UNITED NATIONS UNIVERSITY

UNU-WIDER

World Institute for Development Economics Research

Working Paper No. 2010/29

Agglomeration Index

Towards a New Measure of Urban Concentration

Hirotsugu Uchida[1] and Andrew Nelson[2]

March 2010

Abstract

A common challenge in analyzing urbanization is the data. The United Nations (UN) compiles information on urbanization (urban population and its share of total national population) that is reported by various countries but there is no standardized definition of 'urban', resulting in inconsistencies. This situation is particularly troublesome if one wishes to conduct a cross-country analysis or determine the aggregate urbanization status of the regions (such as Asia or Latin America) and the world. This paper proposes an alternative to the UN measure of urban concentration that we call an agglomeration index. It is based on three factors: …/

Keywords: agglomeration index, urbanization, accessibility map, cost surface

JEL classification: R11, O18

[1] University of Rhode Island, email of corresponding author uchida@uri.edu, [2] IRRI - International Rice Research Institute.

This study has been prepared within the UNU-WIDER project on Beyond the Tipping Point: Development in an Urban World directed by Jo Beall, Basudeb Guha-Khasnobis, and Ravi Kanbur.

UNU-WIDER gratefully acknowledges the financial contributions to the research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

ISSN 1798-7237 ISBN 978-92-9230-264-1

- Population density
- The population of a ‘large’ city centre
- Travel time to that large city centre.

The main objective in constructing this new measure is to provide a globally consistent definition of settlement concentration in order to conduct cross-country comparative and aggregated analyses. As an accessible measure of economic density, the agglomeration index lends itself to the study of concepts such as agglomeration rents in urban areas, the ‘thickness’ of a market, and the travel distance to such a market with many workers and consumers. With anticipated advances in remote sensing technology and geo-coded data analysis tools, the agglomeration index can be further refined to address some of the caveats currently associated with it.

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Typescript prepared by Anne Ruohonen at UNU-WIDER

1 Introduction

One common challenge in analysing urbanization is the data. The United Nations (UN) compiles information on urbanization that is reported by various countries, and makes the data readily available. But one must understand the nature of the data provided. Cohen (2004) summarizes the issues stemming from the limitations in the UN data:

> Although invaluable to those interested in studying urban change, the data in the UN report are somewhat deceptive in their apparent completeness and there is a great deal of misunderstanding and misreporting by nonspecialists about what these data mean and how they should be interpreted. … Most end-users cite the UN data as if it is absolute truth rather than treating them as simply indicative of general broad trends. There is a general underappreciation of the fact that the UN is forced to rely on member countries' existing definitions of what constitutes an urban or a rural area. Not only do these definitions differ widely by country, in many places the traditional urban/rural dichotomy is becoming increasingly inadequate (Cohen 2004: 24–5).

For example, the share of India's population residing in urban areas in 1991 would be 39 per cent, rather than the official figure of 26 per cent, if 113 million inhabitants of 13,376 villages with populations of 5,000 or more were classified as urban.[1] The share would be even higher if the Swedish definition of urban (settlements with more than 200 inhabitants) were applied. In Mauritius in 2000, about one quarter of the population resided in settlements of between 5,000 and 20,000 inhabitants, some of which were district capitals but were not classified as urban. If they were reclassified, the urban percentage would have been more than two thirds, rather than the official 42.7 per cent. A country's definition also can change over time, adding yet another layer of confusion. In China, for example, the urban share in 1999 could have been 24 per cent, 31 per cent, or 73 per cent, depending on the official definition of urban population used. These examples should not be interpreted as asserting that one definition is correct and another incorrect. There is a problem, however, and it is one that is particularly critical when conducting a cross-country analysis.

Researchers in the field are aware of the issues inherent in the UN data. At the heart of the problem is how to measure urban concentration in a consistent and systematic way. Options found in the literature are the Hirschman–Herfindahl index of concentration (for example, Wheaton and Shishido 1981), the Pareto parameter (for example, Rosen and Resnick 1980), and primacy (for example, Ades and Glaeser 1995). The availability of data (for example, the UN's World Urbanization Prospects), and its relation to Zipf's Law in the distribution and ranking of cities based on their population size, has prompted researchers to use primacy as a measure of urban concentration when the analysis calls for wide and long panel data, such as studying the relationship between urbanization and economic growth (for example, Henderson 2000, 2003). Note, however, that these three measures of concentration depend heavily on the definitions of a 'city' and an 'urban area'. Henderson (2005) points out several issues regarding systematic and consistent definitions of a city or metropolitan area, both across

---

1 This and subsequent examples cited here are from Satterthwaite (2007).

countries and over time. Satterthwaite (2007: 13–14) presents examples of how a metropolitan area can be assigned markedly different population sizes depending on how the area is defined. The population of Mexico City, for example, ranges from 1.7 million for the 'central city' to 19.4 million for the 'megalopolis of central Mexico'. Data on national urban populations are also necessary for primacy calculation, which has issues outlined in Cohen (2004).

This paper proposes an alternative to the UN measure of urban concentration that we call 'the agglomeration index (AI)'. It is based on three factors: population density, the size of the population in a 'large' urban centre, and travel time to that urban centre. Each factor used in the index is based on the conceptual framework of agglomeration economies. The index does not define what is urban per se; it does not incorporate urban characteristics such as political status and the presence of particular services or activities. Instead, the index creates a global definition of settlement concentration that could be used to conduct cross-country comparative analyses. The accessibility of this measure of economic density lends itself easily to the study of concepts such as agglomeration rents in urban areas, the 'thickness' of a market, and the travel distance to such a market with many workers and consumers.

This new measure of agglomeration does not suggest that the UN's data are flawed. The matter is analogous to measurements of global poverty and comparisons of poverty levels across countries. Each country has its own definition based on legitimate factors, but the varying definitions among countries make cross-country analysis and aggregation almost impossible. A uniform definition – such as the $1 or $2 a day index – makes such analysis possible. The agglomeration index is aimed at serving as such a systematic and unified measure of settlement concentration.

There are a few caveats inherent to the agglomeration index. It suffers from limited availability of data in the same way as UN data. In addition, there are some issues regarding the key assumptions associated with travel time, discussed later in this paper. Finally, but perhaps most importantly, a critical constraint is that currently the index is available for only one year (2000).

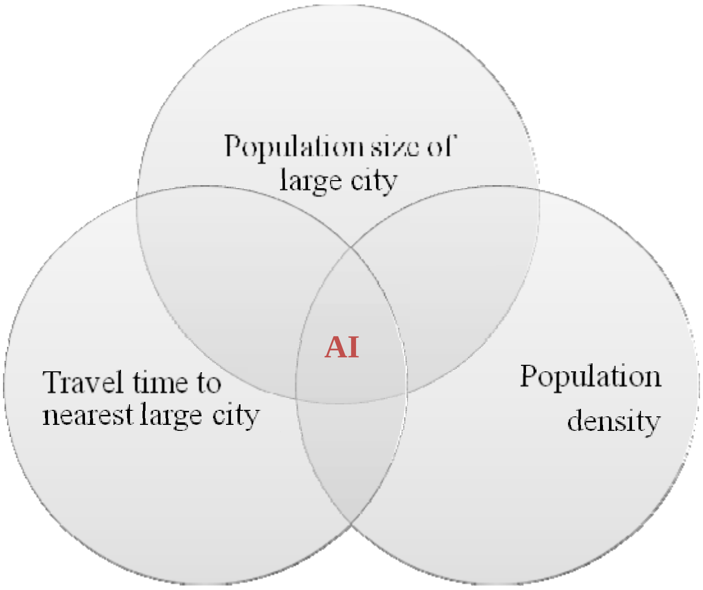
2 Concepts behind the agglomeration index

The proposed agglomeration index does not focus on the conceptual definition of 'urban'; rather, it focuses on the economic significance of urban areas. Residents, workers, and firms typically agglomerate in urban areas, which gives rise to the notion of agglomeration economies. Clearly, people and firms agglomerate because there are benefits from so doing. The literature cites several sources of agglomeration rent, such as the existence of 'thick' markets (both consumer and labour markets), ease of access to these markets, and the resulting, so-called forward and backward linkages associated with large local markets (Marshall 1961; Baldwin *et al.* 2003). We argue that the key indicators of the sources of agglomeration economies and rent can be summarized in the following three indicators: population size, population density, and travel time. More specifically, population size refers to the population of large cities that can be regarded as the focal points of an urban area, and travel time is to the nearest large city.

Consider a location outside the centre of a large city. For that location to take part in agglomeration economies, it must, generally speaking, have both a relatively high population density, which is a proxy for market thickness, and be reasonably near the large city centre, a proxy for market access and lower transportation costs. For example, even with a high population density, a small, isolated community such as a mining town is not likely to generate agglomeration economies. As depicted in Figure 1, locations that satisfy all three indicators are included in calculating the agglomeration index.

The urban area can thus be identified and delineated by the combination of the three indicators. This is a significant step forward, because measures of concentrations of settlement no longer depend on country-specific, and sometimes ad hoc, definitions of a city, an administrative boundary, or an urban area. Rather than treating urban and rural areas as dichotomous and discreet entities, we view the spatial transition as a gradient, a conceptual framework put forth by Chomitz *et al.* (2005).

Figure 1: Key indicators constituting the agglomeration index (AI)



Source: Created by the authors.

The gradient framework is illustrated in Figure 2. Assuming that density is highest at the centre of a large city, and that it gradually decreases as one moves away from the centre, in calculating the agglomeration index we defined some maximum travel time to the city centre as the cut-off point of an urban area (provided that the density criterion is met). Depending on the population density distribution, the density at the boundary varies; this is illustrated in the top panel of Figure 2 by two distributions, one being a mean-preserving spread of the other.

The difference in spatial distributions of population density raises another important point: Why are they different? It is easy to imagine that a region with poor transportation infrastructure will result in a distribution that is concentrated near the centre. A flatter distribution of population density might imply a better transportation infrastructure; accessibility over five miles of poor transportation infrastructure could be equivalent to ten miles over a good infrastructure system. In our agglomeration index, we used road networks as a proxy for transportation infrastructure. We assigned each road network to one of three categories of quality – low, medium, and high – and assigned a realistic travelling speed to each category. As a result, a physical distance identified by 60 minutes of travel time is further from the city centre along a paved road than along a dirt road. The bottom panel of Figure 2 presents the case in which concentric circle-like figures depict travel-time contours or isochrones. Note that further locations along the high-quality road are included in the urban area. For areas without roads, we assumed the mode of transportation is either by rail, via navigable rivers, on foot or by riding on an animal and, in calculating the foot or animal based travel time, incorporated information on land cover (such as forest, grassland, or barren) and the slope of land.

3 Data, and how the agglomeration index is calculated

The procedure to calculate the agglomeration index can be summarized as follows:

(1) Specify a threshold value for each of the three criteria: minimum population density, maximum travel time, and the minimum population size that defines large cities;

(2) Locate the centre of defined large cities from the GRUMP human settlements database;

(3) Determine the border surrounding that large city centre, based on the maximum travel time. This boundary is computed from a cost–distance model that estimates travel time to the city centre over a cost surface. This surface has a spatial resolution of approximately one kilometre and is derived from GIS data on:

- (i) The transport network;
- (ii) Off-road surfaces, derived from land cover data;
- (iii) The slope and estimates of the average travel speeds for each permutation of these data (see Table 3A.1 in the Appendix);

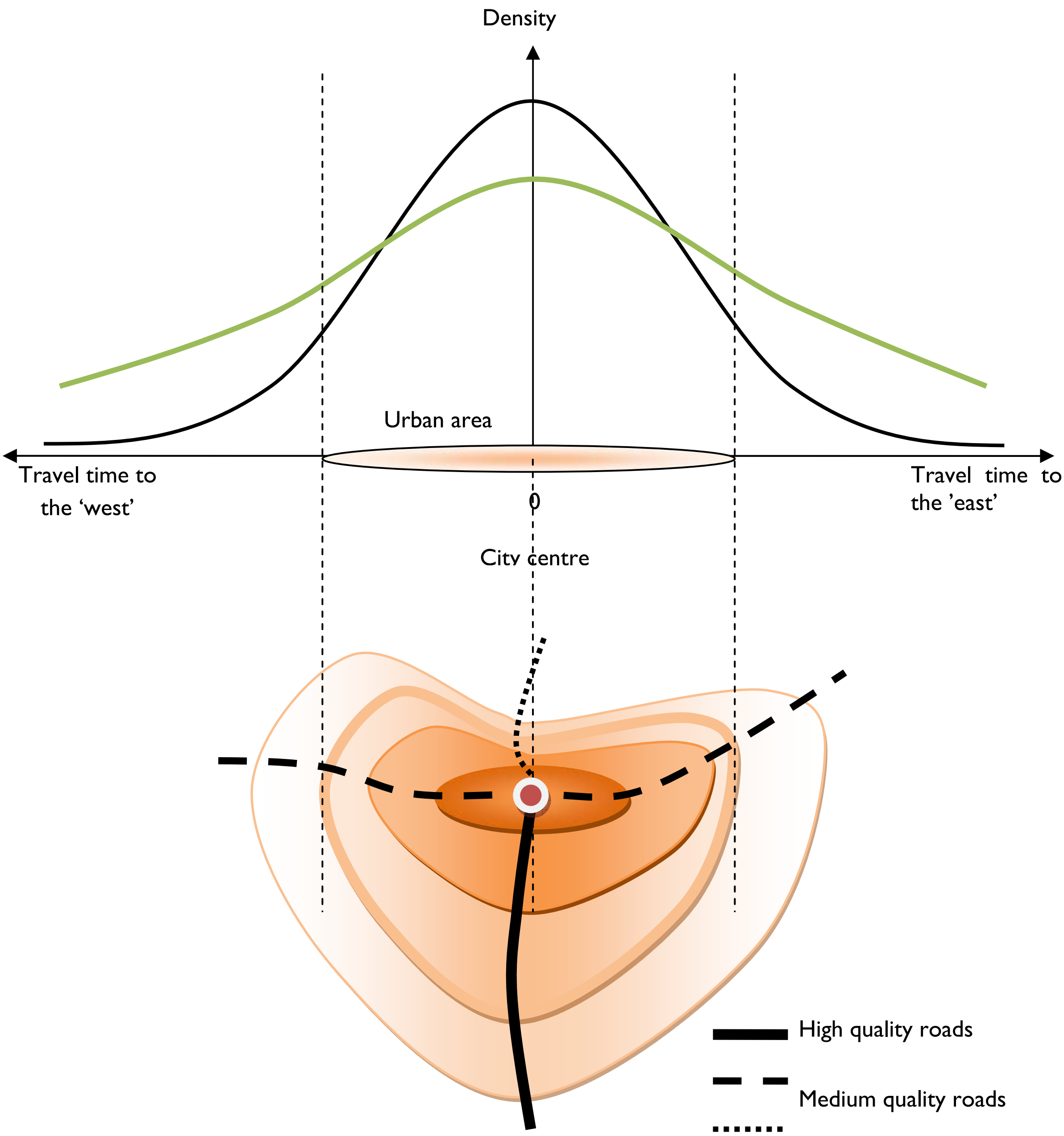
(4) Determine the population and population density at one kilometre resolution within this border. This is derived by averaging two GIS datasets on population for the year 2000: GRUMP and LandScan;[2]

(5) Sum the population in all the grid cells that satisfy all three thresholds. The result is analogous to urban population. The proportion of this number to that country's total population is the agglomeration index.

---

2 GRUMP (Global Rural-Urban Mapping Project) human settlement data is developed by the Center for International Earth Science Information Network (CIESIN) at Columbia University. The data were gathered, primarily, from official statistical offices (census data) and, secondarily, from other web sources, or from specific individual databases when official statistical databases were not available. Based on the data available and applying UN growth rates, population was estimated for 1990, 1995, and 2000 (cf. http://sedac.ciesin.columbia.edu/gpw/index.jsp). This implies that the population growth rate in GRUMP data is identical to that of UN data. LandScan was developed by Oak Ridge National Laboratory; http://www.ornl.gov/sci/landscan/

Figure 2: Concepts of the urban-rural gradient and travel time in the agglomeration index



Source: Created by the authors.

A country's census data form the foundation of the agglomeration index. The GRUMP population estimates are based on population data at the smallest available scale in the national census (state, province, county or district), combined with data from web sources, or from specific individual databases when official statistical databases were not available. GRUMP does not model population distribution within these administrative units. Conversely, the LandScan population estimates are based on population estimates for larger units, but the population within each unit is redistributed across the unit's grid cells based on likelihood coefficients, which are derived from other spatial data such as distance to roads, slope, and land cover. There are strengths and weaknesses in both of these population models. Generally LandScan tends to overestimate the population in urban areas and underestimate them in rural areas; the converse is the tendency for GRUMP. We therefore used both and took the computed population per one kilometre pixel as an average of both sources.

Travel time to large cities is based on estimates of the time required to travel one kilometre over the transport network as well as off-road surfaces. We constructed a cost surface from the data sources listed in Table 3A.1 and applied the estimated travel speeds to each permutation of surfaces (for example, major road on a moderate slope).[3] The cost surface contains the time in minutes to cross each one kilometre cell. The cost surface and the location of the city centres are input to the cost–distance model in order to determine the travel time to each city centre. This cost surface is available for download.[4]

4 Results

4.1 Comparison with the UN data

In order to make a comparison with the UN's urbanization rates, the agglomeration index is calculated based on a minimum density of 150 people per square kilometre, a maximum of 60 minutes for travel time, and a minimum population of 50,000 as the definition of a large city. Each country's agglomeration index and UN urban population share are given in Table 3A.2 in the Appendix. Here, we present the results by world-regions as defined by the World Bank: sub-Saharan Africa (SSA), Middle East and North Africa (MENA), South Asia (SAS), East Asia and the Pacific (EAP), Europe and Central Asia (ECA), and Latin America and the Caribbean (LAC). OECD and non-OECD high-income countries (other high-income economies, OHIE) are excluded from regional groups.

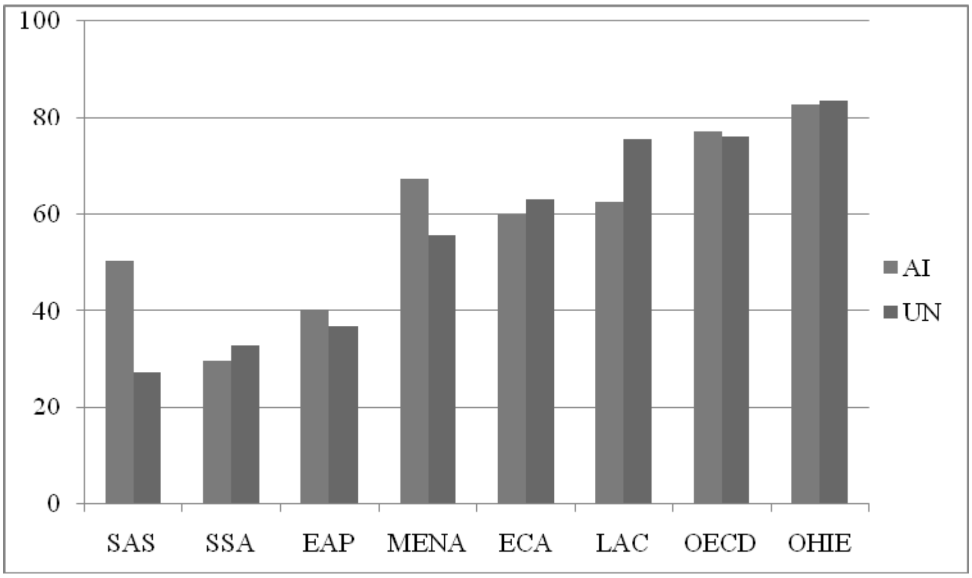
The results of the two calculations are qualitatively consistent, given the issues associated with the UN's data. For example, the agglomeration index for SAS is much higher than the UN result (50.4 per cent as opposed to 27.2 per cent). This suggests that the notion of South Asia as being densely populated but having a low urban population share might not be as paradoxical as it sounds. A similar pattern, where the agglomeration index is higher than the UN figure, is observed for the EAP and MENA regions. The converse is true, however, for regions such as LAC and SSA. The LAC region is of particular interest as this region has often been cited as being the most urbanized among the developing countries. According to the UN figure, urban areas accounted for 75.4 per cent of the region's population in 2000; the agglomeration index puts the urban share at 62.4 per cent. Thus, the LAC region is probably still heavily urbanized but not as much as commonly thought.

---

3 For example, 120km/hr on highways, 60km/hr on major roads, 10km/hr on tracks, and 40km/hr on railways.

4 http://bioval.jrc.ec.europa.eu/products/gam/index.htm

Figure 3: Agglomeration index (AI) and UN urban population share, 2000



Source: AI is calculated by the authors. UN data are compiled by the authors using data from UN (2006).

4.2 Sensitivity of the results to thresholds

The danger of using the agglomeration index to infer the state of urbanization in the world rather than a more conventional urban share is that the results can be manipulated to fit any conclusion by changing the threshold/criteria combination. It is very important, therefore, to have good justification for the combination chosen and to interpret the results as conditional on that choice. This cautionary note raises an interesting question: How sensitive is the agglomeration index to different combinations of thresholds?

We considered three levels of threshold for each of the three indicators and calculated an agglomeration index for every combination. Those levels are:

- Population density: 150, 300, and 500 people per square kilometre;
- Travel time: 30, 60, and 90 minutes;
- Large city's population: 50,000; 100,000, and 500,000.

The base combination is 150/km^2 minimum for density, 60 minutes maximum travel time, and 50,000 minimum population in order to qualify as a large city. The minimum density is what the OECD used for its statistics (Chomitz *et al.* 2005) and is equivalent to placing each person approximately 81.6 metres apart. We viewed this threshold as somewhat sparse, especially for certain parts of the world, which led us to try thresholds of 300 and 500 people per square kilometre. For the large city's population size, inclusion of a minimum population of 20,000 was suggested. However, it became apparent that the GRUMP settlement data do not have sufficient coverage at this low population level in many countries. Thus, we did not use this threshold level.

There are 27 different combinations in total; however, we present only combinations in which a single factor was increased or decreased from the base case and attempt to deduce trends. Figure 4(a) shows the results of changing the population density threshold while holding the other two criteria constant at 60 minutes (travel time) and 50,000 (population size). As expected, with a more stringent threshold the agglomeration index falls, though the magnitude is small. This implies, remarkably, that the results are fairly robust against the changes in population density threshold.

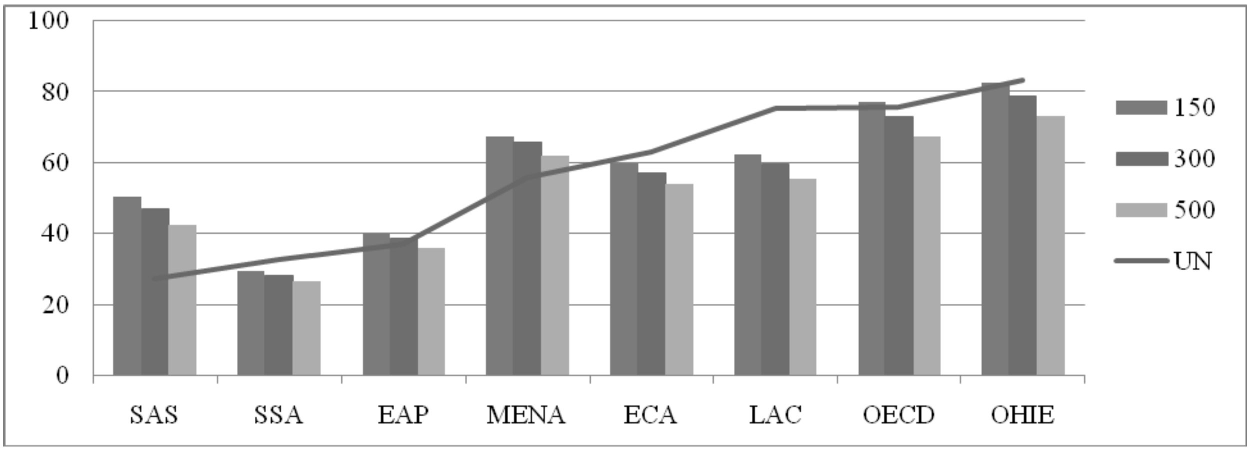
Figure 4(b) shows how the indexes alter as the travel time threshold is changed while holding the other two indicators constant at base-case levels. Unlike the case of the population density, changes in travel time result in large changes in the agglomeration index. The difference is particularly significant in the SAS and EAP regions; with a moderate increase in the threshold (from 60 to 90 minutes), the agglomeration index for the SAS region jumps to 60.8 per cent, more than double the UN figure. The results suggest that the choice of the travel time threshold can easily affect the conclusions about a country's urbanization state.

Figure 4(c) shows the effect when the minimum population size for large cities is changed. This case also generates a substantial change in the values of agglomeration indexes across the board. The drop in the index is particularly steep when the minimum population is increased from 100,000 to 500,000. The figure suggests that a threshold of 500,000 might not be suitable for many regions because it is too restrictive, which drives the agglomeration indexes too low – both in absolute terms and relative to the UN figures.

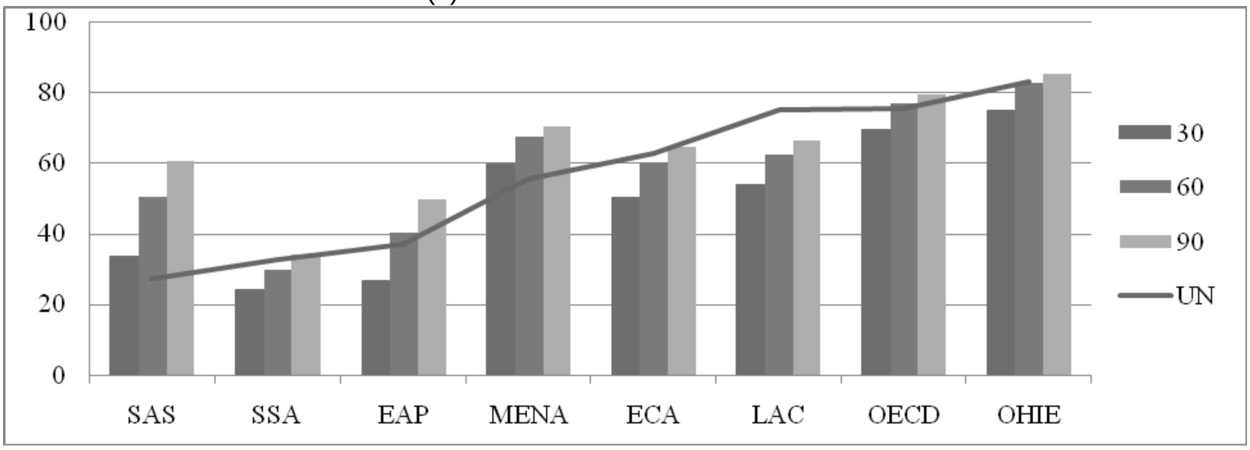
One last note regarding the choice of combinations of thresholds concerns feasibility. The current method estimates travel time from the centre of the nearest large city. However, large cities vary in the degree to which they are dispersed geographically. So, while 30 minutes away from a densely populated city of 50,000 could genuinely be outside of that city's boundary (however defined), this is less likely for a city of 500,000. The bottom line is that threshold combinations – such as a city having a minimum population of 500,000 and a travel time of 30 minutes or less – might simply be implausible.

Figure 4: Sensitivity to indicators

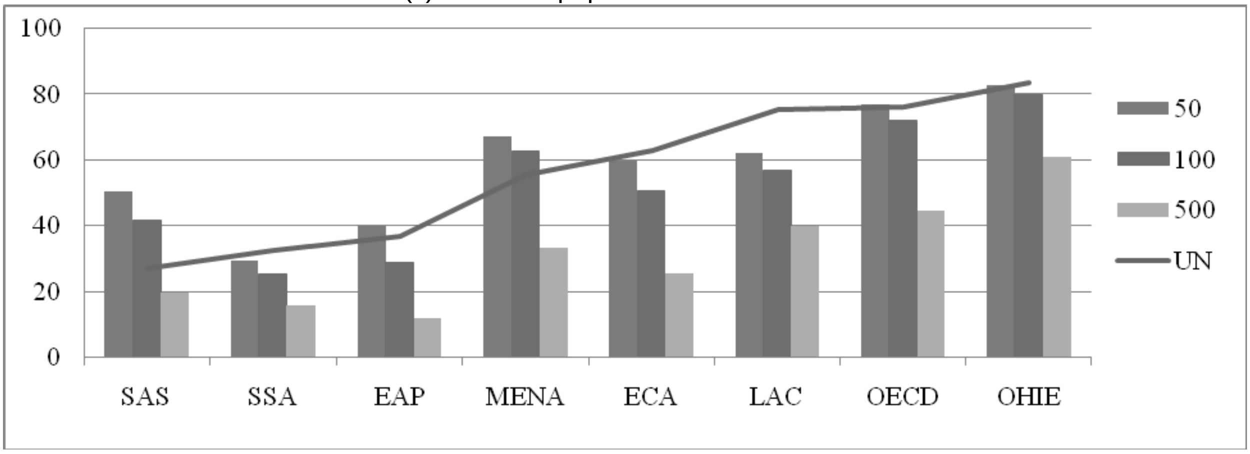
(a): Minimum population density



(b) Maximum travel time



(c) Minimum population size



Source: AI is calculated by the authors. UN data are compiled by the authors using data from UN (2006).

5 Discussions: strength, caveats, and promise

Provided that the conceptual framework and the variables used to calculate the agglomeration index are plausible, the main advantage of this measure is the consistent measurement across the countries that enables direct comparisons between them. The main strength is that the index is not influenced by country-specific definitions of what is urban. While it is not isolated from ad hoc definitions of city boundaries, since it uses city population as a criterion by which to define and locate large cities, the impact is much less severe. Unlike the primacy measure, the agglomeration index uses this information in its calculation, but only to identify the point at which the travel time is measured. As such, the accuracy of population counts is far less critical in the index than in the primacy measure used by Henderson (2003) and others.

The agglomeration index is designed solely to quantify the degree of settlement concentration. Nonetheless, it can make a significant contribution to the debate regarding urbanization. For example, over-concentration of settlement (or urbanization) is at the heart of discussions about issues such as the environmental footprint of a city, congestion problems, and provision of public service infrastructures in densely packed areas. However, the urbanization issues could be much less severe if, instead of a few large cities growing ever bigger, there are many small cities sprouting in what was a sparsely populated area. Since both types of cities will be defined as urban, urbanization data alone cannot effectively differentiate the two situations. In the agglomeration index, population density captures the concentration, and the population size of the nearest largest city distinguishes between the large cities increasing in size from the many small cities that are emerging. The impacts of concentrated settlements will be greater if the population distribution is skewed toward a single point, such as the city centre. This characteristic is captured by travel time. With a globally uniform definition, the agglomeration index might lay to rest some myths about urbanization in various regions of the world.

There are several notable caveats associated with the current version of the agglomeration index. The first is the availability of data related to information for population derived from censuses and road networks, which is the most important factor in constructing the cost surface; this is also the reason why the index is available for only one year. Fortunately, in terms of spatial resolution, census availability, and sub-national growth estimates, the improvements in available population data over the last 10 to 15 years have been immense. This trend is continuing, especially in developing countries. The problems associated with data on road networks are multi-layered. First, the information generally is old. We used data from VMAP0 (also known as the Digital Chart of the World, or DCW), which includes some very old road data.[5] Second, we would like to see more detailed information about the quality of roads. Our current

5 NIMA (2000) Vector Map Level 0 (Digital Chart of the World) edn 5, National Imagery and Mapping Agency. The date of the underlying road information covers a range of sources from the 1960s to the 1990s.

version of the index used three levels of quality; it could, however, be refined with more comprehensive data. Again, fortunately, this issue is being addressed on several fronts.[6]

On the conceptual front, the burning issue is how many people actually travel the roads – in other words, how many people in developing countries drive? We have witnessed newly built, multiple-lane highways in developing countries that are effectively empty because few people own cars. Thus, the existence of a road does not necessarily translate into people living on the outskirts of a city having easy access to the centre. Converse to this issue is how to capture the effect of congestion on roads, and also to incorporate the situation in which the majority of travel is by means of public transportation. Large cities in developed countries (Tokyo and London, for example) have well-developed mass transit system that many commuters and residents use, and these systems could spread quickly to cities in developing countries – especially in light of global warming and climate change. Thus, some measures of accessibility other than roads need to be developed.

Despite these caveats, developments on the horizon – including more extensive remote sensing technologies, and techniques for using and interpreting the resulting data – promise to resolve many of the limitations. With an ever-increasing number of accessible satellite images of high resolution, updating of road network data is within reach. Contiguous areas also can be determined directly using satellite images. Of course, manually determining these characteristics from the images will likely consume a large amount of manpower, which could limit the affordability of such endeavour. As a result, population models such as the GRUMP and LandScan are still needed. However, satellite images and other remote sensing technologies can reduce the reliance on census data for calculating the index.

---

6 An excellent summary of the situation and possible approaches to improving the data on roads can be found on the CIESIN website: http://www.ciesin.org/confluence/display/roads/.

Appendix

Table A.1:List of datasets used to calculate the cost surface

Data layer	Proxy for	Source
VMAP0 Road layer	Major and minor road network	http://earth-info.nga.mil/publications/vmap0.html
VMAP0 Rail layer	Railway network	http://earth-info.nga.mil/publications/vmap0.html
WDBII Rivers	Navigable rivers	www.ngdc.noaa.gov/seg/cdroms/ged_iib/datasets/b14/mw.htm
VMAP0 Country borders	Travel delay for crossing international borders	http://earth-info.nga.mil/publications/vmap0.html
GLWD Water bodies	Navigable water bodies	www.wwfus.org/science/data/globallakes.cfm
SRTM Elevation	Inaccessible areas of very high elevation	www2.jpl.nasa.gov/srtm/
SRTM Slope	Slope factor to reduce travel speed	www2.jpl.nasa.gov/srtm/
GLC2000 Global land cover map	Foot/animal based travel for off road and paths	www-gem.jrc.it/glc2000/defaultglc2000.htm

Travel time is estimated based on the combination of several GIS data layers that are merged into a cost surface layer that represents the time required to cross each pixel. The higher the cost of a pixel, the more time required to cross it.
The friction layer is composed from the following global GIS data layers.

Table A.2: Agglomeration index and UN urban population share comparison

Economy	Income	World region	Agg. Index (a)	Agg. Index (b)	UN Urban share (%)
Afghanistan	LIC	SAS	26.0	23.8	21.3
Albania	LMC	ECA	50.8	28.8	41.8
Algeria	LMC	MENA	56.9	49.8	59.8
American Samoa	UMC	EAP	-	-	87.9
Andorra	HIC	OHIE	-	-	92.4
Angola	LIC	SSA	26.7	26.5	50.0
Antigua and Barbuda	UMC	OHIE	-	-	37.7
Argentina	UMC	LAC	70.5	64.4	89.2
Armenia	LIC	ECA	71.8	68.0	65.1
Aruba (Neth.)	HIC	OHIE	-	-	46.7
Australia	HIC	OECD	75.2	69.9	87.2
Austria	HIC	OECD	58.1	54.0	65.8
Azerbaijan	LIC	ECA	48.7	32.1	50.9
Bahamas, The	HIC	OHIE	70.0	70.0	89.0
Bahrain	UMC	OHIE	99.3	99.3	94.6
Bangladesh	LIC	SAS	42.8	32.3	23.2
Barbados	UMC	OHIE	97.5	-	50.0
Belarus	LMC	ECA	59.8	53.0	70.0
Belgium	HIC	OECD	89.2	88.5	97.1
Belize	LMC	LAC	4.8	4.8	47.9
Benin	LIC	SSA	36.4	26.3	38.4
Bermuda	HIC	OHIE	-	-	100.0
Bhutan	LIC	SAS	7.0	-	9.6
Bolivia	LMC	LAC	54.7	51.8	61.8
Bosnia and Herzegovina	LMC	ECA	33.7	19.3	43.2
Botswana	UMC	SSA	23.8	14.3	53.2
Brazil	UMC	LAC	60.4	53.6	81.2
Brunei Darussalam	HIC	OHIE	71.7	-	71.2
Bulgaria	LMC	ECA	61.3	47.0	68.9
Burkina Faso	LIC	SSA	14.4	9.9	16.6
Burundi	LIC	SSA	17.8	7.6	8.6
Cambodia	LIC	EAP	24.8	19.1	16.9
Cameroon	LIC	SSA	41.2	34.4	50.0
Canada	HIC	OECD	71.0	67.0	79.4
Cape Verde	LMC	SSA	52.6	-	53.4
Cayman Islands	HIC	OHIE	-	-	100.0
Central African Republic	LIC	SSA	21.0	16.6	37.6
Chad	LIC	SSA	12.2	11.3	23.4
Channel Islands	HIC	OHIE	-	-	30.6
Chile	UMC	LAC	74.0	69.5	85.9
China	LMC	EAP	36.2	22.5	35.8
Colombia	LMC	LAC	64.8	59.9	71.2
Comoros	LIC	SSA	-	-	33.9
Congo, Dem. Rep.	LIC	SSA	24.9	23.0	29.8
Congo, Rep.	LIC	SSA	54.7	51.9	58.3
Costa Rica	LMC	LAC	55.4	54.5	59.0
Côte d'Ivoire	LIC	SSA	36.6	30.6	43.0
Croatia	UMC	ECA	34.2	21.8	55.6
Cuba	LMC	LAC	69.0	63.4	75.6
Cyprus	HIC	OHIE	66.6	-	68.7
Czech Republic	UMC	OECD	66.8	45.0	74.0
Denmark	HIC	OECD	58.1	50.2	85.1
Djibouti	LMC	MENA	50.4	50.4	83.4
Dominica	UMC	LAC	-	-	70.5
Dominican Republic	LMC	LAC	72.3	66.5	62.4
Ecuador	LMC	LAC	58.8	55.9	60.3
Egypt, Arab Rep.	LMC	MENA	92.6	90.2	42.5
El Salvador	LMC	LAC	72.3	68.1	58.4
Equatorial Guinea	LMC	SSA	25.2	-	38.8
Eritrea	LIC	SSA	20.9	19.0	17.8
Estonia	UMC	OHIE	49.9	44.4	69.3
Ethiopia	LIC	SSA	10.9	7.8	14.9
Faeroe Islands	HIC	OHIE	-	-	37.0
Fiji	LMC	EAP	21.3	21.3	48.3
Finland	HIC	OECD	52.0	40.0	61.1
France	HIC	OECD	71.4	66.2	75.8
French Guiana	N.A.	N.A.	40.1	-	-
French Polynesia	HIC	OHIE	-	-	52.5
Gabon	UMC	SSA	37.4	37.4	80.2
Gambia, The	LIC	SSA	46.3	-	-
Georgia	LIC	ECA	55.8	47.5	52.7
Germany	HIC	OECD	77.6	73.5	75.1
Ghana	LIC	SSA	36.6	30.2	44.0
Gibraltar	N.A.	N.A.	99.9	99.9	-
Greece	HIC	OECD	59.3	49.9	58.8
Greenland	HIC	OHIE	-	-	82.1
Grenada	UMC	LAC	-	-	30.4
Guadeloupe	N.A.	N.A.	76.5	-	-

Economy	Income	World region	Agg. Index (a)	Agg. Index (b)	UN Urban share (%)
Guam	HIC	OHIE	-	-	93.5
Guatemala	LMC	LAC	34.7	33.8	45.1
Guinea	LIC	SSA	25.3	20.9	31.0
Guinea-Bissau	LIC	SSA	21.9	21.9	29.7
Guyana	LMC	LAC	37.4	37.4	28.6
Haiti	LIC	LAC	36.9	26.2	35.6
Honduras	LMC	LAC	41.4	34.0	44.4
Hong Kong, China	HIC	OHIE	99.8	99.8	100.0
Hungary	UMC	ECA	67.6	51.8	64.6
Iceland	HIC	OECD	65.4	65.4	92.5
India	LIC	SAS	51.9	42.9	27.7
Indonesia	LMC	EAP	57.5	51.1	42.0
Iran, Islamic Rep.	LMC	MENA	62.8	55.4	64.2
Iraq	LMC	MENA	72.2	71.4	67.8
Ireland	HIC	OECD	45.6	39.2	59.1
Isle of Man	HIC	OHIE	-	-	51.9
Israel	HIC	OHIE	87.4	81.2	91.4
Italy	HIC	OECD	77.0	68.5	67.2
Jamaica	LMC	LAC	69.0	51.2	51.8
Japan	HIC	OECD	92.9	90.6	65.2
Jordan	LMC	MENA	79.4	76.5	80.4
Kazakhstan	LMC	ECA	51.3	45.4	56.3
Kenya	LIC	SSA	27.2	21.9	19.7
Kiribati	LMC	EAP	-	-	43.3
Korea, Dem. Rep.	LIC	EAP	51.8	51.8	60.2
Korea, Rep.	UMC	OECD	89.6	88.1	79.6
Kuwait	HIC	OHIE	90.3	89.4	98.2
Kyrgyz Republic	LIC	ECA	34.0	25.8	35.4
Lao PDR	LIC	EAP	12.5	12.2	18.8
Latvia	LMC	ECA	50.7	47.0	68.1
Lebanon	UMC	MENA	86.8	75.7	86.0
Lesotho	LIC	SSA	18.6	10.6	17.8
Liberia	LIC	SSA	20.4	19.0	54.3
Libya	UMC	MENA	83.4	76.6	83.1
Liechtenstein	HIC	OHIE	4.7	-	15.2
Lithuania	LMC	ECA	54.6	53.4	67.0
Luxembourg	HIC	OECD	73.6	1.7	83.9
Macao, China	HIC	OHIE	100.0	100.0	100.0
Macedonia, FYR	LMC	ECA	60.8	32.7	64.9
Madagascar	LIC	SSA	17.8	14.2	26.0
Malawi	LIC	SSA	17.6	15.5	15.1

Economy	Income	World region	Agg. Index (a)	Agg. Index (b)	UN Urban share (%)
Malaysia	UMC	EAP	68.7	66.1	61.8
Maldives	LMC	SAS	14.5	-	27.6
Mali	LIC	SSA	19.0	14.1	27.9
Malta	UMC	OHIE	-	-	93.4
Marshall Islands	LMC	EAP	-	-	65.4
Mauritania	LIC	SSA	26.8	23.1	40.0
Mauritius	UMC	SSA	91.9	90.5	42.7
Mexico	UMC	LAC	66.7	61.0	74.7
Micronesia, Fed. Sts.	LMC	EAP	-	-	22.4
Moldova	LIC	ECA	48.2	44.9	46.1
Monaco	HIC	OHIE	87.4	87.4	100.0
Mongolia	LIC	EAP	34.4	31.3	56.6
Morocco	LMC	MENA	53.6	46.9	55.1
Mozambique	LIC	SSA	24.8	22.7	30.7
Myanmar	LIC	EAP	30.1	22.7	28.0
Namibia	LMC	SSA	12.9	12.9	32.4
Nepal	LIC	SAS	24.4	14.6	13.4
Netherlands	HIC	OECD	89.4	86.8	76.8
Netherlands Antilles	HIC	OHIE	87.0	-	69.3
New Zealand	HIC	OECD	66.0	55.3	85.7
Nicaragua	LIC	LAC	48.4	41.4	57.2
Niger	LIC	SSA	16.5	11.9	16.2
Nigeria	LIC	SSA	43.1	37.3	43.9
Northern Mariana Islands	UMC	EAP	-	-	92.9
Norway	HIC	OECD	51.8	38.0	76.1
Oman	UMC	MENA	72.1	58.1	71.6
Pakistan	LIC	SAS	56.5	49.0	33.1
Palau	UMC	EAP	-	-	68.4
Panama	UMC	LAC	52.9	47.3	65.8
Papua New Guinea	LMC	EAP	7.6	5.1	13.2
Paraguay	LMC	LAC	48.7	45.3	55.3
Peru	LMC	LAC	50.1	46.7	71.6
Philippines	LMC	EAP	56.4	50.9	58.5
Poland	UMC	ECA	63.5	55.1	61.7
Portugal	HIC	OECD	60.0	59.1	54.4
Puerto Rico	UMC	OHIE	87.7	80.2	94.7
Qatar	HIC	OHIE	87.0	87.0	95.0
Romania	LMC	ECA	61.2	49.9	54.6
Russian Federation	LMC	ECA	63.0	56.1	73.4
Rwanda	LIC	SSA	10.6	10.6	13.8
Samoa	LMC	EAP	-	-	22.0

Economy	Income	World region	Agg. Index (a)	Agg. Index (b)	UN Urban share (%)
San Marino	HIC	OHIE	11.1	11.1	92.6
São Tomé and Príncipe	LIC	SSA	58.6	-	53.6
Saudi Arabia	UMC	OHIE	79.3	74.2	79.9
Senegal	LIC	SSA	47.9	44.1	40.6
Serbia and Montenegro		ECA	57.7	42.2	51.6
Seychelles	UMC	SSA	-	-	50.6
Sierra Leone	LIC	SSA	33.1	27.2	37.0
Singapore	HIC	OHIE	100.0	100.0	100.0
Slovak Republic	UMC	ECA	55.4	19.9	56.3
Slovenia	HIC	OHIE	42.3	41.2	50.7
Solomon Islands	LIC	EAP	8.3	-	15.8
Somalia	LIC	SSA	20.7	19.8	33.3
South Africa	UMC	SSA	49.4	45.3	56.9
Spain	HIC	OECD	75.3	71.4	76.3
Sri Lanka	LMC	SAS	44.1	33.7	15.7
St. Kitts and Nevis	UMC	LAC	-	-	32.5
St. Lucia	UMC	LAC	84.0	-	27.9
St. Vincent and the Grenadines	LMC	LAC	-	-	44.8
Sudan	LIC	SSA	30.7	26.4	36.1
Suriname	LMC	LAC	73.2	73.2	72.1
Swaziland	LMC	SSA	19.0	-	23.4
Sweden	HIC	OECD	53.8	41.0	84.0
Switzerland	HIC	OECD	72.8	64.9	73.1
Syrian Arab Republic	LMC	MENA	59.1	56.9	50.1
Taiwan, China	HIC	OHIE	84.4	82.7	-

Economy	Income	World region	Agg. Index (a)	Agg. Index (b)	UN Urban share (%)
Tajikistan	LIC	ECA	36.2	23.3	25.9
Tanzania	LIC	SSA	25.8	20.5	22.3
Thailand	LMC	EAP	36.9	28.6	31.1
Togo	LIC	SSA	31.3	20.1	36.6
Tonga	LMC	EAP	-	-	23.0
Trinidad and Tobago	UMC	OHIE	88.0	87.1	10.8
Tunisia	LMC	MENA	51.9	39.0	63.4
Turkey	LMC	ECA	62.5	52.8	64.7
Turkmenistan	LIC	ECA	43.5	36.2	45.1
Uganda	LIC	SSA	25.0	14.3	12.1
Ukraine	LIC	ECA	61.7	52.8	67.1
United Arab Emirates	HIC	OHIE	75.1	70.5	77.4
United Kingdom	HIC	OECD	84.7	83.5	89.4
United States	HIC	OECD	72.3	65.9	79.1
Uruguay	UMC	LAC	62.2	51.4	91.4
Uzbekistan	LIC	ECA	58.8	49.9	37.3
Vanuatu	LMC	EAP	-	-	22.0
Venezuela, RB	UMC	LAC	78.5	72.9	91.1
Vietnam	LIC	EAP	49.5	41.2	24.3
Virgin Islands (U.S.)	HIC	OHIE	-	-	92.8
West Bank and Gaza	LMC	MENA	90.9	90.4	71.5
Yemen, Rep.	LIC	MENA	25.5	22.8	25.4
Zambia	LIC	SSA	30.5	27.8	34.8
Zimbabwe	LIC	SSA	33.2	30.0	33.8
World			**52.0**	**43.8**	**46.7**

Note: For agglomeration index, column (a) uses largest city size threshold of 50,000 or more, and column (b) uses the threshold of 100,000 or more.
Source: Authors' calculation (for agglomeration index); UN (2006).

References

Ades, F., and Glaeser, E.L. (1995). 'Trade and Circuses: Explaining Urban Giants'. *Quarterly Journal of Economics*, 110: 195–227.

Baldwin, R., Forslid, R., Martin, P., Ottaviano, G., and Robert-Nicoud, F. (2003). *Economic Geography and Public Policy*. Princeton, NJ: Princeton University Press.

Chomitz, K., Buys, P., and Thomas, T.S. (2005). 'Quantifying the Rural–Urban Gradient in Latin America and the Caribbean'. WB Policy Research Working Paper 3634. Washington, DC: World Bank.

Cohen, B. (2004). 'Urban Growth in Developing Countries: A Review of Current Trends and A Caution regarding Existing Forecasts'. *World Development*, 32(1): 23–51.

Henderson, J.V. (2000). 'The Effects of Urban Concentration on Economic Growth'. NBER Working Paper 7503. Cambridge, MA: The National Bureau of Economic Research.

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 P. Aghion and S. Durlauf (eds), *Handbook of Economic Growth*, vol. 1. Amsterdam: North Holland.

Marshall, A. (1961). *Principles of Economics*, 9th edn. London: Macmillan.

NIMA (2000). Vector Map Level 0 (Digital Chart of the World), edn 5. Reston, VA: National Imagery and Mapping Agency.

Rosen, K., and Resnick, M. (1980). 'The Size Distribution of Cities: An Examination of the Pareto Law and Primacy'. *Journal of Urban Economics*, 8: 165–86.

Satterthwaite, D. (2007). *The Transition to a Predominantly Urban World and Its Underpinnings*. London: International Institute for Environment and Development.

UN (United Nations) (2006). *World Urbanization Prospects: The 2005 Revision*. New York: UN, Department of Economic and Social Affairs, Population Division.

Wheaton, W., and Shishido, H. (1981). 'Urban Concentration, Agglomeration Economies, and the Level of Economic Development'. *Economic Development and Cultural Change*, 30: 17–30.