Castells-Quintana, David; Royuela, Vicente

Conference Paper

Malthus living in a slum: Urban concentration, infrastructures and economic growth

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Castells-Quintana, David; Royuela, Vicente (2014) : Malthus living in a slum: Urban concentration, infrastructures and economic growth, 54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/124448

Malthus living in a slum:
Urban concentration, infrastructures and economic growth

(Preliminary version)
August 2014

David Castells-Quintana
(AQR-IREA - Universidad de Barcelona)
dcastells@ub.edu

Abstract:
The link between urban concentration and economic growth at country level is not straightforward, as there are benefits as well as costs associated with urban concentration. Indeed, recent empirical evidence suggests different effects of urban concentration on growth depending on the level of development and on the world region under analysis. This paper revisits the literature on urban concentration and economic growth to shed some light about these previous results. In particular, differences in the process of urbanisation, and in the quality of the urban environment itself, have been suggested as most likely defining the balance between benefits and costs from urban concentration, and are probably behind differences in the its relationship between concentration and growth. However, empirical evidence on this regard remains very limited. The aim of this work is to fill this gap by paying special and explicit attention to differences between world regions in terms of urban infrastructure, essentially access to basic urban services. The main contribution of the paper is therefore to provide empirical evidence on the role that the urban environment plays in the relationship between urban concentration and economic growth.

Keywords:
Agglomeration, urbanisation, urban concentration, infrastructure, congestion diseconomies, growth, Sub-Sahara Africa

JEL classification: O1, O4, R1

Acknowledgements
I thank Marius Brülhart, Markus Brückner, Vernon Henderson and Michiel Gerritse for valuable discussions and comments. I want to specially thank Andrés Rodríguez-Pose and Vicente Royuela for all their help. I am also grateful for comments received at the 53rd European Congress of the Regional Science Association International (RSAI), at the Conference on International Development at the University of East Anglia (2013), at the XXXIX Reunión de Estudios Regionales of the Spanish Regional Science Association, at the AQR-IREA seminars-2014, and at the Economic Geography work-in-progress seminars of the London School of Economics.

1. Introduction

Today more than half of the 7 billion inhabitants of the planet live in urban areas, with this share expected to keep rising. While urbanisation has been long recognised as a fundamental element of the process of economic development, sustainable urbanisation has become one of the main and more pressing challenges for developing countries, where millions live lacking adequate access to basic services like electricity, clean water and sanitation. Building on previous evidence on urban concentration and economic growth, differentiated effects of urban concentration on national economic performance are analysed in this paper. The paper contributes to the literature by providing empirical evidence on how different characteristics of the urban environment - in particular the quality of urban infrastructure - strongly determine whether growth-enhancing benefits of growth-deterring congestion costs prevail in the process of urban concentration, something that previous studies on urban concentration and economic growth have not considered empirically. Looking at different world regions it is found that while increasing urban concentration might have been associated with growth in Asian countries, it seems that congestion diseconomies have prevailed over agglomeration benefits in Sub-Saharan African countries due to their significant deficiencies in terms of access to basic urban services.

The focus on access to basic services lies on two major reasons. The first one relates to magnitude. According to UN-Habitat Reports, today at least 1 billion people worldwide, of whom the vast majority are in the developing world, live in slums lacking access to basic services like electricity, clean water and sanitation. Growing at high rates (higher than 4.5 percent per annum in Sub-Saharan Africa) slums are expected to host 2 billion inhabitants by 2030. The second reason relates to the fact that access to basic services is expect to play a key role in the trade-off between the benefits and the costs that come with urban concentration, especially in developing countries. On the one side low coverage of basic services are likely to handicap the benefits from agglomeration (as specialization, labour pooling and knowledge diffusion) as they hinder physical and social mobility and interaction, information flow and knowledge spillovers and trust. On the other side, deficiencies in terms of access to basic services dramatically increase congestion costs for urban inhabitants in terms of transport costs, but also in terms of disease transmission, pollution, conflict and crime (most likely reducing the capacity of cities to develop and attract talent and investment).

Slums are traditionally considered as a passing phenomenon characteristic of fast-growing economies (REF), and representing a temporary stage in the structural change from rural to industrial activities. However, slums have tended to grow more in poor and stagnant countries where urbanisation and urban concentration does not seem associated with economic growth (Fay and Opal 2000; Kim 2008; Bloom et al. 2008). Indeed as the World Development Report (2011) acknowledges the growth-enhancing benefits from urban concentration, it also warns about the

risks of "rapid urbanisation" in developing countries.[1] With most of their inhabitants having been born in the slum where they live, and with their living standards hardly improving over time, slums in developing countries today are considered a form of poverty trap for a majority of their residents (Marx et al. 2013).[2] In fact, growth of large agglomerations in developing countries today is mostly given in slums, being their growth more the outcome of fast natural growth than the outcome of rural-urban migration: Christiaensen et al. (2013) report a contribution of natural increase to urban growth for 10 African countries from 1950 to 2010 of 2.9%, compared to a contribution of 1.8% due to migration. Even growth driven by migration has been more associated to push than to pull factors (Lipton 1977; Bates 1981; Bairoch 1988; Barrios et al. 2006; Swanson and Buckley 2013), with population being "expulsed" from rural areas rather than attracted to urban areas by the prospects of better living standards.[3] In this line, several authors are now referring to *Malthusian cities*, especially in SSA (Christiaensen et al. 2013; Swanson and Buckley XX).[4] With more than half of the 7 billion inhabitants of the planet living today in urban areas, it is indeed very likely that in many developing countries the Malthusian dilemma of low living standards has in some way moved from the countryside to the main urban centres, where a large proportion of urban dwellers reside under inadequate living conditions and where congestion effects of population growth are expected to dominate the positive effects from urban concentration.

The rest of the paper is organized as follows. The rest of this section sets the theoretical framework of the paper reviewing related strands in the literature and highlighting the contribution of the paper. Section 2 sets a simple economic growth model for the empirical specification to be derived. In section 3 the data used is described along some basic stylized facts. Section 4 discusses estimations and results, while in section 5 some robustness checks are performed. Section 6 focuses on Sub-Sahara African countries. Finally, section 7 concludes and derives policy implications from the results.

---

[1] The UNFPA State of World Population 2011 estimates that there are 60 million new urban inhabitants every year worldwide, most of them in the developing world. Comparing the speed of urbanization processes in Asia and Africa between 1950 and 2010, on the one hand, and in Europe between 1800 and 1910, on the other hand, Jedwab et al. (2014) conclude that developing countries today have experienced the same growth in urbanisation half the time.

[2] Marx et al. (2013) summarize evidence on living standards based on surveys carried-out in slums around the world. According to these surveys, the majority of slum residents were either born in the slum where they live or have been living there most of their live (or moved from a different slum).

[3] Even when driven urban pulled factors, expectations of high returns from moving to urban areas do not necessarily materialize and can lead to additional pressure from new incomers, as the well-know Todaro paradox describes (Todaro 1976). This will be especially true when both rural and urban incomes are close to subsistence levels, as it is the case in SSA.

[4] In a Malthusian equilibrium those societies with greater availability of resources have higher population density but living standards remain low unless productivity is sufficiently increased. Such equilibrium was the rule for most human history (See Ashraf and Galor 2011 for a modern modelling of Malthusian growth equilibrium as well as for transition dynamics out of it and into sustained growth). In its purely rural setting a Malthusian equilibrium has also been considered as a relevant possibility today for many poor countries with large rural populations and largely dependent on low-productivity agriculture and mineral exports (Weil and Wilde 2009).

1.1. Urban concentration and economic growth:

There are at least three main reasons why higher geographical concentration (due to urbanisation and urban concentration) is expected to increase productivity and economic growth: first, due to the reallocation of people and resources from agricultural activities towards industrial activities of higher productivity and value added, which takes places with urbanisation - as in classical models of structural change and economic development (Lewis 1954). Second, due to faster productivity growth linked to the clustering of people and industries and agglomeration economies, which takes place with urban concentration (Spence et al. 2009).[5] Third, due to the fact that concentration enhances economies of scale in the provision of urban infrastructure and public services (Henderson 2003).

However, empirical evidence on the effects of geographical concentration on economic growth at country level has not been conclusive. Bloom et al. (2008) find no empirical link between urbanisation and economic growth suggesting that the absence of such a link lies in the different types of urbanisation observed across countries. While in developed countries urbanisation is expected to be associated with industrialisation and the reallocation of resources to sectors of higher added value and with more growth potential that is not always the case in many developing countries. In fact, there is growing empirical evidence of urban processes in developing countries not necessarily linked to economic development (Firebaugh 1979; Ades and Glaeser 1995; Davis and Henderson 2003; Barrios et al. 2006; Gollin et al. 2014; Behrens and Pholo-Bala 2013).[6] For Africa in particular, negative effects on growth of growing urbanisation have been reported, despite increasing returns from agglomeration (Brückner 2012).[7]

Geographic concentration of economic activity not only allows for growth-enhancing agglomeration economies but it also leads to potential growth-deterring diseconomies of congestion. Moreover, both the benefits and the costs from concentration tend to become significant for large urban agglomerations. In this line several papers focus on long-run effects of urban concentration (Henderson 2003; Bertinelli and Strobl 2007; Brülhart and Sbergami 2009;

[5] Duranton and Puga (2004) and Rosenthal and Strange (2004) provide a good theoretical survey on micro-foundations of agglomeration economies - both of the Marshall type (due to localization and specialization) and of the Jacobs type (due to diversity), and an extensive review of the empirical evidence. Ottavianno and Thisse (2004) describe and explain the forces shaping the geographical distribution of economic activity. More recently, Spence et al. (2009) provide a comprehensive review linking the literature on agglomeration economies with the literature on urbanisation and growth.

[6] Firebaugh (1979) focuses on Latin America and Asia between 1950 and 1970. The rest of these papers, except for David and Henderson (2003), focus on Sub-Saharan Africa.

[7] According Brückner (2012), high ethnic fractionalization, very low economic development and excessive size of primate cities drive negative effects of growing urbanisation in Africa. Brückner suggests that the negative role of the excessive size of primate cities relates to their large squatter settlements with inadequate access to transport, water, sanitation, electricity, and health services, but he does not explicitly examines the role of these services.

Leitão 2103; Castells-Quintana and Royuela 2014).[8] Given benefits and costs, the relationship between urban primacy and economic growth is expected to be non-linear and dependent on the level of development. Hence, according to the Williamson (1965) hypothesis, while increasing urban concentration is desirable and expected in early stages of development, de-concentration eventually occurs as development proceeds. The optimal degree of urban concentration declines as development proceeds as knowledge gets accumulated, lowering the scope from agglomeration economies, and as better infrastructure allows efficient de-concentration to avoid congestion costs. Furthermore, the optimal level of urban concentration is expected to decline with the level of development also as institutional environments improve, allowing for economic growth opportunities from a more diverse urban system. Consequently, depending on their level of development, some countries experience insufficient urban primacy while others experience excessive concentration (Henderson 2003). According to Brülhart and Sbergami (2009) the beneficial net effect of high urban concentration is expected only when income levels are not too high.[9] According to Castells-Quintana and Royuela (2014) for net benefits to arise also income distribution has to remain relatively equal.[10]

1.2 From Optimal to Efficient city size:

The trade-off between the benefits and costs from agglomeration has been also studied in the literature seeking to explain the causes and limits of city growth (von Thunen 1826; Christaller 1933; Alonso 1964), and optimal city size (Mills and De Ferranti 1971; Alonso 1971; Henderson 1974), with some papers aiming to understand the dynamics of rapidly growing megacities around the world (Henderson 1985; Ades and Glaeser 1995), and especially in developing countries (Firebaugh 1979; Kasarda and Crenshaw 1991; Arku 2009; Gollin et al. 2014). Given positive and negative synergies and externalities - location costs and benefits - that cities provide, standard urban economics models predict agglomeration effects increasing with urban size to a given point from which diseconomies of scale, due to congestion, become relevant and decrease the revenue of a given city.[11] In this framework, therefore, urban scale is self-limiting, with the costs of agglomeration otherwise outweighing the benefits (Bertinelli and Black 2004). However, merely

[8] In fact, according to Henderson (2003), "urbanisation represents sectoral shifts within an economy as development proceeds, but is not a growth stimulus per se. However, the form that urbanisation takes, or the degree of urban concentration, strongly affects productivity growth" (Henderson 2003, pp. 67). Furthermore, Henderson highlights that while urbanisation is not fairly well measured across countries urban concentration (as a ratio) is, giving the focus on urban concentration measures an additional advantage over urbanisation measures.

[9] Brülhart and Sbergami (2009) rely on a standard cross-country specification, with growth in GDP per capita as dependent variable. They find a critical level of per capita GDP of US $10.000 (in 2006 prices) from which higher urban concentration becomes detrimental for growth.

[10] Urban concentration and inequality can be understood as two dimensions of concentration of resources (one being geographical while the other personal). High urban concentration along high inequality can therefore represent excessive concentration of resources expected to harm long-run growth.

[11] Thus, while optimal city size refers to the size that maximizes the difference between benefits and costs from agglomeration, the city will tend to grow to the point where benefits and costs cancel each other out.

physical size is not the only determinant of urban agglomeration economies and congestion costs (Richardson 1972). In fact, cities are different from one another - fundamentally as they perform different functions (Henderson 1974, 1985) - and generate a large variety of different externalities as a result of the qualitative characteristics of the urban production environment (Chinitz 1961; Capello and Camagni 2000). Furthermore, cities operate in different national urban systems where they interact with each other (Camagni 1993), which also determines the benefits and costs from agglomeration (Duranton and Puga 2000). Hence, the need to look not only at urban size but also at other city characteristics when analysing optimal city size has recently been highlighted. In particular, rather than focusing on optimal size, one should focus on efficient size, which depends on the functional characteristics of the city and the spatial organization within the urban system (Capello and Camagni 2000; Royuela and Suriñach 2005; Camagni et al. 2013).[12]

1.3. Different processes of urbanisation (and urban concentration) and economic growth: The role of infrastructures

If characteristics of the cities and the national urban system are relevant to define the benefits and costs from agglomeration, these characteristics must be also relevant to define the relationship between urban concentration and national economic growth. Different urban environments, for instance in terms of the quality of urban infrastructures, could indeed explain empirical evidence on relevant heterogeneity across countries in the relationship between urban concentration and growth (as for instance reported in Pholo-Bala 2009 analysing regional-specific effects by continent). Urban infrastructures are not only fundamental *per se* in the process of economic development[13] but also as they define the urban environment, leading to different capacities for cities to benefit from agglomeration economies and to control congestion diseconomies. As Henderson (2005) notes, "public infrastructure affects not just the resources devoted to urban living such as commuting and congestion costs, but also affects production efficiency - the extent to which knowledge spillovers are fully realized and exploited." Bertinelli and Black's (2004) stylized urban economics model indeed suggests an empirically testable prediction; that the growth-enhancing benefits from concentration are significantly affected by the quality of urban infrastructure affecting the urban production technology.[14] And access to basic services, as noted, is expected to play a key role, with

[12] In particular, Capello and Camagni (2000) consider three urban environments that interact with each other generating positive and negative externalities: the physical, economic and social environments. Based on these interactions they build an index for positive externalities within a city (the "city effect") and an index for negative externalities (the "urban overload") and present evidence on how the two indices depend not only on city size, but also on proxies for the type of urban functions and network integration.

[13] Straub (2011) provides a recent survey in the macro-level literature on infrastructure and development. Ayogu (2007) and Calderón and Servén (2010) focus on Africa. In an analysis for Indonesia, Lewis (2014) shows how local governments that invest more heavily in infrastructure are better to cope with the apparent detrimental effects of rapid urbanization on local economic growth.

[14] Bertinelli and Black (2004) introduce dynamic human capital externalities, along traditional congestion externalities in the urban sector, to study how urbanisation influences economic growth at country level. In this framework urbanisation enhances growth by the structural change given by the reallocation of resources,

the WDR (2011) highlighting the relevance of these basic services for the well-functioning of large cities.

While some studies provide empirical evidence on the relevance of infrastructure in economic performance of urban residents - for instance Field and Kremer (2006) focusing on access to basic services in Peru, to the best of my knowledge no paper empirically addresses in a cross-country framework the role that the urban environment plays in the relationship between urban concentration and economic growth.[15] This paper contributes to fill that gap.

2. A simple Economic Growth Model

The empirical analysis is based on a GDP per capita growth framework, following works as Henderson (2000) and Brülhart and Sbergami (2009).[16] To derive an empirical specification we depart from a neoclassical framework of economic growth basis for standard cross-country growth regressions.[17] Economic growth is related to growth due to technological progress and to the gap between the initial level of output and the steady state to which the economy convergences, with the expectation that countries with lower levels grow faster:

$$\gamma_i = t^{-1}(log\ y_i\ - log\ y_{i0}\ ) = g + \beta_i((log\ y_{i0}\ - log(y^E_{i,\infty}) - log(A_{i0})) \quad (1)$$

where γ_i is output per capita average growth rate of economy *i*, $y_{i,0}$ is initial output per capita, $y^E_{i,\infty}$ is its steady state, $A_{i,0}$ is initial efficiency level or technology, and g is the steady-sate growth rate. $\beta_i = -t^{-1}(1 - e^{-\lambda t})$, with λ reflecting the speed of convergence to the steady state. To reach to an estimating equation allowing regression analysis linear in observable variables we follow Mankiw et al. (1992), who extend the Solow (1965) Model of economic growth by introducing human capital accumulation, and assume a simple aggregate Cobb-Douglas function of output:

$$Y_{i,t} = K^{\alpha}_{i,t} H^{\phi}_{i,t} (A_{i,t} L_{i,t})^{1-\alpha-\phi} \quad (2)$$

where $Y_{i,t}$ is output, $K_{i,t}$ is physical capital, $H_{i,t}$ is human capital, and $A_{i,t}L_{i,t}$ is units of efficient labour. Capital (physical and human) is accumulated over time from savings in output (with s_{ki}

---

and through higher human capital accumulation that increases productivity. Thus, "to the extent that urbanisation encourages human capital accumulation, cities become the engines of economic growth."

[15] Sekkat (2013) studies the relationship between urban concentration, poverty and infrastructure in a cross-country setting, but looking at nation-wide, rather than urban-specific, infrastructure.

[16] While Henderson (2000) is based on a GDP per capita growth specification, Henderson (2003) focuses on TFP growth (but also estimates a GDP per capita growth model as robustness). While both analyses are similar, a GDP per capita growth specification allows for the use of a larger dataset.

[17] See Durlauf et al. (2005) for a more detailed explanation of how to derive cross-country growth regressions from neoclassical economic growth theory.

being the fraction of output invested in physical capital and s_{hi} the fraction invested in human capital). Hence, according to the model, income per capita will be:

$$log(y_{it}) = log(A_{i,0}) + gt + \frac{\alpha}{1-\alpha-\phi} log\ s_{ki} + \frac{\phi}{1-\alpha-\phi} log\ s_{hi} - \frac{\alpha+\phi}{1-\alpha-\phi} log\ n + g + \delta \quad (3)$$

where *n* is the rate of population growth, g the rate of technological progress in the steady state, and δ the rate of depreciation of capital. Substituting for the steady state in equation (1):

$$\gamma_i = g + \beta log\ y_{i0} + \beta \frac{\alpha+\phi}{1-\alpha-\phi} log\ n + g + \delta - \beta \frac{\alpha}{1-\alpha-\phi} log\ s_{ki} - \beta \frac{\phi}{1-\alpha-\phi} log\ s_{hi} - \beta log(A_{io}) \quad (4)$$

Interpreting $A_{i,0}$ in a general way not only referring to technology, assumed constant across countries, but also to country-specific factors that influence growth (resources, institutions, location and characteristic of the economic geography),[18] and therefore allowing for heterogeneity in initial conditions but also in growth paths across countries, cross-country differences in output per capita growth finally depend on initial levels of output per capita, factor accumulation and differences in these country-specific factors. In this line, we derive the standard specification of cross-country economic growth taking the following form:

$$\gamma_i = \beta\ logy_{i,0} + \psi X_{i,0} + \pi Z_{i,0} + \varepsilon_i \quad (5)$$

where $X_{i,0}$ is the standard Solow determinants (factor accumulation) plus a constant term, and $Z_{i,0}$ a vector of country-specific factors explaining cross-country differences in efficiency growth (the evolution of technology) or in initial conditions.

Introducing urban concentration as a determinant of growth:

The degree of urban concentration represents one variable that could be considered within the vector $Z_{i,0}$. The degree of urban concentration is a relevant characteristic affecting growth in efficiency (Henderson 2003), as it reflects agglomeration economies that remain unexploited, and therefore offering possibilities for growth, or that become exhausted and subject to congestion:

$$\gamma_i = \beta\ logy_{i,0} + \psi X_{i,0} + \lambda UC_{i,0} + \pi Z_{1i,0} + \varepsilon_i \quad (6)$$

where UC_{i0} is the degree of urban concentration and $Z_{1i,0}$ other remaining relevant country-specific factors. However, further taking into account that the way urban concentration affects

[18] Mankiw, Romer and Weil specifically assume $log(A_{i,0}) = log\ A + e_i$

growth in efficiency depends on specificities of the urban process influencing those possibilities of agglomeration economies or congestion diseconomies, equation (6) extends to:

$$\gamma_i = \beta \ logy_{i,0} \ + \psi X_{i,0} + \lambda_1 UC_{i,0} + \lambda_2 G_{i,0} UC_{i,0} + \ \pi Z_{1i,0} + \varepsilon_i \qquad (7)$$

where $G_{i,0}$ captures specificities of the urban process as the quality of urban infrastructure. Equation (5) is our main equation of analysis.

3. Data and Stylised Facts

3.1. Data:

To study the relationship between urban concentration and growth I rely on panel data for as many countries in the world as possible depending on data availably between 1960 and 2010, covering more countries and a longer time spam than most previous studies on urban concentration and growth. The dependent variable is national economic growth, for which data from the Penn World Tables is used. For *UC* urban primacy, as the most standard measure in the literature on urban concentration, is considered.[19] Data for primacy comes from the World Bank. For the quality of urban infrastructure several measures are considered. Following the World Development Report (2011), I focus on three key indicators: access to improved sanitation, improved water source, and electricity. As data for all these variables is scarce, when I introduce them in the analysis the panel only considers the 1990-2010 period.[20] Finally, as control variables (X_i and Z_{1i} in equation 5) I begin by considering investment, as share of GDP, fertility rates, and average years of secondary and higher education of the adult population, following Henderson (2000) specification. For urban infrastructure variables, as well as for control variables, I rely on a variety of sources. Appendix A lists variables' names, definitions and sources. For robustness a wide variety of other control variables are also considered, following Brülhart and Sbergami (2009) and the literature on cross-country economic growth. In the focus on SSA data on rainfall is used to instrument for economic growth (as explained below). Rainfall data comes from the National Aeronautics and Space Administration (NASA) Global Precipitation Climatology Projects (GPCP), as used in previous papers as Brückner and Ciccone (2011). Additionally, also for the focus on SSA, data on light

[19] Primacy measures consider main metropolitan areas (including core city and satellite cities), which is what we have in main when we think about agglomeration economies and congestion costs. It has been shown that primacy correlates very highly with other measures of concentration (as the Hirschman-Herfindahl index for which there is very limited coverage) and reflects fairly well parameters behind Zipf's law curves (the fact that when we rank cities from largest to smallest, rank times population size is approximately the same constant for all cities). The largest city in the country, therefore, delineates all other city sizes and is sufficient information to calculate any comparative index of national urban concentration (Henderson 2003).

[20] Main results and discussion focus on access to improved sanitation. According to the World Bank, sanitation remains as one of the most off-track Millennium Development Goals (MDG) globally. Access to improved sanitation not only lies at the heart of many other development challenges but the lack of it is also currently holding back economic growth in many less-developed countries. In the robustness section, I discuss results using improved water source and electricity. I further consider infant mortality rates, as a common and basic indicator of health, and access to mass urban transport systems.

density at night is used as robustness for measurement errors in income per capita. This data comes from the Defence Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) and archived by the National Oceanic and Atmospheric Administration (NOAA), and has recently used as proxy for income by several authors (Henderson et al. 2012; Mveyange 2014; Lowe 2014; Henderson et al. 2014).

3.2. Some basic stylized facts:

Before performing econometric analysis, an initial look at urban concentration patterns and economic performance worldwide during the previous decades allows us to highlight some basic but interesting stylized facts. The first of these facts is that while the proportion of urban population living in the primate city (*primacy*) has stayed relatively constant over time at around 40 per cent of urban population, there are important differences between developed and developing countries and across world regions. While the average is about 35 per cent for developed countries, it is higher than 43 for developing countries. Figure 1 shows primacy levels around the world while Table 1 presents descriptive statistics for primacy and economic growth, and the correlation between the two, as well as basic figures related to the urban environment, all for different world regions. Higher values of primacy tend to be concentrated in poorer regions of the globe (as Latin America and the Caribbean -LAC- with average close to 50 per cent, and Sub-Saharan Africa -SSA- with an average above 42). The second fact relates to the fast pace of urbanisation processes in developing countries, and especially the current growth of large agglomeration in these countries. While in 1970 large primate cities in developing countries had on average a similar size of those in developed countries (around 1.2 million inhabitants), in 2010 primate cities in developing countries had on average almost one million inhabitants more (with an average of 3.4 million) than their counterparts in developed countries.[21] The third fact relates to the heterogeneity in the correlation between urban concentration and subsequent economic growth. While there is a negative, although insignificant, correlation for the world sample (-0.03), the picture changes if we consider the correlation by level of development; urban concentration is positively correlated with growth in developed but not in developing countries. By regions the correlation is positive in Europe, Asia, LAC and North Africa, and negative in North America, Oceania and SSA. The final stylized fact, relevant for our analysis, refers to urban infrastructure and the urban environment, where we also find important heterogeneities across countries. In particular, urbanisation in many developing countries indeed appears as characterised by a large proportion of urban inhabitants living under inadequate conditions. While access to basic services was virtually universal in developed countries

[21] I calculate these world averages using World Bank data for the largest agglomeration in 193 countries worldwide and considering only countries with a total population of at least 1 million inhabitants. 150 out of these 193 agglomerations are in developing countries. Also note that averages hide high variability in size. Jakarta, Shanghai and Bombay in Asia, Lagos and Cairo in Africa, Mexico City and Sao Paulo in Latin America, are all above or close to 20 million inhabitants in their respective metropolitan areas with a population still growing at a fast pace.

already in 1990, it was not in developing countries, with important differences among them and particularly significant deficiencies in SSA. These deficiencies in SSA appear as remarkably severe in terms of access to improved sanitation and electricity and remain quite persistent (sanitation increasing since 1990 on average less than 5 percentage points and electricity around 10). Figures 2a and 2b display maps of access to improved sanitation and electricity worldwide. For access to improved sanitation while the average for Asia was close to 85 per cent, it was 44 for SSA (taking average values between 1990 and 2005). In terms of electricity the average coverage in SSA reached only half of the urban population. Similarly, in terms of infant mortality - reflecting access to health services - the average was 11 children per 1000 live births in the developed world, higher than 62 in developing countries, and exceeding 93 in SSA. In terms of transport none of the primate cities in SSA had a massive transport system by 2000.[22] In general, looking at data on urban population living in slums, we find an average of 57 per cent of urban population in developing countries, the figure reaching 77.9 for SSA. These dramatic deficiencies in SSA are not just the consequence of low income levels. As Figure 3 shows for access to improved sanitation, even controlling for income levels SSA countries present significantly lower levels of urban infrastructure.[23] Such deficiencies are in all probability hampering agglomeration benefits while raising congestion costs in African cities.

[Insert Figure 1: Population living in largest city (percentage of urban population)]

[Insert Table 1: Some basic figures]

[Insert Figure 2: Access to improved sanitation and electricity]

[Insert Figure 3: Access to improved sanitation by income levels]

4. Estimations and results

4.1. Urban concentration and economic growth in a panel of countries:

Following the literature on urban concentration and economic growth, I begin by estimating our growth equation based on cross-country panel data (for 137 countries) and without considering differentiated urban patterns across countries. I split the 1960-2010 in 5-year periods.[24] Equations of this type using panel data represent dynamic models. Estimation of these models raises some concerns: reverse causality, unobserved time-invariant country-specific characteristics, and the

[22] Lagos inaugurated a bus rapid transit system in 2008, and Accra has now planned a metro monorail project.
[23] A simple regression analysis yields highly significant lower levels of urban infrastructure for SSA countries (16 percentage points on average for sanitation) compared to other developing countries of same income per capita levels. Ghana presents a gap of almost 50 percentage points in terms of access to sanitation.
[24] I also experimented with 10-years periods in order to reduce any short-term noise from the business cycle, but at the expense of losing observations. Results using 10-years periods are very similar to those presented throughout the paper using 5-years periods.

presence of initial income as a regressor. As it is common in the empirical studies estimating these models, I estimate its dynamic model version by System-GMM, which allows dealing with some of these concerns.[25] For the focus on SSA (section 8) I complement the empirical analysis with Instrumental Variables (IV) estimations taking advantage of the exogenous variability given by rainfall data.

Table 2 presents the result for the first set of estimations of the basic growth model. Columns 1 to 4 present results for different estimation techniques.[26] Control variables have the expected sign reflecting conditional convergence, a positive effect of higher investment and educational levels and a negative effect of higher fertility rates.[27] In column 5 I introduce primacy. Results yield a positive and significant effect (although just at the 10%). But, as reviewed before, there are reasons to expect that the relationship between urban concentration and growth will vary according to the level of development. Following Henderson (2000), column 6 considers a more flexible functional form for the effect of primacy on growth; I introduce not just a linear effect of primacy but also and interaction term with initial income per capita (in logs) and another interaction term with the square of this initial income per capita. Results support the Williamson hypothesis - with a negative coefficient for primacy, a positive for its interaction with income and a negative for the interaction with the square of income (all coefficients significant at the 1%). In Figure 4 this quadratic effect of primacy on growth, depending on income levels, is plotted. At very low levels of development the effect of primacy is negative. It then becomes positive and increasing as income rises up to income levels around $9500 per capita (in PPP converted, at 2005 constant prices) to then start declining. Finally, I take into account the possibility of significant differences across world regions. As column 7 shows, while there seems to be a positive relationship between primacy and growth for the world sample, there is a significantly different relationship for LAC and SSA.[28]

[25] Both Henderson (2003), using first-differences GMM, and Brülhart and Sbegami (2009), using system-GMM, rely on GMM estimations and provide a good explanation on the suitability of these methods for cross-country data on urban concentration and economic growth. In particular, system-GMM (Blundell and Bond, 1998) estimates are expected to be more efficient than any other dynamic GMM estimators, especially when the coefficient of the lagged dependent variable is close to one and the between sample variance is large compared to the within sample variance (as is the case here). For GMM estimations I present standard AR(1), AR(2) and Hansen tests for validity of internal instruments. As Bazzi and Clemens (2013) note, there is yet no reliable and straightforward test for the strength of the instrument set in Sys-GMM estimations. Correlation analysis of our key variables, nevertheless, reveals substantial explanatory power for lagged differences to explain levels and for lagged levels to explain first differences.

[26] OLS, Fixed Effects -FE-, GMM and System GMM -SysGMM- results are presented to ease comparison with previous literature, but I focus throughout the paper on SysGMM results (and IV estimations for SSA).

[27] I also calculate the annual speed of convergence to ease comparability of results with previous papers. The values found are within the range of what is commonly found in the literature, although differing depending on the estimation technique considered.

[28] In fact, when I analyse urban concentration by the different world regions, its effect on growth seems to have been positive and significant only in Asia and Europe. When distinguishing between developed and developing countries, rather than between world regions, while linear effects of primacy are only positive and significant in the former countries, it is in developing countries where the evidence of the Williamson hypothesis is clearer (in line with Bertinelli and Strobl 2007).

[Insert Table 2: Urban concentration and growth in a panel of countries]

[Insert Figure 4: The Williamson hypothesis]

As previous studies have suggested, the panel results confirm the relevance of urban concentration in the economic growth process. But in line with these studies results also confirm the fact that the sign and the form of the relationship are not uniform, as there are benefits, as well as costs, from urban concentration that change with country's characteristics. The relationship is likely to be nonlinear, dependent on country's level of development (the Williamson hypothesis). But the relationship may also be influenced by other factors as heterogeneous results by world regions suggest.

4.2. Positive and negative effects of urban concentration depending on the urban process:

As noted before, the relationship between urban concentration and growth is likely to differ across different urban processes. In this line, we have seen that the quality of urban infrastructure might be fundamental to unleash positive synergies from agglomeration economies or to increase congestion costs, in both cases affecting national productivity.

In Table 3 I present results for estimates of equations like (7), letting the effect of urban concentration to depend on the quality of urban infrastructure. Results are presented using access to improved urban sanitation facilities (*sanitation*) as a proxy for the quality of urban infrastructure.[29] The coefficients for both the direct effect of urban concentration and for its interaction with sanitation are highly significant under OLS (column 1), being negative the first and positive the second. Results are less significant when I estimate by FE (column 2) or SysGMM (column 3). However, as noted in the descriptive analysis, the quality of urban infrastructure substantially differs between developed and developing countries. Accordingly, in columns 4 and 5 I split the sample between developed and developing countries. SysGMM results are now non-significant for developed countries but they are highly significant for developing countries.

The absence of enough variability between developed countries in the variables considered for urban infrastructure could explain their non-significance. As we have seen, access to basic services is very high and quite homogenous among developed countries. However, there is much higher heterogeneity among developing countries, with some of them reaching developed world figures but other lagging behind and with less than half of urban population having access to these services. In the case of developing countries results suggest that while for low levels of sanitation urban concentration is negative associated with economic growth, the association becoming

[29] Below I discus some results (presented in the appendix) using other proxies for the quality of urban infrastructure.

positive as access to sanitation increases.[30] Hence, urban concentration becomes positively associated with growth only when basic services spread to the majority of the urban population.

[Insert Table 3: Urban concentration depending on the urban process]

5. Robustness checks

We can check the robustness of the results found in several additional ways. In first place one could worry that the positive effect of the interaction between primacy and sanitation is due to the fact that higher sanitation is correlated with higher income levels (where urban concentration could have more beneficial effects). Nevertheless, as column 6 of Table 3 shows, main results for developing countries hold when we introduce an interaction between urban concentration and income levels. Likewise, results hold as we control for the Williamson hypothesis, introducing interactions with income levels and their square (column 7 of Table 3). Results are also robust to other regional differences in the relationship between urban concentration and economic growth beyond differences in urban infrastructure (in column 8 the effect of urban concentration is let to vary across world regions). This last result suggests that regional effects do not, therefore, drive the significance of infrastructure in determining the net benefits from urban concentration. A second concern might come from the proxy for urban infrastructure. While access to sanitation is a good proxy and very pertinent for the analysis, there could be different contexts in which the role of other urban infrastructures might be more relevant, for example transport infrastructure (mobility and transport costs being a central issue of congestion analysis in the urban economics literature). In this line, and to expand the analysis, I replicate some of the estimations using other variables for the quality of urban infrastructure. On one side Appendix B presents panel results for access to improved water source (*water*) and access to electricity (*electricity*). Results are non-significant for access to water, but they are for access to electricity. On the other side Appendix C presents some cross-section results. Cross-section analysis is more common in the long-run economic growth literature and, as discussed before, allows us to consider other variables, as transport systems for which there is not enough time variation. Cross-section results for sanitation are in line with panel results. Results also hold when other variables are considered, as *electricity* or *transport_systems*, although the significance is reduced and depends on the controls used.[31] When a composite

[30] I also obtain similar results when considering *growth* in urban concentration and *growth* in sanitation rather than their levels.

[31] Following Brülhart and Sbergami (2009), our cross-section controls expand to include 18 variables found to be robustly associated with long-run growth by Sala-i-Martin et al. (2004) along population growth rate, higher education, fertility, investment share, and population density - to further capture agglomeration between countries. As in the panel analysis, when I analyse by world regions cross-section estimations yield a positive relationship between urban concentration and long-run growth (1990-2010) for Asia, while negative and highly significant for SSA (being robust to all considered controls).

measure for urban infrastructure is considered, rather than just one indicator, estimations yield highly significant results (and robust to all the considered controls).[32]

6. Sub-Saharan Africa

One might still have concerns about reverse causality from growth to primacy and to the quality of urban infrastructure. SysGMM estimations are expected to address endogeneity concerns. However, SysGMM estimations rely on internal instruments (i.e. variables' transformations and lags). Good external instruments for primacy and for the quality of the urban infrastructure are hard to find. Yet, we can find reliable external instruments for economic growth, at least for Sub-Sahara African countries, which give us an additional methodological advantage. Besides, given the particular deficiencies in urban infrastructure and poor performance in terms of economic growth, the focus on Sub-Sahara Africa is interesting in itself. Being still relatively dependent on agriculture and agricultural-dependent activities, economic growth in SSA countries is significantly determined by rainfall.[33] Following Brückner (2012; 2013), I exploit this exogenous variation to construct instrumental variables that allows us to purge the possible effect that economic growth might have on our key variables, urban concentration and sanitation (reverse causality). The use of exogenous instruments allows us to control for simultaneity bias concerns in a more direct way, alternative to SysGMM and without having to rely on internal instruments. Hence, in a first step primacy and sanitation are estimated on economic growth by Two-Stage Least Squares (2SLS) using rainfall and rainfall squared as instruments for growth:

$$UC_{it} = \alpha\gamma_{it-1} + a_i + b_t + \varepsilon_{it} \quad (8)$$

$$G_{it} = \alpha\gamma_{it-1} + a_i + b_t + \varepsilon_{it} \quad (9)$$

where a_i are country fixed effects and b_t are year fixed effects. The introduction of country fixed effects allows us to control for time-invariant country-specific omitted variables, while the introduction of year fixed effects allows us to control for global shocks. Appendix D presents first-stage OLS estimation for growth on rainfall and rainfall squared, and 2SLS estimations of equations (8) and (9). Rainfall (and its square) appears as highly significantly to explain variation in economic growth in SSA. By construction the residual variation on primacy and sanitation from our two-stage least squares estimations of (8) and (7) capture any variation in these variables that is not due to

[32] I simply create a composite measure standardizing *sanitation*, *water*, *electricity* and *transport_systems*, and aggregating them with equal weight.

[33] Higher levels of rainfall are expected to increase agricultural productivity and therefore economic growth in these countries. One should also considered rainfall squared, as too much rainfall can lead to floods detrimental for agriculture. See Miguel et al. (2004), Brückner and Ciccone (2011) and Brückner (2012) for more on the significance of rainfall as an exogenous variable determining economic growth in SSA countries. There is also a relatively recent and increasing literature on the effects of decreasing long-term trends of rainfall, associated with climate change, in Sub-Saharan Africa (see for instance Barrios et al. 2006).

economic growth. In a second step I use these residual variations in primacy and sanitation, as instruments for actual primacy and sanitation, to estimate by 2SLS our economic growth equation (equation 7) for SSA. Appendix E provides a formal proof for why this IV strategy using residual variation can properly address simultaneity bias as long as one has good instruments for the dependent variable (in our case rainfall, and its square, as instruments for economic growth). Table 4 presents the results.[34]

[Insert Table 4: Urban concentration and growth in SSA]

Similar to results in column 7 of Table 2, IV results yield a negative and highly significant coefficient for primacy in SSA (column 1 of Table 4) and in line with Brückner (2012). Regarding the role of urban infrastructure, IV results - column 2 of Table 4 - are also similar to those in Table 3 (with a negative coefficient for primacy and a positive for its interaction with sanitation, both being highly significant). Coefficients are robust to the considered controls as to the introduction of an interaction term between urban concentration and income levels (column 3).[35] They are also highly significant if access to improved water source (column 4) or access to electricity (column 5), rather than sanitation, are considered.

Recent literature has highlighted potential measurement error of income per capita in poor countries, especially sub-Saharan African ones. As I final robustness check I use data on light density at night to measure income (as proposed by Henderson et al. 2012). In Table 5 FE and IV estimations of equation (7) for SSA are replicated using light density at night (as aggregated at the national level by Henderson et al. 2012) and divided per population to proxy for income per capita. Results for primacy and for its interaction with urban infrastructure remain significant. Interestingly, the effect of our interaction term is even larger, while the coefficients for primacy and sanitation remain almost exactly of the same magnitude as those in Table 4.

[Insert Table 5: Estimates for SSA using nigh density at night data]

In sum, results confirm - in this case for SSA - the role of urban infrastructure when it comes to analyse the relationship between urban concentration and economic growth. According to estimates, for urban concentration to have a positive impact on growth access basic services must

[34] Standard tests confirm, on one hand, the relevance and validity of rainfall and its square as instruments for growth in our regressions for primacy or sanitation. Angrist-Pischke F tests and Hansen J tests are reported respectively in Appendix 4. On the other hand, tests also confirm the relevance of the residual variation in primacy and sanitation, once the reverse causality from growth has been removed, as instruments for actual primacy and sanitation in the growth equation. Kleibergen-Paap F and LM tests are reported in Table 4.

[35] We obtain similar results regardless of the estimation technique: OLS, FE or SysGMM. Results also hold if we consider a role for ethnic polarization. As suggested by Brückner (2012), important ethnic divisions increase the severity of negative externalities in urbanised areas. Result available upon request.

at least cover 70 per cent of urban population. But access to basic services is still very deficient in SSA, as we have seen. Only 3 countries out of 34 reached that 70 per cent threshold of urban population with access to sanitation in 1990 (Djibouti, Mauritius and South Africa), three more countries in 2005 (Angola, Botswana and Seychelles).

7. Conclusions and policy implications

Urban concentration plays an important role in the process of economic development. But there are wide heterogeneities across countries in terms of urban processes and urban environments. One aspect of the urban environment that is critical when analysing the relationship between urban concentration and economic growth is the quality of urban infrastructure. The data analysed in this paper indeed reflects important differences across countries in terms of access to basic public services, especially in the developing world. The econometric results provide evidence on the relevance of these differences to explain diverse results found in the literature in what refers to the effect of urban concentration in different regions of the world. The role of access to basic services seems robust to a long list of controls and econometric techniques. In this regard, it has been analysed how urban concentration can be negatively associated with national economic growth under urban environments with deficient urban infrastructure. This situation seems common in Sub-Saharan Africa, where access to improved sanitation and electricity appear as especially deficient and currently hampering structural change as well as the net benefits from urban concentration.

In this line, for large agglomerations in developing countries today a *Malthusian trap* might be a relevant reality, as population growth in these agglomerations exceeds the supply of resources (understood here as urban infrastructure), leading to congestion costs exceeding the benefits from agglomeration. Regarding policy implications previous works have suggested that when urban congestion is due to natural increase rather than due to migration, as seem to be the case for large agglomerations in Sub-Saharan Africa, investments in urban infrastructure are fundamental (Jedwab et al. 2014). Access to basic services, in particular, is not just desirable per se in terms of quality of life for urban residents but also in terms of capital accumulation as well as in terms of economic efficiency at national level, as they allow for the realization of agglomeration economies and the control of congestion costs. Consequently, guaranteeing that adequate urban infrastructure in these large cities (as in all urban areas) keeps pace with their rapid increase in population not only improves living conditions but can also induce a transition away from Malthusian dynamics.

Results suggest that the net benefits from agglomeration can arise in places where that is not the case today if efforts are made to improve the quality of the urban environment, and it should not be different in Sub-Saharan Africa. According to the results provided, the negative effects of urban concentration that the literature has implied in this region can be associated precisely with its severe

lack of adequate basic infrastructure. But as in other regions, improvements in urban infrastructure, leading to upgraded urban environment, can also unleash agglomeration economies while helping control congestion costs in Sub-Sahara African countries. In other words, the lower economic performance of Sub-Saharan Africa can be in part explained by hampered agglomeration economies due to deficient urban infrastructures. Clearly, further research on urban patterns could be of great value to better understand the relationship between urban concentration and national economic performance, an issue of major relevance for developing countries today.

References:

Ades, A. and Glaeser, E. (1995). 'Trade and circuses: Explaining urban giants,' *The Quarterly Journal of Economics* 110(1): 195-227.

Alonso, W. (1964). *Location and land use: Towards a general theory of land rents*. Cambridge, MA: Harvard University Press.

Alonso, W. (1971). 'The economics of urban size,' *Papers and proceedings of the Regional Science Association* 26: 67-83.

Arku, G. (2009). 'Rapidly growing African cities need to adopt smart growth policies to solve urban development concerns,' *Urban Forum* 20: 253-270.

Ayogu, M. (2007). 'Infrastructure and economic development in Africa: A review,' *Journal of African Economies* 16(10): 75-126.

Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present*. Chicago: The university of Chicago Press.

Barca, F., McCann, P., and Rodríguez-Pose, A. (2012). 'The case for regional development intervention: Place-based versus place-neutral approaches,' *Journal of Regional Science* 52(1): 134-152.

Barrios, S. Bertinelli, L., and Strobl, E. (2006). 'Climate change and rural-urban migration: The case of Sub-Saharan Africa,' *Journal of Urban Economics* 60: 357-371.

Bates, R. (1981). *Market and States in Tropical Africa: The Political Basis of Agricultural Policies*. Berkeley: University of California Press.

Bazzi S, Clemens M A (2013) Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth. *American Economic Journal: Macroeconomics* 5(2): 152-186.

Behrens, K. and Pholo Bala, A. (2013). 'Do rent-seeking and interregional transfers contribute to urban primacy in Sub-Saharan Africa?' *Papers in Regional Science* 92: 163-196.

Bertinelli, L., and Black, D. (2004). 'Urbanization and growth,' *Journal of Urban Economics* 56: 80-96.

Bertinelli, L., and Strobl, E. (2007). 'Urbanisation, Urban Concentration and Economic Development,' *Urban Studies*, 44(13): 2499-2510.

Bloom, D. E., Canning, D., and Fink, G. (2008). 'Urbanization and the wealth of nations,' *Science (New York, N.Y.)*, 319(5864): 772-5.

Blundell, R., and Bond, S. (1998). 'Initial conditions and moment restrictions in dynamic panel data models,' *Journal of Econometrics* 87(1): 115-143.

Brückner, M. (2013). 'On the simultaneity problem in the aid and growth debate,' *Journal of Applied Econometrics* 28: 126-150.

Brückner, M. (2012). 'Economic growth, size of the agricultural sector, and urbanization in Africa,' *Journal of Urban Economics* 71: 26-36.

Brückner , M., and Ciccone, A. (2011). 'Rain and the democratic window of opportunity,' *Econometrica* 79: 923-947.

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

Calderón, C., and Servén, L. (2010). 'Infrastructure and economic development in Sub-Saharan Africa,' *Journal of African Economies* 19(1): 13-87.

Camagni, R. (1993). 'From city hierarchy to city network: reflections about an emerging paradigm,' in T. Lashmanan and P. Nijkamp (Eds.), Structure and change in the space economy: 66-90.

Camagni, R., Capello, R., and Caragliu, A. (2013). 'One or infinite optimal city sizes? In search of an equilibrium size for cities,' *Annals of Regional Science* 51: 309-341.

Capello, R., and Camagni, R. (2000). 'Beyond optimal city size: An evaluation of alternative urban growth patterns,' *Urban Studies* 37: 1479.

Castells-Quintana, D., and Royuela, V. (2014). 'Agglomeration, inequality and economic growth,' *Annals of Regional Science* 52(2): 343-366.
Chinitz, B.J. (1961). 'Contrast in agglomeration: New York and Pittsburgh,' *American Economic Review* 51: 279-289.
Christaller, W. (1933). *Die Zentralen Orte in Suddeuschland.* Jena: Gustav Fischer Verlag.
Davis, J., and Henderson, V. (2003). 'Evidence on the political economy of the urbanization process,' *Journal of Urban Economics* 53: 98-125.
Duranton, G. and D. Puga. (2004). 'Micro-Foundations of Urban Agglomeration Economies,' in J.V. Henderson and J-F Thisse (Eds), *Handbook of Urban and Regional Economics*, Edition 1, Vol. 4: 2063-2117.
Duranton, G. and D. Puga. (2000). 'Diversity and specialization in cities: why, where and when does it matter?' *Urban Studies* 37: 533-55.
Durlauf, S., Johnson, P. and Temple, J. 2005. 'Growth Econometrics,' in Philippe Aghion and Steven Durlauf (Eds.), *Handbook of Economic Growth*, Elsevier: 255-677.
Easterly, W., and Levine, R. (1997). 'Africa's growth tragedy: Policies and ethnic divisions,' *Quarterly Journal of Economics* 112(4): 1203-50.
Fay, M. and Opal, C. 2000. Urbanization without growth: a not-so-uncommon phenomenon. World Bank Policy Research Working Paper 2412.
Field, E. and Kremer, M. (2006). 'Impact evaluation for slum upgrading interventions.' *Doing impact evaluation series #3*, World Bank, July 2006.
Firebaugh, G. (1979). 'Structural determinants of urbanization in Asia and Latin America, 1950-1970,' *American Sociological Review* Vol. 44, No. 2: 199-215.
Gallup, J., Sachs, J., and Mellinger, A. (2001). 'Geography Datasets,' *Center for International Development at Harvard University (CID).*
Gollin, D., Jedwab, R., and Vollrath, D. (2014). 'Urbanization with and without industrialization,' Institute for International Economic Policy Working Papers 2014-01.
Henderson, J. V. (1974). 'Optimum city size: The external diseconomy question,' *Journal of political Economy*, Vol. 82(2): pp. 373-88.
Henderson, J.V. (1985). *Economic Theory and the cities.* Orlando, FL: Academic Press.
Henderson, J.V. (2000). 'The effects of urban concentration on economic growth,' *NBER Working Papers Series* No. 7503.
Henderson, J.V. (2003). 'The urbanization process and economic growth*:* The so-what question,' *Journal of Economic Growth* 8: 47-71.
Henderson, J.V., (2005). 'Urbanization and Growth,' in Philippe Aghion and Steven Durlauf (Eds.), Handbook of Economic Growth, Vol. 1B: 1543-1591.
Henderson, J.V., Storeygard, A., and Weil, D. (2012). 'Measuring economic growth from outer space,' *American Economic Review* 102(2): 994-1028.
Henderson, J.V., Storeygard, A., and Deichmann, U. (2014). '50 years of urbanization in Africa,' *World Bank Policy Research Working Paper* 6925.
Jedwab, R. Christiansen, L., and Gindelsky, M. (2014). Rural push, urban pull and… urban push? New historical evidence from developing countries. Institute for International Economic Policy Working Paper Series 2014-4.
Kasarda, J., and Crenshaw, E. (1991). 'Third World Urbanization: Dimensions, Theories and Determinants,' *Annual Review of Sociology* 17: 467-501.
Kim, S. (2008). 'Spatial Inequality and Economic Development: Theories, Facts and Policies.' Working Paper No. 16, *Commission on Growth and Development.*
Leitão, N. (2013). 'A panel data modeling of agglomeration and growth: cross-country evidence,' *Theoretical and Empirical Researchers in Urban Management* 8(1): 67-77
Lewis, B. (2014). 'Urbanization and economic growth in Indonesia: good news, bad news and (possible) local government mitigation,' *Regional Studies* 48(1): 192-207.
Lipton, M. (1977). *Why Poor People Stay Poor: Urban Bias in World Development*. Cambridge: Harvard University Press.
Lowe, M. (2014). Rail revival in Africa? The impact of privatization. Paper presented at the joint RES-SPR Conference on "Macroeconomic Challenges Facing Low-Income Countries."
Malthus, T. (1803). *An Essay on the Principle of Population.* London
Mankiw, G., Romer, D., and Weil, D. (1992). 'A contribution to the empirics of economic growth,' *The Quarterly Journal of Economics* 107(2): 407-437.
Marx, B., Stoker, T., and Tavneet, S. (2013). 'The economics of slums in the developing world,' *Journal of Economic Perspectives* 27(4): 187-210.
Miguel, E., Satyanath, S., and Sergenti, E. (2004). 'Economic shocks and civil conflict: and instrumental variables approach,' *Journal of Political Economy* 112(41): 725-753.
Mills, E., and De Ferranti, D. (1971). 'Market choices and optimum city size,' *The American Economic Review* Vol. 61, No. 2: 340-345.

Mveyange, A. (2014). Night lights and regional income inequality in Africa. Unpublished work.
Ottaviano, G., and Thisse, J-F. (2004). 'Agglomeration and Economic Geography,' in J.V. Henderson and J-F Thisse (Eds), *Handbook of Urban and Regional Economics*, Edition 1, Vol. 4: 2563-2608.
Phola-Bala, A. (2009). 'Urban concentration and economic growth: checking for specific regional effects, *CORE Discussion Paper #2009/38*.
Richardson, H.W. (1972). 'Optimality in city size, systems of cities and urban policy: a skeptic's view,' *Urban Studies*: 29-48.
Rosenthal, S. and Strange, W. (2004) 'Evidence on the Nature and Sources of Agglomeration Economies,' in J.V. Henderson and J-F Thisse (Eds), *Handbook of Urban and Regional Economics*, Edition 1, Vol. 4:
Royuela, V., and Suriñach, J. (2005). 'Constituents of quality of life and urban size, ' *Social Indicators Research* 74: 549-572.
Sala-i-Martin, X., Doppelhofer, G., and Miller, R.I. (2004). 'Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach,' *American Economic Review* 3: 813-835.
Sekkat, K. (2013). 'Does infrastructure mitigate the effect of urban concentration on poverty in developing countries?' *Economic Research Forum Working Paper Series*, WP No. 800.
Solow, R. (1965). 'A contribution to the theory of economic growth,' *Quarterly Journal of Economics* 70(1): 65-94.
Spence, M., Clarke, P., and Buckley, R.M. (Editors). (2009). *Urbanization and growth*. Commission on growth and development.
Straub, S. (2011). 'Infrastructure and development: A critical appraisal of the macro-level literature,' *Journal of Development Studies* 47(5): 683-708.
Swanson, H. and Buckley, R. (2013). Are we creating Malthusian cities in Africa? Conference poster.
Todaro, M. (1969). 'A model of labor migration and rural unemployment in less developed countries,' *American Economic Review* 59(1): 138-148.
United Nations Department of Economic and Social Affairs. (2010). *World Population Prospects*. UN DESA Press.
Weil, D., and Wilde, J. (2009). 'How relevant is Malthus for economic development today?' *The American Economic Review* Vol. 99, No. 2: 255-260.
Williamson, J. (1965). 'Regional inequality and the process on national development,' *Economic Development and Cultural Change* 4: 3-47.
World Bank. (2011). *World Development Report 2011*.

Appendix A: Variables' names, definitions and sources:

Basic growth model	Description	Source
growth	Cumulative annual average per capita GDP growth rate	Constructed with data from PWT 7.1 (Summers and Heston), using real GDP chain data (rgdpch)
primacy	Population living in largest city (percentage of urban population)	World Bank - World Development Indicators
ln(rgdpch)	Per capita GDP (in logs)	Constructed with data from PWT 7.1 (Summers and Heston), using real GDP chain data (rgdpch)
ki	Investment share (percentage of GDP)	PWT 7.1. (Summers and Heston)
fertility	Fertility rates	World Bank - World Development Indicators
schooling23	Average years of secondary and tertiary schooling of adult population	Barro and Lee dataset
Further controls		
primary_edu	Percentage of primary schooling attended in total population	Barro and Lee dataset
higher_edu	Percentage of higher schooling attended in total population	Barro and Lee dataset
pi	Price level of investment	PWT 7.1. (Summers and Heston)
kg	Government consumption (percentage of GDP)	PWT 7.1. (Summers and Heston)
openk	Openness	PWT 7.1. (Summers and Heston)
life_exp	Life expectancy at birth	World Bank - World Development Indicators
dens65c	Density in coastal regions. 1965	Gallup et al. (2001)
tropicar	Proportion of population living in tropical areas	Gallup et al. (2001)
malfal66	Malaria	Gallup et al. (2001)
elf60	Ethno linguistic fractionalization	Easterly and Levine (1997)
buddha	Fraction of Buddhist	Sala-i-Martin et al. (2004). (BACE dataset)
confuc	Fraction of Confucian	Sala-i-Martin et al. (2004). (BACE dataset)
east	Dummy for East Asian countries	Sala-i-Martin et al. (2004). (BACE dataset)
laam	Dummy for Latin American countries	Sala-i-Martin et al. (2004). (BACE dataset)
mining	Percentage of GDP in mining	Sala-i-Martin et al. (2004). (BACE dataset)
muslim00	Fraction of Muslim	Sala-i-Martin et al. (2004). (BACE dataset)
safrica	Dummy for Sub-Sahara African countries	Sala-i-Martin et al. (2004). (BACE dataset)
spain	Dummy for Spanish colony	Sala-i-Martin et al. (2004). (BACE dataset)
pop_density	Population density	World Bank - World Development Indicators
pop_growth	Population growth rate	Constructed with data from PWT 7.1 (Summers and Heston), using data on population
Urban infrastructure		
sanitation	Population with access to improved sanitation facilities (percentage of urban population)	World Bank - World Development Indicators
water	Population with access to improved water source (percentage of urban population)	World Bank - World Development Indicators
electricity	Access to electricity (percentage of urban population)	World Bank - Sustainable Energy for All database
transport_systems	Dummy variable indicating if primate city has a massive transport system (metro, tram or rapid bus)	Constructed by the authors
telephones	Telephone lines (per 1000 inhabitants)	World Bank - World Development Indicators
infant mortality	Infant mortality rates (per 1000 births)	World Bank - World Development Indicators
slums	Population living in slums (percentage of urban population)	UN-Habitat
rainfall	Annual rainfall aggregated at the country level	Global Precipitation Climatology Projects (GPCP)

Appendix B: System GMM results with *water* and *electricity*:

	(1) G=*water*	(2) G=*water*	(4) G=*electricity*	(4) G=*electricity*
	World	Developing	World	Developing
Dependent variable:	growth	growth	growth	growth
UC	0.0256	-0.0536	-0.0224**	-0.0183**
	(0.0455)	(0.0519)	(0.0097)	(0.0081)
G	0.0136	-0.0209	-0.0144**	-0.0057
	(0.0228)	(0.0120)	(0.0072)	(0.0054)
UC*G	-0.0003	0.0005	0.0003**	0.0002**
	(0.0005)	0.0005)	(0.0001)	(0.0001)
Year FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Observations	497	347	540	374
Number of countries	129	91	137	95
AR(1) p-value	0.071	0.087	0.029	0.050
AR(2) p-value	0.203	0.276	0.187	0.179
Hansen test p-value	0.180	0.271	0.118	0.068

Note: Controls include *ln(rgdoch)*, *ki*, *fertility* and *schooling23*. All controls are calculated as averages over 5 years except *ln(rgdoch)* and *schooling23*, which are measured at the beginning of each period. Estimation done by SysGMM. *ln(rgdoch)*, *ki*, *fertility*, *schooling23*, *UC*, *G* and *UC*G* are treated as endogenous using lagged values between 2 and 4 periods as instruments for first differences and variables in first differences lagged between 2 and 4 periods as instruments for variables in levels. Estimations are done with small sample correction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix C: Cross-section results:

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	G=*sanitation*		G=*electricity*		G=*transport_systems*		G=composite	
Dependent variable:	growth	growth	growth	growth	growth	growth	growth	growth
UC	-0.0070***	-0.0070***	-0.0025*	-0.0031**	-0.0004	-0.0033***	-0.0067***	-0.0053***
	(0.0026)	(0.0020)	(0.0014)	(0.0013)	(0.0009)	(0.0011)	(0.0019)	(0.0018)
G	-0.0020	-0.0006	0.0004	0.0009	-0.0532	-0.0880	-0.0019	0.0016
	(0.0015)	(0.0011)	(0.0009)	(0.0012)	(0.0572)	(0.0714)	(0.0012)	(0.0018)
UC*G	0.0001***	0.0001**	0.0001**	0.0000	0.0013	0.0039*	0.0001***	0.0001**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0014)	(0.0020)	(0.0000)	(0.0000)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Further controls	NO	YES	NO	YES	NO	YES	NO	YES
adj R square	0.294	0.637	0.291	0.609	0.231	0.611	0.316	0.674
Obs.	112	87	129	93	129	93	107	84

Note: *growth* is here calculated as cumulative annual average per capita GDP growth rate between 1990 and 2010. In column 7 and 8 composite is calculated combining *sanitation*, *water*, *electricity* and *transport_systems*. Controls include *ln(rgdoch)*, *ki*, *fertility* and *schooling23*. Further Controls include: *primary_edu, higher_edu, pi, kg, yrsopen, life_exp, dens65c, tropicar, malfal66, elf60, buddha, confuc, east, laam, mining, muslim, safrica, spain, pop_dens, ki, fertility* and *pop_growth*. All right-hand variables are measured at the beginning of the period or closest year. Estimations are done by OLS. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix D: First step estimations for SSA:

	(1) OLS	(2) 2SLS	(3) 2SLS
Dependent variable:	growth	primacy	sanitation
rainfall	0.0028**		
	(0.0013)		
rainfall squared	-0.0001***		
	(0.0000)		
growth		-4.1597	-0.4570
		(2.9154)	(1.3064)
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Observations	143	178	143
Number of countries	38	38	38
First-stage F-stat p-value		0.000	0.005
Angrist-Pischke F stat p-value		0.053	0.093
Hansen J stat p-value		0.730	0.944

Note: Columns 2 and 3 use *rainfall* and *rainfall_squared* as instruments for *growth*. 2SLS estimations are done with small sample correction. Angrist-Pischke F tests the significance of excluded instruments. Hansen J tests the null hypothesis of valid instruments. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix E: Adjusting for simultaneity bias, formal proof:

Building on Bruckner's (2013), this appendix briefly formalizes how simultaneity bias can be properly addressed by IV estimations using residual variation in urban concentration and in urban infrastructure that is not driven by economic growth. We start by assuming a possible simultaneous equation model:[36]

$$\gamma = \lambda_1 \ UC \ + u \tag{E.1}$$
$$UC = \lambda_2 \ \gamma \ + e \tag{E.2}$$

where γ is growth and UC is urban concentration. We are interested in the coefficient λ_1. However, if $\lambda_2 \neq 0$ then OLS estimates of λ_1 from equation (a.1) will be biased. Nevertheless, if we can consistently estimate λ_2 in equation (a.2) we can construct a series for UC that is adjusted for the endogenous response (i.e. $res(UC) = UC - \lambda_2(\gamma)$) and use $res(UC)$ as an instrument for actual UC in equation (a.1) to estimate λ_1. The instrumental variables (IV) estimate of λ_1 will not suffer from simultaneity bias:

$$\lambda_1^{IV} = \frac{cov(res\ UC\ ,\gamma)}{cov(res\ UC\ ,UC)} = \lambda_1 + \frac{cov(res\ UC\ ,u)}{cov(res\ UC\ ,UC)} = \lambda_1 + \frac{cov(e,u)}{cov(e,UC)} \tag{E.3}$$

Consistent estimate of λ_2 can only be obtained if one has a valid instrument for γ in equation (a.2) (OLS can not yield a consistent estimate of λ_2 if $\lambda_1 \neq 0$ in equation (a.1)). In our case rainfall, and its square, provide these valid instruments for growth.[37]

In fact, we can identify the adjustment in λ_1 when addressing for simultaneity bias in our growth equation. The first stage estimation, in which actual UC is regressed on $res(UC)$, is:

$$UC = \eth \ res \ UC \quad + v = \eth \ UC - \lambda_2 \gamma \ + v \tag{E.4}$$

the residuals from this stage being:

$$v = UC - \eth(UC - \lambda_2 \gamma) \tag{E.5}$$

We can introduce v as an additional control in our growth equation and estimate by OLS - control function approach (the estimate for λ_1 will be the same than λ_1^{IV} - see Wooldridge 2010 for the equivalence between IV and control function approach estimates in linear models):

$$\gamma_{it,t+1} = \beta \ logy_{it} \ + \psi X_{it} + \lambda_1 UC_{i,t} + \ \theta v + \ \pi Z_{1it} + \varepsilon_{it} \tag{E.6}$$

which equals to:

$$\gamma_{it,t+1} = \beta \ logy_{it} \ + \psi X_{it} + (\lambda_1 + \theta(1 - \eth)) UC_{it} + \theta \lambda_2 \eth \gamma_{it-1,t} + \pi Z_{1it} + \varepsilon_{it} \tag{E.7}$$

and where $\theta(1 - \eth)$ will be the adjustment for simultaneity bias done to an estimate of λ_1 in which simultaneity bias was not addressed (i.e. direct OLS without v). As it can be seen, the adjustment depends on θ, which indicates the role of past growth in explaining current growth, and also on $(1 - \eth)$, which captures the share on the variation of UC that is due to economic growth. If any of the two components, θ or $(1 - \eth)$, is cero then the estimate reduces to the direct OLS estimate (no simultaneity bias).

[36] We formalize here the procedure to adjust for simultaneity bias between primacy and growth. An equivalent procedure is followed to adjust for simultaneity bias between urban infrastructure and growth.

[37] Note that there will still be omitted variables bias in our IV estimate of λ_1 if $cov(e,u) \neq 0$. This bias will, of course, diminish as further controls are taken into account (as well as fixed effect in panel data estimations are included).

Tables and figures:

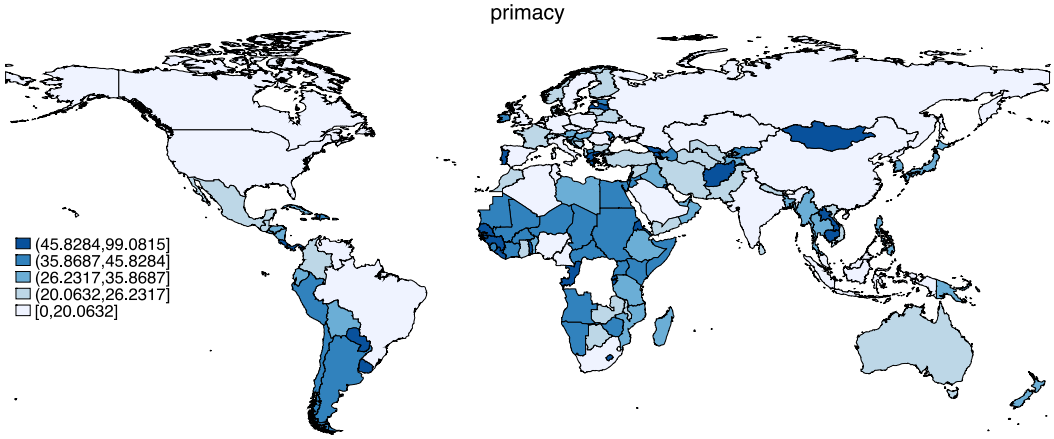
Table 1: Some basic figures

Panel A:	growth			primacy			Correlation	slums		
Region	mean	std dev	sample	mean	std dev	sample		mean	std dev	sample
Developed	1.8	1.1	42	35.0	24.8	44	0.17			
Developing	1.9	2.3	128	43.3	24.2	149	-0.08	57.0	28.9	102
North America	1.5	0.1	3	18.0	6.7	4	-0.10	18.0		1
Europe	2.0	1.5	31	28.3	19.0	38	0.10			
Asia	2.9	2.9	39	38.5	24.1	47	0.07	52.2	24.9	26
Oceania	1.1	0.9	12	72.8	30.8	16	-0.07			
North Africa	2.1	1.2	6	25.8	9.6	6	0.20	39.5	29.0	6
LAC	1.9	1.3	34	49.3	23.7	36	0.08	33.7	23.8	28
SSA	1.3	3.0	45	42.1	16.3	46	-0.13	77.1	19.7	41
World	1.9	2.0	170	41.4	25.5	193	-0.03	57.0	28.9	102

Panel B:	sanitation			Other urban infrastructure measures				
Region	mean	std dev	sample	water	electricity	inf. mort.	tel. lines	transport
Developed	98.6	3.59	40	99.6	96.6	11.0	40.9	69%
Developing	70.6	25.71	142	89.4	79.8	62.4	10.8	20%
North America	100.0	0	2	100.0	100.0	8.1	59.9	100%
Europe	97.5	4.6	35	99.7	99.7	13.1	38.4	74%
Asia	83.8	18.25	45	93.3	92.3	50.2	14.1	33%
Oceania	83.3	18.39	14	93.0	74.3	33.6	16.2	6%
North Africa	87.7	15.15	6	85.9	90.5	56.6	5.8	67%
LAC	84.2	15.68	34	94.8	94.9	35.4	17.6	26%
SSA	44.1	20.26	46	81.1	52.7	93.3	2.2	0%
World	76.7	25.55	182	91.7	83.7	50.7	17.8	32%

Note: *growth* is calculated over 1970-2010. "sample" indicates the number of countries considered (for which we have data for the respective region and variable). *primacy*, *sanitation*, *water*, *electricity* and *tel. lines* are calculated as averages over 1990-2005. *transport* indicates the percentage of countries in the region for which their primate city has a massive transport system (metro, tram or rapid bus).

Figure 1: Population living in largest city (percentage of urban population)



Note: values for *primacy* calculated as averages between 1970 and 2010.

Figure 2a: Access to improved sanitation (percentage of urban population)

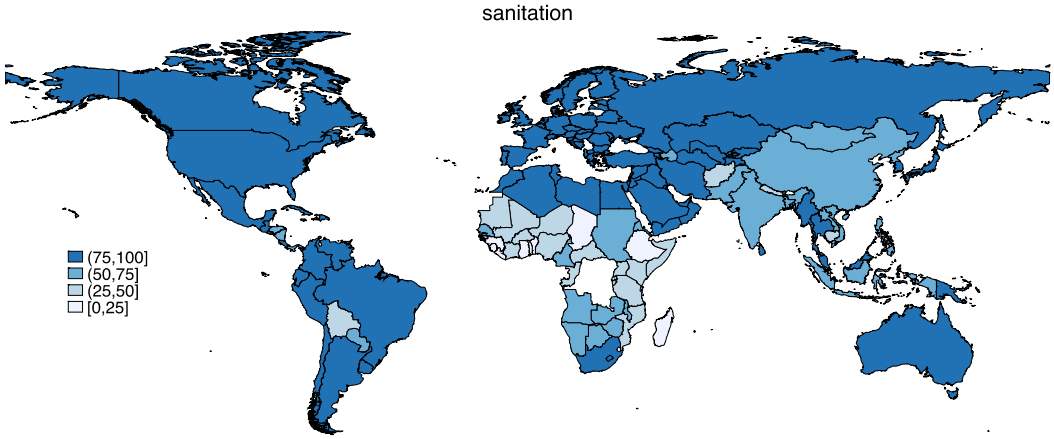
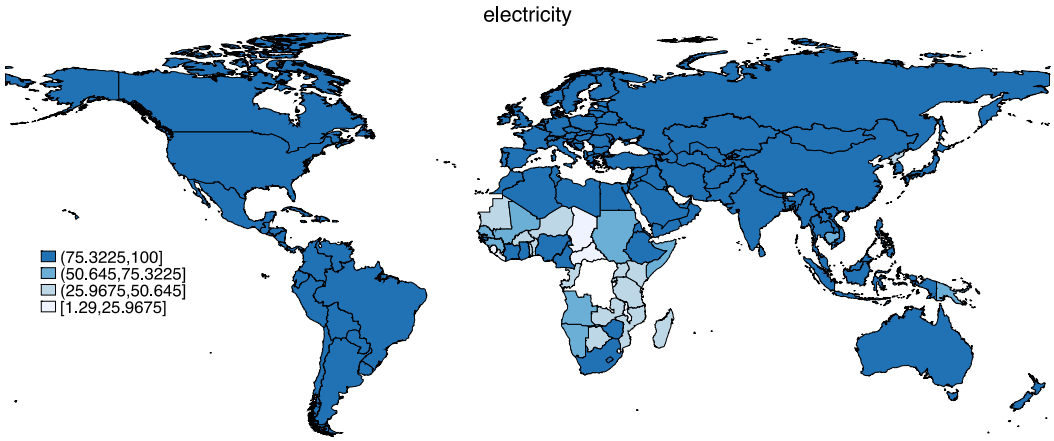


Figure 2b: Access to electricity (percentage of urban population)



Note: values for *sanitation* and *electricity* calculated as averages between 1990 and 2005.

Figure 3: Access to improved sanitation by income levels

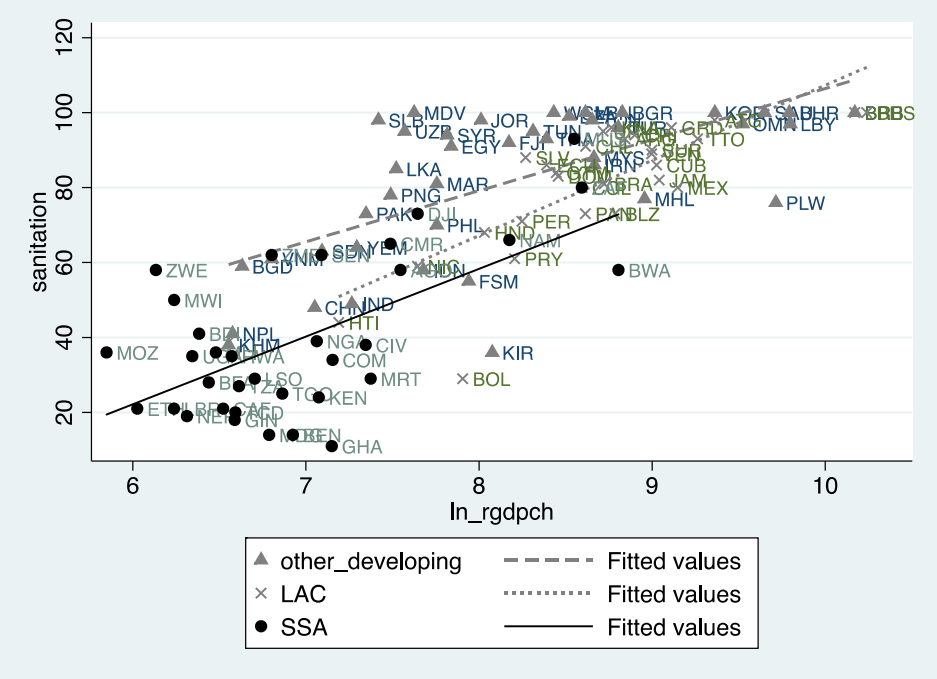
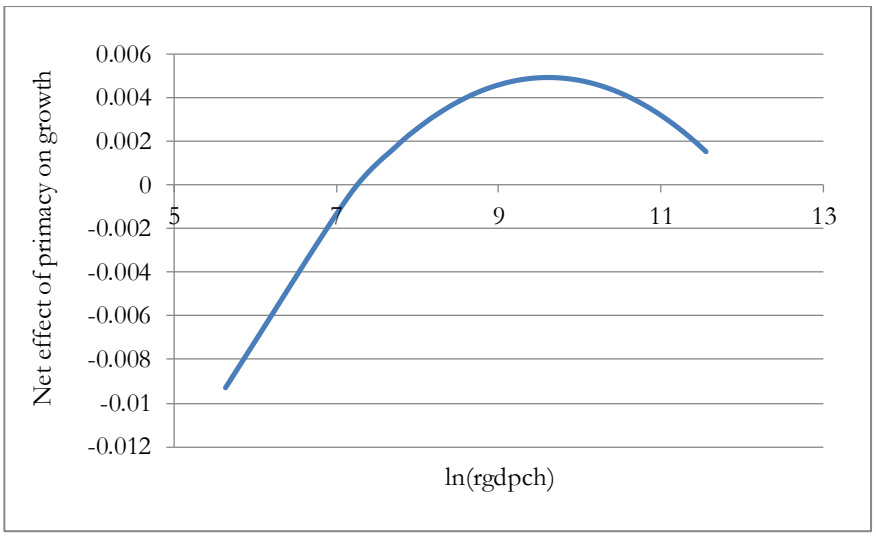


Table 2: Urban concentration and growth in a panel of countries

	(1) OLS	(2) FE	(3) GMM	(4)SysGMM	(5)SysGMM	(6)SysGMM	(7)SysGMM
Dependent variable: Average cumulative annual growth rates of per-capita GDP							
ln(rgdpch)	-0.0996***	-0.4309***	-0.3663***	-0.0290***	-0.0814***	-0.1252***	-0.0715***
	(0.0155)	(0.0683)	(0.1647)	(0.0461)	(0.0507)	(0.0974)	(0.0539)
ki	0.0074***	0.0079***	-0.0026	0.0034*	0.0015	-0.0015	-0.0014
	(0.0012)	(0.0021)	(0.0029)	(0.0018)	(0.0035)	(0.0032)	(0.0036)
fertility	-0.0862***	-0.0546***	0.0205	-0.0580***	-0.0629***	-0.0362**	-0.0448***
	(0.0091)	(0.0182)	(0.0215)	(0.0154)	(0.0170)	(0.0156)	(0.0131)
schooling23	0.0034	0.0129	0.1344*	-0.0311	0.0206	0.0113	-0.0141
	(0.0102)	(0.0279)	(0.0677)	(0.0318)	(0.0388)	(0.0565)	(0.0387)
UC					0.0054*	-0.0782***	0.0049*
					(0.0032)	(0.0269)	(0.0027)
UC*ln(rgdpch)						0.0173***	
						(0.0062)	
UC*(ln(rgdpch))^2						-0.0009***	
						(0.0003)	
UC*LAC							-0.0040***
							(0.0012)
UC*SSA							-0.0070**
							(0.0030)
Year FE	YES	YES	NO	YES	YES	YES	YES
Annual speed of convergence	2.10%	11.27%	9.12%	0.59%	1.70%	2.68%	1.48%
adj R square	0.196	0.217					
Observations	1216	1216	1077	1216	1204	1204	1204
No. of countries	139	139	139	139	137	137	137
AR1 test p-value			0.025	0.001	0.004	0.000	0.002
AR2 test p-value			0.366	0.436	0.437	0.582	0.552
Hansen test p-value			0.040	0.007	0.047	0.166	0.338

Note: *ki*, *fertility* are calculated as averages over 5 years. The time spam goes from 1960 to 2010. All remaining variables are measured at the beginning of the period. For GMM and SysGMM estimations variables in levels lagged between 2 and 4 periods are used as instruments for first differences. For SysGMM estimations variables in first differences lagged between 2 and 4 periods are used as instruments for levels. GMM and SysGMM estimations are done with small sample correction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 4: The Williamson hypothesis



Note: Plot using SysGMM estimation coefficients (column 6 of Table 2).

Table 3: Urban concentration depending on the urban process

	World sample			Developed	Developing	Developing	Developing	Developing
	(1) OLS	(2) FE	(3)SysGMM	(4)SysGMM	(5)SysGMM	(6)SysGMM	(7)SysGMM	(8)SysGMM
Dependent variable:	growth	growth	growth	growth	growth	growth	growth	growth
UC	-0.0171***	-0.0474***	-0.0331	0.0711	-0.0462***	-0.0031	0.1152	
	(0.005)	(0.016)	(0.020)	(0.061)	(0.011)	(0.0211)	(0.0729)	
sanitation	-0.0035	-0.0057	-0.0197	0.0310	-0.0139	-0.0159	-0.0080	-0.0137
	(0.002)	(0.011)	(0.012)	(0.028)	(0.010)	(0.0109)	(0.0089)	(0.0112)
UC*sanitation	0.0002***	0.0004*	0.0004*	-0.0007	0.0005***	0.0005**	0.0004**	0.0005**
	(0.000)	(0.000)	(0.000)	(0.0006)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
UC*ln(rgdpch)						-0.0053	-0.0360**	-0.0354**
						-0.0037	-0.0177	-0.0134
UC*ln(rgdpch)^2							0.0021*	0.0018**
							-0.0011	-0.0009
UC*region								YES
Country FE	NO	YES						
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Observations	500	500	500	144	356	356	356	356
No. of countries	131	131	131	37	94	94	94	94
AR(1) p-value			0.082	0.192	0.071	0.059	0.029	0.043
AR(2) p-value			0.280	0.371	0.569	0.505	0.863	0.711
Hansen test p-value			0.172	0.529	0.424	0.305	0.325	0.272

Note: *UC*region* are interaction terms between *UC* and each world region. Controls include *ln(rgdoch)*, *ki*, *fertility* and *schooling23*. All controls are calculated as averages over 5 years except *ln(rgdoch)* and *schooling23*, which are measured at the beginning of each period. The time spam goes from 1990 to 2010. In SysGMM estimations *ki*, *fertility*, *schooling23*, *UC*, *sanitation* and *UC*sanitation* are treated as endogenous using lagged values between 2 and 4 periods as instruments for first differences and variables in first differences lagged between 2 and 4 periods as instruments for variables in levels. SysGMM estimations are done with small sample correction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Urban concentration and growth in SSA

		G = *sanitation*		G = *water*	G = *electricity*
	(1) IV	(2) IV	(3) IV	(4) IV	(5) IV
Dependent variable:	growth	growth	growth	growth	growth
UC	-0.0287**	-0.0874***	-0.0200	-0.3371***	-0.1754**
	(0.0141)	(0.0146)	(0.0697)	(0.0767)	(0.0752)
G		-0.0638***	-0.0725***	-0.1083***	-0.0361***
		(0.0214)	(0.0229)	(0.0301)	(0.0094)
UC*G		0.0013***	0.0015***	0.0024***	0.0007***
		(0.0005)	(0.0005)	(0.0007)	(0.0002)
UC*ln(rgdpch)			-0.0111	0.0125	0.0170
			(0.0113)	(0.0124)	(0.0118)
Country FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES
Observations	135	103	103	95	108
Number of countries	28	28	28	26	28
Angrist-Pischke F tests p-values	0.000	0.000; 0.000	0.007; 0.000; 0.000; 0.009	0.002; 0.000; 0.001; 0.003	0.001; 0.000; 0.000; 0.002
Kleibergen-Paap F-stat	1253.02	40.15	28.19	18.37	70.07
Kleibergen-Paap LM-stat	6.63**	24.26***	23.38***	12.39***	17.10***

Note: Controls include *ln(rgdoch) ki*, *fertility* and *schooling23,* but also *rainfall* and *rainfall squared*. All controls are calculated as averages over 5 years except *ln(rgdoch)* and *schooling23*, which are measured at the beginning of each period. The time spam goes from 1990 to 2010. For IV estimations, *UC*, *G* and *UC*G* series adjusted for the effect that *growth* has on them are used as instruments. Kleibergen-Paap stats test the null hypothesis that the equation is underidentified. Estimations are done with small sample correction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Estimates for SSA using nigh density at night data

	(1) FE	(2) IV	(3) FE	(4) IV
Dependent variable:	growthavsd	growthavsd	growthavsd	growthavsd
UC	-0.0175*	-0.0127	-0.0624***	-0.0866***
	(0.0094)	(0.0103)	(0.0212)	(0.0198)
sanitation			-0.0258	-0.0601*
			(0.0304)	(0.0304)
UC*sanitation			0.0019**	0.0029***
			(0.0009)	(0.0008)
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Observations	108	108	103	103
Number of countries	28	28	28	28
Angrist-Pischke F tests p-values		0.000		0.000
Kleibergen-Paap F-stat		2745.44		33.02
Kleibergen-Paap LM-stat		3.93**		21.41***

Note: Controls include *ln(avsd) ki*, *fertility* and *schooling23,* but also *rainfall* and *rainfall squared*. All controls are calculated as averages over 5 years except *ln(avsd)* and *schooling23*, which are measured at the beginning of each period. The time spam goes from 1990 to 2010. For IV estimations, *A*, *sanitation* and *A*sanitation* series adjusted for the effect that *growth* has on them are used as instruments. IV estimations are done with small sample correction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1