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

(Preliminary version)
March 2014

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

Abstract:
Today more than half of the 7 billion inhabitants of the planet live in urban areas, with this share expected to keep rising. Whereas in developed countries urbanisation has been a long and slow process, in developing countries this process is now characterised by a really fast pace and a high degree of urban concentration, with urban population tending to concentrate in one or few large metropolitan areas of disproportionate size. 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, in this paper we analyse differentiated effects of urban concentration on national economic performance. We contribute to the literature by providing empirical evidence on how different characteristics of the urban environment - in particular the quality of urban infrastructure - strongly determine the growth-enhancing benefits of urban concentration (something that previous studies on urban concentration and economic growth have not considered empirically). Looking at different world regions we find that while increasing urban concentration might have been associated with growth in Asian countries, it seems that growth-deterring congestion diseconomies have prevailed over agglomeration benefits in most Latin American and Sub-Saharan African countries due to their significant deficiencies in terms of urban infrastructure.

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

JEL classification: O1, O4, R1

Acknowledgements
We thank Marius Brülhart, Markus Brückner, Vernon Henderson and Michiel Gerritse for valuable comments. We are 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) and at the XXXIX Reunión de Estudios Regionales of the Spanish Regional Science Association.

I. Introduction

Back in the late 18th century Malthus presented his theory where population grows faster than subsistence means and suggesting a negative relationship between population growth and living standards (Malthus 1803). This *Malthusian* negative effect of population size is related to the congestion of fixed resources, both productive and environmental, and was originally associated with agricultural output where land is the fixed resource. Malthus' law was a reality for most of human history in most of the world until the industrial revolution, and is still a relevant possibility for many poor countries with large rural populations and largely dependent on agriculture and mineral exports (Weil and Wilde 2009). But the world of the 21st century is an urban world, where more than half of the 7 billion inhabitants of the planet live in urban areas, and where the pressure of population growth has moved from the countryside to the main urban centres. With more than 60 million new urban inhabitants every year worldwide, most cities are today experiencing dramatic increases in their population.[1] A process of urbanization characterized in developing countries not just by a really fast pace, but also by a high degree of urban concentration, with urban population tending to concentrate in one or few large metropolitan areas of disproportionate size. In Asia, the metropolitan areas of Jakarta, Shanghai and Bombay have now more than 20 million inhabitants. In Africa, Lagos and Cairo metropolitan areas are close. In Latin America, Mexico City, Sao Paulo, Buenos Aires and Bogotá are also examples of large urban agglomerations (all above 10 million inhabitants) growing at a fast pace. In fact, while 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. Today at least 1 billion people worldwide live in slums, lacking access to basic services like electricity, clean water and sanitation, and the figure keeps rising.[2]

Hence, it is possible that the Malthusian dilemma has in some way moved from the rural to the urban world, where the relationship between population and resources is in constant readjustment. How would a Malthus living in a slum react to the rising inflow of population to large urban agglomerations? Urban concentration is desirable as it enhances economies of scale in the provision of urban infrastructure and public services. But in the short run urban concentration posses challenges that are of great magnitude today for large agglomerations in developing countries, as resources have to keep pace with fast urban population growth. Access to basic services is not just desired per se in terms of quality of life for urban residents, but also in terms of economic efficiency. Moreover, deficient access to basic services hold back needed structural change and dramatically reduce the productivity of cities. On one side, these deficiencies 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

[1] For more on population trends see UNFPA State of World Population 2011.
[2] The figure expected to be 2 billion by 2030. UN-Habitat Reports.

trust. On the other side, deficiencies 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 to attract talent and investment).

Building on previous evidence on the relationship between urban concentration and economic growth, in this paper we pay special attention to differences between world regions in terms of the quality of the urban environment, with a distinctive focus on Sub-Sahara Africa. We contribute to the literature by providing empirical evidence on how differences in urban infrastructure (in particular access to basic urban services) strongly determine the growth-enhancing benefits of urban concentration. The rest of the paper is organized as follows. In section 2 we present some basic stylized facts. In section 3 we review related strands in the literature. In section 4 we derive an empirical specification and present our data. Section 5 presents our estimations and results. In section 6 we perform some robustness checks and focus on Sub-Sahara African countries. Finally section 7 concludes and derives policy implications from our results.

II. Some basic stylized facts

Looking at urban concentration patterns and economic performance worldwide during the previous decades some basic but interesting facts can be highlighted. The first is that while population living in the primate city 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. As table 1 shows, taking 1990 data, while the average of total population living in primate cites was 35 per cent in developed countries, it was 43 per cent in developing countries, almost reaching 50 per cent in Latin America and the Caribbean (LAC). Moreover, total population living in primate cities (rather than the percentage) has soared from an average (for 193 countries worldwide) of 1.2 millions in 1970 to an average of 2.8 millions in 2010. And the increase has been much more marked in developing countries (with an average of also 1.2 millions in 1970 to one of almost 3.4 millions in 2010).[3] The second relevant stylized fact is a strong correlation between urbanisation and income levels (above 0.8), but not between urban concentration and subsequent economic growth (-0.03). However, the picture changes if we consider the correlation by groups of countries; urban concentration is positively correlated with growth in developed, but not in developing, countries. Likewise, while the correlation is positive in Europe, Asia and LAC, it is negative in North America, Oceania and Africa (especially in Sub-Saharan Africa - 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. While access to basic services was virtually universal in developed countries already in 1990, it was not in developing countries, with important differences among them and particularly significant deficiencies in SSA.

[3] Considering only countries with a total population of at least 1 million inhabitants (for a total of 150 developing countries in 2010).

These deficiencies in SSA appear as remarkably severe in terms of access to improved sanitation and electricity. For access to improved sanitation, for example, the average for Asia was 85 per cent in 1990, while it was less than 40 for SSA, increasing to just above 45 in 2005. 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.[4] 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. As Figure 1 shows for access to improved sanitation, even controlling for income levels SSA countries present significantly lower levels of urban infrastructure.[5]

[Insert table 1: Some basic figures]

[Insert figure 1: access to improved sanitation by income levels]

III. Literature review

From Optimal to Efficient city size:

In the urban economics literature one major goal has been to explain the causes and limits of city growth (von Thunen 1826; Christaller 1933; Alonso 1964), and as an additional goal also to explain optimal city size (Mills and De Ferranti 1971; Alonso 1971; Henderson 1974). Moreover, given current trends, the goal expands to trying to understand the dynamics of rapidly growing megacities around the world (Henderson 1985; Ades and Glaeser 1995), especially in developing countries (Firebaugh 1979; Kasarda and Crenshaw 1991; Arku 2009; Jedwab 2011). Standard urban theories stress the trade-off between positive and negative synergies and externalities - location costs and benefits - that cities provide, being urban size the fundamental determinant of these synergies and externalities. These 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.[6] In this framework, therefore, urban scale is self-limiting, with the costs of agglomeration otherwise outweighing the benefits (Bertinelli and Black 2004). However, neoclassical theories of optimal city size have long been criticized. On the one hand, it is clear that there are other determinants influencing urban agglomeration economies, not merely physical size of cities (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

[4] Lagos inaugurated a bus rapid transit system in 2008, and Accra has now planned a metro monorail project.

[5] A simple regression analysis yields highly significant lower levels of urban infrastructure for SSA countries compared to other countries of same income per capita levels.

[6] 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.

(Chinitz 1961; Capello and Camagni 2000). On the other hand, cities operate in different national urban systems where they interact with each other (Camagni 1993), fact that 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 analyzing 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 on the spatial organization within the urban system (Capello and Camagni 2000; Royuela and Suriñach 2005; Camagni et al. 2013).[7]

Urban concentration and economic growth:

Based on theory and evidence on agglomeration economies, a related goal to that analyzing optimal city size has been to analyse the relationship between geographical concentration of population and economic activity and economic growth *at country level*. 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 reallocating of people and resources from agricultural activities towards industrial activities of higher productivity and value added, which takes places with urbanisation. 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).[8] Third, due to the fact that concentration enhances economies of scale in the provision of urban infrastructure and public services. In this line, there is evidence suggesting a long-run growth-enhancing effect of geographical concentration at country level (Henderson 2003; Bertinelli and Strobl 2007; Brülhart and Sbergami 2009; Leitão 2103; Castells-Quintana and Royuela 2014) with most authors focusing on urban concentration, rather than on urbanisation. The argument being reinforced by the fact that in any case the growth-enhancing effects of urbanisation, related to scale and agglomeration economies, become significant for large urban agglomerations, rather than for small ones, and especially for developing countries.[9] But as

[7] 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.

[8] 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.

[9] 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.

with city efficiency, the same concentration of economic activity that allows for growth-enhancing agglomeration economies, also leads to potential growth-deterring diseconomies of congestion. Hence, while increasing urban concentration is desirable and expected in early stages of development, deconcentration eventually occurs as development proceeds - the Williamson (1965) hypothesis. In this line, Henderson (2003) finds a nonlinear relationship of urban primacy with economic growth suggesting an optimal degree of primacy, which declines with the level of development (with some countries experiencing insufficient urban primacy while others experiencing excessive concentration). 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. In this case, in developing countries "it may well be the case that urban expansion is the only realistic option for overcoming institutional problems promoting growth and development" (Barca et al. 2012). In fact, according to empirical evidence, the beneficial net effect of high urban concentration is expected only when income levels are not too high (Brülhart and Sbergami 2009)[10] and income distribution remains relatively equal (Castells-Quintana and Royuela 2014).

Differentiating processes of urbanisation (and urban concentration) and economic growth:

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 urbanisation (and urban concentration) and national economic growth. 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 urban concentration 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. The process of urbanisation (and urban concentration) can be driven by what Kim (2008) calls "pathological non-economic factors" rather than by agglomeration economies and higher productivity. Different urban processes translate into significantly different urban environments, which could explain empirical evidence on relevant

[10] 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. Results are directly related to the spatial scaled considered; different scales imply different mechanisms at work and, therefore, may yield different results. For small spatial scale, there are positive spillovers associated with clustering activities (mainly knowledge spillovers) and agglomeration may have a positive impact on economic growth even, at probably more importantly, in more developed countries. Their results, however, relate to a higher spatial scale associated with urban concentration, where the agglomeration impact relates to reduction of transaction costs and higher integration of markets.

heterogeneity across countries in the relationship between urban concentration and growth (Bertinelli and Strobl 2007 distinguishing developed from developing countries, and Pholo Bala 2009 analyzing regional-specific effects by continent). In this line, there is now growing empirical evidence of urban processes not necessarily linked to economic development (Firebaugh 1979; Ades and Glaeser 1995; Davis and Henderson 2003; Barrios et al. 2006; Jedwab 2011; Gollin et al. 2012; Behrens and Pholo-Bala 2013).[11] In particular, Brückner (2012) provides evidence of even a negative influence of growing urbanisation on economic growth in African countries, despite increasing returns from agglomeration. The main reasons for this to happen in Africa are a high ethnic fractionalization, very low economic development and excessive size of primate cities. And Brückner suggests that the negative role of the excessive size of primate cities in Africa relates to their large squatter settlements with inadequate access to transport, water, sanitation, electricity, and health services.[12]

Different urban environments allow for different capacities for cities to benefit from agglomeration economies and to control congestion diseconomies. On the theoretical side Bertinelli and Black's (2004) stylized urban economics model suggests how the benefits from urbanisation can significantly be affected by the quality of urban infrastructure affecting the urban production technology.[13] On the policy debate side the WDR (2011), as mentioned above, has highlighted the development of urban infrastructure as essential, putting strengthened emphasis on access to basic services for the well functioning of large cities. 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." And the pace at which both urban concentration and the development of urban infrastructure takes place becomes crucial, especially in developing countries where primate cities grow by thousands inhabitants daily.

IV. Empirical Specification and Data

We base our empirical analysis on a GDP per capita growth framework, following works as Henderson (2000) and Brülhart and Sbergami (2009).[14] To derive an econometric specification that

[11] 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.

[12] Bruckner (2012) analyses drivers of urbanisation, and then relies on such analysis to explain negative effects of urbanisation processes on economic growth. Nevertheless, he considers only African countries and national factors, without considering specific characteristics of the urban environment and how these condition agglomeration and congestion effects.

[13] 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, 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."

[14] 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 for his TFP framework

allows us to analyse the role of urban concentration, we follow the standard neoclassical framework of economic growth basis for standard cross-country growth regressions. In this framework, for a given economy i at time t, $Y_{i,t}$ 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. Output per capita growth between two periods, γ_i, can be derived to depend on growth due to technological progress and growth due to convergence - the gap between initial conditions and the steady-state value of output per capita (measured in efficiency units of labour):[15]

$$\gamma_i = g_i + \beta_i(log y_{i,0} - log y_{i,\infty}^E - log A_{i,0}) \qquad (1)$$

where $y_{i,0} = Y_{i,0}/(A_{i,0}L_{i,t})$ is initial output per capita, $y_{i,\infty}^E = Y_{i,\infty}/(A_{i,\infty}L_{i,\infty})$ is the steady-state value of output per unit of efficient labour, and $A_{i,0}$ is initial efficiency level or technology. The traditional empirical implementation of equation (1) that allows regression analysis linear in observable variables relies on a simple aggregate Cobb-Douglas function of output in an economy (following Mankiw, Romer and Weil 1992):

$$Y_{i,t} = K_{i,t}^{\alpha} H_{i,t}^{\phi} (A_{i,t}L_{i,t})^{1-\alpha-\phi} \qquad (2)$$

where capital (physical and human) is accumulated over time from savings in output - investment - leading to the steady state in equation (1) to depend on investment rates (of physical and human capital), population growth and on technological change, assumed equal for all countries. Taking this into account in equation (1) and 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), 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. Trying to account for these country-specific factors, and therefore allowing for heterogeneity in initial conditions but also in growth paths across countries, the standard specification of cross-country economic growth takes the following form:

$$\gamma_i = \beta\big(log y_{i,0}\big) + \psi X_{i,0} + \pi Z_{i,0} + \varepsilon_i \qquad (3)$$

results). While both analyses are similar, a GDP per capita growth specification allows for the use of a larger dataset.

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

where $y_{i,0}$ is initial per capita GDP, $X_{i,0}$ 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.

The degree of urban concentration represents one variable that could be consider 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. Thus, our basic empirical setting becomes:

$$\gamma_i = \beta\left(log y_{i,0}\right) + \psi X_{i,0} + \lambda U_{i,0} + \pi Z_{1i,0} + \varepsilon_i \qquad (4)$$

where U_{i0} is the degree of urban concentration and $Z_{1i,0}$ other remaining relevant country-specific factors. However, if we further take into account that the way urban concentration affects growth in efficiency depends on specificities of the urban process influencing those possibilities of agglomeration economies or congestion diseconomies, equation (4) extends to:

$$\gamma_i = \beta\left(log y_{i,0}\right) + \psi X_{i,0} + \lambda_1 U_{i,0} + \lambda_2 G_{i,0} U_{i,0} + \pi Z_{1i,0} + \varepsilon_i \qquad (5)$$

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

Data

To study the relationship between urban concentration and growth we 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. Our dependent variable is national economic growth, for which we use data from the Penn World Tables. For our key variable, U, we focus on urban primacy, as the most standard measure in the literature on urban concentration. Data for primacy comes from the World Bank. For the quality of urban infrastructure we consider several measures. Following the World Development Report (2011), we focus on three key indicators: access to improved sanitation, improved water source, and electricity. As data for all these variables is scarce, when we introduce them in the analysis our panel only considers the 1990-2010 period.[16] Finally, as control variables

[16] Our 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, we

(X_i and Z_{1i} in equation 5) we begin by considering investment, as share of GDP, fertility rates, and average years of secondary and higher schooling of the adult population, following Henderson (2000) specification. For urban infrastructure variables, as well as for control variables, we rely on a variety of sources. Appendix 1 lists variables' names, definitions and sources.

In our robustness estimations we further consider a wide variety of other control variables, following Brülhart and Sbergami (2009) and the literature on cross-country economic growth. In our focus on SSA, we also use data on rainfall 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) and Brückner (2012).

V. Estimations and results

Urban concentration and economic growth in a panel of countries

Following the literature on urban concentration and economic growth, we begin by estimating equation (4) based on cross-country panel data (for 137 countries) and without considering differentiated urban patterns across countries. We split the 1960-2010 in 5-year periods.[17] Equations like (4) using panel data represent dynamic models. Estimation of these models raises some concerns: reverse causality, unobserved time-invariant country-specific characteristics, and the presence of initial income as a regressor. As it is common in the empirical studies estimating these models, we estimate the dynamic model version of equation (4) by System-GMM, which allows us to deal with some of these concerns.[18] For our focus on SSA we complement our empirical analysis with Instrumental Variables (IV) estimations taking advantage of the exogenous variability given by rainfall data.

Table 2 presents the result for our first set of estimations of our basic growth model. Columns 1 to 4 present our results for different estimation techniques.[19] Control variables have the expected sign reflecting conditional convergence, a positive effect of higher investment and educational levels and

discuss our results using improved water source and electricity. We further consider infant mortality rates, as a common and basic indicator of health, and access to mass urban transport systems.

[17] We also experimented with 10-years periods in order to reduce any short-term noise from the business cycle, but at the expense of loosing observations. Results using 10-years periods are very similar to those presented throughout the paper using 5-years periods.

[18] 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 our case). For GMM estimations we present standard AR(1), AR(2) and Hansen tests for relevance and validity of internal instruments.

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

a negative effect of higher fertility rates.[20] In column 5 we 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), columns 6 considers a more flexible functional form for the effect of primacy on growth. We 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 2 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, we 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 our world sample, there is a significantly different relationship for LAC and SSA.[21]

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

[Insert figure 2: The Williamson hypothesis]

As previous studies have suggested, our panel results confirm the relevance of urban concentration in the economic growth process. But in line with these studies our 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.

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

---

[20] We also calculate the annual speed of convergence to ease comparability of our 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.

[21] In fact, when we analyse urban concentration by the different world regions, its effect on growth seems to have been positive and significant only in Asia and Europe. If we distinguish 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).

be fundamental to unleash positive synergies from agglomeration economies or to increase congestion costs, in both cases affecting national productivity.

In Table 3 we present result for estimates of equations like (5), letting the effect of urban concentration to depend on the quality of urban infrastructure. We present results using access to improved urban sanitation facilities (*sanitation*) as a proxy for the quality of urban infrastructure.[22] 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 we 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 we 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 our variables 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 our results suggest that while for low levels of sanitation urban concentration is negative associated with economic growth, the association becoming positive as access to sanitation increases.[23] Hence, urban concentration becomes positively associated with growth only when basic services spread to the majority of the urban population (in the case of sanitation when its coverage exceeds about 70 per cent).

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

VI. Robustness and focus on SSA

We can check the robustness of our results in several additional ways. In first place we could worry that the positive effect of our 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, our results for developing countries hold when we introduce an interaction between urban concentration and income levels. A second concern might come from our proxy for urban infrastructure. While access to sanitation

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

[23] We also obtain similar results when we consider *growth* in urban concentration and *growth* in sanitation rather than their levels.

seems very pertinent for our 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 our analysis, we replicate some of our estimations using other variables for the quality of urban infrastructure. On one side, Appendix 2 presents our 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 3 presents some cross-section results. Cross-section analysis is more common in the long-run economic growth literature and, as discussed before, allow us to consider other variables, as transport systems for which we do not have enough time variation. Results for sanitation are in line with our panel results. Results also hold when we use other variables as *electricity* or *transport_systems*, although the significance is reduced and depends on the controls used.[24] When we consider a composite measure for urban infrastructure, rather than just one indicator, estimations yield highly significant results (and robust to all our considered controls).[25]

Sub-Sahara African countries:

Finally one might still have concerns about our 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. Good external instruments for primacy and for the quality of the urban infrastructure are hard to find. However, we can find reliable external instruments for economic growth, at least for Sub-Sahara African countries, which gives as 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.[26] Following Brückner (2012; 2013), we exploit this exogenous variation to construct instrumental variables (IV) that purge the possible effect that economic growth might have on our key variables, urban concentration and sanitation. The use of exogenous instruments allows us to control for reverse

[24] 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 our panel analysis, when we analyse by world regions our 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).

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

[26] 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).

causality in a more direct way (alternative to SysGMM and without having to rely on internal instruments). Hence, in a first step we estimate primacy and sanitation on economic growth by two-stage least squares using rainfall and rainfall squared as instruments for economic growth:

$$U_{it} = \alpha\gamma_{it-1} + a_i + b_t + \varepsilon_{it} \quad (6)$$

$$G_{it} = \alpha\gamma_{it-1} + a_i + b_t + \varepsilon_{it} \quad (7)$$

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 4 presents our first-stage OLS estimation for growth on rainfall and rainfall squared, and our estimations of equations (6) and (7). By construction, the residual variation on primacy and sanitation from our two-stage least squares estimations of (6) and (7) capture any variation in these variables that is not due to economic growth. As in Brückner (2012; 2013), in a second step we use these residual variations in primacy and sanitation as instruments for actual primacy and sanitation to estimate by two-stage least squares our economic growth equation (equation 5) for SSA. Table 4 presents these results for our IV as well as for FE and SysGMM estimations.[27]

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

Our estimations yield similar results to those in Table 3 (with highly significant coefficients, negative for primacy and positive for its interaction with sanitation) and regardless of the estimation technique (FE, SysGMM or IV). Our key coefficients are robust to the considered controls as to the introduction of an interaction term between urban concentration and income levels (column 4). They are also highly significant when we consider access to improved water source (column 5) or access to electricity (column 6), rather than sanitation, and confirm - in this case for SSA - the role of urban infrastructure when it comes to analyse the relationship between urban concentration and economic growth. This role is particularly relevant in SSA where access to basic services is still very deficient, as we have seen. As an example, only 3 countries, out of 34, reached that 70 per cent threshold of urban population with access to improved sanitation in 1990 (Djibouti, Mauritius and South Africa), three more countries in 2005 (Angola, Botswana and Seychelles).

[27] We report first-stage F-statistics and Angrist-Pischke F tests for relevance and validity of excluded instruments. According to the tests, rainfall and rainfall squared are not just relevant but also valid instruments for economic growth in SSA countries. Likewise, similar tests confirm the relevance and validity of the residual variation of primacy and sanitation as instruments for actual primacy and sanitation in our growth equation.

VI. 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. Our economtric 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. In this regard, we have provided empirical evidence on how urban concentration can be negatively associated with national economic growth under urban environments with deficient urban infrastructure, indicating congestion costs that exceed the benefits from agglomeration. 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.

The policy implications of our results are straightforward. The Malthusian trap, by which higher living standards lead to population growth and higher congestion of resources, has become a relevant reality for many of the large urban agglomerations of today's developing world. Hence, to properly exploit the benefits from agglomeration and sustain or increase their productivity, guaranteeing that adequate urban infrastructure in these large cities keeps pace with their rapid increase in population appears as an imperative necessity. In this line, access to basic services emerges not just as desirable per se in terms of quality of life for urban residents, but also in terms of economic efficiency at national level. The good news from our results are that 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 Latin America or Sub-Saharan Africa. According to our results, the negative effects of urban concentration that the literature has implied in these regions can be associated precisely with the 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 Latin American and Sub-Sahara African countries. 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.

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.

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.

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* (forthcoming).

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.

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. (2012). 'Urbanization without structural transformation: Evidence from Sub-Saharan Africa.'

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.

Jedwab. JMP. (2011). 'Why is African urbanization different? Evidence from resource exports in Ghana and Ivory Coast,'

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

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.

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.

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.

Spence, M., Clarke, P., and Buckley, R.M. (Editors). (2009). *Urbanization and growth.* Commission on growth and development.

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 1: 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 2: System GMM panel 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
U	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)
U*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*, *U*, *G* and *U*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 3: 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
U	-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)
U*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 9 and 10 composite is calculated combining *sanitation*, *water*, *electricity* and *transport_systems*. Controls include *ln(rgdoch)*, *ki*, *fertility* and *schooling23*. Further Controls include: *primary_edu, higjher_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 4: First step estimations for SSA:

	(1) OLS	(2) 2SLS	(3) 2SLS
Dependent variable:	growth	primacy	sanitation
rainfall	0.028**		
	(0.0013)		
rainfall squared	-0.0001***		
	(0.0000)		
growth		-4.1597	-0.4570
		(3.2665)	(1.9904)
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Observations	143	178	143
Number of countries	38	38	38
First-stage F-stats p-value		0.053	0.091
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. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Tables and figures:

Table 1: Some basic figures

Panel A:	Growth			Primacy			Correlation	Slums		
Region	mean	sample	std dev	mean	sample	std dev		mean	sample	std dev
Developed	1.8	42	1.1	34.9	44	24.7	0.17			
Developing	1.9	128	2.3	43.3	149	25.1	-0.08	57.0	102	28.9
North America	1.5	3	0.1	18.0	4	7.4	-0.10	18.0	1	
Europe	2.0	31	1.5	28.3	38	19.3	0.10			
Asia	2.9	39	2.9	37.8	47	23.9	0.07	52.2	26	24.9
Oceania	1.1	12	0.9	73.1	16	31.2	-0.07			
North Africa	2.1	6	1.2	24.9	6	10.5	0.20	39.5	6	29.0
LAC	1.9	34	1.3	49.1	36	23.9	0.08	33.7	28	23.8
SSA	1.3	45	3.0	43.0	46	18.6	-0.13	77.1	41	19.7
World	1.9	170	2.0	41.4	193	25.0	-0.03	57.0	102	28.9

Panel B:	Sanitation			Other urban infrastructure measures				
Region	mean	sample	std dev	Water	Electricity	Inf. Mort.	Tel. lines	Transport
Developed	99.2	36	2.39	99.8	96.3	11.0	32.7	69%
Developing	69.8	115	25.59	89.3	78.1	62.4	6.8	20%
North America	100.0	2	0	100.0	100.0	8.1	50.3	100%
Europe	98.9	27	2.74	99.7	99.6	13.1	31.0	74%
Asia	85.5	37	19.44	94.4	91.3	50.2	10.2	33%
Oceania	83.1	13	19.32	92.1	74.1	33.6	11.7	6%
North Africa	87.7	6	13.7	87.3	90.5	56.6		67%
LAC	82.5	31	16.61	93.7	94.1	35.4	11.3	26%
SSA	39.9	35	20.95	79.9	48.4	93.3	1.3	0%
World	76.8	151	26.99	91.6	82.3	50.7	13.3	32%

Note: Sample indicates the number of countries considered (for which we have data for the respective region and variable). 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: Access to improved sanitation by income levels

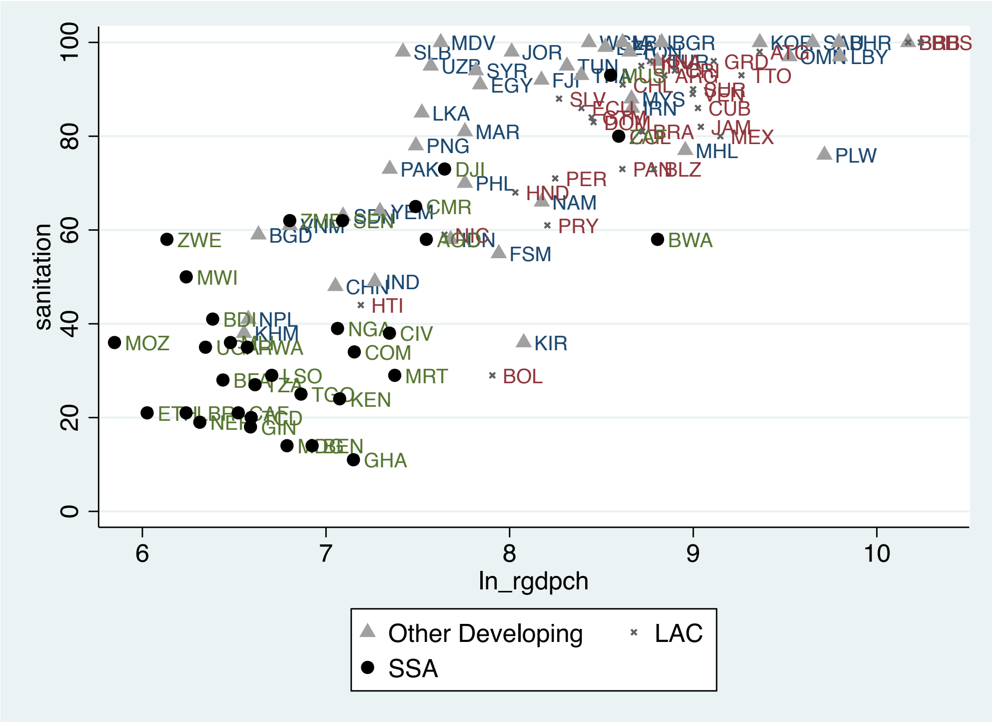
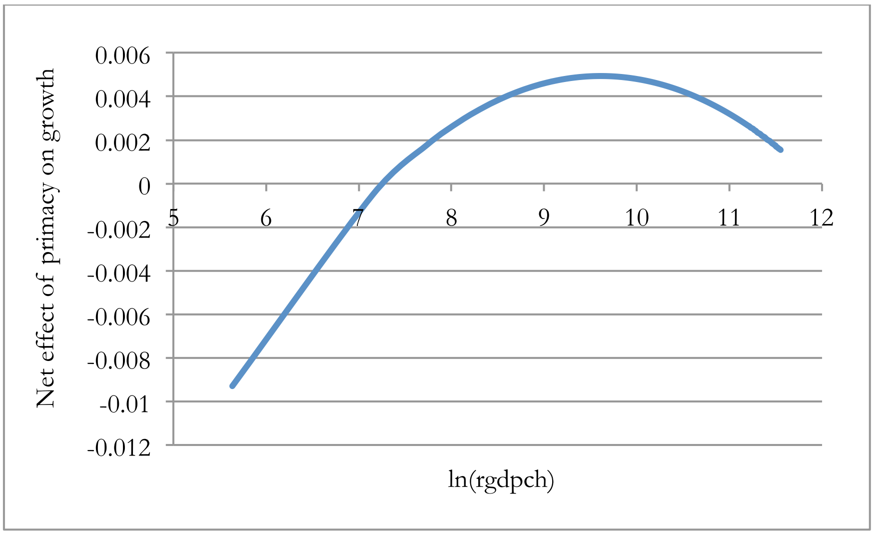


Table 2: Urban concentration and growth in a panel of countries

	(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)
U					0.0054*	-0.0782***	0.0049*
					(0.0032)	(0.0269)	(0.0027)
U*ln(rgdpch)						0.0173***	
						(0.0062)	
U*(ln(rgdpch))^2						-0.0009***	
						(0.0003)	
U*LAC							-0.0040***
							(0.0012)
U*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%		1.34%	
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			0.030	0.001	0.004	0.000	0.002
AR2 test p-value			0.267	0.436	0.437	0.582	0.552
Hansen test p-value			0.000	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 2: The Williamson hypothesis



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

Table 3: Urban concentration depending on the urban process

	World			Developed	Developing	Developing
	(1) OLS	(2) FE	(3)SysGMM	(4)SysGMM	(5)SysGMM	(6)SysGMM
Dependent variable:	growth	growth	growth	growth	growth	growth
U	-0.0171***	-0.0474***	-0.0331	0.0711	-0.0462***	-0.0031
	(0.005)	(0.016)	(0.020)	(0.061)	(0.011)	(0.0211)
Sanitation	-0.0035	-0.0057	-0.0197	0.0310	-0.0139	-0.0159
	(0.002)	(0.011)	(0.012)	(0.028)	(0.010)	(0.0109)
U*Sanitation	0.0002***	0.0004*	0.0004*	-0.0007	0.0005***	0.0005**
	(0.000)	(0.000)	(0.000)	(0.0006)	(0.0002)	(0.0002)
U*ln(rgdpch)						-0.0053
						-0.0037
Country FE	YES	YES				
Year FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Observations	500	500	500	144	356	356
Number of countries	131	131	131	37	94	94
AR(1) p-value			0.082	0.192	0.071	0.059
AR(2) p-value			0.280	0.371	0.569	0.505
Hansen test p-value			0.172	0.529	0.424	0.305

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. The time spam goes from 1990 to 2010. In SysGMM estimations *ki*, *fertility*, *schooling23*, *U*, *sanitation* and *U*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) FE	(2)SysGMM	(3) IV	(4) IV	(5) IV	(6) IV
Dependent variable:	growth	growth	growth	growth	growth	growth
U	-0.0788***	-0.0562***	-0.0874***	-0.0200	-0.1420**	-0.1754***
	(0.0172)	(0.0198)	(0.0145)	(0.0788)	(0.0671)	(0.0620)
G	-0.0549**	-0.0409*	-0.0638***	-0.0725***	-0.0099*	-0.0361***
	(0.0201)	(0.0239)	(0.0191)	(0.0221)	(0.0058)	(0.0106)
U*G	0.0011**	0.0010**	0.0013***	0.0015***	0.0019***	0.0007***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0002)
U*ln(rgdpch)				-0.0111	0.0123	0.0170*
				(0.0129)	(0.0101)	(0.0094)
Country FE	YES		YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
adj R square	0.848					
Observations	103	103	103	103	95	108
Number of countries	28	28	28	28	26	28
AR(1) p-value		0.481				
AR(2) p-value		0.811				
Hansen test p-value		0.906				
First-stage F-stats p-values			0.000; 0.000; 0.000	0.000; 0.000; 0.000	0.000; 0.000; 0.000	0.000; 0.000; 0.000
Angrist-Pischke F tests p-values			0.000; 0.000; 0.000	0.000; 0.000; 0.000	0.000; 0.000; 0.000	0.000; 0.000; 0.000

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, *U*, *G* and *U*G* series adjusted for the effect that *growth* has on them are used as instruments. In SysGMM estimations *ki*, *fertility, schooling23*, *U*, *G* and *U*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. IV and SysGMM estimations are done with small sample correction. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1