Do rent-seeking and interregional transfers contribute to urban primacy in Sub-Saharan Africa? *Papers in Regional Science, 92*(1), 163-195. doi:10.1111/j.1435-5957.2011.00400.x

Bertinelli, L., & Strobl, E. (2007). Urbanisation, urban concentration and economic development. *Urban Studies, 44*(13), 2499-2510. doi:10.1080/00420980701558442

Bleakley, H., & Lin, J. (2012). Portage and Path Dependence. *Quarterly Journal of Economics, 127*(2), 587-644. doi:10.1093/qje/qjs011

Bockstette, V., Chanda, A., & Putterman, L. (2002). States and markets: The advantage of an early start. *Journal of Economic Growth, 7*(4), 347-369. doi:10.1023/a:1020827801137

Bosker, M. (2007). Growth, Agglomeration and Convergence: a Space-time Analysis for European Regions. *Spatial Economic Analysis, 2*(1), 91-100. doi:10.1080/17421770701255237

Brinkhoff, T. (no date). City Population. Retrieved from http://www.citypopulation.de

Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*(1), 48-63. doi:10.1016/j.jue.2008.08.003

Castells-Quintana, D. (2016). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*. doi:10.1016/j.jue.2016.02.003

Castells-Quintana, D., & Royuela, V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science, 52*(2), 343-366. doi:10.1007/s00168-014-0589-1

Castells-Quintana, D., & Royuela, V. (2015). Are Increasing Urbanisation and Inequalities Symptoms of Growth? *Applied Spatial Analysis, 8*(3), 291-308. doi:10.1007/s12061-015-9146-2

Cheshire, P. (1999). Chapter 35 - Trends in sizes and structures of urban areas *Handbook of Regional and Urban Economics* (Vol. 3, pp. 1339-1373). Amsterdam: North-Holland.

Christiaensen, L., Weerdt, J., & Todo, Y. (2013). Urbanization and poverty reduction: the role of rural diversification and secondary towns. *Agricultural Economics, 44*(4-5), 435-447. doi:10.1111/agec.12028

Crozet, M., & Koenig, P. (2007). The Cohesion vs Growth Tradeoff: Evidence from EU Regions (1980-2000). Paris: Paris School of Economics.

Cuervo, L. M., & Cuervo, N. (2013). Urban Primacy and Regional Economic Disparities in Latin America. In J. R. Cuadrado-Roura & P. Aroca (Eds.), *Regional Problems and Policies in Latin America* (pp. 135-162). Berlin: Springer.

Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063-2117). Burlington: Elsevier.

Durlauf, S. N., Johnson, P. A., & Temple, J. R. W. (2005). Chpater 8 - Growth Econometrics. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part A, pp. 555-677). Amsterdam: Elsevier.

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: cities, regions and international trade*. Cambridge Mass.: MIT Press.

Fujita, M., & Thisse, J. F. (2003). Does geographical agglomeration foster economic growth? And who gains and loses from it? *Japanese Economic Review, 54*(2), 121-145. doi:10.1111/1468-5876.00250

Gardiner, B., Martin, R., & Tyler, P. (2011). Does spatial agglomeration increase national growth? some evidence from Europe. *Journal of Economic Geography*, *11*(6), 979-1006. doi:10.1093/jeg/lbq047

Giesen, K., & Südekum, J. (2011). Zipf's law for cities in the regions and the country. *Journal of Economic Geography, 11*(4), 667-686. doi:10.1093/jeg/lbq019

Graham, D. J. (2009). Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science*, *88*(1), 63-84. doi:10.1111/j.1435-5957.2008.00166.x

Hausman, J., Stock, J. H., & Yogo, M. (2005). Asymptotic properties of the Hahn-Hausman test for weak-instruments. *Economics Letters, 89*(3), 333-342. doi:10.1016/j.econlet.2005.06.007

Henderson, J. V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth, 8*(1), 47-71. doi:10.1023/A:1022860800744

Henderson, J. V. (2005). Chapter 24 - Urbanization and Growth. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part B, pp. 1543-1591). Amsterdam: Elsevier.

Henderson, J. V. (2010). Cities and Development. *Journal of Regional Science*, *50*(1), 515-540. doi:10.1111/j.1467-9787.2009.00636.x

Ioannides, Y. M., & Overman, H. G. (2003). Zipf's law for cities: an empirical examination. *Regional Science and Urban Economics, 33*(2), 127-137. doi:10.1016/S0166-0462(02)00006-6

Krugman, P. (1991). Increasing Returns and Economic-Geography. *Journal of Political Economy*, *99*(3), 483-499. doi:10.1086/261763

Makido, Y., Dhakal, S., & Yamagata, Y. (2012). Relationship between urban form and CO2 emissions: Evidence from fifty Japanese cities. *Urban Climate, 2*, 55-67. doi:10.1016/j.uclim.2012.10.006

Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review, 42*(4), 947-968. doi:10.1111/1468-2354.00141

Martin, R. (2008). National growth versus spatial equality? A cautionary note on the new 'trade-off' thinking in regional policy discourse. *Regional Science Policy & Practice, 1*(1), 3-13. doi:10.1111/j.1757-7802.2008.00003.x

Mohajeri, N., Gudmundsson, A., & French, J. R. (2015). CO2 emissions in relation to street-network configuration and city size. *Transportation Research Part D, 35*, 116-129. doi:10.1016/j.trd.2014.11.025

Moomaw, R. L. (2004). An empirical analysis of competing explanations of urban primacy evidence from Asia and the Americas. *Annals of Regional Science, 38*(1), 149-172. doi:10.1007/s00168-003-0137-x

Motamed, M. J., Florax, R., & Masters, W. A. (2014). Agriculture, transportation and the timing of urbanization: Global analysis at the grid cell level. *Journal of Economic Growth, 19*(3), 339-368. doi:10.1007/s10887-014-9104-x

Nunn, N., & Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics, 94*(1), 20-36. doi:10.1162/REST_a_00161

Oyvat, C. (2016). Agrarian Structures, Urbanization, and Inequality. *World Development, 83*, 207. doi:10.1016/j.worlddev.2016.01.019

Pholo-Bala, A. (2009). Urban concentration and economic growth: checking for specific regional effects. *Center for Operations Research and Econometrics (CORE). Université Catholique de Louvain. Discussion Paper 2009/38*. Retrieved from https://www.uclouvain.be/cps/ucl/doc/core/documents/Chapter2COREDP.pdf

Portes, A., & Roberts, B. (2005). The free-market city: Latin American urbanization in the years of the neoliberal experiment. *Studies in Comparative International Development, 40*(1), 43-82. doi:10.1007/BF02686288

Puga, D., & Venables, A. J. (1996). The spread of industry: Spatial agglomeration in economic development. *Journal of the Japanese and International Economies, 10*(4), 440-464. doi:10.1006/jjie.1996.0025

Putterman, L., & Bockstette, V. (2012). State Antiquity Index ("Statehist") Version 3.1. Retrieved from http://www.econ.brown.edu/ fac/louis_putterman/ antiquity%20index.htm

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy, 94*(5), 1002-1037. doi:10.1086/261420

Rosen, K. T., & Resnick, M. (1980). The size distribution of cities: An examination of the Pareto law and primacy. *Journal of Urban Economics, 8*(2), 165-186. doi:10.1016/0094-1190(80)90043-1

Rosenthal, S., & Strange, W. (2004). Chapter 49 - Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2119-2171). Burlington: Elsevier.

Sbergami, F. (2002). Agglomeration and Economic Growth: Some Puzzles. *HEI Working Paper No: 02/2002*. Retrieved from http://repec.graduateinstitute.ch/pdfs/Working_papers/HEIWP02-2002.pdf

Sekkat, K. (2016). Urban Concentration and Poverty in Developing Countries. *Growth and Change*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/grow.12166/full

Short, J. R., & Pinet-Peralta, L. M. (2009). Urban Primacy: Reopening the Debate. *Geography Compass, 3*(3), 1245-1266. doi:10.1111/j.1749-8198.2009.00234.x

Soo, K. T. (2005). Zipf's Law for cities: a cross-country investigation. *Regional Science and Urban Economics, 35*(3), 239-263. doi:10.1016/j.regsciurbeco.2004.04.004

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557-586. doi:10.2307/2171753

United Nations. (2014). World Urbanization Prospects: The 2014 Revision. Retrieved from http://esa.un.org/unpd/wup/

Venables, A. J. (2005). Spatial disparities in developing countries: cities, regions, and international trade. *Journal of Economic Geography, 5*(1), 3-21. doi:10.1093/jnlecg/lbh051

Wheaton, W. C., & Shishido, H. (1981). Urban Concentration, Agglomeration Economies, and the Level of Economic Development. *Economic Development and Cultural Change, 30*(1), 17-30. doi:10.1086/452537

Williamson, J. G. (1965). Regional Inequality and the Process of National Development - A Description of the Patterns. *Economic Development and Cultural Change, 13*(4), 1-84. doi:10.1086/450136

World Bank. (2009). *World Development Report 2009. Reshaping Economic Geography*. Washington, D.C.: Author.

Zipf, G. K. (1949). *Human behavior and the principle of least effort: an introduction to human ecology*. Cambridge, Mass.: Addison Wesley.

Appendix 1 – City population dataset

Country	HHI50	HHI100	HHIrank	Number of cities above 10K	Restrictions on inclusion in regressions
Argentina	Yes	Yes	Yes	179 - 180	
Australia	Yes	Yes	Yes	102	
Austria	Yes	Yes	Yes	63 - 76	
Azerbaijan	Yes	Yes	Yes	31 - 35	
Bangladesh	Yes	Yes	Yes	123 - 247	
Belgium	Yes	Yes	-	18 - 18	
Benin	Yes	Yes	Yes	21 - 60	
Bolivia	Yes	Yes	Yes	19 - 44	
Brazil	-	-	Yes	53 - 56	
Bulgaria	Yes	Yes	Yes	48 - 48	
Cambodia	Yes	Yes	Yes	31 - 33	1995 – 2010
Czech Republic	Yes	Yes	Yes	68	
Chile	Yes	Yes	Yes	59 - 61	
China	Yes	Yes	Yes	1649 - 1686	
Colombia	Yes	Yes	Yes	125 - 153	
Cote d`Ivoire	Yes	Yes	Yes	43 - 80	1985 – 2000
Denmark	Yes	Yes	Yes	43 - 58	
Dominican Republic	Yes	Yes	Yes	24 - 71	
Ecuador	Yes	Yes	Yes	40 - 78	
Egypt	Yes	Yes	Yes	69 - 179	
Ethiopia	Yes	Yes	Yes	138	
Finland	Yes	Yes	Yes	52 - 54	
France	Yes	Yes	Yes	393 - 475	
Germany	Yes	Yes	Yes	226 - 226	1990 – 2010
Ghana	Yes	Yes	Yes	31 - 71	
Greece	Yes	Yes	Yes	58 - 66	
Honduras	Yes	Yes	Yes	17 - 38	
Hungary	Yes	Yes	Yes	69 - 69	
India	Yes	Yes	Yes	1996 - 2332	
Ireland	Yes	Yes	Yes	25 - 44	
Italy	-	-	Yes	31 - 31	
Jordan	Yes	Yes	Yes	40 - 41	1990 – 2010
Kazakhstan	Yes	Yes	Yes	74 - 87	
Kyrgyzstan	Yes	Yes	Yes	23 - 27	
Mali	Yes	Yes	Yes	17 - 46	
Malaysia	Yes	Yes	Yes	113	1990 - 2010
Mexico	-	Yes	Yes	92 - 93	
New Zealand	Yes	Yes	Yes	54	
Mongolia	Yes	Yes	-	19 - 21	
Morocco	Yes	Yes	Yes	75 - 100	
Mozambique	Yes	Yes	Yes	60 - 82	1995 – 2010

Country	HHI50	HHI100	HHIrank	Number of cities above 10K	Restrictions on inclusion in regressions
Nepal	Yes	Yes	Yes	21 - 58	
Niger	Yes	Yes	Yes	8 - 32	
Nigeria	-	-	Yes	37 - 42	
Norway	Yes	Yes	Yes	49 - 49	
Pakistan	Yes	Yes	Yes	135 - 138	1985 – 2000
Panama	Yes	Yes	Yes	13 - 25	
Paraguay	Yes	Yes	Yes	10 - 19	
Peru	Yes	Yes	Yes	54 - 117	
Philippines	Yes	Yes	Yes	130 - 265	
Poland	Yes	Yes	Yes	226 - 234	
Portugal	Yes	Yes	Yes	51 - 54	
Romania	Yes	Yes	Yes	129 - 137	
Russia	Yes	Yes	Yes	665 - 761	
Senegal	Yes	Yes	Yes	19 - 52	
Slovak Republic	Yes	Yes	Yes	43 - 43	
Spain	Yes	Yes	Yes	59 - 59	
Sweden	Yes	Yes	Yes	117 - 117	
Switzerland	Yes	Yes	Yes	50 - 50	
Tajikistan	Yes	Yes	Yes	27 - 41	
Turkey	Yes	Yes	Yes	135 - 279	
Uganda	Yes	Yes	Yes	11 - 86	
United Kingdom	-	-	Yes	28 - 28	
Ukraine	Yes	Yes	Yes	210 - 216	
Uruguay	Yes	Yes	Yes	26 - 31	
United States	Yes	Yes	Yes	928 - 929	
Vietnam	Yes	Yes	Yes	39 - 145	
Zambia	Yes	Yes	Yes	24 - 48	

Appendix 2 – Data sources for control variables

Variable	Source
GDPpc	Natural logarithm of "Real GDP at constant 2005 national prices "/ "Population" Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Private Investment	Five-year averages of "Share of gross capital formation at current PPPs" Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Government Consumption	Five-year averages of "Share of government consumption at current PPPs" Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Land Area	Natural logarithm of "Land area (sq. km)" World Development Indicators http://data.worldbank.org/data-catalog/world-development-indicators
Population	Natural logarithm of "Population" World Development Indicators http://data.worldbank.org/data-catalog/world-development-indicators
Urbanization rate	% of population living in urban areas World Development Indicators http://data.worldbank.org/data-catalog/world-development-indicators
State Antiquity	Variable "statehistn50v3" State Antiquity Index (Statehist) database http://www.econ.brown.edu/fac/louis_putterman/antiquity%20index.htm
Ruggedness	Variable "rugged_slope" Nathan Nunn & Diego Puga dataset http://diegopuga.org/data/rugged/
Soil	% of fertile soil Nathan Nunn & Diego Puga dataset http://diegopuga.org/data/rugged/

Appendix 3 – Descriptive statistics

Variable	Mean	Standard Deviation			Maximum	Minimum	Obs.
		Overall	Between	Within			
HHI50	.2686185	.1920574	.1893	.037806	.8484542	.0142202	345
HHI100	.3609915	.2616493	.2573724	.055666	1	.0046886	357
HHrank	.204461	.1014418	.1007964	.0134064	.4592862	.0533313	369
Growth	.086799	.1705554	.0810117	.1517875	.6918678	-1.052486	369
Ln(GDPpc)	8.64969	1.29269	1.27156	.2332912	10.98374	5.7789	369
Private investment	.1949938	.0734472	.0602916	.0441101	.4988933	.0304211	369
Government consumption	.1865482	.0809189	.0643119	.053734	.6061998	.056944	369
State history	.7349033	.1646811	-	-	.9995915	.3611112	369
LN (Population)	2.975971	1.348927	1.330788	.1456344	7.158376	.6693815	369
LN (Land area)	12.87616	1.395671	-	-	16.61218	10.58479	369
Urbanization rate	55.08527	22.28428	22.0232	3.5355	93.319	6.091	369

Appendix 4 – Fixed effects results

Dependent variable: GDP per capita growth in five-year intervals – 1985 - 2010

	World Sample			Developed Countries			Developing Countries		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HHI50	0.00529			4.426***			-0.220		
	(0.533)			(1.249)			(0.509)		
HHI50 squared	-0.213			-5.865***			0.00755		
	(0.372)			(1.993)			(0.388)		
HHI100		0.222			2.238**			0.175	
		(0.330)			(0.953)			(0.307)	
HHI100 squared		-0.0964			-1.905			-0.0954	
		(0.258)			(1.386)			(0.240)	
HHIrank			0.153			6.261*			0.139
			(1.760)			(3.046)			(2.064)
HHIrank squared			-1.084			-14.83**			-0.468
			(3.052)			(6.658)			(3.766)
Initial GDPpc	-0.609***	-0.526***	-0.514***	-0.754***	-0.658***	-0.742***	-0.625***	-0.536***	-0.522***
	(0.0716)	(0.0883)	(0.0875)	(0.0766)	(0.125)	(0.0893)	(0.0732)	(0.0930)	(0.0947)
Private investment	0.577**	0.720***	0.661**	0.450**	0.322	0.529***	0.403	0.575**	0.460
	(0.238)	(0.247)	(0.257)	(0.192)	(0.202)	(0.179)	(0.254)	(0.279)	(0.298)
Government consumption	-0.511*	-0.544**	-0.593**	-1.245***	-1.330***	-1.234***	-0.152	-0.225	-0.281
	(0.261)	(0.270)	(0.265)	(0.175)	(0.250)	(0.233)	(0.255)	(0.253)	(0.234)
LN (Population)	-0.295**	-0.281**	-0.271*	-0.168	-0.219	-0.205	-0.891***	-0.907***	-0.917***
	(0.138)	(0.139)	(0.147)	(0.174)	(0.165)	(0.196)	(0.243)	(0.249)	(0.261)
Urbanization rate	-0.00101	0.00372	0.00400	-0.00852**	-0.00667*	-0.00642**	-0.00769	-0.00247	-0.00302
	(0.00436)	(0.00522)	(0.00513)	(0.00387)	(0.00363)	(0.00310)	(0.00536)	(0.00597)	(0.00615)
Constant	6.351***	5.243***	5.251***	8.361***	7.538***	8.174***	8.227***	7.335***	7.557***
	(0.853)	(1.019)	(1.058)	(0.720)	(1.022)	(0.777)	(1.038)	(1.205)	(1.334)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	352	364	370	130	130	136	222	234	234
R-squared	0.595	0.544	0.539	0.747	0.719	0.719	0.691	0.635	0.635
Number of countries	64	66	67	23	23	24	41	43	43

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

Appendix 5 – First-stage IV regression results

	World Sample			Developed Countries			Developing Countries		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	HHI50	HHI100	HHIrank	HHI50	HHI100	HHIrank	HHI50	HHI100	HHIrank
Ruggedness	**-0.0126*** **	**-0.0122*** **	**-0.00712*** **	**-0.0174*** **	**-0.0161*** **	**-0.00930*** **	**-0.0141*** **	**-0.0140*** **	**-0.00698*** **
	(0.00254)	**(0.00281)**	**(0.00164)**	**(0.00387)**	**(0.00367)**	**(0.00233)**	**(0.00338)**	**(0.00437)**	**(0.00190)**
Soil				**0.00333*** **	**0.00344*** **	**0.00202*** **			
				(0.000598)	**(0.000492)**	**(0.000368)**			
Initial GDPpc	0.0504***	0.0547**	0.0141	0.110***	0.146***	0.0378	0.00232	-0.00784	0.0140
	(0.0176)	(0.0218)	(0.0115)	(0.0359)	(0.0339)	(0.0246)	(0.0321)	(0.0385)	(0.0171)
Private investment	-0.0508	-0.0803	-0.117*	0.0396	0.327	0.0745	-0.0797	-0.121	-0.254***
	(0.153)	(0.164)	(0.0694)	(0.246)	(0.226)	(0.159)	(0.188)	(0.202)	(0.0775)
Government consumption	-0.00749	0.205	-0.145**	-0.569**	-0.0359	-0.302*	0.134	0.348	-0.101
	(0.128)	(0.163)	(0.0658)	(0.268)	(0.223)	(0.156)	(0.175)	(0.228)	(0.0803)
State History	0.222***	0.292***	0.146***	0.278***	0.142**	0.280***	0.220**	0.444***	0.0803*
	(0.0538)	(0.0646)	(0.0307)	(0.0714)	(0.0625)	(0.0401)	(0.0913)	(0.124)	(0.0465)
LN (Population)	-0.124***	-0.161***	-0.0260***	-0.167***	-0.225***	-0.0746***	-0.118***	-0.152***	-0.0136*
	(0.0111)	(0.0119)	(0.00522)	(0.0200)	(0.0163)	(0.0104)	(0.0146)	(0.0157)	(0.00714)
LN (Land area)	-0.00101	2.85e-06	-0.0229***	0.0358***	0.0564***	0.0210**	-0.00466	-0.0150	-0.0253***
	(0.00965)	(0.0116)	(0.00519)	(0.0131)	(0.0128)	(0.00884)	(0.0142)	(0.0179)	(0.00744)
Urbanization rate	-0.00459***	-0.00682***	-0.000377	-0.00658***	-0.00715***	-0.00373***	-0.00270*	-0.00435**	0.000194
	(0.000851)	(0.00105)	(0.000528)	(0.00107)	(0.00104)	(0.000960)	(0.00143)	(0.00175)	(0.000723)
Constant	0.371***	0.578***	0.473***	-0.689**	-1.244***	-0.232	0.669***	1.001***	0.510***
	(0.126)	(0.158)	(0.0806)	(0.313)	(0.277)	(0.211)	(0.240)	(0.280)	(0.127)
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	340	352	364	130	130	136	210	222	228
R-squared	0.565	0.621	0.384	0.658	0.779	0.522	0.522	0.576	0.364

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

2 Average City Size and Economic Growth

2.1 Introduction

The role of cities has been at the forefront of development policy debates over the past two decades. The rapid urbanization process in many developing countries paired with the ever increasing size of cities are simultaneously hailed as key drivers of productivity and growth, as well as big challenges for the developing and emerging worlds. The prominence of the topic is not a surprise given the sheer magnitude and speed of the changes. Recent statistics (United Nations, 2014) illustrate well these developments. While there were only 3 cities with more than 10 million inhabitants in 1960, there are 29 today. Similarly, the number of cities with 5 to 10 million inhabitants increased from 9 to 44. Hence, not only do more people live in cities, they also live on average in far larger cities than 50 years ago.

This is particularly true for the developing world, where the most dramatic changes have occurred in the past decades. Massive mega-cities have sprung up virtually anywhere, including countries as diverse as Pakistan, Peru, the Philippines, the Democratic Republic of the Congo and Bangladesh, driving up average city sizes everywhere in the developing world. Once the largest cities were found in high-income countries; today it is the developing world that accounts for the majority of the global urbanization and for most mega cities (McCann & Acs, 2011; United Nations, 2014).

While many researchers and policy makers have voiced concerns about these trends and their social (i.e. increased urban congestion), environmental (i.e. increased pollution) and economic (i.e. rising interpersonal and interspatial inequality) consequences, the 2009 World Development Report summarizes well an often dominant view in the economic development policy sphere: "No country has grown to middle-income without industrializing and urbanizing. None has grown to high-income without vibrant cities. The rush to cities in developing countries seems chaotic, but it is necessary. It seems unprecedented, but it has happened before" (World Bank, 2009, p. 24). Consequently, the rapid urbanization in the developing world, with its rising average city size and sprawling megacities, is seen as inevitable and beneficial from an economic perspective.

A wealth of theoretical and empirical literature lends support to this notion. The New Economic Geography School (NEG) emphasises the benefits of agglomeration and growing cities for economic growth in particular at low levels of economic development (Fujita & Thisse, 2003; Henderson, 2003; P. Martin & Ottaviano, 2001). Similarly, urban economists stress the static and dynamic productivity gains from increased city size (Duranton, 2015; Duranton & Puga, 2004; Rosenthal & Strange, 2004).

However, some caveats in the existing literature raise questions about the universal applicability of the benefits of increases in average city size. First, most empirical research focuses, with few exceptions, on developed countries, and on the US and the UK in particular. Recent literature (Gollin, Jedwab & Vollrath, 2016; Jedwab & Vollrath, 2015) suggests, however, that the drivers of urbanization differ significantly between developed and developing countries. Such differences in turn may impact on the expected productivity gains from increases in average city size. Furthermore, the sheer size of many cities in developing countries, differences in the industrial structure, lower institutional capacity, and limited infrastructure can reduce the benefits developing countries can extract from rapid increases in average city size. More empirical evidence from developing countries is therefore sorely needed.

Second, the analytical focus of the empirical literature leaves room for interpretation with regards to the impact of recent urbanization trends on aggregate economic growth. Studies in urban economics have been mainly concerned with comparisons of productivity levels and changes involving cities of different sizes (Duranton & Puga, 2004; Rosenthal & Strange, 2004), but do not address the question at an aggregate level. Some NEG research delves into the link between levels of urban concentration and aggregate growth (Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014; Henderson, 2003). However, while frequently cited in the literature as driving growth in countries with large cities, the indicators used are only indirectly linked with city size and do not reflect the size aspect of current urbanization trends. It is therefore pertinent to specifically look at the relationship between an aggregate measure of city size, i.e. average city size, and national economic growth.

In this paper, we address these two gaps by empirically examining the question of whether the average size of a country's cities affects economic growth at the national level. The analysis covers average city size in 114 countries for the period between 1960 and 2010, specifically distinguishing between industrialized and developing countries. Different estimation techniques – including system generalized method-of-moments (GMM) estimator and an instrumental variable design (IV) – are used in order to address potential endogeneity concerns arising from the intimate relationship between urbanization and economic development. Our results suggest that there is indeed a positive effect of average city size on economic growth, however only for developed countries. We do not find evidence that the presence of large cities is growth inducing in developing countries – to the contrary, the IV results suggest a negative impact of city size on growth at the national level.

The remainder of the paper is structured as follows. The following section describes the theoretical and empirical literature that explores the link between city size and economic growth. The third section introduces the methodology, indicators used, and presents the dataset. The results and a test of their robustness by means of an IV approach are included in the fourth section, and section five discusses them in the context of the existing literature. The final part concludes and proposes areas for further research.

2.2 Average city size and economic growth: from theory to empirical evidence

The question of whether a country's cities influence economic growth has been addressed by several streams of literature, most prominently in recent years by the New Economic Geography School (NEG) and urban economics. The underlying assumption is that cities, in particular larger ones, create agglomeration economies and thereby make people more productive. This in turn increases the level of economic development at any given level of inputs (Duranton, 2008).

The NEG School emphasizes the economic efficiency-related benefits of agglomeration. In the traditional NEG framework, centripetal forces, such as localized knowledge spill-overs, pooled labour markets and forward and backward linkages, make companies and people more productive if they concentrate in one area. Centrifugal forces, such as immobile factors, increasing rents and congestion in the prime area, however, incentivize people and firms to locate elsewhere. The relative strength of these two forces shapes the economy's spatial structure and hence if people concentrate in one large city or, by contrast, spread out to smaller ones (Fujita, Krugman, & Venables, 1999). Several authors combine this basic framework with an endogenous growth model to analyse the effect of agglomeration on economic growth and vice versa. While the approaches vary in the specific channels used as agglomeration and dispersion forces, they generally conclude that more agglomeration and thus larger cities are beneficial for economic growth (Fujita & Thisse, 2003; P. Martin & Ottaviano, 2001).

A number of empirical studies confirm this relationship. Henderson (2003), Bertinelli and Strobl (2007), Brülhart and Sbergami (2009), and Castells-Quintana and Royuela (2014) use urban primacy, the percentage of the urban population which lives in the largest city, as a measure of agglomeration and test its influence on national economic growth. They all find that primacy has a positive effect on economic growth, but that the positive effect decreases as the level of economic development rises. Brülhart and Sbergami (2009) also use the percentage of the urban population living in cities above 750,000 inhabitants as an alternative measure of urban concentration and come to the same conclusions. Evidence emerging from the theoretical and empirical NEG literature thus suggests that a more concentrated urban structure with larger cities spurs economic growth, in particular at low levels of economic development. Consequently, recent urbanization trends in the developing world are considered beneficial for economic development in these countries.

The urban economics literature also emphasizes productivity gains stemming from increases in city size but gives no unambiguous answer regarding the effect of average city size on national economic performance. Duranton and Puga (2004) describe a number of channels – similar to the NEG drivers of growth – which make people in cities more productive: the sharing and the matching of infrastructure, inputs, suppliers and labour as well as learning through the generation, diffusion, and accumulation of knowledge. Urban economics also underlines the importance of a dynamic effect of cities on worker's productivity through learning, which increases over time (Duranton, 2008). The dominant view is that agglomeration economies increase with city size.

This sort of productivity gains have been often documented at the city level. Rosenthal and Strange (2004), for example, indicate that a doubling of city size leads to a productivity increase of 3 – 8%. Melo, Graham, and Noland (2009) confirm this positive relationship in their meta-analysis of 34 studies, despite uncovering important regional differences. In the same vein, Duranton (2015) reviews the studies examining developing countries and concludes that productivity increases are even higher in developing countries than in the industrialized world. Differences in terms of productivity gains for different sectors are also in evidence in this type of research, with high-tech sectors and service industries exhibiting the strongest agglomeration economies (Graham, 2009; Henderson, 2010).

A rise in city size also leads, however, to negative externalities such as congestion, higher rents and commuting time which undermine the benefits of agglomeration (Duranton & Puga, 2004). People's productivity within a city, therefore, does not rise *ad infinitum* with increases in city size. It follows an inverted U-shape function: productivity increases up to a certain threshold of city population, after which congestion costs outweigh the benefits from agglomeration and productivity starts to decrease. Beyond the said threshold, workers and firms would be better off relocating to a different city. Through this process, a system of cities arises within a country which – if adequately functioning – efficiently allocates people to cities and maximises the productivity of a country's urban population, as well as its national growth (Duranton & Puga, 2004; Henderson, 1974).

In practice, however, this process can be obstructed by a number of factors. Coordination failure may prevent migration from the prime to secondary cities, as a single actor – i.e. an individual employee or company – cannot internalize the external benefits it creates for others by moving (Duranton, 2008; Venables, 2005). Furthermore, the political sway of the primary city, openness to trade, and a lack of an adequate intercity transport infrastructure network can all hamper the emergence of secondary cities (Ades & Glaeser, 1995; Duranton, 2008; Puga, 1998). A prime city whose size is beyond optimal can therefore emerge, perhaps in combination with small or virtually non-existing secondary cities. Venables (2005) argues that this may result in a low economic development trap. The larger-than-optimum city reduces the productivity of workers and firms which, in turn, curtails economic growth (Venables, 2005). Low growth makes the possibility of starting a new city more difficult, often leading to an ever-growing expansion of the prime city, even after it has exceeded the optimal size threshold.

Hence, from a theoretical urban economics perspective whether recent dramatic increases in city size in the developing world are growth-inducing or detrimental to economic growth depends crucially on where on the productivity curve a country's cities are and on whether a functioning system of cities has emerged. Whether this is the case needs to be determined empirically.

As described above, the existing empirical evidence from both NEG and urban economics is large and points relatively uniformly to a growth promoting effect of increasing city size. However despite the wealth of studies, some important limitations remain. First, most research addressing these issues, in particular from the empirical urban economics literature, focuses on the developed world. Only a handful of developing countries are covered. This is problematic for a number of reasons related to the underlying urbanization process, likely differences in the balance between agglomeration economies and diseconomies, as well as the absolute size of the cities in the developing world. With regards to the urbanization process, it is frequently assumed that developing countries follow the same path industrialized countries did in the past (as described in the World Development Report 2009). There is, however, increasing evidence that questions this assumption. Economists traditionally explain the urbanization process with a structural shift from an agricultural to a manufacturing-based economy in which higher (expected) wages in the urban industrial sectors stimulate migration from rural areas to cities (Harris & Todaro, 1970; Lewis, 1954). Urbanization in today's developed countries is considered to have been strongly tied to industrialization and economic growth. In contrast, many developing countries are urbanizing in the presence of low growth and without a strong, accompanying industrialization process. Fay and Opal (2000) described this situation of "urbanization without growth" and pointed to a wide array of factors driving urbanization in developing countries. These included push-factors in the rural countryside such as conflict, negative agricultural shocks, and rural poverty as well as pull-factors from cities, such as better urban living conditions through improved access to public services and other urban amenities. Glaeser (2014) also suggested that cities have grown larger in the developing world due to their ability to import food. In developed countries in contrast, cities are historically deemed to have increased in population as the agricultural hinterland became more productive and could feed a larger non-food producing urban population. Furthermore, Jedwab and co-authors (Gollin et al., 2016; Jedwab & Vollrath, 2015) recently reported that urbanization in developing countries is strongly linked to natural resource exports as opposed to industrialization. As a consequence, the share of urban dwellers working in the non-tradable service sector is much higher than in developed countries, where urban population grew as employment in manufacturing and in the tradable service sectors. This has important implications for the magnitude of agglomeration economies that a city generates. As Gollin et al. (2016) have stressed, cities growing as a consequence of the expansion of resource exports do not create the same push in productivity as industrial cities.

Related to the previous point are differences in urban infrastructure endowments and in the industry composition of the cities in developing countries. As pointed out above, the balance between agglomeration economies and diseconomies determines the benefits from increasing city size. The fast growth of cities in the developing world, together with a lack of public resources for infrastructure investments, and economies based on low technology sectors imply that cities in the developing world face decreasing productivity levels at a lower city size than cities in the developed world. Limited institutional capacity to deal with the challenges aggravates this situation (Glaeser, 2014).

Finally, cities in the developing world are frequently much larger than their counterparts in industrialized countries. Empirical evidence stemming from what are smaller cities in the developed world is unlikely to be a good *explanans* for the situation in developing countries, in particular bearing in mind the U-shaped productivity function of cities. All these aspects − urbanization path, the balance between agglomeration economies and diseconomies as well as overall city size − strongly call for caution when applying evidence generated on the basis of analysis conducted in developed countries to the developing world. Empirical evidence, which focuses on differences between developed and developing countries, is, in this respect, still much needed.

The second limitation relates to the analytical focus of the existing empirical literature. On the one hand, empirical studies with an urban economics lens take a city perspective, i.e. they provide an answer to the question if people in large cities are more productive than in small cities. They do, however, not tackle the question at the national level. Following the system of cities approach it is not necessarily evident that because one city is more productive than another implies an overall productivity maximising effect at the country level if all resources are pooled in the larger city. For example, increasing the population in 'city one' through migration from 'city two' may make 'city one' more productive, but the reduction in size of 'city two' may result in a larger decrease in productivity there. The result will then be suboptimal at the national level. As most countries include more than one city, looking purely at this question from a city-level perspective may not be very revealing.

On the other hand, a relatively large number of NEG studies address the issue at the national level. However, the focus of this literature is not the actual size of a country's cities, but how concentrated the urban structure is. While these studies are frequently cited as evidence to support the notion of fostering agglomeration within developing countries, they do not say much about the size-related effect of cities. Most studies use urban primacy, which − if at all − is negatively correlated with the size of a country's cities (e.g. there is a very low primacy in India despite its cities being large, but a very high primacy in most island states which have very small cities). The results therefore are more likely to reflect the benefits of being able to focus public spending in one place as a result of a concentrated urban population as opposed to agglomeration economies arising through city size.

We aim to address these two limitations by studying the effect of increased average city size on economic growth at an aggregate level. In other words we ask the question of whether countries grow faster if the urban population lives on average in larger cities? We furthermore specifically focus on possible differences in this respect between developed and developing countries.

2.3 Model and data

In order to test these two issues, we follow the dominant approach of the empirical literature on urban concentration and national economic growth (Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014; Henderson, 2003). We build a simple GDP per capita growth equation based on the extended Solow growth model (Durlauf, Johnson, & Temple, 2005). A country's growth rate in 5 year-periods is estimated as a function of GDP per capita at the beginning of the period and a set control variables which reflect both variables related to the accumulation of factors, as well as a set of other characteristics influencing national growth and the size of a country's cities. Rather than a measure for the level of urban concentration, as has frequently been the case in the above-mentioned literature, we include an indicator depicting the average size of a country's cities. If living in larger cities boosts people's productivity, economic growth should be higher while holding the other inputs constant. The model takes the following form:

$$\Delta GDPpc_{ip} = \alpha + \beta \; citysize_{ip} + \gamma \; GDPpc_{ip} + \delta \; \boldsymbol{X}_{ip} + \mu_p + \varepsilon_{ip}$$

Where

p denotes five-year intervals;

$\Delta GDPpc_{ip}$ is the GDP per capita growth rate of country *i* in period *p*;

$\mathrm{citysize_{ip}}$ is an aggregate indicator for the average size of cities in country i;

$\mathrm{GDPpc_{ip}}$ represents the natural logarithm of GDP per capita of country *i* at the beginning of period *p*;

$\mathrm{X_{ip}}$ depicts a set of control variables for country i, measured either at the beginning or as an average of period *p*;

μ_p represents time fixed effects; and

ε_{ip} is the error term.

The variable of interest is $citysize_{ip}$. Testing if the size of a country's cities has an impact on national economic performance requires an aggregate measure of city size at the country level. For this purpose, we calculate the population-weighted average city size as the sum of the absolute number of each city's population multiplied by its share of the urban population. This indicator reflects the average agglomeration (dis-)economies which a typical urban dweller experiences.

The source of data is the 2014 revision of the World Urbanization Prospects [WUPS] (United Nations, 2014). WUPS 2014 includes, among other data, information about (i) the population of every city above 300,000 inhabitants, (ii) the share of the urban population living in cities below 300,000 and (iii) the total urban population. It covers a total of 199 countries from 1960 to 2010.[12]

We multiply the exact population size of each city above 300K with its share of the urban population. As there is no information available about the exact size of cities below 300K, we use a proxy for the average city size of cities below 300K and multiply it by the share of the urban population living in cities below 300K.[13] These fractions are subsequently added in order to obtain the population weighted average city size of a country.

How the population-weighted average city size indicator works is reflected in the following examples. Let us take country A which has two cities, each with 1M inhabitants. In this case, the population weighted average city size is: 1M*50% + 1M*50% = 1M (as would be the simple average). Country B, by contrast, has one city of 1.9M inhabitants and a second city of 0.1M inhabitants. The resulting population weighted average city size is: 1.9M*95% + 0.1M*5% = 1.81. In contrast to the simple average which as in country A would be 1M, this number reflects the fact that the vast majority of people live in the larger city of 1.9M inhabitants and therefore experience the agglomeration economies and diseconomies of a city of such size. Finally, country C has one city of 1.15M inhabitants and a second city of 0.2M inhabitants. The population weighted average city size is with roughly 1M similar to that of

[12] One challenge with data on cities and urbanization is that the definition of what constitutes a city varies across countries. While some countries count villages over 1,000 inhabitants as cities, others only include cities starting from 10,000 inhabitants. Similarly, some consider the administrative city, while others have measures that are closer to the agglomeration. This makes comparisons between countries challenging. While WUPS 2014 relies on the local definitions, it aims to smooth out these issues as far as possible by correcting for agglomeration size and standardizing urban definitions. This system is not without problems, but despite these caveats, WUPS remains the best available dataset for the purpose of our analysis.

[13] In order to obtain a reasonable proxy for the cities below 300K inhabitants, we use a complementary dataset based on census data sourced from citypopulation.de, which includes population numbers for each city (including those below the 300K threshold) in a country during the period between 1985 and 2010. The correlation between the proxy based on the WUPS data and the more finely-tuned average from the second dataset is 0.94. This is a clear indication that the proxy works well. We resort to the WUPS dataset rather than the citypopulation.de dataset because of its longer time horizon and larger country coverage.

country A. This takes into account that, in spite of the differences between the two countries, the majority of the urban population lives in a city of a similar size.

The population-weighted average city size indicator also differs considerably from primacy, the traditional measure used by the literature concerned with urban concentration. A country in which city 1 has a population of 0.3M and a city 2 has a population of 0.5M has the same primacy level (62.5%) as a country with one city of 3M and another city of 5M inhabitants. This difference in absolute size is, however, reflected in our average city size measure.

Figure 2-1: Evolution of the median of average city size 1960 – 2010 (in 1000s)

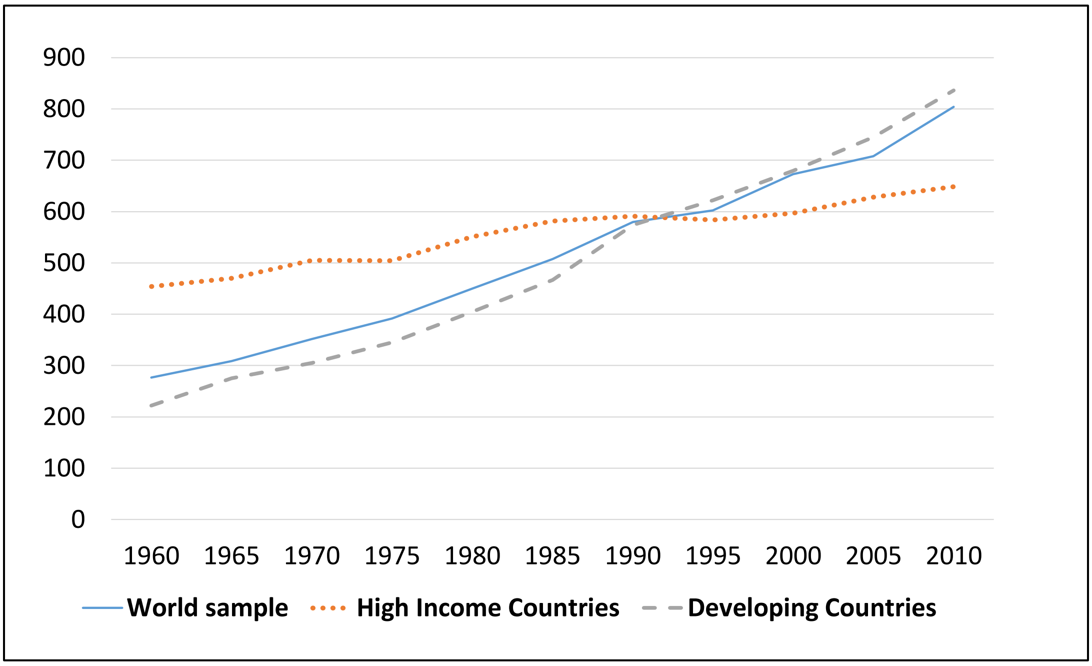


Figure 2-1 shows the evolution of the average city size as the median for the countries included in our regressions. The numbers are consistent with the urbanization trends addressed in the introduction of this paper. Overall city size increased from a median of approximately 280,000 in 1960 to 800,000 in 2010, signalling almost a tripling of average city size. Figure 2-1 also illustrates the diverging trends between the industrialized and the developing world. Average city size in developing countries increased from 220,000 to 845,000 during the 50 year period considered, surpassing average city size in high-income countries around the year 1990. In the same period average city size in developed countries only increased from 500,000 to 650,000.

The control variables are those typically employed in cross-country growth regressions: initial GDP per capita (GDPpc) to control initial levels of wealth and for conditional convergence, i.e. countries are expected to grow faster if starting from a lower level, holding other factors constant (Durlauf et al., 2005); years of schooling as a proxy for human capital, since a more educated work force is assumed to be more productive and drive up growth (Lucas, 1988; Romer, 1986); and private investment and government consumption as a percentage of GDP, due to their contribution to a country's capital stock – the crucial growth driver in the basic neoclassical Solow model.

Additionally, the model includes a number of controls which may be directly linked to the effect of city size on growth. The first of these is openness, which is measured as the sum of national exports and imports as a percentage of GDP. A country's openness is believed to influence city size via its effect on the balance between centripetal and centrifugal forces (Krugman & Elizondo, 1996). Openness is also understood to be directly related to growth (Sachs, Warner, Åslund, & Fischer, 1995). Failure to control for openness may therefore result in an omitted variable bias. The second is national population. A country's population is expected to be closely linked to its average city size (i.e. India has bigger cities than, say, Switzerland). Furthermore, a country's size and, by extension its market potential, can be also envisaged to affect national economic performance (Alesina, Spolaore, & Wacziarg, 2005). Controlling for population size also ensures that the results are not driven by a handful of specific cases, such as China or India, both of whom have experienced extremely strong growth in past decades. The final control relates to a country's political system. Certain political systems such as dictatorships have been shown to increase city size (Ades & Glaeser, 1995) and may at the same time impact on a country's growth performance. We, therefore, include five-year averages of the widely used 'polity indicator' which rates countries on a scale from -10 (autocracy) to 10 (consolidated democracy).

The data for the controls is sourced from the eighth edition of the Penn World Tables with the exception of the years of schooling indicator and the measure for the political systems. These come from the Barro and Lee database and the Polity IV project dataset respectively. A more detailed description of the indicators and sources as well as descriptive statistics are included in Appendix 1 and Appendix 2. Depending on the specifications, our analysis covers a maximum of 114 countries.

We estimate the model using pooled 2SLS as a baseline as well as country fixed effects and system GMM to take advantage of the panel structure of the dataset. All regressions include time fixed effects and robust standard errors which are clustered by country in the pooled 2SLS regression. Average city size is instrumented with its second lag in the pooled 2SLS. Furthermore, we employ system GMM to address the issue of reverse causality that inevitably arises when studying the relationship between city size, urbanization, and economic development, to test if city size is in fact a result of economic growth as opposed to being a driver of it. Furthermore, system GMM is appropriate for dynamic panels in which the dependent variable is influenced by its lagged values. This is the case for our estimates because of the inclusion of the lagged value of GDPpc as an explanatory variable and the assumption that growth rates are influenced by GDPpc. System GMM addresses both points by creating a system of equations in which the levels of endogenous variables are instrumented with lagged differences and the first differences are instrumented with past levels (Roodman, 2009). We present these results alongside the 2SLS and fixed effects estimates for all our regressions. As an additional robustness test to address the issue of endogeneity, we conduct an instrumental variable approach in a separate section, which resorts to historical data for the construction of the instrument.

2.4 Results

2.4.1 Main results

Table 2-1 displays the results for the 114 countries making the world sample – columns 1, 3 and 5 for the regressions with the simple term only; columns 2, 4 and 6 contain the squared term of average city size in order to account for possible non-linearity.

Table 2-1: World sample – Dependent variable: GDP per capita growth, 1960 – 2010

VARIABLES	(1) Pooled 2SLS	(2) Pooled 2SLS	(3) Fixed Effects	(4) Fixed Effects	(5) System GMM	(6) System GMM
Average city size	6.15e-05	0.0237**	0.0122	0.0461**	-0.0321*	-0.0312
	(0.00388)	(0.0104)	(0.0126)	(0.0215)	(0.0167)	(0.0348)
Average city size squared		-0.00171***		-0.00221**		0.001000
		(0.000584)		(0.000946)		(0.00185)
Initial GDPpc	-0.0360***	-0.0391***	-0.219***	-0.222***	-0.014***	-0.018***
	(0.00979)	(0.00986)	(0.0354)	(0.0352)	(0.0509)	(0.0404)
Years of schooling	0.0166***	0.0172***	0.00667	0.00568	0.00608	0.00906
	(0.00365)	(0.00373)	(0.0112)	(0.0110)	(0.0175)	(0.0150)
Private investment	0.511***	0.512***	0.797***	0.786***	0.784**	0.806***
	(0.117)	(0.114)	(0.141)	(0.142)	(0.357)	(0.311)
Government consumption	0.122	0.144	0.101	0.0981	0.472	0.473
	(0.111)	(0.111)	(0.110)	(0.111)	(0.296)	(0.338)
Openness	-0.00539	-0.00796	-0.0126	-0.0121	-0.0545*	-0.0548**
	(0.0154)	(0.0140)	(0.0160)	(0.0154)	(0.0284)	(0.0268)
Population	0.0162***	0.00926	-0.0945**	-0.122***	0.102*	0.0753*
	(0.00508)	(0.00597)	(0.0450)	(0.0449)	(0.0544)	(0.0401)
Political system	-0.000979	-0.000850	-0.00160	-0.00182	-0.00254	-0.00340
	(0.000937)	(0.000948)	(0.00117)	(0.00116)	(0.00521)	(0.00458)
Constant	0.240***	0.260***	1.759***	1.828***	-0.239	-0.158
	(0.0718)	(0.0724)	(0.338)	(0.334)	(0.329)	(0.273)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	-	-	Yes	Yes	Yes	Yes
Observations	971	971	1,058	1,058	1,058	1,058
R-squared	0.202	0.207	0.270	0.275		
Number of countries			114	114	114	114
Sargans	-	-	-	-	25.44	32.66
					(0.062)	(0.018)
AR1	-	-	-	-	2.27	-2.69
					(0.007)	(0.007)
AR2	-	-	-	-	1.06	1.07
					(0.290)	(0.285)

Notes: Robust standard error clustered by country in parentheses: *** p<0.01, ** p<0.05, * p<0.1
System GMM uses second and third lag 2-step estimator

The model works well and the controls show the expected signs (Table 2-1). The coefficient of GDP per capita at the beginning of the period is negative and strongly significant in all estimations, indicating conditional convergence. Private investment also displays, as expected, positive and highly significant coefficients across all estimates. Years of schooling is significant in the 2SLS model only. This is in line with previous literature: Henderson (2003), for instance, points out that education is rarely robust in these types of regression as changes in years of schooling are more likely to have an effect over a longer time horizon than in the short-term (Durlauf et al., 2005). Population size is significant, but with different signs depending on the estimation technique. Again this is not surprising as in a cross-section comparison large countries such as China and India have grown faster than smaller ones. If we look at the within changes, however, countries which experienced a strong population growth tend to be countries with low-income, in particular in Africa. Many of those countries have also experienced low growth in the past decades.

The results for our variables of interest in the global sample provide mixed evidence. The coefficients for average city size – in those regressions including only average city size and not its squared term – are insignificant in the pooled 2SLS and fixed effects analyses, but negative and significant at the 10% level in the system GMM estimate (Table 2-1, Regressions 1, 3, and 5). Once we consider the squared term in the regression analysis the picture changes. The pooled 2SLS and fixed effects coefficients for average city size are positive and significant, while the system GMM results insignificant. The squared terms coefficients are negative and significant at the 1% and 5% level in the 2SLS and fixed effects estimates respectively (Table 2-1, Regressions 2, 4, and 6). This would indicate a positive connection between city size and economic growth, which diminishes as the average size of cities increases. Overall, these results send contrasting messages about the link between average city size and national economic growth. If anything they may signal, as hinted in the theoretical section, widely differing realities about the association between city size and economic growth in developed and developing countries.

Table 2-2: By income group – Dependent variable: GDP per capita growth, 1960 – 2010

	High-income Countries						Developing Countries					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	Pooled 2SLS	Pooled 2SLS	Fixed Effects	Fixed Effects	System GMM	System GMM	Pooled 2SLS	Pooled 2SLS	Fixed Effects	Fixed Effects	System GMM	System GMM
Average city size	-0.000278	0.0172*	0.0119	0.120***	0.00785	0.0945***	-0.000215	0.0248	0.00956	0.01000	-0.0355	-0.108
	(0.00241)	(0.00939)	(0.0125)	(0.0434)	(0.0158)	(0.0281)	(0.00945)	(0.0244)	(0.0133)	(0.0437)	(0.0243)	(0.0698)
Average city size		-0.00112**		-0.00485**		-0.00494***		-0.00440		-6.30e-05		0.00910
squared		(0.000523)		(0.00180)		(0.00141)		(0.00348)		(0.00502)		(0.00967)
Initial GDPpc	-0.115***	-0.109***	-0.267***	-0.299***	-0.038***	-0.050***	-0.0238*	-0.0268*	-0.224***	-0.224***	0.027***	0.016***
	(0.0144)	(0.0136)	(0.0543)	(0.0521)	(0.0885)	(0.0702)	(0.0138)	(0.0147)	(0.0406)	(0.0404)	(0.0506)	(0.0490)
Years of schooling	0.00657*	0.00628*	0.0118	0.00903	-0.0201	-0.0147	0.0172***	0.0177***	0.00296	0.00297	0.00686	0.00829
	(0.00348)	(0.00336)	(0.0138)	(0.0117)	(0.0356)	(0.0305)	(0.00437)	(0.00455)	(0.0146)	(0.0145)	(0.0105)	(0.0108)
Private investment	0.574***	0.556**	0.427*	0.383*	0.225	-0.105	0.340***	0.344***	0.831***	0.831***	-0.143	0.0447
	(0.221)	(0.219)	(0.225)	(0.206)	(0.779)	(0.572)	(0.123)	(0.123)	(0.170)	(0.168)	(0.468)	(0.406)
Government	0.238	0.272	-0.0756	-0.0352	0.722	0.749	0.0980	0.106	0.0631	0.0629	0.0996	0.227
consumption	(0.171)	(0.173)	(0.355)	(0.369)	(0.805)	(0.814)	(0.115)	(0.115)	(0.114)	(0.119)	(0.281)	(0.316)
Openness	0.0309*	0.0270	0.0871*	0.0736	0.107	0.0942	-0.0155*	-0.0170*	-0.0208	-0.0208	-0.00803	-0.0291
	(0.0185)	(0.0192)	(0.0477)	(0.0469)	(0.105)	(0.0773)	(0.00916)	(0.00923)	(0.0132)	(0.0132)	(0.0323)	(0.0344)
Population	0.0134*	0.00813	0.0732	-0.0506	0.0313	-0.0423	0.0189***	0.0145*	-0.240***	-0.240***	0.0610**	0.106***
	(0.00696)	(0.00765)	(0.0873)	(0.118)	(0.0669)	(0.0583)	(0.00647)	(0.00808)	(0.0823)	(0.0826)	(0.0279)	(0.0374)
Political system	0.000788	0.00114	0.00193	0.000729	-0.00345	0.00455	-0.00184	-0.00194*	-0.00267*	-0.00267*	-0.00407	-0.00427
	(0.00177)	(0.00170)	(0.00209)	(0.00210)	(0.0109)	(0.00579)	(0.00115)	(0.00117)	(0.00149)	(0.00150)	(0.00472)	(0.00486)
Constant	0.966***	0.914***	2.158***	2.645***	0.392	0.581	0.172*	0.190*	1.856***	1.857***	-0.185	-0.217
	(0.172)	(0.171)	(0.645)	(0.611)	(0.817)	(0.724)	(0.0985)	(0.105)	(0.351)	(0.344)	(0.279)	(0.275)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	-	-	Yes	Yes	Yes	Yes	-	-	Yes	Yes	Yes	Yes
Observations	317	317	361	361	361	361	654	654	697	697	697	697
R-squared	0.494	0.499	0.468	0.497			0.180	0.178	0.261	0.261		
Number of countries			38	38	38	38			76	76	76	76
AR1	-	-	-	-	-2.47	-2.79	-	-	-	-	-2.32	-2.36
					(0.017)	(0.005)					(0.020)	(0.017)
AR2	-	-	-	-	-0.64	-0.14	-	-	-	-	0.92	1.00
					(0.524)	(0.888)					(0.356)	(0.317)
Sargans	-	-	-	-	30.18	28.05	-	-	-	-	19.32	20.48
					(0.017)	(0.061)					(0.252)	(0.306)

Notes: Robust standard error clustered by country in parentheses: *** p<0.01, ** p<0.05, * p<0.1
System GMM uses second and third lag 2-step estimator

In order to assess whether this is the case, we divide the sample into high-income and low- and middle-income countries, using the World Bank's classification. Columns 1-6 of Table 2-2 show the results for 38 high-income and columns 7-12 for 76 low- and middle-income (developing) countries. The results differ sharply between the two samples, but provide consistent results within the two groups. For high-income countries, there is now strong evidence that average city size does indeed drive national economic growth. In all regressions where the squared term for average city size is included (Table 2-2, Regressions 2, 4, and 6), the coefficient for the main term is positive and significant in all estimates, including the system GMM results. The negative and significant coefficient of the squared term in all estimates indicates that the positive effect of city size on national growth diminishes as the average size of a city grows. The coefficients are similar between the fixed effects and system GMM, but lower in the 2SLS estimate (Table 2-2, Regressions 2, 4, and 6).

In order to understand the magnitude of the effect, we calculate the fitted values of the partial association between average city size and economic growth based on the system GMM coefficients. Raising average city size from 1M to 1.1M increases the five-year growth rate by 0.84%. When average city size increases from 5 to 5.1M, growth rises by 0.45%. Values become close to 0 around 7.5M and turn negative at approximately 9.7M. This last value should be interpreted with caution, as there are very few observations in our sample with such high average city sizes. The evidence, thus, strongly suggests that high-income countries indeed grow faster when the population on average lives in larger cities, although the impact is relatively modest.

The contrary is true for developing countries, as shown in columns 7 to 12 of Table 2-2. No evidence at all can be found of an impact of average city size on economic growth. Neither the main term of average city size nor the squared term is significant in any of the estimations. We have also experimented by further dividing the sample into a low- and a middle-income group. The results remain insignificant for both groups. It is worth noting that the R-squared drops somewhat in the regressions for developing countries compared to the global sample, while it significantly increases for the high-income group. This is reflective of the fact that in general developing countries' growth does not seem to follow the same patterns and rules as in developed countries.

2.4.2 Instrumental variable approach

A frequent concern in regressions addressing the relationship between city size and economic growth is reverse causality. Is city size in fact a mere result of economic growth as opposed to a driver of it? In the previous section we partially addressed this issue by including system GMM estimations alongside the other estimation techniques. To further test the robustness of the results, we employ an additional instrumental variable design.

The challenge is to find a suitable instrument for city size, i.e. a factor that is correlated to the size of a country's cities but not to national economic growth, in order to isolate the exogenous component of city size. We resort to historic data for this purpose and employ a measure of the number of years that the current capital city of a country has been the capital since the 1 A.D.[14] Countries where the current capital has been an important place of political (and economic) power for a long time are more likely to have developed a more concentrated urban structure with a more dominant and larger prime city. Our population weighted average city size indicator should thus also be larger in a country where the current capital has been the main political centre for longer time periods. This hypothesis is confirmed by a brief examination of our dataset. Countries with large average city sizes, such as the United Kingdom, France, Peru, Mexico, Russia, or China also have longstanding capital cities. The first-stage F-stat supports the relevance of our instrument.[15]

The question that remains concerns the exogeneity of the instrument. It could be argued that the 'years as capital' has an impact on the level of economic development and this, in turn, affects growth performance. Countries with long-established capitals may be states where

[14] We rely on Pierskalla, Schultz and Wibbels (2014) for this purpose.
[15] The first stage results are included in Appendix 4.

sound political institutions – which are widely regarded as growth enhancing – have had more time to develop. Once again, a brief examination of the dataset suggests that no such correlation exists. There are countries with relatively young capitals, such as Germany and Australia, which are highly developed. At the same time there are countries where the capital was established a long time ago, such as Peru, Nepal, or Mozambique, with much lower levels of economic development. The two-way scatterplot of GDP per capita in 1960 and the instrument confirms this assertion (Appendix 3). No obvious correlation can be found between both factors. 'Years as a capital' has also an insignificant coefficient when regressed on GDP per capita in 1960. Finally, it is highly unlikely that the existence of a longstanding capital has a direct influence on the growth performance of a country over a specific short period of time.

Table 2-3: IV estimates – Dependent variable: GDP per capita growth, 1960 – 2010

VARIABLES	(1) High-income countries	(2) Developing countries
Average city size	0.0729**	-0.232**
	(0.0359)	(0.102)
Initial GDPpc	-0.075***	0.047***
	(0.0242)	(0.0318)
Years of schooling	0.00685	0.0121**
	(0.00418)	(0.00585)
Private investment	0.298**	0.306**
	(0.151)	(0.143)
Government consumption	0.287*	-0.151
	(0.170)	(0.175)
Openness	0.0190	0.00734
	(0.0171)	(0.0169)
Population	-0.0113	0.111***
	(0.0134)	(0.0431)
Political system	0.00272	-0.000589
	(0.00260)	(0.00140)
Constant	0.658***	-0.353
	(0.223)	(0.234)
Time fixed effects	Yes	Yes
Observations	258	579
R-squared	0.973	0.955
First stage F-stat	10.07	9.62

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

The model is estimated using pooled 2SLS. For the sake of simplicity, we only include the main term in this estimate (Table 2-3). The results for high-income countries are confirmed.

A country's average city size is positively and significantly associated with economic growth (Table 2-3, Regression 1). The coefficient is similar to the system GMM estimate reported in Table 2-2, roughly confirming its accuracy. Every 100,000 population increase in average city size raises the five-year growth rate by approximately 0.7%.

For developing countries, the coefficient of average city size is negative and significant at the 5% (Table 2-3, Regression 2). This contradicts the results in the previous section where no impact of average city size on economic growth for developing countries was reported. This estimate provides evidence of a detrimental impact of large cities on economic growth for developing countries, suggesting a 2.3% decrease in five-year growth rates for a 100,000 inhabitant increase in average city size.

2.4.3 Discussion

Our results provide novel insights to the debate about the economic impact of mega cities and overall increases in city size around the world. First, they lend support to the hypothesis by Gollin et al. (2016) that many cities in developing countries do not generate the same productivity gains as cities in developed countries. This is evidenced by both the positive results for high-income countries, which suggest the presence of productivity gains associated with larger cities, as well as the insignificant (and negative) IV-results for developing countries. These results are explained by the underlying mechanisms driving urbanization which are fundamentally different for developed and developing countries. Hence, developing country cities have a high share of workers in sectors that do not benefit from agglomeration economies. Simultaneously, developing countries' cities are now much larger than their counterparts were in the developed world when they had reached the same level of economic development. The greater size of developing world cities has been facilitated, among other factors, by the possibility of importing more food, aid inflows, and improved public service provision (Fay & Opal, 2000; Glaeser, 2014). This in turn has contributed to intensify the urban diseconomies present in developing countries' cities. Instead of being the place for industrialization and productivity growth as in the developed world, they are increasingly becoming loci for the concentration of the poor and those at risk of severe poverty in the emerging world.

Second, the evidence presented in this paper complements past empirical research on urban concentration and economic growth (Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014; Henderson, 2003). At first sight, our insignificant results for developing countries may seem at odds with the findings of this literature, which suggest a particularly important role for agglomeration at early stages of development. However, they may simply reflect different coverage and a focus on different aspects. Urban concentration, even when the actual urban population is small, provides benefits for the provision of public infrastructure, in particular in the face of limited resources. As countries grow richer and have more resources, this advantage loses some of its relevance. The urban concentration literature may, to a considerable degree, also capture this aspect. Our average city size measure in turn reflects the agglomeration (dis-) economies which arise through the actual size of the cities. Industrialized countries with a strong tradable service sector and high-tech manufacturing benefit from larger cities, while in developing countries diseconomies of scale and negative externalities may prevail.

Third, from a system of cities perspective, our results are also telling. The positive results for high-income countries suggest that people and firms are more prone to relocate in high-income countries, once the primary city reaches the tipping point on the productivity curve. Furthermore, cities in developed countries are able to overcome some of the diseconomies by innovating their function within the system of cities (Camagni, Capello, & Caragliu, 2015). In contrast, our results indicate that this is not the case for developing countries and that many of the cities may even be already in a low economic development trap, as suggested by Venables (2005). This finding is consistent with observations by other authors that show that mature manufacturing sectors are still located within the prime cities in developing countries as opposed to secondary cities, as should be expected following the systems of cities approach (Duranton, 2015).

Finally, the results also point to the need to adapt theoretical models and their assumptions, developed based on high-income countries, more strongly to the realities of the developing world. While the analysis for high-income countries is broadly in line with the predictions of the NEG models (Fujita & Thisse, 2003; P. Martin & Ottaviano, 2001), the insignificant coefficients for the developing countries sample are clearly not. One explanation may be the already mentioned balance between urban economies and diseconomies, which is probably different for developing and high-income countries. NEG models may be well calibrated to capture the situation for high-income countries, while they may underestimate urban diseconomies and over-emphasize the benefits of agglomeration in developing countries' context. Some authors (Henderson, 2010; R. Martin, 2008) support this notion and argue that the treatment of urban diseconomies is still limited in the literature.

2.5 Conclusion

In this paper we have analysed whether countries grow faster if their population lives on average in larger cities, as has been frequently implied by recent economic theories and development policy alike. For this purpose, we have used a panel of 114 countries for the period between 1960 and 2010. We have specifically tested for varying effects of average city size on economic growth in high-income and in developing countries, as, until now, most evidence has been based on data from the developed world. A raft of estimation techniques – including pooled 2SLS, fixed effects, system GMM, and a novel instrumental variable – has been used in order to establish the relationship between average city size and economic growth. The system GMM and the IV analyses allow us to address statistically the concern of reverse causality when considering questions of urbanization, city size, and economic growth.

The analysis has revealed that any statement about a uniform relationship between average city size and economic growth does not hold. For the whole sample, the evidence of such a relationship is inconclusive, with results differing depending on the estimation technique. However, there is consistent evidence of a growth promoting effect of average city size in high-income countries throughout all our estimation techniques. The robustness of the effect is confirmed by the IV-results. An increase in the average city size of 100,000 inhabitants boosts a country's five-year growth performance by up to 0.84%. This effect decreases as the average city size increases. Conversely, in developing countries all coefficients of the average city size indicator are insignificant. The IV results even suggest a negative impact of average city size on economic growth. The results for developing countries thus imply that city size is not a driver of economic growth. If it has an impact at all, it is negative.

Whether recent urbanization trends with increasing average city size are a growth driver or not seems highly context dependent and the answer may very well be negative for many developing countries. Heterogeneity in the underlying urbanization patterns between developed and developing countries, a differing balance between agglomeration economies and diseconomies, and structural barriers to the creation of a functioning system of cities in developing countries may result in an urban environment which does not make its inhabitants more productive, as would be expected based on the literature for developed countries.

The analysis opens several avenues for future research. In order to better understand how city size shapes economic growth at an aggregate level, greater protagonism needs to be awarded to understand the underlying urbanization paths and how these ultimately influence the balance between urban economies and diseconomies. In particular, there is a clear need to analyse in greater detail the extent and possibilities of managing urban diseconomies in a developing country context. Another line of analysis should explore the structural barriers to the creation of a functioning system of cities in developing countries.

Finally, our analysis underlines the need to re-evaluate the increasingly widespread policy view that bigger cities spur economic growth. A more nuanced view of how urban policies impinge on overall economic growth, especially in the developing world, is required. It should be based on new empirical analysis, as well as on theoretical approaches that are more attuned to the realities of emerging countries. Otherwise the risk of coming to simplistic and, in some cases, perhaps overly harmful policy recommendations may increase based on the wrong assumption that developing countries simply follow the same urbanisation path which was previously followed by developed countries.

References

Ades, A. F., & Glaeser, E. L. (1995). Trade and Circuses - Explaining Urban Giants. *Quarterly Journal of Economics, 110*(1), 195-227. doi:10.2307/2118515

Alesina, A., Spolaore, E., & Wacziarg, R. (2005). Chapter 23 - Trade, Growth and the Size of Countries. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part B, pp. 1499-1542). Amsterdam: Elsevier.

Bertinelli, L., & Strobl, E. (2007). Urbanisation, urban concentration and economic development. *Urban Studies, 44*(13), 2499-2510. doi:10.1080/00420980701558442

Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*(1), 48-63. doi:10.1016/j.jue.2008.08.003

Camagni, R., Capello, R., & Caragliu, A. (2015). The Rise of Second-Rank Cities: What Role for Agglomeration Economies? *European Planning Studies, 23*(6), 1069-1089. doi:10.1080/09654313.2014.904999

Castells-Quintana, D., & Royuela, V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science, 52*(2), 343-366. doi:10.1007/s00168-014-0589-1

Duranton, G. (2008). Viewpoint: From cities to productivity and growth in developing countries. *Canadian Journal of Economics,, 41*(3), 689-736. doi:10.1111/j.1540-5982.2008.00482.x

Duranton, G. (2015). Growing through Cities in Developing Countries. *World Bank Research Observer, 30*(1), 39-73. doi:10.1093/wbro/lku006

Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063-2117). Burlington: Elsevier.

Durlauf, S. N., Johnson, P. A., & Temple, J. R. W. (2005). Chpater 8 - Growth Econometrics. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part A, pp. 555-677). Amsterdam: Elsevier.

Fay, M., & Opal, C. (2000). Urbanization without Growth: A Not-So-Uncommon Phenomenon. *The World Bank Policy Research Working Paper Series, 2412*. Retrieved from http://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-2412 doi:10.1596/1813-9450-2412

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: cities, regions and international trade*. Cambridge Mass.: MIT Press.

Fujita, M., & Thisse, J. F. (2003). Does geographical agglomeration foster economic growth? And who gains and loses from it? *Japanese Economic Review*, *54*(2), 121-145. doi:10.1111/1468-5876.00250

Glaeser, E. L. (2014). A world of cities: the causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association, 12*(5), 1154-1199. doi:10.1111/jeea.12100

Gollin, D., Jedwab, R., & Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth, 21*(1), 35-70. doi:10.1007/s10887-015-9121-4

Graham, D. J. (2009). Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science*, *88*(1), 63-84. doi:10.1111/j.1435-5957.2008.00166.x

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development - a two-sector analysis. *American Economic Review, 60*(1), 126-142.

Henderson, J. V. (1974). Sizes and types of cities. *American Economic Review, 64*(4), 640-656.

Henderson, J. V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth, 8*(1), 47-71. doi:10.1023/A:1022860800744

Henderson, J. V. (2010). Cities and Development. *Journal of Regional Science, 50*(1), 515-540. doi:10.1111/j.1467-9787.2009.00636.x

Jedwab, R., & Vollrath, D. (2015). Urbanization without growth in historical perspective. *Explorations in Economic History, 58*, 1-21. doi:10.1016/j.eeh.2015.09.002

Krugman, P., & Elizondo, R. L. (1996). Trade policy and the Third World metropolis. *Journal of Development Economics, 49*(1), 137-150. doi:10.1016/0304-3878(95)00055-0

Lewis, A. W. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School, 28*(2), 139-191.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3-42. doi:10.1016/0304-3932(88)90168-7

Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review, 42*(4), 947-968. doi:10.1111/1468-2354.00141

Martin, R. (2008). National growth versus spatial equality? A cautionary note on the new 'trade-off' thinking in regional policy discourse. *Regional Science Policy & Practice, 1*(1), 3-13. doi:10.1111/j.1757-7802.2008.00003.x

McCann, P., & Acs, Z. J. (2011). Globalization: Countries, Cities and Multinationals. *Regional Studies, 45*(1), 17-32. doi:10.1080/00343404.2010.505915

Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics, 39*(3), 332-342. doi:10.1016/j.regsciurbeco.2008.12.002

Puga, D. (1998). Urbanization patterns: European versus less developed countries. *Journal of Regional Science, 38*(2), 231-252. doi:10.1111/1467-9787.00090

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy, 94*(5), 1002-1037. doi:10.1086/261420

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal, 9*(1), 86-136.

Rosenthal, S., & Strange, W. (2004). Chapter 49 - Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2119-2171). Burlington: Elsevier.

Sachs, J. D., Warner, A., Åslund, A., & Fischer, S. (1995). Economic Reform and the Process of Global Integration. *Brookings Papers on Economic Activity, 1995*(1), 1-118. doi:10.2307/2534573

United Nations. (2014). World Urbanization Prospects: The 2014 Revision. Retrieved from http://esa.un.org/unpd/wup/

Venables, A. J. (2005). Spatial disparities in developing countries: cities, regions, and international trade. *Journal of Economic Geography, 5*(1), 3-21. doi:10.1093/jnlecg/lbh051

World Bank. (2009). *World Development Report 2009. Reshaping Economic Geography*. Washington, D.C.: Author.

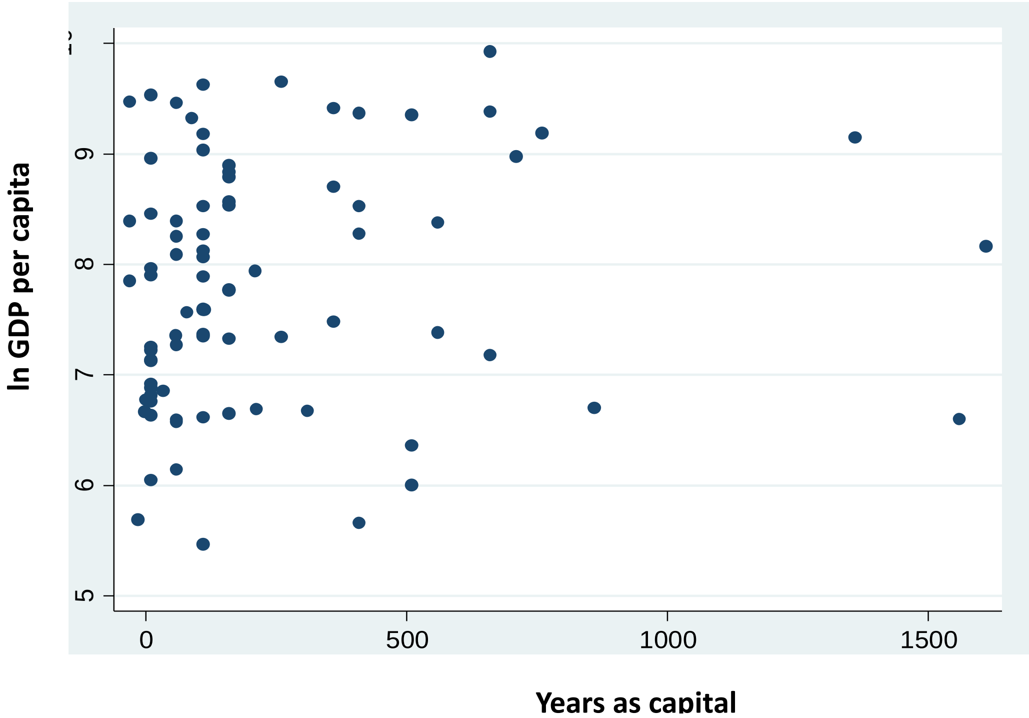
Appendix 1 – Data sources

Variable	Source
Average City Size	Population weighted average city size at the beginning of the five-year period Calculated based on World Urbanization Prospects, the 2014 revision http://esa.un.org/unpd/wup/
GDPpc	Natural logarithm of “Real GDP at constant 2005 national prices “/ “Population” Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Years of Schooling	Years of schooling at beginning of the five-year period Barro & Lee dataset http://www.barrolee.com/
Private Investment	Five-year averages of “Share of gross capital formation of GDP at current PPPs” Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Government Consumption	Five-year averages of “Share of government consumption of GDP at current PPPs” Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Openness	Five-year averages of sum of “Share of merchandise exports of GDP at current PPPs”and “Share of merchandise imports of GDP at current PPPs” Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Population	Population in millions at the beginning of the five-year period Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Political System	Five-year averages of “POLITY2” indicator Polity IV: Regime Authority Characteristics and Transitions Datasets http://www.systemicpeace.org/inscrdata.html

Appendix 2 – Descriptive statistics

Variable	Mean	Standard Deviation			Maximum	Minimum	Obs.
		Overall	Between	Within			
Average city size	1.143712	1.70483	1.488533	.6116051	16.82135	.15	1014
Growth	.0996111	.1669525	.0834671	.148743	1.199445	-1.259275	1014
Ln(GDPpc)	8.42027	1.307601	1.252394	.3871163	11.61566	4.980256	1014
Years of schooling	5.579546	3.265375	3.096604	1.554405	13.27	.15	1014
Private investment	.2008379	.0895442	.0718894	.0550774	.6267554	.0159164	1014
Government consumption	.1853325	.0935007	.0720912	.0626505	.9272356	.0237472	1014
Openness	.4507186	.5670825	.4185368	.3827675	11.06161	.0021284	1014
LN (Population)	2.378241	1.444446	1.405376	.3235726	7.158376	-1.815018	1014
Political System	1.772023	7.329566	6.05452	4.161617	10	-10	1014

Appendix 3 – Scatterplot GDPpc and years as capital in 1960



Appendix 4 – First-stage IV-results

Dependent variable: average city size

VARIABLES	(1) High-income countries	(2) Developing countries
Years capital	**0.549***** **(0.173)**	**-0.331***** **(0.107)**
GDPpc	-619.7*** (149.1)	317.4*** (41.01)
Years of schooling	34.15 (37.95)	-31.84* (16.60)
Private investment	3,375*** (890.2)	-319.4 (300.7)
Government consumption	-2,822*** (996.8)	-1,222*** (273.6)
Openness	250.1* (141.4)	126.6*** (25.80)
Population	297.7*** (49.96)	476.5*** (29.49)
Political system	-29.07** (14.53)	6.092 (4.378)
Constant	4,948*** (1,151)	-2,122*** (311.1)
Time fixed effects	Yes	Yes
Observations	258	579
R-squared	0.451	0.575
F-stat of excluded instrument	10.07 (0.0017)	9.62 (0.0020)

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

3 Big or Small Cities? Does City Size Matter for Growth?

3.1 Introduction

For a long time in world history, building an empire around a city or having a large city within a country was crucial for economic success and/or political dominance: Alexandria around 200BC, Rome in 1AD or Chang'an in China in 800AD are clear examples of this (Chandler, 1987): big cities brought about development and prosperity; and development, in turn, made the growth of big cities possible. This rule was still very much in evidence in the 1950s. At that time, the majority of the world's biggest cities – 20 out of 30 – were located in high-income countries (United Nations, 2014). The few exceptions to the rule were cities in large developing countries, such as China, Mexico, and Brazil. The poorest countries in the world at the time lacked, in general, very large cities.

Partially inspired by this observation, policy makers and economic theorists have long stressed the role of urbanization and cities for economic efficiency and growth (Duranton, 2015; Glaeser, Kallal, Scheinkman, & Shleifer, 1992; World Bank, 2009). The size of a city is considered an important driver of economic development. Based mainly on empirical evidence from the US and the UK, research has stressed how large cities lead to greater productivity and economic growth (Melo, Graham, & Noland, 2009; Rosenthal & Strange, 2004).

Over the past few decades, however, the world has changed significantly and these changes have put the link between large cities and economic success under greater scrutiny. Countries at low levels of economic development have rapidly urbanized and very large cities are no longer predominantly found in rich and/or dynamic countries (United Nations, 2014). Moreover, cities have grown larger. This has challenged the city size/productivity link. Recent empirical evidence suggests that the largest cities in the world today are not necessarily the most productive, but mid-size agglomerations (McCann & Acs, 2011). Furthermore, factors that are often independent of city size, such as urban infrastructure, institutional capacity and industry composition, are increasingly regarded as more important determinants of whether a city can be judged as a motor for economic growth or not (Camagni, Capello, & Caragliu, 2015; Castells-Quintana, 2016; Glaeser, 2014).

The overarching aim of this paper is to delve in greater detail into the link between the size of cities and aggregate economic growth in order to advance our understanding of the factors that may influence this relationship. More specifically, we explore empirically: 1) whether in recent decades certain city sizes can be considered as more growth enhancing than others and 2) how additional factors highlighted in the literature affect the city size growth relationship. The main contribution of our paper is twofold: 1) its adds to the relatively scarce empirical literature on the relationship between city size and aggregate growth (as opposed to an individual city or urban concentration perspective) and 2) it further explores how this relationship is shaped by other factors, something rarely tested in the empirical literature to date.

The remainder of the paper is structured as follows: the next section describes the theoretical and empirical literature that explores the link between city size and economic growth. The third section addresses our methodology, indicators used and presents the dataset. The results and discussion are included in the fourth section. The final part concludes and proposes areas for further research.

3.2 City size and economic growth

Economists have long taken an interest in the relationship between economic development, productivity, urbanization and cities (Hoselitz, 1953; Lewis, 1954; Marshall, 1890). Over the past two decades, the New Economic Geography (NEG) and urban economics strands were among those to examine the mechanisms which govern the relationship. Both schools of thought take different but complementary angles in examining the link, although the overall tenor, in particular from the empirical literature, is that city size and agglomeration are fundamental drivers of economic growth through their impact on people's productivity.

The NEG literature generally sets out to analyse the link between agglomeration and economic performance at the national level. For this purpose the standard NEG framework is combined with endogenous growth models. In the NEG framework of industry location (Krugman, 1991), external scale economies, which are created through the co-location of population and economic activity, make people and companies more productive. The mechanisms through which external economies increase productivity are threefold: a) knowledge spill-overs between workers enable learning and spur innovation; b) forward and backward linkages between companies, suppliers and buyers make interactions between economic actors more efficient; and c) a pooled labour market allows for an easier matching between firms and employees. Simultaneously, centrifugal forces such as high land rents, pollution, and congestion work in the opposite direction as population concentration grows and thus decrease people's productivity. Combining this framework with an endogenous growth model allows to assess the link between a country's level of agglomeration and national growth. The conclusion of these theoretical exercises tends to be that more agglomeration of people and firms and thus a more concentrated urban structure with large cities at its apex, increase people's and companies productivity and is, thus, tantamount to greater economic growth (Fujita & Thisse, 2003; Martin & Ottaviano, 2001).

Empirical studies, using the percentage of the urban population living in cities above 750,000 or 1 million inhabitants as an indicator of agglomeration, support this claim (Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014)[16]. The larger the percentage of the population living in cities above these thresholds, the better the economic performance of countries, particularly at low levels of economic development. Henderson (2003) additionally claims that optimal city size has to be considered in relation to the overall size of a country. One drawback of this literature is, however, that it ignores the vast differences in the sizes of

[16] A number of studies such as Henderson (2003), Brülhart & Sbergami (2009) and Castells (2016), also or primarily use Primacy – the percentage of the urban population that lives in the largest city – as an indicator of agglomeration. In this study, we are however interested in the city size aspect and not in how the urban population is distributed across different cities.

cities beyond a certain population threshold – i.e. over 1 million inhabitants. While any threshold is arbitrary, it is by no means obvious that cities should have a uniform relationship with economic growth beyond this size.

Urban economists adopt a slightly different perspective. Instead of analysing the impact of city size or agglomeration on economic growth at the national level, they are concerned with the impact of city size on the productivity of the urban worker. In this framework, the mechanisms which determine people's productivity are similar to those in the NEG literature. On the one side, Duranton and Puga (2004) describe forces related to the sharing of local infrastructure, matching of skills, suppliers and markets, and learning between people as the micro-foundations behind the localized scale economies that make people in cities more productive. They particularly emphasize the cumulative nature of knowledge generation and diffusion. On the other side, larger city size leads to increased rents and commuting time which reduces people's productivity. Taken all together, a city's productivity is therefore believed to follow an inverted U-shape function of total urban employment. Productivity increases with city size up to a certain population threshold, beyond which the disadvantages of agglomeration overshadow its benefits.

An array of empirical studies find evidence for these productivity gains. Rosenthal & Strange (2004) find that labour productivity increases by 3-8% through a doubling of city size. Productivity gains tend to be greater in cities with high-tech sectors and service industries (Graham, 2009; Henderson, 2010). Meta-analyses confirm this positive relationship, although often underscoring the presence of important regional differences (e.g. Melo, Graham, and Noland, 2009). This view is reinforced by Duranton (2015) who reports that productivity gains are even larger in emerging countries' cities than in cities located in high-income countries (e.g. productivity increases for China between 10% and 12% as a result of doubling of city size).

Despite their differing approach and focus of analysis, both NEG literature and urban economics have contributed to forge a prevailing view: agglomerating people in larger cities increases people's and companies' productivity and, consequently, drives economic growth. In recent years, however, this assertion has increasingly been called into question by a raft of empirical studies (Dijkstra, Garcilazo, & McCann, 2013; ESPON, 2012; McCann & Acs, 2011). Several factors contributed to what seems an increasing need to re-examine the link between city size and aggregate economic growth. First, many developing countries' cities have outgrown their counterparts in the developed world: out of the 30 largest agglomerations in 2015, 23 were located in low- and-middle income countries. Countries as different as Bangladesh, the Democratic Republic of Congo, Mexico, Peru, or Indonesia are represented in this list (United Nations, 2014). This explosion of large cities in many developing countries, which often went hand in hand with the emergence of large inner city slums and significant congestion, has cast increasing doubt on the relationship between city size and productivity for developing countries.

Second, cities have generally grown much larger than before. While New York was the largest city in 1950 with roughly 12 million inhabitants, today Tokyo, the largest city in the world, has 38 million inhabitants. Many developing country cities have grown to become mega cities: Delhi and Shanghai, with 26 million and 24 million inhabitants respectively are clear examples of this trend (United Nations, 2014). The ever growing size of megacities has brought the linearity of the relationship between increased city size and economic growth at these city dimensions into question, especially given the urban economics u-shaped productivity function. Recent empirical research for European countries has provided evidence for this. OECD (2006) for example, finds support for a positive effect of city size on income at the city level, but only if some of the largest cities are excluded from the sample. When restricting the sample to megacities of more than 7 million inhabitants, the coefficient turns negative. Similarly, McCann and Acs (2011) indicate that among the 75 most productive cities in the world 29 cities have a population of less than 3 million inhabitants, while 32 range between 3 and 7 million. Considering the size of today's megacities in particular, these are by no means the largest cities. Dijkstra, Garcilazo, and McCann (2013) explore the contribution of different

European cities to national economic growth. Again, they find that not only have midsize cities grown faster than very large cities, but also that their contribution to national economic growth has been greater. Research by Parkinson and co-authors covering 31 European capitals and 124 second-tier cities (ESPON, 2012) comes to a similar conclusion. The evidence emerging from this literature thus suggests a mixed picture in which the role played by city size is far from uniform. This research has, thus, raised pertinent questions about the types of city sizes that tend to be growth maximising.

Finally, researchers have increasingly gone beyond city size and emphasized the role of other factors behind changes in city-level productivity (Camagni, Capello, & Caragliu, 2013; Camagni et al., 2015). Industry specific aspects such as industrial composition (Au & Henderson, 2006; Graham, 2009) and company size (Faggio, Silva, & Strange, 2014; Rigby & Brown, 2015), context specific elements including government capacity (Ahrend, Farchy, Kaplanis, & Lembcke, 2014; Glaeser, 2014), and urban infrastructure (Castells-Quintana, 2016) as well as network integration (Sassen, 1991; Scott, 2001) and borrowed city size (Alonso, 1973) have featured prominently in these analyses. The importance of industrial composition, in particular, has attracted considerable attention. Tradable services and high-tech manufacturing benefit more from large cities than other sectors (Graham, 2009), meaning that economies with a strong presence of such sectors benefit more from the presence of large cities than countries with a larger and/or more mature manufacturing industry base. Similarly, the provision of an adequate urban infrastructure is increasingly regarded as key for productivity and growth. Castells-Quintana (2016), for instance, shows empirically that in countries where the urban infrastructure, as proxied by sanitation, is not sufficiently developed, urban concentration can be growth inhibiting. Glaeser (2014) and Ahrend et al. (2014) furthermore highlight the importance of institutional capacity, in particular in developing countries, in order to reduce negative externalities which arise through rapid city growth. An empirical analysis of the city size growth link should therefore ideally account for these factors.

3.3 Model and data

In the following sections, we empirically test the relationship between city size and aggregate growth. As our interest lies in the relationship between city size and growth at the national level and not growth at the city level, the approach of the empirical NEG literature, that examines the link between urban concentration and national economic growth, is most suited for our purpose (Brülhart & Sbergami, 2009; Castells-Quintana & Royuela, 2014).

The econometric model is built around a simple GDP per capita growth equation based on the extended Solow growth model (Durlauf, Johnson, & Temple, 2005). We estimate a country's growth rate in five-year periods as a function of an indicator for the size of a country's cities as well a set of control variables commonly used in growth regressions. The model takes the following simple form:

$$\Delta GDPpc_{ip} = \alpha + \beta \; citysize_{ip} + \gamma \; GDPpc_{ip} + \delta \; \boldsymbol{X}_{ip} + \mu_p + \varepsilon_{ip}$$

Where

p denotes five-year intervals;

$\Delta GDPpc_{ip}$ is the GDP per capita growth rate of country i in period p;

$\mathrm{citysize}_{\mathrm{ip}}$ is the percentage of the urban population of country i living in cities of certain size at the beginning of period p;

$\mathrm{GDPpc}_{\mathrm{ip}}$ represents the ln GDP per capita of country i at the beginning of period p;

X_{ip} depicts a set of control variables for country i, measured either at the beginning or as an average of period p;

μ_p represents time fixed effects; and

ε_{ip} is the error term.

For the variable of interest, $citysize_{ip}$, we rely on data from the 2014 World Urbanization Prospects (WUP) which provides information on the percentage of the urban population living in cities of different size categories (United Nations, 2014). We hypothesise that if cities of a certain size contribute to maximising the productivity of workers, a growing share of the urban population living in cities of that size should promote economic growth. To test this hypothesis, we combine different city-size categories available from the WUP to fit the OECD classification used in McCann and Acs (2011) as well as define further categories below and above the established thresholds in order to add greater nuance. We also take into consideration different well-established thresholds of what is considered a megacity: 7 million (OECD, 2006) and 10 million (United Nations, 2014). The resulting categories are the percentage of the urban population living in cities 1) below 500,000, 2) 500,000 to 1 million, 3) 1 million to 3 million, 4) 3 million to 7 million, 5) 7 million to 10 million and 6) 10 million and more.[17] An interaction term with the natural logarithm of the urban population is also included in the analysis, as it is conceivable that the effect depends on the size of the country (Henderson, 2003).

As controls, we resort to the standard variables mainly used in cross-country growth regressions as well as a set of variables that may influence the size of a country's cities and growth simultaneously. For the first set, we include a) initial GDP per capita to control for conditional convergence – growth is assumed to be faster in countries which start from a lower level (Durlauf et al., 2005); b) years of schooling to measure a country's human capital (Lucas, 1988; Romer, 1986); and c) private investment as a percentage of GDP as it is assumed to be a key growth driver in the Solow model.

With regards to the controls which may be directly linked to city size as well as a country's economic performance, we include the following variables. The first is openness, which is believed to influence a country's spatial structure and city size (Krugman & Elizondo, 1996)

[17] Ultimately, any size threshold is arbitrary. Our goal is to provide a more nuanced analysis than studies thus far that rely on a unique threshold of 750,000 or 1,000,000. Resorting to the widely used city category definitions from OECD and the UN provides a helpful reference point for this purpose.

as well as to directly promote economic growth (Sachs, Warner, Åslund, & Fischer, 1995). Including it ensures that the coefficient of our city size categories does not wrongly capture this latter effect. The second is a country's physical land area. A large surface area such as in India or Brazil is expected to be closely linked to the size of a country's cities. This may also be the case for a country's urban population. Similarly, a country's size (either measured in physical size or population) also influences its market potential, which in turn may have an effect on aggregate economic performance (Alesina, Spolaore, & Wacziarg, 2005). It is therefore pertinent to include it in the regression. Furthermore, a number of dummy variables are considered in the analysis to account for the particularities in the growth and urbanization path that certain groups of countries may have experienced: a) continental dummies to account for systematic differences in the urbanization and growth paths of regions; b) economies which are strongly dependent on oil exports and c) states which formed part of the former Soviet Union. Finally, we include time fixed effects to ensure results are not driven by period specific trends.

The data for the controls are derived from different sources. The years of schooling indicator comes from the Barro and Lee dataset, urban population from the 2014 edition of the World Urbanization Prospects, physical land area and the dummy for oil dominated economies is based on the World Development Indicators. The remainder of the data is sourced from the eighth edition of the Penn World Tables. A more detailed description of the indicators and sources as well as descriptive statistics are included in Appendix 1 and Appendix 2. Depending on the different specifications, a maximum of 113 countries is covered in the analysis.

The model is estimated using the Hausman-Taylor (HT) estimator. The HT has the advantage over a fixed effects estimator that it allows for the inclusion of time-invariant variables in a panel setting, i.e. land area as well regional and oil dummies in our case (Baltagi, Bresson, & Pirotte, 2003). This approach also permits the introduction of a number of interaction terms to address our second research question. For the time-varying variables, the HT works analogue to the fixed effects models as it uses the within transformation of time-varying variables to estimate consistent coefficients. In contrast to the FE model however, the HT also uses individual means of the time variant regressors as instruments for the time invariant variables and therefore allows their inclusion in the model.

3.4 Results

3.4.1 Size categories and economic growth

In order to analyse whether the percentage of the urban population living in cities of different size categories affects national economic growth, we include the indicators for the different size categories as well as their interaction terms with urban population one by one. Table 3-1 presents the results for these estimates.

The basic model (regression 1) works well. The control variables show the expected signs and remain stable when the variables of interest – the different city size categories – are introduced in the analysis (regressions 2 through 13). Initial GDPpc is negative and significant at the 1% level throughout all our estimates, indicating conditional convergence. Private investment is the key driver for growth in the model: the coefficient is positive and highly significant in all estimations. Urban population has a negative association with economic growth in most regressions. This result is intuitive given that urban population is highly correlated to the overall population. Increasing population while holding other factors such as capital and years of schooling constant decreases the factor endowment per person and should, hence, have a negative impact on GDP per capita growth. The positive coefficients of the physical land area are also in line with expectations: holding other factors constant, adding another unit of land increases factor endowment per person and growth. This result does not hold when we control for the percentage of the urban population living in small cities. Furthermore, the dummy for countries whose exports are dominated by oil has – as expected – a positive and significant coefficient throughout and the dummy for the countries which formed part of the former Soviet Union is negative and significant at the 10% level in all but two regressions.

Finally, years of schooling and openness are not significant in our results. While these coefficients do not necessarily match expectations, they are not entirely surprising. Increases in years of schooling are unlikely to show an immediate effect on growth (Durlauf et al., 2005). Likewise the results for different indicators of openness and trade integration have varied between studies with results frequently being insignificant (Durlauf et al., 2005).

Let us now turn to the results for our variable of interest: the share of the urban population living in cities of certain size categories (regressions 2 - 13). Each indicator is inserted in the model individually, as there is a clear interdependence between them – first alone and then with the interaction term. Three out of the six size categories have a significant coefficient when introduced alone; four of them once the interaction term of urban population is considered.

Table 3-1: Dependent variable: GDPpc growth in five-year intervals, 1980 – 2010

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
% Below 500K		0.00364***	0.00431***				
		(0.000667)	(0.000695)				
* Urban population			-0.00166***				
			(0.000527)				
% 500K to 1M				-0.00114**	-0.00199***		
				(0.000498)	(0.000697)		
* Urban population					0.00146*		
					(0.000846)		
% 1M to 3M						-0.00107**	-0.00353***
						(0.000499)	(0.000888)
* Urban population							0.00228***
							(0.000681)
Initial GDPpc	-0.475***	-0.491***	-0.503***	-0.477***	-0.478***	-0.481***	-0.488***
	(0.0285)	(0.0279)	(0.0280)	(0.0284)	(0.0284)	(0.0286)	(0.0284)
Years of schooling	0.00450	0.00422	-0.00283	0.00449	0.00129	0.00400	0.000227
	(0.0129)	(0.0126)	(0.0127)	(0.0129)	(0.0130)	(0.0129)	(0.0128)
Private investment	0.792***	0.807***	0.793***	0.794***	0.785***	0.821***	0.806***
	(0.124)	(0.121)	(0.120)	(0.124)	(0.123)	(0.124)	(0.123)
Openness	0.000916	-0.00384	-0.000798	-0.00341	-0.00177	0.00258	0.00426
	(0.0158)	(0.0154)	(0.0153)	(0.0159)	(0.0159)	(0.0158)	(0.0156)
Urban population	-0.118***	-0.0570	0.0727	-0.112***	-0.115***	-0.110***	-0.124***
	(0.0384)	(0.0391)	(0.0565)	(0.0384)	(0.0383)	(0.0385)	(0.0384)
Land area	0.0859***	0.0554	0.0400	0.0814**	0.0761**	0.0812**	0.0755**
	(0.0333)	(0.0341)	(0.0338)	(0.0332)	(0.0333)	(0.0335)	(0.0337)
Oil dummy	0.411***	0.457***	0.477***	0.417***	0.413***	0.428***	0.424***
	(0.159)	(0.164)	(0.160)	(0.158)	(0.158)	(0.161)	(0.163)
Soviet union dummy	-0.329*	-0.289	-0.226	-0.312*	-0.299*	-0.322*	-0.313*
	(0.171)	(0.177)	(0.174)	(0.171)	(0.171)	(0.173)	(0.175)
Constant	2.468***	2.667***	2.877***	2.551***	2.634***	2.578***	2.713***
	(0.434)	(0.440)	(0.437)	(0.433)	(0.435)	(0.439)	(0.443)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113	113

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
Oil dummy: Oil rents on average more than 15% of GDP per capita since data is available for a given country

Table 3-1: (continued)

VARIABLES	(8)	(9)	(10)	(11)	(12)	(13)
% 3M to 7M	0.000181	0.00301				
	(0.000768)	(0.00210)				
* Urban population		-0.00135				
		(0.000935)				
% 7M to 10M			0.000138	-0.0139		
			(0.00117)	(0.0100)		
* Urban population				0.00426		
				(0.00302)		
% 10M +					0.00134	-0.0196*
					(0.00142)	(0.0109)
* Urban population						0.00587*
						(0.00304)
Initial GDPpc	-0.475***	-0.478***	-0.474***	-0.483***	-0.477***	-0.485***
	(0.0286)	(0.0287)	(0.0286)	(0.0292)	(0.0286)	(0.0289)
Years of schooling	0.00441	0.00375	0.00456	0.00368	0.00400	0.00209
	(0.0129)	(0.0129)	(0.0130)	(0.0130)	(0.0129)	(0.0129)
Private investment	0.797***	0.813***	0.792***	0.786***	0.780***	0.783***
	(0.126)	(0.127)	(0.124)	(0.124)	(0.125)	(0.124)
Openness	0.000691	-0.000541	0.000837	0.00134	0.00212	0.00365
	(0.0159)	(0.0159)	(0.0158)	(0.0158)	(0.0159)	(0.0158)
Urban population	-0.118***	-0.122***	-0.118***	-0.122***	-0.117***	-0.117***
	(0.0385)	(0.0386)	(0.0386)	(0.0387)	(0.0385)	(0.0384)
Land area	0.0860***	0.0935***	0.0861***	0.0866**	0.0842**	0.0801**
	(0.0333)	(0.0340)	(0.0333)	(0.0337)	(0.0333)	(0.0334)
Oil dummy	0.411***	0.413**	0.410***	0.422***	0.418***	0.439***
	(0.159)	(0.161)	(0.159)	(0.162)	(0.159)	(0.159)
Soviet union dummy	-0.328*	-0.338*	-0.330*	-0.330*	-0.324*	-0.311*
	(0.172)	(0.174)	(0.171)	(0.174)	(0.171)	(0.172)
Constant	2.471***	2.407***	2.465***	2.534***	2.505***	2.621***
	(0.434)	(0.440)	(0.435)	(0.442)	(0.435)	(0.439)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

Notes: Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
Oil dummy: Oil rents on average more than 15% of GDP per capita since data is available for a given country

The results here do not necessarily go along expectations. First, the percentage of the urban population living in cities below 500,000 inhabitants is positive and significant at the 1%. This indicates a growth promoting effect of a large share of people living in relatively small cities. This result holds when we add the interaction term with the urban population, however it is nuanced by a negative sign on the interaction term. This points to the existence of an urban population threshold: the connection between the percentage of the urban population living in cities below 500,000 inhabitants is positive up to a total urban population of 12.3 million. Beyond that level, it turns negative. Hence, the growth promoting effect of small cities holds for the great majority of countries in our sample. 77 countries out of the 113 countries included in our regressions had urban populations of 12.5 million or less in 2010.

In contrast, the coefficients for the urban population share in cities between 500,000 and 1 million inhabitants as well as between 1 million and 3 million are negative and significant at the 5% level. Countries with a higher share of their urban population in mid-sized cities have thus had a weaker economic performance. The introduction of the interaction term strengthens this relationship: the coefficients becoming significant at the 1% level. Both interaction terms are, however, positive and significant, meaning that the effect of the share of the urban population in cities with a population between 1 and 3 million turns positive for countries with an overall urban population of 4.8 million and at 3.8 million for cities between 500,000 and 1 million inhabitants. The influence of mid-sized cities is, thus, heavily affected by the country's size.

Beyond the 3 million threshold, the percentage of the urban population living in very large to megacities seems irrelevant for economic growth. All the coefficient for cities between 3 and 7 million inhabitants and from 7 million to 10 million are insignificant (Regressions 8-11). For the percentage of the urban population living in megacities of more than 10 million, the coefficient is also not relevant (Regression 12) and only becomes significant, once the interaction term is introduced (Regression 13). In this case, the association is negative: the bigger the share of population living in very large megacities, the lower the economic growth. The positive coefficient of the interaction term signals that this relationship is, however, not always pernicious for development. In countries with urban populations larger than 28.2 million – 25 out of the 113 countries in the sample in 2010 – the presence of megacities becomes a positive force for growth. However, the low level of significance – both coefficients at the 10% level – warrants some caution in the interpretation of these results.[18]

[18] If we run regressions with alternative size categories, i.e. the percentage of the urban population living in cities above 1 million as well as the percentage of the urban population living in cities above 3 million, the results suggest that the growth inducing or inhibiting effect of city size is indeed captured by the different categories employed in our regressions. The coefficients for 1M+ and its interaction with urban population mirror those for our 1M to 3M category results, while the 3M+ results are insignificant.

Taken together, our results highlight two important facts. First, in contrast to expectations, most of the action takes place at the lower rather than at the upper end of the urban scale: small cities with a population of less than 500,000 inhabitants have been a source for economic dynamism. By contrast, a large share of the urban population living in cities between 500,000 and 3 million is detrimental for economic growth. Beyond that threshold, the size of cities in a country does not seem to affect its economic performance in a positive or a negative way. Second, the size of the country matters for the link between population living in cities of different sizes and growth: the positive connection between small cities (below 500,000) and national growth levels works best in relatively small countries while for countries with more than 12.3 million urban citizens, the relationship turns negative. The negative effect of the 500,000 to 3 million categories is only present in countries of up to 4 million urban inhabitants. Megacities beyond 10 million only fuel growth in countries with more than 28.5 million urbanites. Roughly speaking, small cities work best in relatively small countries, megacities in big countries. Medium-sized cities are only detrimental to economic growth in very small countries.

Our results hence neither validate current wisdom about the relationship between different city sizes and growth nor refute it completely. As in Henderson (2003), the analysis points to a strong dependence of the impact of city size on the country's overall population. As mentioned, large cities are more efficient in larger countries, small ones in smaller ones. However, the positive economic role of cities with a population of less than 500,000 in the analysis in the large majority of countries in the sample is something which has been overlooked by a literature which has overwhelmingly tended to place the emphasis on large cities (Martin & Ottaviano, 2001; Rosenthal & Strange, 2004). The results of the analysis stress that megacities become a force for aggregate economic dynamism only in the most populous countries, which make a small part of our sample. In fact, the greatest economic benefits have been extracted from the presence of relatively small cities of up to 3 million inhabitants rather than from the growth of megacities for most countries in the world.

A number of potential explanations can be put forward. First, small cities may simply make it easier to manage urban diseconomies, arising through increasing city size in most parts of the world. Congestion, high rents, and pollution do not arise in small- and medium-sized cities to the same degree as in larger cities. The solutions to address diseconomies of scale are also often easier to implement. At the same time, small- and medium-sized cities still often deliver many of the benefits of large cities, for instance in terms of provision of public goods and efficiency of market interactions. Hence, smaller cities may on average maximise productivity without generating negative externalities. This is probably especially true in many developing countries, which struggle to provide an adequate urban infrastructure.

A second explanation may relate to the levels of urban concentration that a larger share of the urban population in small cities imply. The results suggest a positive effect of small cities, which decreases the larger the urban population and eventually turns negative. We also detect a negative effect of cities between 500,000 and 3 million, which turns positive beyond an urban population threshold of roughly 4 million. Put together, these factors suggests that a relatively balanced urban structure within a given country is growth promoting, but that levels of urban concentration should not fall beyond a certain threshold as the positive interaction terms with urban population for cities between 500,000 and 3 million and the negative one for the cities below 500,000 suggests.

Finally, the peculiarities of the data require some caution when interpreting the insignificant results for larger city categories. Only a limited number of countries in the sample have cities with a population of 3 million inhabitants or more (43 out of the 113). Hence, there are many observations where the percentage of the urban population in the size categories above 3 million is zero which may make it more difficult to obtain significant coefficients. However, it is worth pointing out that the results remain insignificant even if we combine these categories, i.e. the percentage of the urban population in cities above 3 million. This suggests that there is high heterogeneity in cities above this threshold: while some may be true motors of economic growth, as indicated by McCann and Acs (2011), others may be less so.

3.4.2 City size, context and economic growth

Recent literature has shown that a variety of factors may shape the relationship between city size and economic growth (see literature section above). The sectoral composition of the economy, a country's infrastructure endowment and the capacity of the national and city governments have been particularly emphasized in this regard. To explore these interactions, we rely on the same specifications as in the previous section and check through the introduction of additional variables whether a) the sectoral composition of economies, b) access to urban infrastructure and c) the institutional capacity affect how cities mould the economic growth of a country. While many other factors, such as borrowed city size, international linkages and the maturity of firms, may also influence the relationship, we analyse those factors for which information is more readily available

3.4.2.1 Sectoral composition

To test for the potential influence of a country's sectoral composition on how cities affect economic growth, we resort to export statistics for industries which, according to the literature, generate the largest agglomeration economies, namely market-oriented services and high-tech industries (Graham, 2009; Henderson, 2010). Countries where these industries are important should benefit more from larger cities than countries where these industries are more trivial.

Table 3-2: Sectoral composition – Dependent variable: GDPpc growth in five-year intervals, 1980 – 2010

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
% Below 500K	0.00235***					
	(0.000675)					
* Service exports	0.00492					
	(0.00711)					
* Urban population	-0.00137***					
	(0.000424)					
% 500K to 1M		-0.00266***				
		(0.000730)				
* Service exports		0.0130**				
		(0.00596)				
* Urban population		0.00104				
		(0.000683)				
% 1M to 3M			-4.13e-05			
			(0.00101)			
* Service exports			-0.00611*			
			(0.00329)			
* Urban population			0.000732			
			(0.000604)			
% 3M to 7M				-0.00182		
				(0.00254)		
* Service exports				0.00764*		
				(0.00456)		
* Urban population				0.000195		
				(0.000933)		
% 7M to 10M					-0.00453	
					(0.00797)	
* Service exports					0.00166	
					(0.0193)	
* Urban population					0.00143	
					(0.00236)	
% 10M +						-0.00632
						(0.00867)
* Service exports						-0.0265
						(0.0217)
* Urban population						0.00261
						(0.00232)
Service exports	0.0844	0.153	0.439**	0.293	0.334	0.366*
	(0.491)	(0.221)	(0.216)	(0.212)	(0.211)	(0.210)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	528	528	528	528	528	528
Number of countries	88	88	88	88	88	88

	(7)	(8)	(9)	(10)	(11)	(12)
% Below 500K	0.00425***					
	(0.000696)					
* High-tech dummy	0.00166					
	(0.00156)					
* Urban population	-0.00165***					
	(0.000540)					
% 500K to 1M		-0.00190***				
		(0.000705)				
* High-tech dummy		-0.000408				
		(0.00173)				
* Urban population		0.00139				
		(0.000855)				
% 1M to 3M			-0.00318***			
			(0.000916)			
* High-tech dummy			-0.00173			
			(0.00112)			
* Urban population			0.00226***			
			(0.000681)			
% 3M to 7M				-0.000628		
				(0.00291)		
* High-tech dummy				0.00349**		
				(0.00177)		
* Urban population				-0.000366		
				(0.00109)		
% 7M to 10M					-0.0131	
					(0.0101)	
* High-tech dummy					-0.000404	
					(0.00269)	
* Urban population					0.00402	
					(0.00307)	
% 10M +						-0.0184*
						(0.0110)
* High-tech dummy						-0.00142
						(0.00279)
* Urban population						0.00565*
						(0.00309)
High-tech Dummy	-0.0611	0.0552	0.0918**	0.0536	0.0552	0.0575
	(0.108)	(0.0449)	(0.0446)	(0.0406)	(0.0403)	(0.0413)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

For market-oriented services, UNCTAD provides export statistics for 88 countries in our sample since the year 1980. Based on this, we calculate the share of service exports of total GDP as a measure of the intensity of services in the economy in each five-year period. Information on high-tech exports, namely the percentage of high-tech of overall manufacturing exports, can be obtained from the World Development Indicators since the year 1990. In order to maintain comparability with our base case results (1980-2010), we define a dummy variable to identify countries in which high-tech manufacturing plays an important role. The dummy takes the value of 1 if the share of high-tech goods of manufactured exports in a given five-year interval is larger than the 80th percentile of the overall observations (16.2%).[19] For the vast majority of the countries the value remains constant throughout the period with data (1990-2010). We therefore assume them to be the same for the two five-year periods in the 1980s without data. In some special cases such as China and Costa Rica, where the export structure has changed significantly since the early 1990s, we account for this by assigning a 0 in the early periods and a 1 for later periods. [20]

The export measures are introduced into the regressions together with their interaction term with the city size categories. Table 3-2 shows the results: columns 1 – 6 for the service sector indicator and 7 – 12 for high-tech industries.[21]

In both cases, the results for the city category below 500,000 remains unchanged with the main term being positive and the interaction term with urban population negative. For the 500,000 to 1 million category, the interaction term with service exports is positive, whereas the main term remains negative. This indicates that while the main negative effect holds, a strong service sector actually mitigates the negative connection of this type of cities and economic growth. In contrast, high-tech industries do not seem to affect this category.

[19] The average export share of high-tech exports of manufacturing exports remains relatively unchanged for the period with data, i.e. 8.9% in the 5 year period from 1991 to 1995 and 9.1% for 2006 to 2010.
[20] This is the case in only 14 out of the 113 countries, including China, Costa Rica, Iceland and Hungary, where there is a clear trend with little high-tech exports in the early years and strong presence in the later.
[21] The full results are included in Appendix 3.

Similarly the 1M to 3M impact remains unaffected by the accounting for high-tech industries, while controlling for the service sector renders the impact insignificant. The most interesting effect can be observed in the 3 to 7 million category which is insignificant in our base case results. The interaction term for both services and high-tech is positive, while the main term and the interaction with urban population stay insignificant. This implies a growth promoting effect of cities between 3 and 7 million inhabitants provided that the service sector or the high-tech manufacturing industries play an important role in the economy. There are no changes for the 7 to 10 million category compared to the base case. The over 10 million becomes insignificant when including service exports in the analysis. We thus find that while most results from our base case hold when accounting for industry composition, countries can benefit from larger cities, i.e. a higher share of the urban population in cities between 3 and 7 million, provided that the country's economy has a strong presence in sectors that benefit from agglomeration economies. This suggests that larger city size can indeed be growth promoting, but only under certain conditions.

3.4.2.2 Context

For the analysis of how context factors shape the relationship between city size and aggregate growth, we consider urban infrastructure and government capacity. As in Castells-Quintana (2016), the share of the urban population with access to improved sanitation facilities is used as a proxy for urban infrastructure. We build a dummy which takes the value of 1 if more than 90% of the urban population has access to sanitation in each five-year period. We cannot rely on the actual data since, as with high-tech exports, data is only available from 1990 onwards.[22] Countries with a good urban infrastructure should be able to benefit more from larger cities. For government capacity, we rely on the World Governance Indicators. We specifically use the sub-indicator "governance effectiveness" which captures the quality of public services on a scale from -2.5 to 2.5. Data is available since 1996, so we calculate the average score for

[22] For the period 1980-1990, we assume that the dummy takes a value of 0 if a country does not pass the 90% in 1990.

each country from 1996 to 2010. Based on this, we assign a dummy which equals one if the average score is 0.97 or above[23] and zero if it is below.[24] As with the sanitation dummy, countries with a higher level of government effectiveness are expected to be able to capture more of the benefits of larger cities. Table 3-3 presents the results: column 1 – 6 accounting for sanitation and 7 – 12 for government effectiveness.[25]

The base case results remain largely unaffected by the introduction of the two new control variables with two notable exceptions. Similarly to the industrial composition results, taking into account sanitation and government effectiveness makes cities between 3 to 7 million inhabitants a driver of growth as indicated by the positive and significant interaction term. This result is as anticipated: larger cities can be growth promoting provided that their urban infrastructure is sufficiently developed and the government is effective, including implementing those policies aimed at mitigating the negative externalities which come with large city size.

The second notable difference is the interaction term between the percentage of the urban population living in cities with less than 500,000 inhabitants and sanitation. It is negative and significant at the 5% level. This implies that despite a main positive effect of a large share of people in small cities, this effect is smaller in countries with a well-developed urban infrastructure. The urban population threshold at which a larger share of the urban population in small cities becomes negative is also lower than in the base analysis.

[23] 0.97 equals the 80th percentile of the observations.
[24] Most countries' scores have not varied dramatically over the years. There are some cases where the difference between minimum and maximum score is larger than 1. However, these cases do not pass our defined threshold of 0.97 in any of the years. Thus, we believe the classification, while simplified, is accurate enough for our purpose.
[25] The full results are included in Appendix 2.

Table 3-3: Contextual factors – Dependent variable: GDPpc growth in five-year intervals, 1980 – 2010

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
% Below 500K	0.00448***					
	(0.000696)					
* Sanitation	-0.00350**					
	(0.00141)					
* Urban population	-0.00135**					
	(0.000540)					
% 500K to 1M		-0.00196***				
		(0.000713)				
* Sanitation		-0.000310				
		(0.00138)				
* Urban population		0.00150*				
		(0.000858)				
% 1M to 3M			-0.00369***			
			(0.000933)			
* Sanitation			0.000605			
			(0.00102)			
* Urban population			0.00224***			
			(0.000687)			
% 3M to 7M				-0.00165		
				(0.00323)		
* Sanitation				0.00361*		
				(0.00190)		
* Urban population				-7.66e-05		
				(0.00115)		
% 7M to 10M					-0.0134	
					(0.0103)	
* Sanitation					-0.000592	
					(0.00239)	
* Urban population					0.00417	
					(0.00303)	
% 10M +						-0.0155
						(0.0130)
* Sanitation						-0.00196
						(0.00331)
* Urban population						0.00511
						(0.00330)
Sanitation	0.157	0.0269	-0.00649	0.0241	0.0208	0.0254
	(0.104)	(0.0877)	(0.0880)	(0.0875)	(0.0878)	(0.0875)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

	(7)	(8)	(9)	(10)	(11)	(12)
% Below 500K	0.00438***					
	(0.000702)					
* Government effectiveness	-0.00294					
	(0.00310)					
* Urban population	-0.00149***					
	(0.000547)					
% 500K to 1M		-0.00191***				
		(0.000701)				
* Government effectiveness		0.00262				
		(0.00209)				
* Urban population		0.00100				
		(0.000914)				
% 1M to 3M			-0.00321***			
			(0.000920)			
* Government effectiveness			-0.00158			
			(0.00114)			
* Urban population			0.00237***			
			(0.000687)			
% 3M to 7M				-0.00229		
				(0.00340)		
* Government effectiveness				0.00412**		
				(0.00205)		
* Urban population				0.000251		
				(0.00124)		
% 7M to 10M					-0.0146	
					(0.0101)	
* Government effectiveness					-0.00222	
					(0.00352)	
* Urban population					0.00453	
					(0.00305)	
% 10M +						-0.0196*
						(0.0110)
* Government effectiveness						-0.00128
						(0.00373)
* Urban population						0.00593*
						(0.00305)
Government effectiveness	0.982**	0.768*	0.847**	0.829*	0.827*	0.815*
	(0.464)	(0.414)	(0.430)	(0.430)	(0.425)	(0.417)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

This finding is in line with Castells-Quintana (2016) who shows that, in particular in developing countries, city growth without the adequate provision of infrastructure results in inefficiencies, as urban diseconomies prevail over agglomeration benefits. Furthermore, they provide support for the hypothesis that the positive effect of cities below 500,000 results more often than not from the lack of capacity of numerous countries to manage the diseconomies of scale associated to larger cities.

These results clearly need to be interpreted as directional as the dummies do not capture all the variation between countries and over time. However, they still emphasize the need to take each country's context into account when analysing the link between city size and growth.

3.5 Conclusion

In this paper we have analysed the link between city size and aggregate growth. Using a panel of 113 countries between 1980 and 2010, we have specifically explored whether there are certain city sizes that are growth enhancing and how additional factors highlighted in the literature impact the city size growth relationship. Our results suggest a non-linear relationship which is greatly dependent on the country's size. A larger share of the urban population in small cities with less than 500,000 inhabitants is growth promoting in small- and medium-size countries (those with an urban population below 12.3 million). Conversely, the share of urban population living in medium-sized cities – defined as those between 500,000 and 3 million – limits economic growth in most small countries (those with an urban population roughly below 4 million), but facilitates economic dynamism beyond this threshold. Large populations living in megacities – defined as those beyond 10 million inhabitants – help growth only in relatively large countries, with an urban population of 28.5 million and more. Finally, cities between 3 and 10 million do not seem to have a systematic influence on growth at all.

The analysis also showed that this relationship is highly context dependent: a high share of industries that benefit from agglomeration economies, a well-developed urban infrastructure and an adequate level of governance effectiveness allow countries to take advantage of agglomeration benefits from larger cities, i.e. between 3 and 7 million inhabitants. It also reduces the benefits of a large share of the urban population in small cities below 500,000 inhabitants. Our results thus suggest that larger cities, albeit not the very large metropoli, can indeed be drivers of growth, but only if the context is favourable. For many countries, smaller cities in fact provide a better balance between the advantages of cities and the diseconomies arising through growing cities.

The results of our analysis should be interpreted as somewhat directional as data availability limits a more accurate measurement of the enabling sectoral composition and contextual factors. Similarly, the city size categories, while chosen to fit the OECD and other frequently used classifications, are still arbitrary and potentially camouflage greater complexity. Despite these caveats, our study adds important further nuance to the scarce empirical literature which links city size to aggregate economic performance and which tends to emphasize the benefits of increased city size (Brülhart & Sbergami, 2009; Rosenthal & Strange, 2004). By using more detailed indicators for the size of a country's cities, it becomes evident that the benefits of increasing city size are not without limits and for many countries, smaller cities seem to offer a better balance. Furthermore, the findings complement the growing stream of literature that goes beyond city size to consider additional factors that shape the relationship between cities and economic growth (Camagni et al., 2015) and again emphasize the context dependency of the relationship.

These findings are particularly pertinent in the current context in which – in contrast to previous centuries – the largest cities are now in developing countries and not in the developed world. A continued focus on the benefits of large cities over smaller ones without considering the context from policy makers and academics may lead to simplified and non-adequate policy recommendations, in particular in a developing country context.

References

Ahrend, R., Farchy, E., Kaplanis, I., & Lembcke, A. C. (2014). *What Makes Cities More Productive? Evidence on the Role of Urban Governance from Five OECD Countries*. Retrieved from Paris: http://dx.doi.org/10.1787/5jz432cf2d8p-en

Alesina, A., Spolaore, E., & Wacziarg, R. (2005). Chapter 23 - Trade, Growth and the Size of Countries. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part B, pp. 1499-1542). Amsterdam: Elsevier.

Alonso, W. (1973). Urban zero population growth. *Daedalus, 102*(4), 191-206.

Au, C. C., & Henderson, J. V. (2006). Are Chinese Cities Too Small? *Review of Economic Studies, 73*(3), 549-576. doi:10.1111/j.1467-937X.2006.00387.x

Baltagi, B. H., Bresson, G., & Pirotte, A. (2003). Fixed effects, random effects or Hausman-Taylor? A pretest estimator. *Economics Letters, 79*(3), 361-369. doi:10.1016/s0165-1765(03)00007-7

Brülhart, M., & Sbergami, F. (2009). Agglomeration and growth: Cross-country evidence. *Journal of Urban Economics, 65*(1), 48-63. doi:10.1016/j.jue.2008.08.003

Camagni, R., Capello, R., & Caragliu, A. (2013). One or infinite optimal city sizes? In search of an equilibrium size for cities. *Annals of Regional Science, 51*(2), 309-341. doi:10.1007/s00168-012-0548-7

Camagni, R., Capello, R., & Caragliu, A. (2015). The Rise of Second-Rank Cities: What Role for Agglomeration Economies? *European Planning Studies, 23*(6), 1069-1089. doi:10.1080/09654313.2014.904999

Castells-Quintana, D. (2016). Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*. doi:10.1016/j.jue.2016.02.003

Castells-Quintana, D., & Royuela, V. (2014). Agglomeration, inequality and economic growth. *Annals of Regional Science, 52*(2), 343-366. doi:10.1007/s00168-014-0589-1

Chandler, T. (1987). *Four thousand years of urban growth: an historical census*. Lewiston, N.Y., U.S.A: St. David's University Press.

Dijkstra, L., Garcilazo, E., & McCann, P. (2013). The Economic Performance of European Cities and City Regions: Myths and Realities. *European Planning Studies, 21*(3), 334-354. doi:10.1080/09654313.2012.716245

Duranton, G. (2015). Growing through Cities in Developing Countries. *World Bank Research Observer, 30*(1), 39-73. doi:10.1093/wbro/lku006

Duranton, G., & Puga, D. (2004). Chapter 48 - Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063-2117). Burlington: Elsevier.

Durlauf, S. N., Johnson, P. A., & Temple, J. R. W. (2005). Chpater 8 - Growth Econometrics. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1, Part A, pp. 555-677). Amsterdam: Elsevier.

ESPON. (2012). *Second Tier Cities and Territorial Development in Europe: Performance, Policies and Prospects*. Retrieved from http://www.espon.eu/export/sites/default/Documents/Projects/AppliedResearch/SGPTD/SGPTD_Final_Report_-_Final_Version_27.09.12.pdf

Faggio, G., Silva, O., & Strange, W. (2014). *Heterogenous Agglomeration*. Retrieved from http://eprints.lse.ac.uk/58426/1/sercdp0152.pdf

Fujita, M., & Thisse, J. F. (2003). Does geographical agglomeration foster economic growth? And who gains and loses from it? *Japanese Economic Review, 54*(2), 121-145. doi:10.1111/1468-5876.00250

Glaeser, E. L. (2014). A world of cities: the causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association, 12*(5), 1154-1199. doi:10.1111/jeea.12100

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy, 100*(6), 1126-1152. doi:10.1086/261856

Graham, D. J. (2009). Identifying urbanisation and localisation externalities in manufacturing and service industries. *Papers in Regional Science, 88*(1), 63-84. doi:10.1111/j.1435-5957.2008.00166.x

Henderson, J. V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth, 8*(1), 47-71. doi:10.1023/A:1022860800744

Henderson, J. V. (2010). Cities and Development. *Journal of Regional Science, 50*(1), 515-540. doi:10.1111/j.1467-9787.2009.00636.x

Hoselitz, B. F. (1953). The Role of Cities in the Economic Growth of Underdeveloped Countries. *Journal of Political Economy, 61*(3), 195-208. doi:10.1086/257371

Krugman, P. (1991). Increasing Returns and Economic-Geography. *Journal of Political Economy, 99*(3), 483-499. doi:10.1086/261763

Krugman, P., & Elizondo, R. L. (1996). Trade policy and the Third World metropolis. *Journal of Development Economics, 49*(1), 137-150. doi:10.1016/0304-3878(95)00055-0

Lewis, A. W. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School, 28*(2), 139-191. doi:10.1111/j.1467-9957.1954.tb00021.x

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22*(1), 3-42. doi:10.1016/0304-3932(88)90168-7

Marshall, A. (1890). *Principles of economics*. London, New York: Macmillan and Co.

Martin, P., & Ottaviano, G. I. P. (2001). Growth and agglomeration. *International Economic Review*, *42*(4), 947-968. doi:10.1111/1468-2354.00141

McCann, P., & Acs, Z. J. (2011). Globalization: Countries, Cities and Multinationals. *Regional Studies, 45*(1), 17-32. doi:10.1080/00343404.2010.505915

Melo, P. C., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics*, *39*(3), 332-342. doi:10.1016/j.regsciurbeco.2008.12.002

OECD. (2006). *Competitive Cities in the Global Economy*. Retrieved from Paris: http://www.oecd.org/gov/regional-policy/oecdterritorialreviewscompetitivecitiesintheglobaleconomy.htm

Rigby, D. L., & Brown, W. M. (2015). Who Benefits from Agglomeration? *Regional Studies, 49*(1), 28-43. doi:10.1080/00343404.2012.753141

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy, 94*(5), 1002-1037. doi:10.1086/261420

Rosenthal, S., & Strange, W. (2004). Chapter 49 - Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2119-2171). Burlington: Elsevier.

Sachs, J. D., Warner, A., Åslund, A., & Fischer, S. (1995). Economic Reform and the Process of Global Integration. *Brookings Papers on Economic Activity, 1995*(1), 1-118. doi:10.2307/2534573

Sassen, S. (1991). *The Global City*. Princeton, NJ: Princeton University Press.

Scott, A. J. (2001). *Global City-regions Trends, Theory, Policy*. Cambridge: Oxford University Press.

United Nations. (2014). World Urbanization Prospects: The 2014 Revision. Retrieved from http://esa.un.org/unpd/wup/

World Bank. (2009). *World Development Report 2009. Reshaping Economic Geography*. Washington, D.C.: Author.

Appendix 1 – Data sources

Variable	Source
City Size Categories	Percentage of urban population that lives in cities of a certain size calculated based on World Urbanization Prospects, the 2014 revision http://esa.un.org/unpd/wup/
GDPpc	Natural logarithm of "Real GDP at constant 2005 national prices "/ "Population" Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Years of Schooling	Years of schooling at beginning of the five-year period Barro & Lee dataset http://www.barrolee.com/
Private Investment	Five-year averages of "Share of gross capital formation at current PPPs" Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Openness	Five-year averages of sum of "Share of merchandise exports at current PPPs" and "Share of merchandise imports at current PPPs" Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Urban Population	Natural logarithm of the urban population at the beginning of the five-year period (in millions) Penn World Tables 8.0 http://www.rug.nl/research/ggdc/data/pwt/
Land Area	Natural logarithm of the land area in square kilometer World Development Indicators http://data.worldbank.org/indicator
Oil Dummy	Oil rents on average more than 15% of GDP per capita since data is available for a given country World Development Indicators http://data.worldbank.org/indicator
Services Exports	Service exports as % of total GDP Service exports in current US$: UNCTAD http://unctadstat.unctad.org/wds/TableViewer/tableView.aspx?ReportId=17648 GDP in current US$: World Development Indicators http://data.worldbank.org/indicator
High-Tech-Dummy	Dummy based on share of high-tech goods of overall manufacturing exports; equals 1 if larger than the 80th percentile of the overall observations, which is 16.2%. World Development Indicators http://data.worldbank.org/indicator
Sanitation Dummy	Dummy takes value of 1 if 90% or more of the population have access to sanitation. Access to sanitation is sources from World Development Indicators http://data.worldbank.org/indicator
Government Effectiveness	Dummy takes value of 1 if a country's average score for government effectiveness is 0.97 or larger (80^{th} percentile of countries) since data is available World Development Indicators http://data.worldbank.org/indicator

Appendix 2 – Descriptive statistics

		Standard Deviation					
Variable	Mean	Overall	Between	Within	Maximum	Minimum	Obs.
% Below 500K	61.36863	20.32682	18.54052	8.502813	100	0	652
% 500K to 1M	11.99106	14.9533	10.97207	10.53413	65.1741	0	652
% 1M to 3M	16.08785	18.9301	15.67486	10.75994	99.98664	0	652
% 3M to 7M	6.338234	14.21833	12.21128	7.047835	100	0	652
% 7M to 10M	1.873779	6.570864	4.359323	4.866689	40.38609	0	652
% 10M +	2.340507	8.529494	7.414783	4.034982	54.08545	0	652
Growth	.0689286	.1891929	.0852382	.1689796	1.199445	-1.259275	652
Ln(GDPpc)	8.424952	1.320437	1.289883	.2256611	10.99794	4.980256	652
Years of schooling	6.08204	3.191872	3.099505	.9629458	13.19	.28	652
Private investment	.1990553	.0830768	.0667662	.0493473	.6267554	.0159164	652
Openness	.5074964	.6697758	.5455402	.3805213	11.06161	.0060162	652
LN (Urban population)	1.778303	1.49284	1.467248	.2757214	6.303228	-1.720663	652
Ln (Land)	12.44327	1.753231	-	-	16.61218	6.507277	652
Service exports	.0702727	.0595708	.0547594	.0240528	.2861891	0	528

Appendix 3 – Full results sectoral composition and context

Dependent variable: GDPpc growth in five-year intervals from 1980 – 2010 – Sectoral Composition (1/2)

VARIABLES	(1) growth	(2) growth	(3) growth	(4) growth	(5) growth	(6) growth
Initial GDPpc	-0.321***	-0.300***	-0.303***	-0.305***	-0.298***	-0.302***
	(0.0272)	(0.0267)	(0.0275)	(0.0276)	(0.0283)	(0.0277)
Years of schooling	-0.00298	0.00194	0.000810	0.00134	0.00374	0.00295
	(0.0106)	(0.0107)	(0.0106)	(0.0108)	(0.0107)	(0.0106)
Private investment	0.592***	0.584***	0.602***	0.603***	0.582***	0.563***
	(0.0979)	(0.0980)	(0.101)	(0.102)	(0.0999)	(0.0996)
Openness	-0.00415	-0.00851	-0.00613	-0.00417	-0.00437	-0.00256
	(0.0115)	(0.0117)	(0.0118)	(0.0117)	(0.0118)	(0.0117)
Urban population	0.191***	0.0606*	0.0444	0.0513	0.0512	0.0526
	(0.0473)	(0.0323)	(0.0327)	(0.0327)	(0.0330)	(0.0326)
Land area	-0.0370	-0.0197	-0.00874	-0.00488	-0.00900	-0.0126
	(0.0246)	(0.0230)	(0.0233)	(0.0237)	(0.0230)	(0.0230)
Oil dummy	0.463***	0.409***	0.391***	0.410***	0.396***	0.410***
	(0.130)	(0.118)	(0.121)	(0.121)	(0.117)	(0.118)
Service exports	0.0844	0.153	0.439**	0.293	0.334	0.366*
	(0.491)	(0.221)	(0.216)	(0.212)	(0.211)	(0.210)
% Below 500K	0.00235***					
	(0.000675)					
* Service exports	0.00492					
	(0.00711)					
* Urban population	-0.00137***					
	(0.000424)					
% 500K to 1M		-0.00266***				
		(0.000730)				
* Service exports		0.0130**				
		(0.00596)				
* Urban population		0.00104				
		(0.000683)				
% 1M to 3M			-4.13e-05			
			(0.00101)			
* Service exports			-0.00611*			
			(0.00329)			
* Urban population			0.000732			
			(0.000604)			
% 3M to 7M				-0.00182		
				(0.00254)		
* Service exports				0.00764*		
				(0.00456)		
* Urban population				0.000195		
				(0.000933)		
% 7M to 10M					-0.00453	
					(0.00797)	
* Service exports					0.00166	
					(0.0193)	
* Urban population					0.00143	
					(0.00236)	
% 10M +						-0.00632
						(0.00867)
* Service exports						-0.0265
						(0.0217)
* Urban population						0.00261
						(0.00232)
Constant	2.442***	2.316***	2.178***	2.154***	2.145***	2.219***
	(0.311)	(0.292)	(0.295)	(0.296)	(0.292)	(0.295)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	528	528	528	528	528	528
Number of countries	88	88	88	88	88	88

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Note: Dummy for Soviet Union dropped because of collinearity

Dependent variable: GDPpc growth in five-year intervals from 1980 – 2010 – Sectoral Composition (2/2)

VARIABLES	(1) growth	(2) growth	(3) growth	(4) growth	(5) growth	(6) growth
Initial GDPpc	-0.505***	-0.482***	-0.495***	-0.492***	-0.488***	-0.490***
	(0.0281)	(0.0286)	(0.0286)	(0.0290)	(0.0294)	(0.0290)
Years of schooling	-0.00146	0.00115	-0.00335	-0.00154	0.00278	0.00170
	(0.0128)	(0.0131)	(0.0129)	(0.0131)	(0.0130)	(0.0130)
Private investment	0.808***	0.795***	0.839***	0.840***	0.794***	0.788***
	(0.120)	(0.124)	(0.124)	(0.126)	(0.124)	(0.124)
Openness	-0.00161	-0.00215	0.00189	-0.000873	0.000994	0.00338
	(0.0153)	(0.0159)	(0.0157)	(0.0158)	(0.0158)	(0.0158)
Urban population	0.0783	-0.112***	-0.127***	-0.122***	-0.119***	-0.114***
	(0.0575)	(0.0385)	(0.0386)	(0.0385)	(0.0388)	(0.0384)
Land area	0.0348	0.0749**	0.0770**	0.0951***	0.0853**	0.0789**
	(0.0340)	(0.0331)	(0.0338)	(0.0342)	(0.0335)	(0.0332)
Oil dummy	0.477***	0.426***	0.433***	0.443***	0.434***	0.450***
	(0.160)	(0.157)	(0.164)	(0.164)	(0.161)	(0.158)
Soviet union dummy	-0.214	-0.282*	-0.288	-0.308*	-0.310*	-0.293*
	(0.173)	(0.170)	(0.177)	(0.177)	(0.174)	(0.171)
High-tech dummy	-0.0611	0.0552	0.0918**	0.0536	0.0552	0.0575
	(0.108)	(0.0449)	(0.0446)	(0.0406)	(0.0403)	(0.0413)
% Below 500K	0.00425***					
	(0.000696)					
* High-tech dummy	0.00166					
	(0.00156)					
* Urban population	-0.00165***					
	(0.000540)					
% 500K to 1M		-0.00190***				
		(0.000705)				
* High-tech dummy		-0.000408				
		(0.00173)				
* Urban population		0.00139				
		(0.000855)				
% 1M to 3M			-0.00318***			
			(0.000916)			
* High-tech Dummy			-0.00173			
			(0.00112)			
* Urban Population			0.00226***			
			(0.000681)			
% 3M to 7M				-0.000628		
				(0.00291)		
* High-tech dummy				0.00349**		
				(0.00177)		
* Urban population				-0.000366		
				(0.00109)		
% 7M to 10M					-0.0131	
					(0.0101)	
* High-tech dummy					-0.000404	
					(0.00269)	
* Urban population					0.00402	
					(0.00307)	
% 10M +						-0.0184*
						(0.0110)
* High-tech dummy						-0.00142
						(0.00279)
* Urban population						0.00565*
						(0.00309)
Constant	2.949***	2.675***	2.759***	2.503***	2.578***	2.665***
	(0.438)	(0.434)	(0.445)	(0.444)	(0.441)	(0.439)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: GDPpc growth in five-year intervals from 1980 – 2010 – Contextual factors (1/2)

VARIABLES	(1) growth	(2) growth	(3) growth	(4) growth	(5) growth	(6) growth
Initial GDPpc	-0.505***	-0.476***	-0.486***	-0.483***	-0.483***	-0.484***
	(0.0280)	(0.0285)	(0.0286)	(0.0288)	(0.0293)	(0.0289)
Years of schooling	-0.00243	0.00158	0.000750	0.00231	0.00399	0.00288
	(0.0127)	(0.0131)	(0.0130)	(0.0130)	(0.0131)	(0.0131)
Private investment	0.809***	0.787***	0.798***	0.836***	0.784***	0.785***
	(0.120)	(0.124)	(0.124)	(0.127)	(0.125)	(0.124)
Openness	-0.00168	-0.00158	0.00510	0.000119	0.00152	0.00342
	(0.0153)	(0.0159)	(0.0157)	(0.0158)	(0.0159)	(0.0159)
Urban population	0.0511	-0.115***	-0.122***	-0.119***	-0.121***	-0.118***
	(0.0570)	(0.0386)	(0.0387)	(0.0386)	(0.0389)	(0.0387)
Land area	0.0405	0.0761**	0.0750**	0.0899***	0.0862**	0.0814**
	(0.0339)	(0.0331)	(0.0337)	(0.0337)	(0.0335)	(0.0334)
Oil dummy	0.496***	0.407***	0.421***	0.421***	0.418***	0.432***
	(0.163)	(0.157)	(0.163)	(0.160)	(0.162)	(0.159)
Soviet union dummy	-0.196	-0.302*	-0.312*	-0.327*	-0.334*	-0.320*
	(0.176)	(0.170)	(0.175)	(0.173)	(0.174)	(0.172)
Sanitation	0.157	0.0269	-0.00649	0.0241	0.0208	0.0254
	(0.104)	(0.0877)	(0.0880)	(0.0875)	(0.0878)	(0.0875)
% Below 500K	0.00448***					
	(0.000696)					
* Sanitation	-0.00350**					
	(0.00141)					
* Urban population	-0.00135**					
	(0.000540)					
% 500K to 1M		-0.00196***				
		(0.000713)				
* Sanitation		-0.000310				
		(0.00138)				
* Urban population		0.00150*				
		(0.000858)				
% 1M to 3M			-0.00369***			
			(0.000933)			
* Sanitation			0.000605			
			(0.00102)			
* Urban population			0.00224***			
			(0.000687)			
% 3M to 7M				-0.00165		
				(0.00323)		
* Sanitation				0.00361*		
				(0.00190)		
* Urban population				-7.66e-05		
				(0.00115)		
% 7M to 10M					-0.0134	
					(0.0103)	
* Sanitation					-0.000592	
					(0.00239)	
* Urban population					0.00417	
					(0.00303)	
% 10M +						-0.0155
						(0.0130)
* Sanitation						-0.00196
						(0.00331)
* Urban population						0.00511
						(0.00330)
Constant	2.881***	2.620***	2.706***	2.484***	2.529***	2.594***
	(0.439)	(0.435)	(0.443)	(0.439)	(0.441)	(0.440)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: GDPpc growth in five-year intervals from 1980 – 2010 – Contextual factors (2/2)

VARIABLES	(1) growth	(2) growth	(3) growth	(4) growth	(5) growth	(6) growth
Initial GDPpc	-0.508***	-0.480***	-0.494***	-0.492***	-0.490***	-0.490***
	(0.0282)	(0.0286)	(0.0286)	(0.0292)	(0.0296)	(0.0291)
Years of schooling	-0.00325	-0.00108	-0.00271	0.00165	0.00259	0.00162
	(0.0127)	(0.0131)	(0.0130)	(0.0130)	(0.0130)	(0.0130)
Private investment	0.804***	0.777***	0.828***	0.855***	0.791***	0.788***
	(0.121)	(0.124)	(0.124)	(0.128)	(0.125)	(0.125)
Openness	-0.00199	-0.00345	0.00165	-0.00131	0.000988	0.00311
	(0.0154)	(0.0159)	(0.0158)	(0.0159)	(0.0159)	(0.0159)
Urban population	0.0526	-0.127***	-0.142***	-0.131***	-0.132***	-0.126***
	(0.0580)	(0.0389)	(0.0394)	(0.0390)	(0.0392)	(0.0389)
Land area	0.0463	0.0829***	0.0818***	0.0940***	0.0899***	0.0832***
	(0.0311)	(0.0309)	(0.0312)	(0.0316)	(0.0311)	(0.0308)
Oil dummy	0.636***	0.553***	0.566***	0.583***	0.573***	0.585***
	(0.156)	(0.155)	(0.161)	(0.162)	(0.160)	(0.157)
Soviet union dummy	0.0802	0.00121	-0.00302	-0.0113	-0.0179	-0.00292
	(0.214)	(0.215)	(0.223)	(0.223)	(0.220)	(0.216)
Government effectiveness	0.982**	0.768*	0.847**	0.829*	0.827*	0.815*
	(0.464)	(0.414)	(0.430)	(0.430)	(0.425)	(0.417)
% Below 500K	0.00438***					
	(0.000702)					
* Government effectiveness	-0.00294					
	(0.00310)					
* Urban population	-0.00149***					
	(0.000547)					
% 500K to 1M		-0.00191***				
		(0.000701)				
* Government effectiveness		0.00262				
		(0.00209)				
* Urban population		0.00100				
		(0.000914)				
% 1M to 3M			-0.00321***			
			(0.000920)			
* Government effectiveness			-0.00158			
			(0.00114)			
* Urban population			0.00237***			
			(0.000687)			
% 3M to 7M				-0.00229		
				(0.00340)		
* Government effectiveness				0.00412**		
				(0.00205)		
* Urban population				0.000251		
				(0.00124)		
% 7M to 10M					-0.0146	
					(0.0101)	
* Government effectiveness					-0.00222	
					(0.00352)	
* Urban population					0.00453	
					(0.00305)	
% 10M +						-0.0196*
						(0.0110)
* Government effectiveness						-0.00128
						(0.00373)
* Urban population						0.00593*
						(0.00305)
Constant	2.832***	2.589***	2.691***	2.501***	2.541***	2.616***
	(0.399)	(0.402)	(0.409)	(0.411)	(0.408)	(0.405)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	652	652	652	652	652	652
Number of countries	113	113	113	113	113	113

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4 Special Economic Zones: What Makes Them Truly Special? A quantitative analysis of the drivers of SEZ performance

4.1 Introduction

“Special Economic Zones” (SEZs) have been on the policy agenda for a considerable amount of time. The usually stated objectives of SEZs are not only centred around the aim of transforming local economies, by attracting inward investment and creating new jobs, but also by scaling the technology and know-how ladder. Developing countries have been particularly active on this front in recent years. SEZs have been promoted there with the intention of boosting exports, diversifying the economy, and generating direct and indirect jobs. Developed economies have also resorted to SEZs as a way to foster economic development in their lagging regions. The early success of some SEZs in parts of the developed world, as well as some early cases in China and the Asian Tigers has contributed to enhance the appeal of SEZs amongst policymakers in less developed regions and countries as a development tool.

The popularity and importance of SEZs has mainly rocketed in the last two decades. While there were 176 zones in 47 countries in 1986, the International Labour Office (ILO) database registered 3,500 in 130 countries in 2006 (Singa Boyenge, 2007). The Foreign Investment Advisory Service (2009) estimates that, in the mid-2000s, SEZs accounted for almost 20% of exports and employed more than 60 million people in developing countries.

Most SEZs share a number of features: 1) they generally have a regulatory and incentive framework that is different from the rest of the country; 2) zones tend to provide dedicated infrastructure services; and 3) their area of activity is clearly delineated by physical boundaries (Asian Development Bank, 2015; FIAS, 2009; World Bank, 2011). However, zones differ greatly in the application of these features, meaning that a wide range of different types of SEZs have emerged in different parts of the world over the years. Even within countries, it is not infrequent for different forms of SEZs to coexist, each displaying a different mix of incentive schemes, services, industries, and target markets.

Despite often overstated claims about the impact of the zones and the wide diversity of economic zone policies, there is limited empirical evidence that systematically analyses how differences in the set-up of the zones impact on their performance. Hence, our knowledge as to which types of SEZs and which incentive schemes have been more successful in contributing to further the goals of the zones remains highly imperfect and partial.

Mainly as a consequence of the limited availability of cross country data to measure SEZs' outcomes and characteristics, most of the literature which has delved into the analysis of the impact of SEZs, has adopted a case study approach (see for example Asian Development Bank, 2015; Engman, Onodera, & Pinali, 2007; Nel & Rogerson, 2013; World Bank, 2011). Many of these cases represent solid analyses of the economic dynamism and influence of individual zones and provide interesting insights about their viability and the characteristics that make them successful. However, the case-study nature of such analyses is also not without problems. More often than not research has focused on the most successful cases, raising questions about the capacity to generalize the factors behind the success of a specific SEZ across economic, social, political, and legal contexts that often differ widely from those that have contributed to make a particular case successful. Replicating policy and incentive models is also tricky. As a consequence and despite providing very interesting policy insights, extracting wide-ranging policy implications from this type of analyses remains risky.

The purpose of this paper is to overcome this gap in our knowledge, by shedding more light on the drivers of successful SEZ policies across countries from a comparative perspective. In order to do so, we rely on two new datasets, which were developed in collaboration with the World Bank's Trade and Competitiveness Practice Department.[26] The first one maps SEZs in 22 different countries – mostly in the emerging world – assessing the characteristics of the zones, the incentives and enticements provided either at the zone or national level for the establishment of SEZs, as well as the socio-economic and institutional characteristics of the regions and countries in which a zone is located. The second dataset uses nightlights data as

[26] A special thanks to the WB team, which helped to assemble the data set, in particular to Egle Pugaciauskaite, Elliot Rasmuson, Keith Garett and Le Duy Binh.

SEZ performance proxy to overcome the lack of reliable economic indicators when measuring SEZ performance.

The paper is structured as follows. The next section provides an overview of the existing literature on SEZ performance determinants. Section 3 introduces the SEZ dataset and discusses the use of nightlights data to proxy SEZ performance. Section 4 lays out the methodology of the econometric analysis. A first descriptive account and the results of the econometric analysis examining the factors behind SEZ performance are presented in Section 5 and 6. The final section summarizes the conclusions and policy implications.

4.2 SEZ performance drivers

SEZ programmes and zones differ along many dimensions and depend on the characteristics of the country and regional context in which they are located. Factors internal and external to the SEZ programme and zones are likely to affect a zone's ability to attract investors, create employment, and facilitate firm performance and economic growth. Three sets of factors can be highlighted in this regard.

The first set of factors linked to SEZ performance is related to the set-up and design of the overall **SEZ programme**. This includes the incentives package, requirements imposed on firms to benefit from the incentives as well as the organizational set up of the programme itself. Traditionally, fiscal incentives have been at the core of any SEZ policy. The underlying reasoning is to provide companies with an advantageous cost-reducing, fiscal environment. Fiscal incentives vary from country to country and from zone to zone, but frequently include a mix of exemptions from import duties on machinery and inputs to reductions or exemptions on corporate and other local taxes. Many programmes also offer subsidized utilities to companies, either through VAT exemptions or explicit subsidies (Asian Development Bank, 2015). Studies have come to differing conclusions about the effectiveness of these tax-breaks. While Rolfe, Woodward, and Kagira (2004) and Aggarwal (2005) underline the importance of the incentive package from an investor's point of view, Farole (2011) – in one of the few attempts to quantitatively assess SEZ performance drivers – does not find any correlation between the tax holidays offered to companies and zone success in terms of employment generation and exports. Similarly, a recent report by the Asian Development Bank (2015) concludes that while many countries feel the need to offer tax incentives, their effectiveness may be limited and well below those of other pull-factors. In the worst-case scenario, tax exemptions, subsidies and other incentive packages may rear a rent-seeking behaviour by firms in the zone, undermining the entire viability of the SEZs scheme (Rodríguez-Pose & Arbix, 2001; World Bank, 2011).

Exemptions from national labour regulations and the facilitation of administrative services through national one-stop-shops is another popular way of providing non-fiscal benefits to companies (Asian Development Bank, 2015; OECD, 2009). While the reduction of labour protection is frequently seen with concern with regards to its social impacts (Jauch, 2002), several authors claim that more flexible labour regulations have contributed to the success of many SEZ policies (Aggarwal, 2005; Madani, 1999; Watson, 2001). Administrative facilitation through one-stop-shops, by contrast, is generally approved of and considered best practice by many international institutions (Asian Development Bank, 2015; Farole & Kweka, 2011).

Programmes also differ in the requirements needed by companies in order to benefit from the incentive packages. As the aim of many programmes is to attract foreign direct investment, some programmes specifically target foreign companies, meaning that often only firms either partially or fully owned by foreign investors benefit from the incentives schemes. Similarly, as the aim of many policies is to increase a country's export performance, some policies impose a minimum level of exports. Finally, certain programmes also require minimum investment or minimum of employment thresholds in order for the company to access the tax breaks.

Despite the proliferation of tax breaks and holidays, incentives, and subsidies that inevitably accompany the formation and development of a SEZ, few studies have looked into the question if and how these incentives and requirements may impact zone performance. Many of those studies focus on specific incentives and not necessarily on the entire package, when providing policy recommendations. The OECD (2009), for instance, advocates the need to remove minimum export requirements in order to avoid a bias against local firms and to ensure compliance with WTO regulations. But even in this case, it does however not say much about how such a measure would impact on SEZ performance.

The organizational set-up of the SEZ programme has also been linked to the success of the policies. An independent zone regulator – expected to be shielded from political pressures as well as equipped with sufficient resources – is commonly considered to facilitate an efficient overview of SEZ programme development and implementation (Farole & Kweka, 2011; OECD, 2009). As a consequence, independent regulators may lead to better economic outcomes at the zone level. Furthermore, the time and period in the cycle of the establishment of the programmes may influence zone performance. Some studies have pointed to a first mover advantage. Countries that moved early in order to establish SEZ programmes are often regarded as more likely to have been successful than late comers (Watson, 2001). Start-up periods for successful policy implementation are also considered to matter in the success of the zones. The experience a country has in the implementation of the SEZ policies may also contribute to the effectiveness of the zone programme.

The second set of factors is made of **SEZ characteristics**, that is, characteristics that are exclusively related to the structure and layout of the zone. SEZ characteristics are generally linked to the dimension of the zone, the sectors targeted, its location as well as to the services and infrastructure provided within the zone. In recent years, there has been a shift in the literature and among policy-makers to highlighting the importance of these factors as opposed to purely relying on the incentive package provided in the overall SEZ programme (UNCTAD, 2015). Furthermore and in contrast to contextual factors, zone characteristics can be influenced and/or modified relatively easily. Hence, it is reasonable to expect that the SEZ-specific characteristics will affect the economic performance of the zone.

Among the most prominent SEZ characteristics is the maturity of the zone. Several case studies have highlighted the challenge for zones to kick-start growth as well as to maintain it after the initial years of success, as competition from other countries for FDI increases (Henderson & Phillips, 2007; World Bank, 2011). The technological content of the zone is another factor that may make a difference for economic success. Many zones in less developed areas have increasingly aimed to attract investors in the high-tech sector as opposed to the low-tech manufacturing on which many initially successful zones relied (Asian Development Bank, 2015). High-tech zones are regarded as a faster and more illustrious way to achieve employment creation and economic growth than low-tech, low cost and often massive production zones. However, questions have been raised about the viability of high-tech zones in less-developed environments, as they are not always very successful (Luger & Goldstein, 1991; Quintas, Wield, & Massey, 1992). The nature of the operator has also been identified as a success driver with best practice guides frequently emphasizing the advantages of private operators over publicly run zones (Farole & Kweka, 2011; OECD, 2009; Watson, 2001). Farole (2011) does however not find any correlation between the type of zone operator and SEZ performance.

An important question also concerns the location choice. SEZ policies frequently have an explicit spatial aspect, i.e. they aim to promote the economic development of certain regions. At the same time, a strategic location close to ports, consumer markets, and the labour pool are elements many firms actively consider when deciding on location (Aggarwal, 2005). Several studies have stated that closeness to ports or large cities is more likely to spur zone dynamism than locating a SEZ in more remote areas (Asian Development Bank, 2015; Madani, 1999).

The type of services provided within the zones may also affect the economic dynamism of the zone factor. Traditionally many zones have provided services aimed at easing infrastructural and other challenges in the country. These services range from the existence of a dedicated customs-office to, among others, the provision of more reliable utilities – electricity supplies in particular. Increasingly, zones also offer other, “softer” services such as human resources, restaurants, housing services and one-stop-shop facilities onsite to deal with administrative processes for the companies within the zone (Farole, 2011; World Bank, 2011).

Finally, the **regional and country context** in which the SEZs are located also matter for the success of the zone. While the aim of many SEZ programmes is to overcome the challenges that companies face in the local environment, SEZs do not operate in a void and are likely to be heavily influenced by the socioeconomic characteristic, market potential and general business climate of the host country. The country and regional context, in which a SEZ operates, have therefore been highlighted as key for a successful SEZ policy implementation (Farole, 2011). A number of authors stress the importance of the national investment environment and institutions for FDI (Daude & Stein, 2007; Portugal-Perez & Wilson, 2012) and thus the success of SEZs. Aggarwal (2005) and Farole (2011) specifically emphasize a strong correlation between SEZ outcomes and the general business climate. Moreover, the attractiveness of a host country is enhanced/diminished by its proximity/distance and access (or lack of it) to large markets (Madani, 1999; Rolfe et al., 2004; Watson, 2001) as well as by its industrial structure. Trade between countries decreases as distance and trade costs increase (Disdier & Head, 2008). Hence, proximity to a large national or international market is an attractive feature for efficiency seeking investors. Favourable national industrial structures with a solid pre-existing manufacturing base also increase a host country’s attractiveness (Hidalgo & Hausmann, 2009). Economies, primarily reliant on agricultural production, will in all likelihood, have a more difficult time convincing investors of their capabilities to produce manufacturing goods at a large scale than countries with a pre-existing industry base. Finally, a country’s overall socio-economic context may be an important stimulus/deterrent for investors. Efficiency seeking investors in labour intensive sectors require a sufficiently large and cheap workforce and are therefore prone to look for cheaper locations with an abundant supply of labour. Human capital endowments, affecting productivity, are also assumed to play a role in making places more or less attractive for firms, in particular in the process of upgrading to higher value added products (Larrain, Lopez-Calva, & Rodriguez-Clare, 2009; World Bank, 2011).

As this overview shows, a large amount of factors, both internal and external to the zones and to SEZ policies, are on the table as potential drivers of zone performance. While much has already been written about the impact of these factors from a case study approach, a more systematic quantitative analysis of whether these factors apply universally is still missing.

4.3 Gathering data on SEZ characteristics and performance

In order to identify which of the above factors has a systematic bearing on SEZ performance, a dataset is needed that reflects both zone-specific characteristics as well as the diversity of programme related factors, such as legislative provisions, institutional frameworks and fiscal incentives. For this purpose, a new SEZ dataset was assembled in collaboration with the World Bank's Trade and Competitiveness Practice. Furthermore, to overcome the lack of comparable SEZ performance data, we resort to nightlights data as a SEZ performance proxy. The following section provides details about how the data was collected and the resulting database, before turning to the econometric analysis.

4.3.1 SEZ database

In order to collect data about SEZ characteristics, first, a workable definition of SEZs has to be established. The definition has to embrace not only conceptual but also practical considerations, linked to the suitability of the zone for the use of nightlights data as a proxy for its performance as well as data availability. Consequently, all zones included in the dataset fulfil the following five criteria:

- A *differentiating regulatory framework and/or incentive scheme* for the SEZ is the essential differentiator in order to define what constitutes a SEZ. This is in line with most literature and allows establishing the all-important distinction between SEZs and other types of industrial parks.

- A *focus on manufacturing or services* within the zone with the objective of singling out and eliminating zones that are primarily commercial and logistics hubs.
- The presence of *clear territorial boundaries* in order to be able to better delimit performance using nightlight data. This implies that some SEZ schemes, such as single factory zones or large wide zones, are excluded from the analysis.
- A *minimum size of 50ha* in order to increase the reliability of the nightlights measurement as a proxy for zone performance. This is determined by the size of the grid-cells in the nightlights dataset; the data is furthermore restricted to zones with a *maximum size of 1000ha* to ensure better comparability of the zones.
- The SEZs had to be *operational by the year 2007*, meaning that at least one company had started operations within the SEZ by then. This criterion ensures that a reasonable variation in the nightlights can be detected between start of operations and 2012, which is the last year for which nightlight data is available.

For the selection of the countries covered in the database, a number of factors were considered such as geography, income levels, and maturity of zone programmes. The objective was to allow for a considerable variation in SEZ experiences to be represented in the sample. The selection was furthermore guided by more practical considerations regarding data availability for a given country, time of establishment of the SEZ policy as well as type and number of SEZs.

In each targeted country, the aim was to identify the entire population of zones and filter them based on their fulfilment of the five criteria. For each of the qualifying zones, information was collected for SEZ-specific and programme-related variables using a variety of sources. These included information available online from SEZ and public authority homepages, reports from international organization and related sources. We, furthermore, reached out to SEZ authorities and zones over email and phone to verify and complement the data collected online.

Table 4-1: Overview SEZs per country

Countries	Number of zones
East Asia & Pacific	255 (73%)
China	33
Philippines	29
Malaysia	6
South Korea	64
Thailand	20
Vietnam	103
Europe & Central Asia	40 (10%)
Turkey	36
Russia	4
Middle East and North Africa & Sub-Saharan Africa	6 (2%)
Ghana	1
Jordan	1
Kenya	1
Lesotho	1
Nigeria	1
South Africa	1
Latin America & Caribbean	26 (7.5%)
Argentina	4
Chile	3
Colombia	6
Dominican Republic	10
Honduras	3
South Asia	19 (5%)
Bangladesh	8
India	8
Pakistan	3
Total	346 (100%)

The resulting sample includes 346 zones in 22 countries across the developing world and South Korea. Table 4-1 provides an overview of the resulting country coverage and number of zones per country. As indicated in the table, the sample covers countries from all over the developing world. It is, however, biased towards countries in the East Asia and Pacific region. On the one hand, this reflects the strong proliferation of SEZ policies in this region. On the other hand, this is driven by the fact that many Latin American zones do not fulfil the size requirements (e.g. out of the more than 60 zones in the Dominican Republic, only 10 have the required size to be included in the sample). Furthermore, many countries introduced their zone programme only recently and thus have a more limited number of zones which fulfil the time criterion.

Table 4-2: Characteristics of SEZs included in the dataset

Period of establishment	
Before 1990	61 (18%)
1990 to 1999	105 (30%)
Since 2000	180 (52%)
Average size	
Below 100ha	70 (20%)
Between 100ha and 200ha	130 (38%)
Between 200ha and 500ha	113 (33%)
Above 500ha	33 (9%)
Sector focus	
Manufacturing	241 (70%)
Services	1 (0.3%)
Mixed	104 (30%)
Technology intensity of industry	
Low and medium technology	274 (79%)
High-technology	72 (21%)
Zone operator	
Public	142 (41%)
PPP	116 (34%)
Private	85 (25%)

Table 4-2 provides an overview of some key characteristics of the SEZs, i.e. the time of establishment of the zones, the sector focus, technology intensity and size. The majority of zones have become operational since the year 2000 (52%), 30% in the nineties and 18% before 1990. This reflects their increasing popularity as a policy tool. There is a wide variety of zones according to size: 20% of zones are smaller than 100ha, 38% range between 100ha and 200ha, 33% between 200ha and 500ha, and the remainder (9%) is above 500ha. The largest zone included is 998ha and the smallest 51ha. In terms of the sector, 70% of the total are fundamentally manufacturing zones, 30% are mixed and one zone is purely service focused. The near absence of service oriented zones is due to the fact that they tend to be much smaller in area and thus fall through the filter. Approximately 21% of zones used in the analysis contain a sectoral focus on high-technology manufacturing.

The type of zone operator is distributed between public, private, or public-private partnership (PPP), depending on the set-up of the management company. 41% of all zones are entirely publicly managed, while 25% are privately run. 34% are PPPs, in which both the private and the public sector are involved.

4.3.2 SEZ performance proxy

Ideally, SEZ performance should be measured using indicators such as FDI inflows into the zones, exports from the zone and employment generation (both direct and indirect). However, as mentioned in the introduction, the lack of such data for a large amount of SEZs and countries requires an alternative approach. The use of night-time light data as a proxy for the economic performance of an individual SEZ provides a suitable and increasingly popular alternative to those cases where direct economic data are not readily available (Chen & Nordhaus, 2011; Elvidge et al., 2007; Ghosh et al., 2010; Henderson, Storeygard, & Weil, 2012).

Stemming from the field of remote-sensing, economists and other social scientists have increasingly resorted to light data as a proxy for economic activity (Ebener, C., Tandon, & Elvidge, 2005; Elvidge, Baugh, Anderson, Sutton, & Ghosh, 2012; Florida, Gulden, & Mellander, 2008; Henderson et al., 2012). Nightlight data is available from the Defense Meteorological Satellite Program (DMSP) for the period 1992 – 2012. The dataset provides the average luminosity created by human activity going from 0 to 63 in cells of roughly one square kilometre, covering the majority of the world's land area. Given the small size of the grid cells, reliable measurements can be obtained for almost any geographical area. Mellander, Lobo, Stolarick, and Matheson (2015) demonstrate that the correlation between the luminosity and alternative data for economic activity is high even at a very small scale. They use data on employment and number of firms from the Swedish Statistics Bureau, which is geocoded in cells of 250m x 250m, and compare it to the nightlights data. They find a high correlation between the two. Similarly, Levin and Duke (2012) conclude in a study for Israel that the data is highly appropriate to proxy the extent of human activity at a small scale.

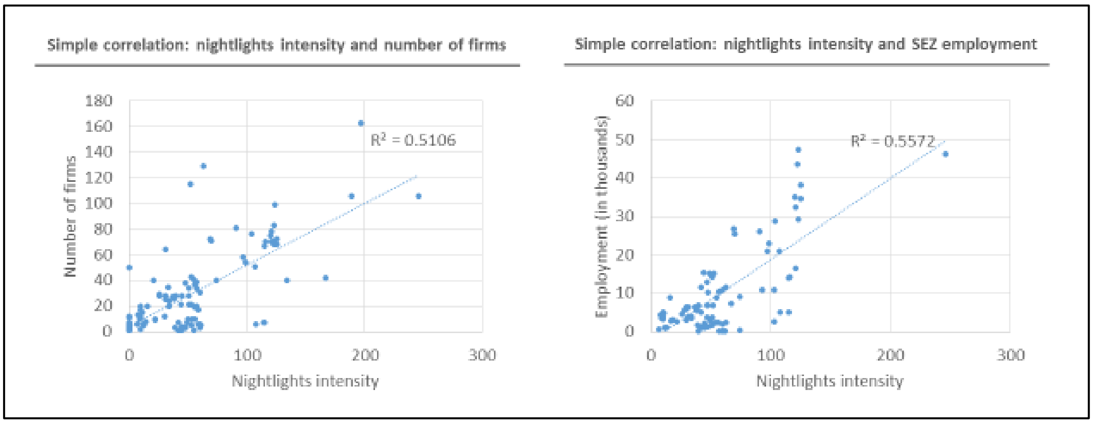
However, the use of night-time lights data is not without caveats. Recent critics advise caution on the use of this type of indicator as a proxy for GDP. In particular, while luminosity could be considered, as we have seen, a good proxy for GDP when employed at country level, the risks increase when artificial night-time light data is considered at the subnational level. As indicated by Keola, Andersson, and Hall (2015), studies resorting to nightlights data to calculate economic activity have a tendency "to underestimate economic activities that emit less or no additional night-time light as they grow". This is particularly problematic for areas with a high dependence on agricultural activities. For cases such as Burundi, Cambodia and Laos, this would lead to an overestimation of regional inequalities: lights in these countries are concentrated in the capital cities, but the agricultural and mining activity which supports more peripheral regions is not captured by luminosity data (Keola et al., 2015). Given that most SEZs in our sample (see Table 4-2) either focus on manufacturing or are mixed, our estimates should be less affected by this. Differences between different manufacturing sectors are however, while probably less pronounced, still plausible; hence, any results of the analyses will have to be considered in light of these caveats.

But how reliable are night-time light data when measuring the economic performance of SEZs? In order to use the nightlights data as an SEZ performance proxy, we calculate the luminosity for each zone. For this purpose, we identify the size, location and centroid of each SEZ using google maps satellite imagery as well as online sources from the national SEZ authorities and the SEZ homepages. We then draw a circle around the centroid as a proxy for the area of the SEZ. The night-lights within the area of the circle are assumed to reflect the economic activity within the zone.

To determine the length of the radius used to draw the circle, it is assumed that the SEZ has a squared shape identical to the overall surface of the SEZ. A circle is then drawn around the centroid which touches each corner of the square. The resulting surface of the circle is consequently slightly larger than the actual zone. Experiments with other radius lengths were conducted, but it was found that this method provided the best fit with the alternative performance measurements. This circle is then overlaid onto the nightlights raster file. If the circumference of the circle (which represents the zone) intersects or passes beyond the centroid of the night-lights cell (such that the centre of the pixel is within the buffer), the pixel is included in that count. If the buffer covers more than one nightlights cell, the values of both cells are added up.

To test the suitability of the nightlights as a SEZ performance proxy, we collect data on the number of companies and employment for a number of SEZs, as an alternative performance measure. Figure 4-1 shows a simple scatterplot between the number of firms and employment and the nightlights proxy. Both graphs display a clear positive association between the alternative measures and our proxy.

Figure 4-1: Correlation between nightlights, number of firms and SEZ employment



To further test the fit, we run two simple regressions, in which the number of firms and SEZ employment are the dependent variables and our nightlights proxy the explanatory variable, controlling for country fixed effects. In both cases, the nightlights are a highly significant predictor of the number of firms and the employment within the zone (see appendix 1 for the regression results). It can therefore be concluded that, in spite of the caveats, nightlights represent a good proxy for the economic performance of a zone.

It is, however, fair to state that while our proxy on average is a good predictor of SEZ performance, there is some spread around the trend line. To identify potential sources of this heterogeneity in the fit of the night-light as a proxy for SEZ performance, the satellite images of the outliers visible in the scatterplot were inspected. Location in densely populated areas, next to large highways and/or directly on the coastline were the factors that affected the accuracy of the proxy. Zones located in densely populated areas or next to highways were reflecting a higher amount of lights from outside the zone. This is in line with Levin and Duke (2012), who find that a significant amount of the lights reflected in the nightlights imagery stems from streets. In order to minimize this reflection, the level of population density around the zone (on a scale from 1 to 3) is identified for each zone. Information as to whether specific SEZs are located next to a large water body or a highway is also recorded. These three factors are included in the regressions as structural nightlights controls (see following section).

4.4 Methodology

Equipped with these two datasets, we operationalize the three sets of SEZ performance determinants described in the literature section of this paper using the following simple econometric model:

$$\Delta y_{i,t} = \alpha_1 + \beta_1 SEZ\ related\ factors_{i,t0} + \ \beta_2\, SEZ\ \text{programme factors}_{i,t0}$$
$$+ \beta_3 Country/regional\ level\ endowments_{i,t0}$$
$$+ \beta_4\ Structural\ nightlights\ controls + \ \epsilon_i$$

where

- Δy_{it} is the dependent variable, a measure of success of an individual SEZ (*i*) at time *t*,
- *SEZ-related factors*: characterizing the dimension of the zone, location, the type of sectors targeted, and services provided within the zones. These are zone specific variables;
- *SEZ-regulatory variables*: linked to the incentives offered, the requirements imposed, and the organizational set-up of the programme. These variables are either national level variables or SEZ specific, in those cases where multiple SEZ regimes exist within a country;

- *Country/ regional level endowments:* reflecting economic, social, political, and institutional factors at the country and regional level that may impact SEZ performance as well as proximity to markets;

- *Structural nightlights controls:* factors to improve the fit of nightlights as a SEZ performance proxy (see previous section)

- ϵ_i is the robust standard error clustered at the within-country region level

In order to maximize the number of zones included in the analysis – especially taking into account that the SEZ phenomenon has really taken off in emerging countries in the last few years – the main period of analysis covers the years between 2007 and 2012 for which all variables are available. To add further nuance to these findings, we also present two complementary sets of results. First, we run regressions on the same cross country dataset. In this section, however, we look at each zone's growth performance during the first 5 years after the zone became operational and not during the fixed period from 2007 to 2012. The aim of this exercise is to uncover the factors that facilitate the success of SEZs during the initial years of operation, regardless of when they were founded. This implies that the period of analysis covers the first five years in the life of a zone and varies by zone. The sample for this analysis is reduced as data is not available for every zone in the sample in the period immediately after their establishment. The reduction in the dataset fundamentally concerns older zones. Second, we present the results of a 'deep dive' into the performance of the Vietnamese zones. Taking this within country perspective allows to focus on zone characteristics, since both the contextual environment and policies are the same for all zones within the country.

4.4.1 SEZ performance

The dependent variable, Δyit, is the proxy for SEZ performance based on nightlights data as described in the previous section. We use two variations in the different sections: (1) the growth rate of the nightlights emitted from the SEZ in the period of analysis and (2) the ratio of the change of the nightlight emissions within the zone compared to the change in nightlights in the entire country. While the first indicator, the growth rate of nightlights in the zone, provides an indication of absolute growth and is our main dependent variable, the second indicator is a relative performance measure and captures whether a zone has grown faster than the national average. This allows to tease out differences in national growth across countries, as less dynamic zones in rapidly growing countries may often have – as a consequence of the overall dynamism of the country – higher rates of growth than very dynamic zones in low growth countries. This relative indicator is expected to reflect better the capacity of the SEZs to act as a motor of national growth within a country and is used as a robustness check in the main regressions.

4.4.2 SEZ variables

A number of key characteristics of the zone are taken into account in order to determine if zone-related factors can make a difference for SEZ performance. First, the size of the zone is included to test if there are potential differences depending on zone extension. Years operating represents the number of years the zone had been in operation in 2007, which helps to understand if zone performance is affected by zone maturity. A dummy for high-technology is included if the zone focuses on attracting firms in the high-tech sector.[27] To understand whether the type of zone management makes a structural difference for zone performance, the analysis considers the nature of management, distinguishing between whether a zone is operated by the public sector, as public-private partnership (PPP), or private entity.[28] The attractiveness of the location of the zone is measured using the road distance to the largest city.[29] And finally to reflect a zone's infrastructure, two dummies capture if a zone offers a one-stop-shop onsite and if it has its own power sub-station to ensure a reliable electricity supply. These last variables are only included in the deep dive section for Vietnam. A detailed

[27] The dummy takes the value of 1, if the zone either 'self-proclaims' on their advertising material that it specifically targets high-tech sectors or if companies established are in high-tech sectors, as defined by OECD.
[28] This indicator also takes into account the development stage of the zone. For instance, if a zone was developed by a public entity, but is operated privately, the indicators reflect this as PPP.
[29] We also experimented with the road distance to the closest city with at least 500K and 300K inhabitants as well as the road distance to the closest major port. As both indicators yielded insignificant results, they are not presented here.

list of variables for each zone is available in the Appendix 2 of this report. The information is sourced from the dataset described in the previous section.

4.4.3 SEZ programme variables

To capture the diversity of different SEZ policies, as described in the conceptual framework, the analysis includes information on the incentive package provided to companies, requirements imposed on the firms in order to be able to locate within the SEZ, and a number of factors depicting the institutional set up of the zone programme. The level of corporate tax breaks is calculated as an index based on the level of tax exemption and the number of years granted over a 20 years horizon. This index can take values from 20 – reflecting a company that is 100% exempt from paying corporate income tax over the entire 20 year horizon – to 0 – indicating 0% exemption in any of the years. A dummy which takes the value of 1, if firms within the SEZ benefit from subsidized utilities is also included in the dataset. Non-fiscal incentives are captured using two dummies which reflect whether firms are exempt from following certain labour regulations that normally apply within the country; and if there is a national one-stop-shop available to companies to facilitate administrative processes. The existence of a minimum investment requirement is also included as an explanatory variable as is the level of foreign ownership required from companies. Both variables reflect the potential presence of restrictions on companies in order to participate in the zone. Finally, the institutional set-up of the SEZ programme is considered in the dataset. As pointed out above, having an independent zone regulator is regarded as best practice and is therefore added as a potential driver of zone performance. As is the case with the SEZ related variables, the data stem from the newly built dataset and details can be found in Appendix 2.

4.4.4 Contextual factors – country and regional level endowment

A set of variables reflecting the country and regional endowments is used as a base model in order to control for the contextual factors that may influence the SEZ. At the country level, controls for the proximity of a country to large markets, the level of industrialization, GDP per capita, and the general business environment, as reflected in institutional variables, are included in the dataset. The indicator for proximity to large markets is calculated based on the inverse distance of the country, in which the SEZ is located, to the US and Europe. The distance is calculated using information sourced from http://www.distancefromto.net/. The higher this indicator, the closer the country is to these markets. Given the importance of access to markets for companies, we would expect this coefficient to be positive. Level of industrialization is the GDP generated by a country's manufacturing sector as a percentage share of the overall GDP at the beginning of the period of analysis. The data is sourced from the World Development Indicators. A higher share of pre-existing industry reflects the inherent capacity of the host country to produce manufacturing goods (Hidalgo & Hausmann, 2009). A higher value, keeping other things equal, should therefore be attractive to companies, leading to a positive coefficient. The natural logarithm of GDP per capita reflects a country's overall level of development and also provides an indication of the wage level. We do not have a strong prior on the sign of this coefficient: on the one hand, companies may require a minimum level of development in order to be attracted to an area. The sign may thus be positive. On the other hand, provided that salary levels are lower in poorer countries, zones in less wealthy countries may be particularly attractive for firms searching to reduce costs. Different variables are tested in order to capture the general institutional and business environment in the host country. The Rule of Law estimate is based on Kaufmann, Kraay and Mastruzzi's (2011) Worldwide Governance Indicators and is also sourced from the World Development Indicators. The values of this indicator range from -2.5 to 2.5. The value for the beginning of the period of analysis is used. The higher the value, the better the rule of law. A higher score, reflecting a more stable institutional environment, should be positively correlated with SEZ performance. We also test the Ease of Doing business rank, however do not include the results in the main body of the paper as the results are insignificant. Finally, a country's nightlights growth is also included in the regression to control for the overall change in economic performance in the country. This allows us to single out whether a zone's performance was actually driven by the other characteristics included or simply followed national growth.

These country variables are complemented by a proxy, which reflects the level of development and socio-economic characteristics of the within-country region the zone is located in. As mentioned above, for political and social reasons, zones are frequently located in lagging regions within a country to stimulate economic activity in these areas. The natural logarithm of the ratio of the regional GDPpc over national GDPpc provides an indication of how well-off a region is, in comparison to the national average. Values over 0 indicate that the zone has a higher GDPpc than the national average and is thus likely to be endowed with better socio-economic characteristics, but also higher salaries. Values below 0 indicate the opposite case. The variable, thus, allows us to test if zones in lagging regions are performing better or worse than those located in the economic cores. This variable is sourced from Gennaioli, La Porta, Lopez De Silanes, and Shleifer (2014) and reflects the level at the beginning of the period of analysis.[30] It is also worthwhile noting that complementing the national controls with the ratio allows to control for the immediate geographical context of the SEZ, which, particularly in large countries, may be very different from the national average. Details for all variables are included in Appendix 2.

4.5 Descriptive analysis of SEZ performance

4.5.1 Zone performance 2007 - 2012

Before turning to the econometric analysis, we examine the performance of the 346 SEZs in our dataset for the main period of analysis (2007-2012). The average of the absolute growth rate across all SEZs is 14.7% over the entire period. A median growth performance of 2.8% and a standard deviation of 28% indicate a vast spread in growth among the SEZs. Looking at the relative SEZ performance (zone growth compared to the national growth) presents further interesting insights. An average ratio of 0.98 shows that zones on average have grown roughly at the same speed as the countries they are located in. Similarly to the absolute performance, the median is, at 0.95, lower than the national growth level. Again, there is a considerable

[30] http://scholar.harvard.edu/shleifer/publications/growth-regions

spread with a standard deviation of 0.22. SEZ growth performance has on average, thus, been rather moderate – and far from displaying the stellar performance that often drove the design and launch of SEZs. Consequently, the ambitious goals of SEZ policies have been far from fulfilled during the period of analysis. There is, moreover, a large diversity in zone performance. Appendix 3 shows further details of the summary statistics per country.

For the purpose of a first descriptive analysis, the zones are grouped into different performance categories. For the absolute performance the following three groups are used: 1) shrinking, 2) stable, and 3) growing. 'Shrinking' includes those zones, whose absolute light emissions shrank by more than 5% over the period of 2007 to 2012; 'stable' zones are those that remained within a +/- 5% range over the entire period; and the 'growing' group includes those SEZs with an increase in the absolute nightlights emissions of more than 5%.[31] The zones performance relative to that of national growth is captured by the following categories: 1) slower; 2) equal; and 3) faster. The 'slower' group includes those zones, whose ratio between zone and national growth is less than 0.9, the 'equal' group, those which are between 0.9 and 1.1, while the 'faster' group, all zones with a ratio larger than 1.1.

Figure 4-2 shows the number of zones in each of the groups. The numbers reflects the large variability in zone performance. While only 33 of the zones considered in the analysis shrank from 2007 to 2012, 150 remained relatively stable and 163 grew. This implies that only half of the zones exhibited a positive growth performance. Looking at the performance relative to national growth paints an even less optimistic picture: only 65 zones grew considerably faster than the national average, while the vast majority of zones grew at the speed of the national economy. One out of four zones grew well below the national average.

[31] Growth rates refer to the entire period of analysis, not the yearly growth rate.

Figure 4-2: Number of SEZs in each performance category, 2007 – 2012

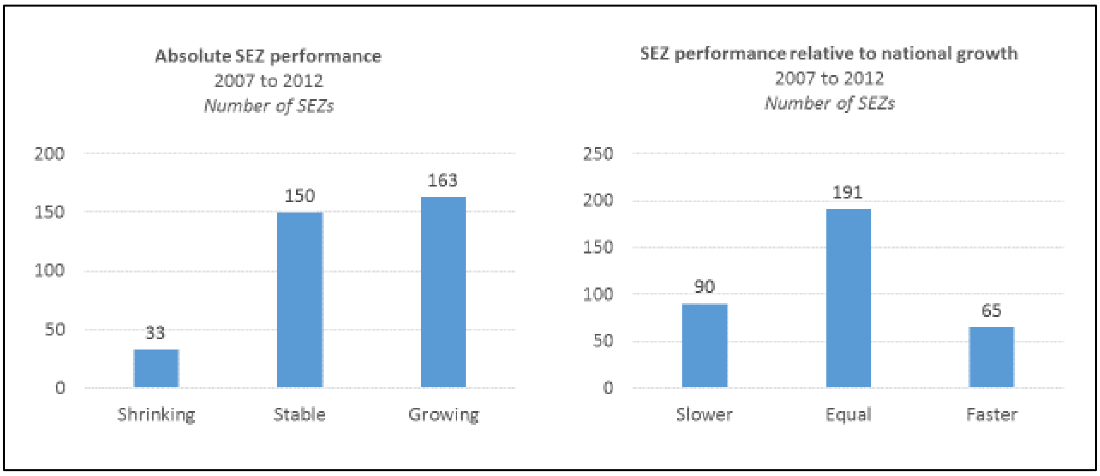
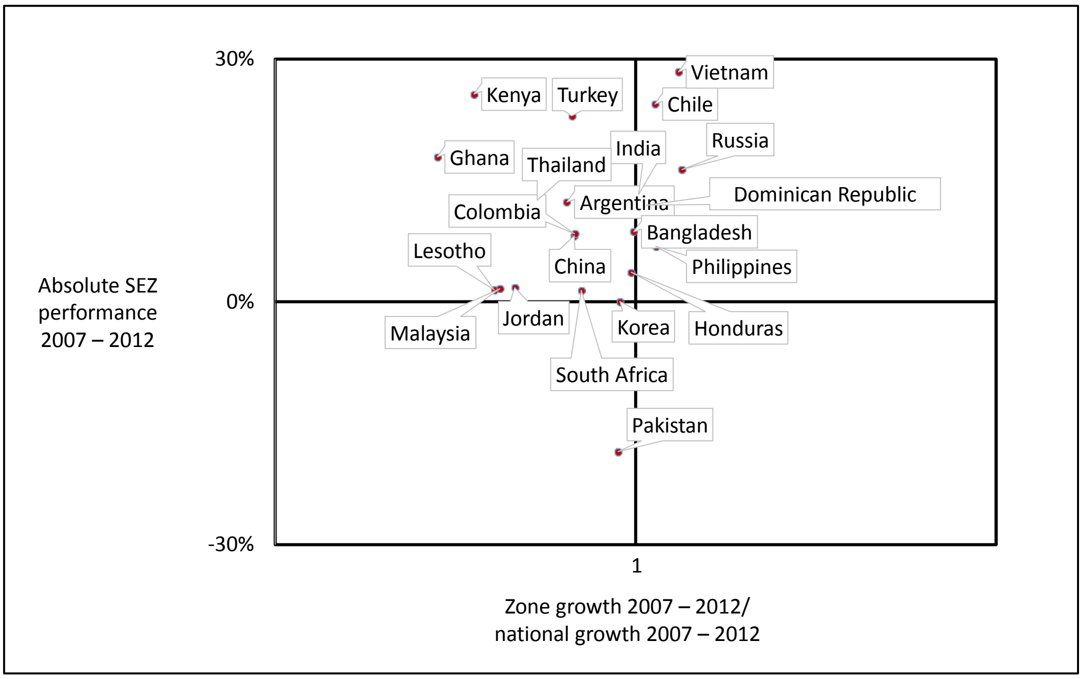


Figure 4-3 depicts the average SEZ growth performance per country. The y-axis plots the absolute growth performance, while the x-axis shows the ratio of zone over national growth. Countries above the average horizontal line had SEZs that performed better than the average, those below the line shrank. Among those countries included in the study, only zones in Pakistan experienced absolute negative growth rates during the period of analysis. Zones in Malaysia, Lesotho, Jordan, South Africa, and Korea remained on average relatively stable. The remaining countries display a strong increase in nightlights within the zones.

Figure 4-3: National averages of absolute and relative SEZ growth, 2007 – 2012



The picture is, however, less favourable when we consider the relative growth performance of the zones. The majority of countries has an average ratio below 1 and is therefore positioned to the left of the vertical line, indicating that nightlights in the zones grew slower than in the country as a whole. Even countries where the absolute zone growth was dynamic, such as in Kenya, Turkey, and Ghana, SEZ growth was lower than overall national growth. In other countries with a high absolute growth, such as Vietnam and Russia, zones did grow faster than the national average. However the ratio of average zone growth relative to national economic growth never exceeds 1.06.

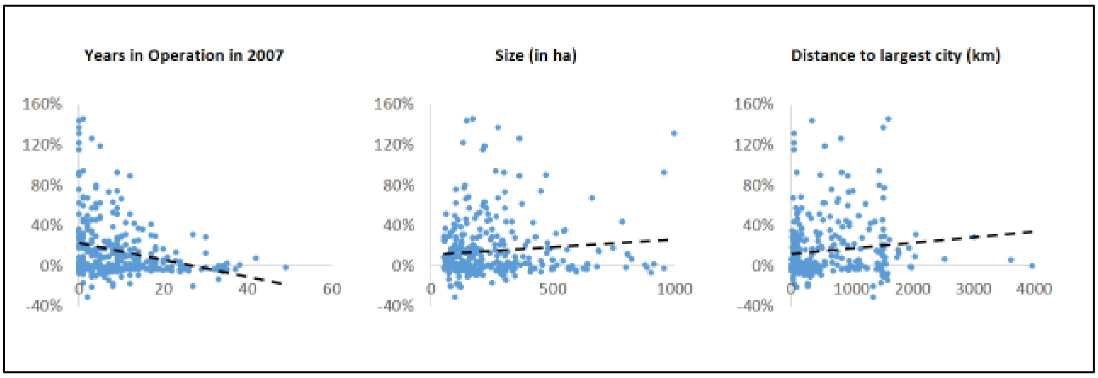
There is no clear pattern in what concerns the geographical distribution of the success of SEZs. Figure 4-3 shows that successful zone programmes, both in absolute and relative terms, can be found in different parts of the world. Zone performance within countries (Appendix 3) also displays considerable heterogeneity. Taking the case of Vietnam, which figures positively in terms of absolute and relative growth, zones with stellar performance combine with others whose economic growth levels have left a lot to be desired. Zones in the country fall almost evenly in the faster (33), equal (38), and slower (32) categories. A similar, albeit more positive, picture emerges for absolute growth: 70 Vietnamese zones grew during the period of analysis, while only 11 shrank, and 22 remained stable. A standard deviation of 37% for absolute growth demonstrates the significant differences between zones.

4.5.2 Zone performance, SEZ characteristics and policies

To obtain a first understanding of how SEZ characteristics and policies may relate to SEZ performance, we plot absolute SEZ performance (the SEZ growth rate) against some of the explanatory variables included in the econometric analysis. Figure 4-4 compares, by means of simple scatterplots, zone characteristics (on the x-axis) – the year the zone became operational, its size, and the distance to the largest city in the country – with the economic performance of the zone (on the y-axis).

The years in operation figure displays a clear negative correlation with zone growth from 2007-2012. More recently-established zones seem to perform better than older ones, although the spread is large even among zones established relatively recently. The association between zone size, on one side, and zone performance, on the other, is positive. The last scatterplot, which relates the distance to the largest city with the performance, reveals a slight positive correlation, which counterintuitively would suggest that zones which are located further away from the largest city in the country are more dynamic. Overall, the evidence arising from the scatterplots for a correlation between certain zone characteristics and performance is limited. They rather underline that there is considerable variation in zone performance depending on which the zone characteristic is taken into account.

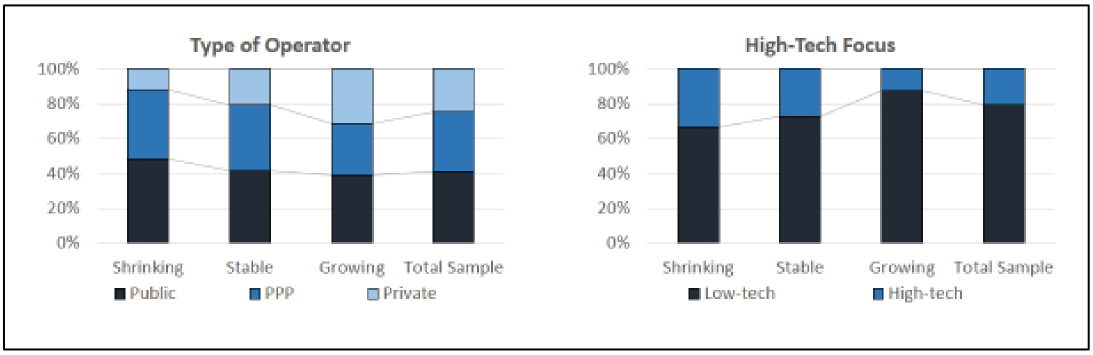
Figure 4-4: SEZ growth and SEZ maturity, size and location, 2007 – 2012



Moving on to the zone operator and sector focus, Figure 4-5 shows how these two characteristics, which are frequently considered to be of importance for zone performance, combine. For this purpose, we compare the share of the public, PPP and privately operated zones in the overall sample with their representation in the shrinking, stable and growing sub-groups (left side of Figure 4-5). Among the growing zones, privately operated zones are overrepresented compared to their share in the total sample while they are underrepresented in the shrinking group. The share of publicly run zones is larger in the shrinking group than in the overall sample. This, in principle, confirms the general perception in the literature that publicly operated zones tend to be less successful.

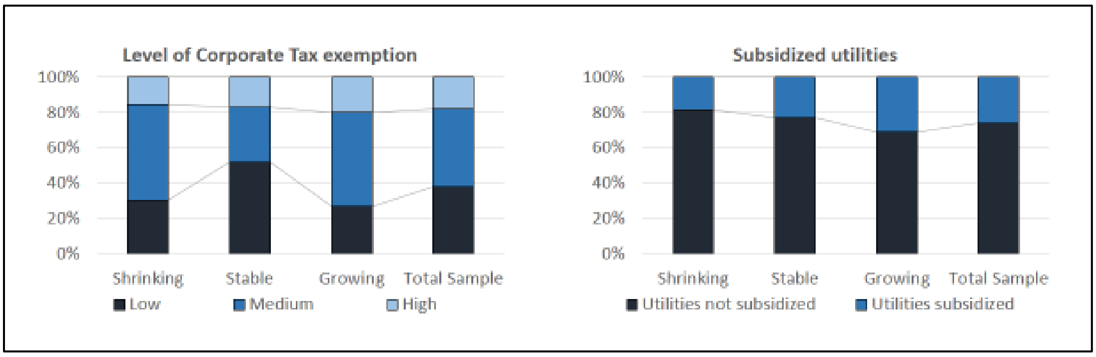
We apply the same procedure to the sector focus (right hand side of Figure 4-5). A clear pattern can be detected: among the shrinking zones there is a strong presence of zones focused on high-technology sectors. This first descriptive account suggests that low-tech manufacturing firms tend to do better than their high-tech counterparts.

Figure 4-5: SEZ growth, nature of operator and sector focus, 2007 – 2012



Regarding the programme variables and specifically some of the incentives commonly provided to companies, Figure 4-6 shows how the level of corporate tax exemption, and availability of subsidized utilities is related to SEZ performance. For this purpose, the level of corporate tax exemption is grouped into a low, medium, and high category using the corporate tax exemption indicator as defined in the previous section. Zones with a value between 0 and 7 for this indicator are in the low incentive category; those from 7 to 14 in the medium one; and those above 14 in the high category. Again, we compare the representation of these groups in the overall sample with their shares in the three performance groups. Among the growing SEZs, the medium incentives category is overrepresented compared to its share in the overall sample. The same is true for the shrinking group. High levels of tax incentives can be found in the shrinking group in a similar share than in the overall sample. Thus, no clear pattern can be detected overall. This may suggest that generous tax exemptions alone do not guarantee zone success. The picture for subsidized utilities lends some support that subsidized utilities are zone growth promoting: zones with subsidized utilities are overrepresented in the growing category, while they are slightly underrepresented in in the shrinking group.

Figure 4-6: Fiscal incentive package and SEZ performance, 2007 – 2012



4.6 Econometric analysis

4.6.1 Period of analysis 2007 - 2012

The correlations presented above, although interesting, only give a very partial picture of what factors drive SEZ performance, as they do not control for additional factors that may influence what drives the economic dynamism of SEZs. When internal and external factors are considered together, some indicators may display a greater connection than others or even limit the association of other factors with zone economic growth. Hence, in order to be able to give a more accurate and complete picture of what drives SEZ performance, we conduct a simple OLS econometric analysis of Model (1). The model combines both SEZ-related factors and programme factors (the factors internal to the zone) together with country- and regional-level endowments (the factors external to the zone). The analysis is conducted for 343 zones considered over the period between 2007 and 2012. Table 4-3 provides an overview of the results taking into account the two dependent variables: a) the overall performance of the zone and b) how well does the zone do relative to the economic performance of the country where it is located.

Table 4-3: Dependent variable: SEZ performance 2007 – 2012

VARIABLES	(1) Zone growth	(2) Zone growth	(3) Zone growth	(4) Zone growth	(5) Zone/national growth	(6) Zone/national growth	(7) Zone/national growth	(8) Zone/national growth
SEZ specific variables								
Initial lights in zone	-0.000988***	-0.000990***	-0.000986***	-0.000992***	-0.000800***	-0.000804***	-0.000799***	-0.000803***
	(0.000179)	(0.000166)	(0.000167)	(0.000164)	(0.000150)	(0.000139)	(0.000140)	(0.000138)
Years in operation	-0.00303**	-0.00330***	-0.00393***	-0.00439***	-0.00262**	-0.00278***	-0.00332***	-0.00365***
	(0.00143)	(0.00125)	(0.00138)	(0.00141)	(0.00120)	(0.00104)	(0.00117)	(0.00120)
Size	0.0931***	0.0943***	0.0924***	0.0937***	0.0751***	0.0763***	0.0746***	0.0756***
	(0.0199)	(0.0186)	(0.0185)	(0.0182)	(0.0166)	(0.0155)	(0.0154)	(0.0152)
High-tech focus	-0.0400*	-0.0485**	-0.0318	-0.0372*	-0.0320	-0.0376**	-0.0241	-0.0280
	(0.0239)	(0.0214)	(0.0223)	(0.0222)	(0.0197)	(0.0179)	(0.0186)	(0.0185)
Operator								
PPP	-0.00974	-0.0190	-0.00566	-0.00288	-0.00856	-0.0166	-0.00671	-0.00470
	(0.0342)	(0.0330)	(0.0331)	(0.0329)	(0.0283)	(0.0272)	(0.0275)	(0.0273)
Private	0.0102	-0.0158	-0.0237	-0.0283	0.00890	-0.0109	-0.0191	-0.0225
	(0.0428)	(0.0329)	(0.0379)	(0.0384)	(0.0349)	(0.0270)	(0.0316)	(0.0319)
Distance largest city	-0.00725***	-0.00456*	-0.00477*	-0.00556**	-0.00591***	-0.00374*	-0.00383*	-0.00440**
	(0.00271)	(0.00262)	(0.00248)	(0.00253)	(0.00224)	(0.00215)	(0.00206)	(0.00209)
SEZ programme variables								
Corporate Tax exemption			0.00255	-0.0787**			0.00236	-0.0562**
			(0.00351)	(0.0311)			(0.00278)	(0.0261)
* ln GDP pc				0.00918**				0.00662**
				(0.00357)				(0.00297)
Subsidized utilities			-0.0595	-0.0240			-0.0486	-0.0230
			(0.0429)	(0.0447)			(0.0352)	(0.0378)
National one-stop-shop			-0.0201	0.0295			-0.0121	0.0237
			(0.0411)	(0.0370)			(0.0344)	(0.0310)
Foreign owner. requirement			-0.414**	-0.438**			-0.339**	-0.357**
			(0.187)	(0.188)			(0.161)	(0.162)
Independent zone regulator			-0.0233	-0.0116			-0.0143	-0.00583
			(0.0279)	(0.0265)			(0.0225)	(0.0216)
Contextual factors								
Reg. / nat. GDPpc	-0.107***	-0.0848***	-0.0900***	-0.0926***	-0.0849***	-0.0659**	-0.0704**	-0.0722***
	(0.0378)	(0.0313)	(0.0338)	(0.0328)	(0.0309)	(0.0255)	(0.0277)	(0.0270)
Proximity to large markets		0.0104***	0.0111***	0.00939***		0.00744***	0.00810***	0.00686**
		(0.00327)	(0.00347)	(0.00350)		(0.00273)	(0.00287)	(0.00297)
Industry (% of GDP)		0.375**	0.346**	0.374**		0.261**	0.244*	0.264**
		(0.158)	(0.166)	(0.157)		(0.125)	(0.135)	(0.132)
Rule of law		0.0145	-0.0282	-0.0474		0.00820	-0.0244	-0.0382
		(0.0392)	(0.0388)	(0.0367)		(0.0325)	(0.0332)	(0.0313)
GDPpc 2007		-0.0268	-0.00127	-0.0711*		-0.0182	0.00279	-0.0476
		(0.0243)	(0.0275)	(0.0380)		(0.0201)	(0.0227)	(0.0331)
Country nightlights growth		0.301***	0.317**	0.101		-0.516***	-0.496***	-0.652***
		(0.113)	(0.147)	(0.140)		(0.0921)	(0.115)	(0.124)
Constant	1.295***	0.710***	0.498	1.400***	1.045***	1.533***	1.341***	1.992***
	(0.0802)	(0.225)	(0.308)	(0.446)	(0.0660)	(0.188)	(0.253)	(0.391)
Country dummies	Yes	-	-	-	Yes	-	-	-
Structural nightlights controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	343	343	343	343	343	343	343	343
R-squared	0.422	0.388	0.401	0.408	0.372	0.336	0.349	0.354

Robust standard errors in parentheses, clustered at within country regional level
*** p<0.01, ** p<0.05, * p<0.1
NOTE: Structural controls are the population density around the zone and whether the zone is located directly next to a highway or a water body

Columns 1 – 4 present the results for the regressions with absolute zone growth as dependent variable, columns 5 – 8 use the zone performance relative to its host country. We start by testing the effect of SEZ specific characteristics, using country fixed effects and then sequentially add contextual and SEZ policy specific variables. To examine the impact of zone specific variables, zone characteristics together with the initial levels of lights are introduced in the analysis. The regressions including only country fixed effects are presented in columns 1 and 5, while columns 2 and 6 substitute the country fixed with more specific regional and national contextual controls which may affect the economic performance of the zone as well as that of the region and country where it is located. SEZ programme variables are finally included in columns 3 – 4 and 7 – 8. Given the large number of independent variables, the estimations are as parsimonious as possible. Of the six SEZ specific variables included in the regressions, four show a consistently significant correlation with SEZ performance and one displays a significant correlation in 4 out of the 8 regressions.

First, the results of the analysis display, as expected, a certain convergence in the zone growth. The initial level of lights within the zone is negatively correlated with zone economic performance in all regressions, regardless of the level of controls included. This implies that more established zones in 2007, which in most cases display a high level of nightlights, grew at a slower pace than younger zones and that zones that were created at the beginning of the period of analysis. Not surprisingly, SEZs grow faster in the initial years of their life, with their economic dynamism plateauing as they mature.

A second factor that confirms that older, more established zones tend to be less dynamic is that the coefficient for the number of years the zone had been in operation by 2007 is consistently negative and statistically significant in all eight regressions. This result is robust to the inclusion of initial level of lights in the estimation. Hence, the coefficient cannot be considered as driven by lower levels of initial light for newer zones. This points to the fact that the success of zones is relatively short-lived. Growth is higher (as also indicated in Figure 4-4 and reinforced by the degree of convergence in zone growth) in the early years of the zone and peters out with time. More established zones in our sample, once everything else is controlled for, are less economically dynamic. This is in line with much of the literature that has stressed the challenge of maintaining economic performance after initial success (World Bank, 2011).

Third, zone size matters. The size of the SEZ is positively and significantly correlated with zone performance. Larger zones have an advantage over smaller ones when it comes to growth potential.

Fourth, the results provide consistent evidence that the distance to the largest city is negatively correlated with zone performance. Zones located further away from the main city in the country are less dynamic, holding other things constant. This is in line with the large body of literature, which emphasizes the strategic role of zone location (Asian Development Bank, 2015; Madani, 1999). Alternative city distances were also considered, including distance to the closest city of either more than 1 million or 500 thousand inhabitants and distance to the closest major port, but the results of the analysis provide no consistent evidence to support the role of large cities – beyond that of the main city in the country – or to ports. SEZs have therefore benefited from proximity to the largest and often more accessible agglomeration in the country, but the benefits of greater agglomeration and accessibility do not expand beyond the primary city.

Fifth, the more successful SEZs in emerging countries during the period of analysis have been those with a low technological component. The indicator depicting the presence of high-technology zones displays negative and statistically significant coefficients in four out of the eight regressions. This lends support to the notion that, in this type of countries, more successful zones are generally those specialized in low-tech, low-cost manufacturing products and not those that have aimed and succeeded in attracting sectors with a higher technological component and value added. This result reflects the challenge zones face when aiming to move away from more standard manufacturing and up in the value chain, in particular when they are located in areas with inauspicious conditions for the development of high-tech sectors (Asian Development Bank, 2015; World Bank, 2011). It also reflects the risks of technology driven shortcuts to economic development in many parts of the world where the conditions for the rapid development and assimilation of new technology are simply not there (Rodríguez-Pose & Hardy, 2014).

Finally, the nature of the zone management does not seem to matter as much as frequently assumed: the results of the analyses do not support the idea that private operators are more effective than public ones or vice versa. This is in line with the findings by Farole (2011) and is likely to reflect a strong context dependency for this variable: it is frequently the case that whether zones are operated by the private or public sector is dependent on country level policy-making and legislation.

In brief, the results of the zone specific variables point to a number of structural features that are closely connected to zone performance. First, zone growth is difficult to sustain over time and the largest benefits accrue shortly after start of operations and wane as the zone matures. Second, larger SEZs seem to have an advantage over smaller ones. Third, despite a recent push to upgrade SEZs from being purely labour intensive “sweatshops” of standard, low value-added manufacturing products to locations for industries with a greater technology component, it is the low-tech manufacturing zones that have done well in the period of analysis. Furthermore, a strategic location in close proximity to the largest city in the country is beneficial for zone performance. The insignificant results of the other variables likely reflect a large degree of variability and context dependency in terms of these characteristics and their impact on growth.

Programme variables (Table 4-3, Columns 3 – 4 and 7 – 8) tend to have a more limited association with zone performance than zone-specific characteristics. Only two of the five programme variables related to incentive packages, programme requirements, and set up are significant. This would suggest that specific aspects of the programme design of the zones by themselves – which have been the object of considerable attention in past research – are not sufficient in and of themselves to explain zone level growth.

There seems to be a limited connection between incentive packages and SEZ economic performance. Both variables for corporate tax exemption and subsidized utilities have an insignificant coefficient in columns 3 and 7 of Table 4-3. This implies that incentives on their own do not play an important role in explaining zone performance. When we test for a varying effect of corporate tax breaks depending on the level of development, the results, however, become highly significant (Table 4-3, Columns 4 and 8): the main term is negative and its interaction with GDP per capita is positive. The impact of corporate tax holidays thus seems to depend to a large extent on the level of development, as the impact is negative for poorer countries, but becomes positive for wealthier ones. The tipping point is at a GDP per capita of roughly 5100US$ where corporate tax exemptions start to have a positive correlation with zone performance. Tax breaks may thus be an effective tool to attract investments in more developed countries, but not in less developed ones.

The second significant result is the negative correlation between the foreign ownership requirement and SEZ performance. This suggests that imposing a minimum participation of foreign firms on SEZ companies hinders SEZ dynamism. This finding lends support to best practice guides that frequently advocate the removal of foreign ownership requirements in order to minimize the distortions created by favouring foreign companies over local ones (OECD, 2009).

The remaining programme variables – availability of a one-stop shop and the independence of the zone regulator – display insignificant coefficients throughout. They thus do not seem to be drivers of SEZ performance. These results go counter claims in many best practice guides, which have underlined the importance of programme characteristics for the viability of SEZs (Asian Development Bank, 2015; OECD, 2009).

From a programme design perspective, we can therefore conclude that corporate tax exemptions can play an important role in stimulating growth in SEZs, but only under certain circumstances. Furthermore, interventions such as imposing foreign ownership requirements are likely to be detrimental for SEZ performance. By contrast, the type of programme set-up and other benefits play a less vital role than anticipated.

Last but not least, examining the results for the contextual factors provide some interesting insights. Proximity to large markets delivers significant and positive coefficients, pointing to a beneficial effect of being close to the customer base, as is the case of the previous industrialization level. This is in line with much of the case study literature that emphasizes the importance of "traditional" locational advantages (Madani, 1999; Rolfe et al., 2004; Watson, 2001). It also highlights the challenge that countries with an economic structure dominated by agriculture, face if they attempt to industrialize through SEZ policies.

In contrast to previous studies stressing the salience of the general business environment (Aggarwal, 2005; Daude & Stein, 2007; Farole, 2011), rule of law is insignificant in our analysis. Experimentation with alternative measures of the quality of institutions at a national level, such as the Ease of Doing Business Rank, also delivers insignificant results. The business environment therefore seems to have limited sway over the performance of SEZs. This may also be related to the low-tech, low value added dimension behind the success of many SEZs. When the main factors of SEZ success are related to low labour costs, proximity to large markets, and some background in industry, the quality of national institutions may matter less than when the more complex networks and value chains related to high-tech manufacturing are required to be in place.

The ratio between regional GDP per capita and national GDP per capita is negative and highly significant throughout, further underlining the importance of low-cost environments for the success of SEZs. Consequently, SEZs in poorer areas of the country – albeit with reasonable accessibility to the main city – have performed better than those in better off regions. Traditional wage-based advantages therefore remain of great importance for firms seeking a location in a SEZ in an emerging country.

Finally, while GDP per capita levels in 2007 are insignificant apart from one regression (column 4), the growth of lights from 2007 to 2012 in the whole country is strongly significant throughout. It displays a positive correlation with SEZ performance when we use the absolute level of SEZ growth as dependent variable. Not surprisingly this relationship turns negative once the dependent variable is the relative performance. This positive correlation suggests that zones grow faster in rapid growth environments. At the same time, when using the relative performance measure, it is more difficult for a zone to outperform national growth in the presence of high growth rates, hence the negative correlation in columns 5 to 8.

The analysis of the contextual factors, hence, indicates that firms in SEZs still seek, overall, low cost locations in less developed areas of the countries and in close proximity to the main city, and with easy access to North American and European markets. Previous industrialization also plays a role in the success of zones. By contrast, institutional factors, seem to be less relevant for SEZ economic dynamism.

4.6.2 Five-year growth rates

The analysis for the period between 2007 and 2012 contains zones at different stages of development: some nascent, some more mature. As the results of Table 4-3 show, the maturity of zone has a bearing on its overall performance, and limits the perception of what drives the success of SEZs start-ups. Hence, in order to get a clearer picture of the factors behind the zone take-off, we analyse what determines a zone performance in the first five years after the start of operations. This implies that the period of analysis is different for each zone, covering the phase between t0 (start year) and t5 (five years after the establishment of the zone). This can only be done for a reduced sample, as the founding of the SEZ needs to take place after 1992, when the nightlights data becomes available. The sample for this analysis contains 252 zones, in contrast to the 343 considered in Table 4-3.

Furthermore, the SEZ dataset currently only contains information for the policies applicable in the years from 2007 onwards. We therefore have to exclude the SEZ programme related explanatory variables from the five-year growth regressions. All other explanatory variables remain the same as in the previous section with one exception: as each zone's performance is measured from its start date, the years in operation variable is substituted by a variable which reflects the year the zone became operational. This allows to control for the fact that zones started operating at different points in time and therefore may have been exposed to different economic environments. Appendix 4 includes the summary statistics per country.

Table 4-4: Dependent variable: SEZ growth after 5 years from the start of operations

VARIABLES	(1) Zone growth	(2) Zone growth	(3) Zone growth	(4) Zone/national growth	(5) Zone/national growth	(6) Zone/national growth
SEZ specific variables						
Initial lights in zone	-0.00156***	-0.00136***	-0.00129***	-0.00118***	-0.00116***	-0.00111***
	(0.000246)	(0.000278)	(0.000277)	(0.000242)	(0.000252)	(0.000317)
Year established	-0.00180	0.00588	-0.0103	0.0101*	0.00485	0.00996
	(0.00864)	(0.00792)	(0.00762)	(0.00592)	(0.00655)	(0.00631)
Size	0.00145***	0.00132***	0.00108***	0.00115***	0.00114***	0.00107***
	(0.000254)	(0.000302)	(0.000312)	(0.000257)	(0.000282)	(0.000365)
High-tech focus	-0.0754	-0.0494	-0.0756	-0.0468	-0.0470	-0.0609
	(0.0544)	(0.0426)	(0.0502)	(0.0389)	(0.0359)	(0.0370)
Operator						
PPP	-0.00806	0.124	0.138	0.135*	0.117	0.176*
	(0.0647)	(0.0954)	(0.0946)	(0.0740)	(0.0790)	(0.103)
Private	-0.0386	0.00619	-0.0960	0.0345	0.0168	-0.0419
	(0.0591)	(0.0545)	(0.0778)	(0.0473)	(0.0501)	(0.0649)
Distance largest city	-7.84e-05	-9.10e-05**	-0.000101**	-9.11e-05***	-7.89e-05**	-8.38e-05**
	(6.16e-05)	(4.56e-05)	(4.85e-05)	(3.33e-05)	(3.94e-05)	(3.72e-05)
Contextual factors						
Ratio regional/national GDPpc	-0.0263	-0.0393	0.00699	-0.0791***	-0.0357	-0.0756***
	(0.0192)	(0.0246)	(0.0229)	(0.0170)	(0.0225)	(0.0273)
Industry (% of GDP)		-0.00108			0.000690	
		(0.00505)			(0.00467)	
Proximity to large markets		-0.478*			-0.324	
		(0.278)			(0.278)	
Rule of law		-0.0133			-0.0355	
		(0.0687)			(0.0612)	
GDPpc in year operational		-0.0218			-0.0191	
		(0.0493)			(0.0428)	
Country nightlights growth		0.557***			-0.430***	
		(0.122)			(0.106)	
Constant	3.948	-11.27	20.52	-19.14	-8.341	-19.09
	(17.27)	(15.77)	(15.21)	(11.83)	(13.01)	(12.61)
Structural nightlights controls	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	-	-	Yes	-	-	Yes
Observations	252	252	252	252	252	252
R-squared	0.240	0.355	0.413	0.247	0.302	0.305

Robust standard errors in parentheses, clustered at regional level
*** p<0.01, ** p<0.05, * p<0.1
NOTE: Structural controls is whether the zone is located directly next to a water body.

Table 4-4 provides an overview of the results. As in the previous section, we use two dependent variables: the absolute growth of the zone (Table 4-4, Columns 1 – 3) and zone growth relative to national growth (Table 4-4, Columns 4 – 6). Columns 1 and 4 show the results taking into account only the SEZ characteristics. In columns 2 and 5 the contextual controls are introduced, whereas country dummies substitute those in columns 3 and 6. Country dummies have the advantages in this context that they pick up some of the effects of the SEZ policies, which cannot be included individually in this section.

The results lend further support to some of the findings for SEZ specific characteristics presented in Table 4-3. Zone size remains positively correlated with zone performance, indicating a stronger growth performance of larger zones in the first five years of establishment. The negative impact of distance to the largest city is also confirmed. Zones which are located further away tend to be less dynamic than those closer to the largest city. The coefficient for the high-tech dummy remains negative throughout, but is not significant. Furthermore, neither the year of zone establishment nor the nature of the operator seem to make a difference for zone performance. We find no evidence of either an early mover advantage or of a 'learning-from-past-errors' effect, as more recent zones have not had a better economic performance in their first five years of life than those that were founded earlier.

In terms of the contextual factors, most indicators are insignificant with the exception of country nightlights growth and the ratio between regional and national GDPpc. The latter is, however, only significant in two out of the six regressions. The national growth of nightlights displays the same dynamics as those reported in Table 4-3: it is strongly positively correlated with absolute zone growth, while it is negatively correlated with the relative growth rate. Proximity to markets is negatively correlated but only in one regression and at the 10% significance level. This suggests that the result should not be emphasized overly. The remainder of the contextual controls is insignificant.

While these results should be interpreted with some caution due to the lower number of observations, they lend further support to the notion, found in the analysis for the period 2007 to 2012 that larger zones in closer proximity to the largest city, but in relatively cheap locations tend to display the best overall performance.

4.6.3 Case study: Vietnam

In order to get a closer insight into the results of the analysis, we look into one specific case study of an emerging country that has been particularly active in the promotion of SEZs. Vietnam introduced its SEZ policy in 1996, allowing first the establishment of industrial zones, export processing zones, and economic zones. This was followed by the introduction of high-tech parks in 2003. Since then, different types of SEZs have proliferated throughout the country. The advantage of analysing zones within one country that has been active in promoting this type of intervention, such as Vietnam, is that the contextual factors related to the SEZ policy, institutional set-up, and the country endowment in terms of socio-economic characteristics apply to all zones. This allows to delve deeper into whether the role of SEZ specific characteristics and potential interactions may become enhanced – once the noise stemming from the inclusion of a variety of national, institutional and regulatory variables has been reduced.

Table 4-5 presents the results of the case study analysis. As we are dealing with only one country, only the results of the absolute performance of the zone are reported (Columns 1 – 3). The results in Column 1 include regional dummies as controls, in Columns 2 and 3 we include the regional GDP per capita to control for the socio-economic characteristics of the areas in which the zones are located.

The results, by and large, support the findings of the previous sections: zone performance within Vietnam is much more related to the size of the zone, its (low-tech) dimension and location than to the specific programme characteristics. The positive connection between SEZ size and zone performance, and the negative coefficients for the maturity of the zone and distance to the largest city are confirmed. An interesting nuance to the previous high-tech findings is presented in Column 3 where we interact the high-tech dummy with the regional GDP per capita. The main effect of a high-tech focus remains negative but turns highly significant, whereas the interaction term displays a positive and highly significant coefficient. This result is intuitive: high-tech zones in remote areas struggle as they lack the basic local capabilities and endowments to make SEZs viable. A high-tech focused zone in a more developed area of Vietnam – i.e. in close proximity to Hanoi or Ho Chi Minh City – in contrast has a greater chance of becoming successful.

Table 4-5: Case study Vietnam. Dependent variable: SEZ performance 2007 – 2012

VARIABLES	(1) Zone growth	(2) Zone growth	(3) Zone growth
Initial lights in zone	-0.00193***	-0.00203***	-0.00220***
	(0.000356)	(0.000315)	(0.000309)
Size	0.00149***	0.00151***	0.00159***
	(0.000265)	(0.000206)	(0.000201)
Years operating	-0.0218*	-0.0249**	-0.0231**
	(0.0111)	(0.0106)	(0.0103)
High-tech focus	-0.00997	-0.0463	-4.669***
	(0.141)	(0.131)	(1.177)
* ln (regional GDPpc)			0.586***
			(0.150)
Operator			
PPP	-0.146	-0.0272	-0.0220
	(0.0882)	(0.0990)	(0.0984)
Private	-0.0646	-0.0589	-0.0411
	(0.0794)	(0.0661)	(0.0664)
Distance largest city	-0.000264	-0.000177**	-0.000179**
	(0.000329)	(7.70e-05)	(7.71e-05)
One-stop-shop on-site	0.0221	-0.0635	-0.0370
	(0.107)	(0.0994)	(0.101)
Power substation in zone	0.0682	0.00912	0.0142
	(0.116)	(0.0992)	(0.100)
Ln (regional GDPpc)		-0.000668	0.00221
		(0.0652)	(0.0644)
Constant	-42.37*	-49.38**	-45.86**
	(22.25)	(21.34)	(20.80)
Structural nightlights controls	Yes	Yes	Yes
Regional dummies	Yes	-	-
Observations	100	100	100
R-squared	0.692	0.437	0.451

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

Two new indicators to reflect the infrastructure and services offered within the zone were also included in the analysis. Both, the existence of a dedicated sub-power station and a one-stop-shop within the zone, are not correlated with zone performance. Similarly, the regional GDP per capita is not significant in any of the regressions. This differs from the results in the previous sections, where it was found that zones in less developed regions performed better than in more developed ones.

4.7 Conclusions

The aim of this paper was to analyse the factors driving SEZ performance in emerging countries. We relied on an entirely new dataset with information on SEZ characteristics and programmes as well as contextual factors across 346 zones in 22 emerging countries, which were operational by or before 2007. To overcome the challenge of limited data availability for SEZ outcomes, nightlights data have been used to proxy for SEZ performance.

While there is certainly no shortage of research which has focused on the lessons learnt from SEZ policies around the world using case study approaches, the analysis conducted in this paper is the first to deal with the economic dynamism of SEZs from a quantitative perspective, covering a large number of zones across developing countries. The change in approach and method has delivered results that to a certain extent confirm, but also refute part of the dominating knowledge about the viability, success, and influence of SEZs on economic development in the emerging world.

First and foremost, SEZs on the whole cannot be considered as a growth catalyst in emerging countries. Despite considerable variation in their performance across and within countries, their overall economic dynamism does not exceed that of the countries where they are located, casting doubts about claims that portray them as a panacea for growth.

Moreover, the results of the zone specific econometric analysis point to some crucial structural features behind the economic success – or lack of it – of SEZs. Key results include that a) zone growth is difficult to sustain over time; that b) trying to upgrade the technological component or value-added of SEZs is challenging, as zones focused on high-tech sectors have performed worse than those in low-cost, labour intensive sectors; and that c) size matters: larger zones seem to have an advantage in terms of growth potential.

Country- and regional-specific context further determine SEZ performance. Zones in relatively poor areas, but not too far away from the largest city in the country and in countries with relatively easy access to the main developed markets of the world have displayed the greatest economic dynamism. Zones in countries with a history of pre-existing industrialization have also been favoured.

Incentive packages to attract firms to SEZs and ownership and management schemes, by contrast, have had limited influence in the success of the zones. Factors such as the type of operator of the zone i.e. private, public, or PPP, corporate tax exemptions, or sundry subsidized utilities do not seem to have largely affected the success of zones across the emerging world. The backbone of most SEZ policies, corporate tax breaks, also seem to have played a relatively minor role in zone dynamism, which has been limited to the more developed countries in the sample. Hence, the role of factors such as tax breaks, the presence of an independent zone regulator, or non-fiscal benefits, such as the availability of a national one-stop-shop, seems to be much more context dependent than hitherto thought and there is no guarantee that the provision of these sort of support, incentives, and/or subsidies bears fruit in terms of zone dynamism.

The findings of the analysis have important policy implications. They point to the fact that SEZ policies in emerging countries do not take place in a vacuum and certain pre-conditions need to be met for these policies to maximize the returns of SEZs. Closeness to attractive markets is essential as is the pre-disposition of the economy. A country dominated by agriculture will have difficulty to industrialize through SEZ policies alone. Furthermore, a cost advantage through a low cost labour base is likely to remain an attractive feature for firms and continue to affect the dynamism of zones and their surrounding areas. Policy-makers should therefore consider carefully whether a SEZ programme can credibly achieve the desired outcomes in a given country context. Moreover, even in places where zone programmes have a greater potential to succeed, the effects are likely to be limited both in time. Hence, SEZ policies cannot substitute for wider structural reforms aimed at enhancing the potential for the development of economic activities, as well as the absorptive capacity in the country. Finally, there is a high degree of context dependency for SEZ policies. Whether a country requires an independent zone regulator, a private or a public operator, or certain services that are more or less needed in a specific zone, depends essentially on the precise context where the zone operates. Different combinations may be effective in different contexts.

The research presented here represents an important change in approach with respect to previous analyses about what determines the economic dynamism of SEZs. However, it is certainly not without limitations. First, the analysis measures economic growth based on nightlights data. Nightlights are an increasingly common alternative in economics for economic activity in those areas of the world where economic data either do not exist or are not reliable. However, their use is not exempt from controversy. Second, the definition of SEZs – in part conditioned by the use of nightlights as a proxy for economic growth – discards a large number of small SEZs, as well as those that, despite being planned, did not take off or became operational after 2007. The sample remains highly dependent on data availability in in some specific geographical areas of the world (e.g. East Asia vs. Africa). Finally, the gathering of data about the characteristics, programmes and incentives associated to the SEZs is limited to the type of information that can be readily quantified. This implies a loss of information particularly regarding 'soft' aspects, such as those relating to the quality of services provided at zone level or about the political will driving zone implementation, both at the zone and at the national level. Consequently, while, on the whole, the approach represents a considerable step forward in our understanding towards what makes SEZs across the emerging world work, given the caveats associated to the approach, the results must be considered with some caution.

References

Aggarwal, A. (2005). Performance of export processing zones: a comparative analysis of India, Sri Lanka and Bangladesh *Working Paper No.155*. New Delhi: Indian Council for Research on International Economic Relations.

Asian Development Bank. (2015). *Asian Economic Integration Report 2015. How can Special Economic Zones catalyze economic development?* Manilla, Philippines: Author.

Chen, X., & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences of the United States of America, 108*(21), 8589-8594. doi:10.1073/pnas.1017031108

Daude, C., & Stein, E. (2007). The qualilty of institutions and foreign direct investment. *Economics & Politics, 19*(3), 317–344. doi:10.1111/j.1468-0343.2007.00318.x

Disdier, A.-C., & Head, K. (2008). The puzzling persistence of the distance effect on bilateral trade. *Review of Economics and Statistics, 90*(1), 37-48. doi:10.1162/rest.90.1.37

Ebener, S., C., M., Tandon, A., & Elvidge, C. C. (2005). From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery. *International Journal of Health Geographics, 4*(1).

Elvidge, C., Safran, J., Tuttle, B., Sutton, P., Cinzano, P., Pettit, D., Arvesen, J., & Small, C. (2007). Potential for global mapping of development via a nightsat mission. *GeoJournal, 69*(1), 45-53. doi:10.1007/s10708-007-9104-x

Elvidge, C., Baugh, K., Anderson, S., Sutton, P., & Ghosh, T. (2012). The night light development index (NLDI): a spatially explicit measure of human development from satellite data. *Social Geography, 7*(1). doi:10.5194/sg-7-23-2012

Engman, M., Onodera, O., & Pinali, E. (2007). Export Processing Zones: Past and future role in trade and development. Paris: OECD Publishing.

Farole, T. (2011). *Special Economic Zones in Africa: Comparing Performance and Learning from Global Experience*. Washington D.C.: The World Bank.

Farole, T., & Kweka, J. (2011). *Institutional Best Practices for Special Economic Zones: An Application to Tanzania*. Retrieved from http://siteresources.worldbank.org/INTRANETTRADE/Resources/SEZTanzaniaPolicynoteFinal.pdf

FIAS. (2009). Special Economic Zones. Performance, Lessons Learned, and implications for Zone Development. Washington D.C.: The World Bank Group.

Florida, R., Gulden, T., & Mellander, C. (2008). The rise of the mega- region. *Cambridge Journal of Regions, Economy and Society, 1*(3), 459-476. doi:10.1093/cjres/rsn018

Gennaioli, N., La Porta, R., Lopez De Silanes, F., & Shleifer, A. (2014). Growth in regions. *Journal of Economic Growth, 19*(3), 259-309. doi:10.1007/s10887-014-9105-9

Ghosh, T., Powell, R. L., Elvidge, C., Baugh, K. E., Sutton, P. C., & Anderson, S. (2010). Shedding Light on the Global Distribution of Economic Activity. *The Open Geography Journal, 3*, 147-160. doi: 10.2174/1874923201003010147

Henderson, J., & Phillips, R. (2007). Unintended consequences: social policy, state institutions and the 'stalling' of the Malaysian industrialization project. *Economy and Society, 36*(1), 78-102. doi:10.1080/03085140601089853

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review, 102*(2), 994-1028. doi:10.1257/aer.102.2.994

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America, 106*(26), 10570-10575. doi:10.1073/pnas.0900943106

Jauch, H. (2002). Export processing zones and the quest for sustainable development: a Southern African perspective. *Environment and Urbanization, 14*(1), 101-113. doi:10.1177/095624780201400109

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and Analytical Issues. *Hague Journal on the Rule of Law, 3*(2), 220-246. doi:10.1017/S1876404511200046

Keola, S., Andersson, M., & Hall, O. (2015). Monitoring Economic Development from Space: Using Nighttime Light and Land Cover Data to Measure Economic Growth. *World Development, 66*, 322. doi:10.1016/j.worlddev.2014.08.017

Larrain, F., Lopez-Calva, L. F., & Rodriguez-Clare, A. (2009). *Intel: A Case Study of Foreign Direct Investment in Central America*. Retrieved from https://www.hks.harvard.edu/centers/cid/publications/faculty-working-papers/cid-working-paper-no.-58

Levin, N., & Duke, Y. (2012). High spatial resolution night-time light images for demographic and socio-economic studies. *Remote Sensing of Environment, 119*, 1-10. doi:10.1016/j.rse.2011.12.005

Luger, M. I., & Goldstein, H. (1991). *Technology in the garden: research parks and regional economic development*. Chapel Hill: University of North Carolina Press.

Madani, D. (1999). *A Review of the Role and Impact of Export Processing Zones. World Bank Policy Research Working Paper.* Washington D.C.: The World Bank.

Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-Time Light Data: A Good Proxy Measure for Economic Activity? *Plos One*(10), 18. doi:10.1371/journal.pone.0139779

Nel, E., & Rogerson, C. (2013). Special Economic Zones in South Africa: Reflections from International Debates. *Urban Forum, 24*(2), 205-217. doi:10.1007/s12132-012-9184-7

OECD. (2009). Towards Best Practice Guidelines for the Development of Economic Zones. Retrieved from https://www.oecd.org/mena/competitiveness/44866585.pdf

Portugal-Perez, A., & Wilson, J. S. (2012). Export Performance and Trade Facilitation Reform: Hard and Soft Infrastructure. *World Development, 40*(7), 1295-1307. doi:10.1016/j.worlddev.2011.12.002

Quintas, P., Wield, D., & Massey, D. (1992). Academic- industry links and innovation: questioning the science park model. *Technovation*, *12*(3), 161-175. doi:10.1016/0166-4972(92)90033-E

Rodríguez-Pose, A., & Arbix, G. (2001). Strategies of Waste: Bidding Wars in the Brazilian Automobile Sector. *International Journal of Urban and Regional Research*, *25*(1), 134-154. doi:10.1111/1468-2427.00302

Rodríguez-Pose, A., & Hardy, D. (2014). *Technology and industrial parks in emerging countries: panacea or pipedream?* Heidelberg and New York: Springer.

Rolfe, R. J., Woodward, D. P., & Kagira, B. (2004). Footloose and tax free: Incentive preferences in Kenyan export processing zones. *South African Journal of Economics, 72*(4), 784-807. doi:10.1111/j.1813-6982.2004.tb00134.x

Singa Boyenge, J.-P. (2007). *ILO database on export processing zones (Revised)*. Retrieved from Geneva: http://staging.ilo.org/public/libdoc/ilo/2007/107B09_80_engl.pdf

UNCTAD. (2015). *Enhancing the Contribution of Export Processing Zones to the Sustainable Development Goals*. Retrieved from http://unctad.org/en/PublicationsLibrary/webdiaepcb2015d5_en.pdf

Watson, P. (2001). *Export Processing Zones: Has Africa Missed the Boat? Not yet!* Retrieved from Washington, D.C.: http://siteresources.worldbank.org/EXTEXPCOMNET/Resources/2463593-1213887855468/15_EPZsHasAfricaMissedtheBoat.pdf

World Bank. (2011). *Special Economic Zones. Progress, Emerging Challenges and Future Directions*. Washington, D.C.: Author.

Appendix 1 – Regressing nightlights on SEZ firms and employment

VARIABLES	(1) SEZ employment	(2) Number of Firms
Nightlights within zone	177.3***	0.363***
	(23.77)	(0.0567)
Country dummies	Yes	Yes
Constant	-7,859***	9.740
	(1,909)	(9.635)
Observations	104	135
R-squared	0.556	0.524

Robust standard errors in parentheses, clustered at the regional level
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2 – Variable description

Variable	Description	Source
SEZ performance		
Absolute SEZ performance	$(Yi_{,1} - Yi_{,0}) / Yi_{,0}$: Growth rate of the sum of nightlights of the pixels that compose the SEZ surface over period of analysis	Based on National Centers for Environmental Information https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
Relative SEZ performance	Ratio of change in SEZ light intensity $(Yi_{,1}/Yi_{,0})$ over change of country light intensity $(Y_{country,1}/ Y_{country,0})$	Based on National Centers for Environmental Information https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
SEZ related variables		
Years in Operation	Number of years zone has been operating in 2007	All SEZ related variables are from the newly assembled dataset. Information obtained as described in the body of this paper.
Size	SEZ size in hectars	
High-Tech Focus	Dummy = 1 if the zone either 'self-proclaims' on their advertising material that they specifically target high-tech sectors or if companies established are within high-tech sectors, as defined by OECD	
Operator	Nature of zone operator: 0 = public, 1 = PPP, 3= private Variable takes into account whether the public sector is involved in the development of the zone and/ or provides the land	
Electricity Sub-Power Station	Dummy = 1 if SEZ has its own Sub-power station onsite	
One-stop Shop Onsite	Dummy = 1 if SEZ provides ones-top-shop services onsite	
Distance Largest City	Road distance in kilometres to the largest city in the country	
Distance closest Major Port	Road distance in kilometres to the closest major port	
Distance closest City with at least 500k Inhabitants	Road distance in kilometres to the closest city with at least 500,000 inhabitants	
Distance closest City with at least 300k Inhabitants	Road distance in kilometres to the closest city with at least 300,000 inhabitants	
SEZ programme variables		
Corporate Tax Exemption	Index based on the level of tax exemption and the number of years granted over a 20 years horizon. This index can take values from 20 – reflecting a company that is 100% exempt from paying corporate income tax over the entire 20 year horizon – to 0 – indicating 0% exemption in any of the years.	All regulatory variables are from the newly assembled dataset. Information obtained as described in the body of this paper.
Subsidized Utilities	Dummy = 1 if firms within the SEZ benefit from subsidized utilities	
National One-stop-shop	Dummy = 1 if one-stop-shop services are available to companies within the SEZ from a national authority	All regulatory variables are from the newly assembled dataset. Information obtained as described in the body of this paper
Foreign Ownership Requirement	% of firm ownership required to be hold by foreign company in order for firm to be able to locate within SEZ	

Variable	Description	Source
Independence of Zone Regulator	Dummy = 1 if zone regulator is an independent entity	
Free trade domestic market	Dummy = 1 if firms within SEZ can trade with the local market without paying import and export duties or other restrictions	
Contextual factors		
Ratio regional / national GDPpc	Natural logarithms of Regional GDP per capita / Country GDP per capita	Regional dataset sourced from Gennaioli, LaPorta, Lopez-de-Silanes & Shleifer http://scholar.harvard.edu/shleifer/publications?page=1
Proximity to Large Markets	Sum of the inverse distances from each country to the US and European Union	Based on distances from http://www.distancefromto.ne
Industry (% of GDP)	Industry, value added (% of GDP) in the beginning of the period of analysis	World Development Indicators
Rule of Law	Rule of Law indicator in the beginning of the period of analysis. Values range from -2.5 to 2.5.	Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues"
Political Stability	Political Stability indicator in the beginning of the period of analysis. Values range from -2.5 to 2.5.	Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi (2010). "The Worldwide Governance Indicators: Methodology and Analytical Issues"
GDPpc	Natural logarithm of the GDP per capita in the beginning of the period of analysis (constant 2010 US$)	World Development Indicators
Country nightlights growth	Growth rate of the sum of lights within the country in the period of analysis	Based on National Centers for Environmental Information https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
Regional GDPpc	Natural logarithm of the GDPpc in the within-country region the SEZ is located in	Regional dataset sourced from Gennaioli, LaPorta, Lopez-de-Silanes & Shleifer http://scholar.harvard.edu/shleifer/publications?page=1
Structural nightlights controls		
Population density around SEZ	Population density in immediate vicinity of the zone: 1 = isolated,; 2 = sparsely populated; 3 = densely populated	Based on visual inspection of SEZ sites in googlemaps satellite view
Waterbody	Dummy = 1 if zone is located directly next to a waterbody	
Highway	Dummy = 1 if zone is located directly next to a highway	

Appendix 3 – SEZ growth 2007 – 2012

Summary statistics for absolute SEZ nightlights growth 2007 - 2012

Country	# SEZs	Mean	Min	Max	StdDev
Argentina	4	0.123127	0.079646	0.2160494	0.063017898
Bangladesh	8	0.086411614	-0.17241379	0.25	0.128250978
Chile	3	0.244405533	0.127572	0.3157895	0.102008353
China	33	0.081519892	-0.11031175	0.9285714	0.256483049
Colombia	6	0.08383643	-0.04761905	0.2972973	0.150838106
Dominican Republic	10	0.119304864	-0.03225806	0.3714286	0.142807016
Ghana	1	0.1787709	0.1787709	0.1787709	-
Honduras	3	0.036355767	0.0204082	0.0535714	0.016617926
India	8	0.132515369	-0.03174603	0.4213836	0.177547738
Jordan	1	0.0173913	0.0173913	0.0173913	-
Kenya	1	0.2564103	0.2564103	0.2564103	-
Korea	64	0.000157947	-0.09128631	0.5128205	0.095994402
Lesotho	1	0.0147059	0.0147059	0.0147059	-
Malaysia	6	0.01654229	-0.02564103	0.111399	0.04894747
Nigeria	1	0.6321839	0.6321839	0.6321839	-
Pakistan	3	-0.18533668	-0.31034483	-0.04958678	0.130710478
Philippines	29	0.0682588	-0.19565217	0.4615385	0.144810764
Russia	4	0.163269225	0.0086207	0.3064516	0.151793801
South Africa	1	0.0140845	0.0140845	0.0140845	-
Thailand	20	0.125659473	-0.03174603	0.8915663	0.23944778
Turkey	36	0.229802393	-0.0625	1.1904762	0.291811076
Vietnam	103	0.284063211	-0.20454545	1.4615385	0.377889949
Grand Total	346	0.146868494	-0.31034483	1.4615385	0.282058067

Summary statistics for relative SEZ nightlights growth 2007 – 2012

Country	# SEZs	Mean	Min	Max	StdDev
Argentina	4	0.90429435	0.8692853	0.9791115	0.05073932
Bangladesh	8	0.997298475	0.7597033	1.147469	0.117731233
Chile	3	1.0263724	0.9300092	1.085249	0.084135502
China	33	0.914741112	0.7524914	1.631171	0.216931408
Colombia	6	0.915561033	0.8045153	1.09588	0.127418987
Dominican Republic	10	1.01467387	0.8772789	1.24323	0.129457635
Ghana	1	0.7246067	0.7246067	0.7246067	-
Honduras	3	0.993996233	0.9787006	1.010508	0.015938535
India	8	1.003846625	0.8582475	1.259896	0.157376007
Jordan	1	0.8321588	0.8321588	0.8321588	-
Kenya	1	0.7757092	0.7757092	0.7757092	-
Korea	64	0.978339575	0.8888901	1.479818	0.093900244
Lesotho	1	0.80386	0.80386	0.80386	-
Malaysia	6	0.811284067	0.7776183	0.8869875	0.0390641
Nigeria	1	1.424929	1.424929	1.424929	-
Pakistan	3	0.975099367	0.8254728	1.137583	0.156451821
Philippines	29	1.027987041	0.7740251	1.406441	0.139351601
Russia	4	1.06371485	0.9223014	1.194643	0.138802985
South Africa	1	0.9246849	0.9246849	0.9246849	-
Thailand	20	0.86354643	0.7427932	1.45111	0.183691788
Turkey	36	0.911232989	0.6946489	1.623053	0.216220009
Vietnam	103	1.059035955	0.6560541	2.030163	0.311666184
Grand Total	346	0.984477436	0.6560541	2.030163	0.224244824

Appendix 4 – SEZ five-year growth rate

Summary statistics for absolute SEZ five-year nightlights growth

Country	# SEZs	Mean	Min	Max	StdDev
Argentina	4	-0.07602095	-0.3050848	0.16	0.194771585
Bangladesh	7	-0.035036886	-0.375	0.75	0.39563221
Chile	0				
China	32	0.158902197	-0.1988304	0.928571	0.293642116
Colombia	3	0.241249667	0	0.511628	0.257054784
Dominican Republic	2	0.4151585	0.176471	0.653846	0.3375551
Ghana	1	0.036585	0.036585	0.036585	-
Honduras	2	-0.0161118	-0.0526316	0.020408	0.051646796
India	3	0.344177	0.083333	0.676471	0.30295544
Jordan	1	0.084112	0.084112	0.084112	-
Kenya	1	0.580645	0.580645	0.580645	-
Korea	45	0.143176873	-0.1884058	1.023256	0.254785274
Lesotho	1	0.021739	0.021739	0.021739	-
Malaysia	2	0.984837	0.318584	1.65109	0.942224029
Nigeria	1	-0.1518987	-0.1518987	-0.1518987	-
Pakistan	2	-0.1760606	-0.2121212	-0.14	0.05099739
Philippines	19	0.344530279	-0.475	1.318182	0.546229908
Russia	3	0.072497567	-0.1256983	0.270777	0.198237663
South Africa	1	0.358108	0.358108	0.358108	-
Thailand	6	0.191734317	-0.0666667	0.578125	0.252919611
Turkey	31	0.155172355	-0.4133334	0.925	0.285877618
Vietnam	85	0.597637024	-0.2441314	1.571429	0.476264011
Grand Total	252	0.315885587	-0.475	1.65109	0.440788337

Summary statistics for relative SEZ five-year nightlights growth

Country	# SEZs	Mean	Min	Max	StdDev
Argentina	4	0.876521975	0.6581042	1.077775	0.174224155
Bangladesh	7	0.987825186	0.5963211	1.83394	0.415574403
Chile	0				
China	32	0.92287325	0.6922644	1.631171	0.241013158
Colombia	3	1.222506	1.092802	1.410803	0.16690121
Dominican Republic	2	1.5301508	0.9581916	2.10211	0.808872458
Ghana	1	1.131593	1.131593	1.131593	-
Honduras	2	0.90150405	0.8243075	0.9787006	0.109172408
India	3	1.0433606	0.9444895	1.22534	0.157795732
Jordan	1	0.8414701	0.8414701	0.8414701	-
Kenya	1	1.218322	1.218322	1.218322	-
Korea	45	1.070589516	0.7813815	1.650469	0.175391175
Lesotho	1	1.027075	1.027075	1.027075	-
Malaysia	2	1.6027185	1.068876	2.136561	0.754967304
Nigeria	1	0.7914019	0.7914019	0.7914019	-
Pakistan	2	0.81882245	0.674718	0.9629269	0.203794468
Philippines	19	1.156766153	0.5583654	2.122151	0.466711879
Russia	3	0.7674731	0.6256452	0.909361	0.14185791
South Africa	1	1.368062	1.368062	1.368062	-
Thailand	6	0.94279455	0.5595368	1.162033	0.222323678
Turkey	31	0.979659939	0.6013014	1.409558	0.21887086
Vietnam	85	1.225770053	0.5948063	2.233004	0.42690143
Grand Total	252	1.093289317	0.5583654	2.233004	0.355346135