Difference Within Peers: The Infrastructure Stock in the Least Developed Countries

Lisa Borgatti

UNCTAD, Geneva Switzerland

Email: Lisa.borgatti@unctad.org

Abstract:

The decrepit state of infrastructure in many least developed countries (LDCs) constitutes a major obstacle for these countries to grow and develop. Yet, the differences in the infrastructure stocks within the group of LDC countries are too often overlooked. Using a cluster analysis technique, this paper shows that LDCs with poor infrastructure tend to be large African LDCs with a low population density, while the group with the best infrastructure stocks is composed of small LDCs with high population and urbanization rates. The paper highlights the worrying trend that less than a handful of LDCs managed to improve their infrastructure over time. Most of them are still grappling with the same infrastructure stock they had 13 years earlier.

Keywords: infrastructure, cluster analysis, least developed countries

JEL: C69, L91, L98, O55

1. Introduction

Available data on transport, energy and telecommunications indicates that the least developed countries (LDCs) have the worst infrastructure stock in the world and the lowest quality of infrastructure services. In 1999, the length of roads per square kilometre and per capita were about half the level in other developing countries, and only 22% of LDC roads were paved, compared to 43% in other developing countries. In 2003, fixed and mobile phone densities were 11% of the level in other developing countries. In 2002, electricity consumption per capita in the LDCs was a mere 7% that of the other developing countries (UNCTAD, 2006). While it is well known that the LDCs have the lowest level of infrastructure in the world, the real and concrete differences in the stocks of infrastructure available to them are too often overlooked. This paper groups LDCs according to their state of infrastructure in order to highlight the existing gaps within and between them.

Rather than discussing single variables, this paper assesses the available infrastructure stocks at the country level using both an ad hoc infrastructure index and cluster analysis techniques. The infrastructure index is used (i) to show the extent of the gap between the infrastructure stocks of the LDCs and that of the other developing countries, and (ii) to assess the evolution of the infrastructure stock in the LDCs over time.

The goals of this paper are twofold. First, it attempts to highlight the differences that exist within the LDCs in terms of infrastructure stocks. Second, it investigates whether the infrastructure stocks of the individual LDCs have improved, worsened or stagnated over the past 13 years.

The data for the analysis was taken from the infrastructure database compiled by Estache and Goicoechea (2005). Notwithstanding this recent data gathering effort, data unavailability seriously limits the extent of the analysis, which is only based on 31 LDCs.

The paper is organized as follows: Section 2 shows the extent of the gap between the infrastructure stocks available to the LDCs and other developing countries. Section 3 highlights the differences in terms of infrastructure stocks within the LDCs and groups them according to whether they have good, average or poor infrastructure. Section 4 contains an analysis of whether the infrastructure stocks of the individual LDCs have improved, stagnated or worsened over time. Section 5 concludes.

2. Infrastructure stocks in the developing countries

Using the internationally-comparable database compiled by Estache and Goicoechea (2005), an infrastructure index was created for 114 developing countries (31 of which are LDCs) for which data is available. The unavailability of data seriously limits the extent of the analysis as the construction of the index requires full data for three out of the four variables used. The selection of the variables was therefore made on the basis of the widest country coverage. Four variables have been selected and used to construct the index, as they properly represent the existing stock of physical infrastructure as well as the quality of it.

The variables considered are: (i) the percentage of roads paved (PvRoads), (ii) the road density measured as the km of roads per 1000 people, (iii) the percentage of population with access to electricity, and (iv) the phone-density, which measures the total number of mobile and fixed phone subscribers per 1,000 people.

The above four variables have been normalized to zero mean and variance equal to one, *N(0,1)*, as in Limao and Venables (2001). The linear average over the four variables is taken to obtain a single indicator per country and over time. Only those countries that have data for a minimum three out of four variables are retained and the remaining missing observations are ignored.

The infrastructure index obtained from the normalization is visually represented in chart 1. High income developing countries are at the forefront with the best infrastructure indices, while all LDCs are located below average (μ) and below $\mu - s.d.$ (standard deviation). Some Asian LDCs (Bhutan, Laos and Yemen) stand out as having the best infrastructure indices of the LDCs and having better infrastructure than some other developing countries, such as Zimbabwe, Cote d'Ivoire, Gabon, to cite a few. The majority of the African LDCs (with the notable exceptions of Mauritania and Senegal) have the worst infrastructure indicators of the entire 114 countries sample. Out of the 30 developing countries with the worst infrastructure indicators, 24 are LDCs, mostly located in Africa. Furthermore, quite a few non-LDC African countries have an infrastructure index that is not too different from that of the African LDCs. This is rather unfortunate as it limits the regional benefits that can be gained by neighbouring a more developed country with better infrastructure. Mozambique, for example, should be benefiting greatly from the regional infrastructure projects that have been set up with South Africa.

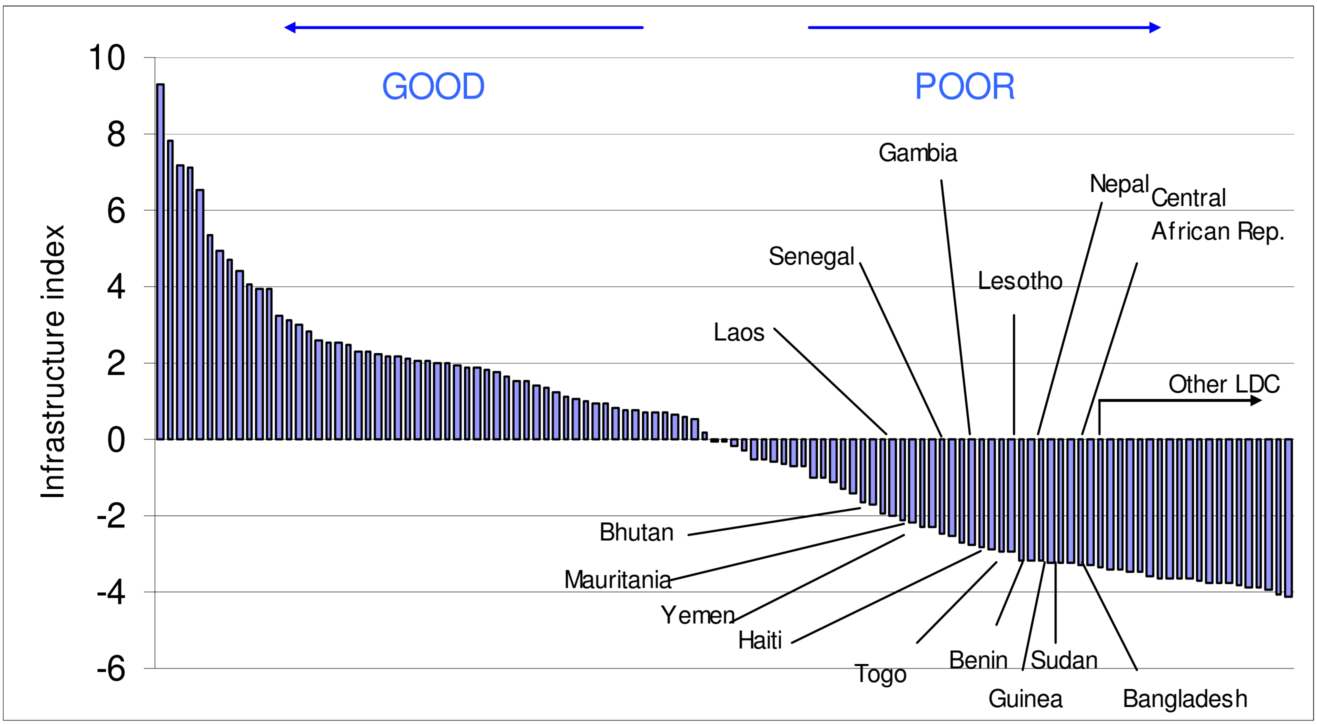
Such a result was somewhat expected, as the infrastructure stock tends to be highly correlated with countries' incomes. Chart 1 also shows that not all LDCs have the same infrastructure endowment: Some are better (worse) endowed than others. The next section attempts to group the LDCs according to the similarities in their infrastructure stocks using cluster analysis techniques.

3. Infrastructure stocks within the LDCs: differences and similarities

A non-hierarchical K-mean cluster analysis (see Annex 1 for details) is used to group the LDCs according to the similarity in their infrastructure stocks. This analysis relies on the same four variables discussed in section 2 and focuses on the 31 LDCs for which data is available. Table 1 summarises the results of the cluster analysis and groups the LDCs accordingly.

The interpretation of the results of this cluster analysis involves a certain degree of subjectivity. The cluster of LDCs with poor infrastructure, or Group III, contains 6 LDCs that have on average, the highest road density, while having the lowest phone-density, share of paved roads as well as the electrification rate. This cluster only has just a quarter of the electrification rate, the phone-density and the paved roads of the LDCs with the best infrastructure level, or Group I. The stock of infrastructure for those 14 LDCs with average infrastructure (or Group II) is slightly better than that of Group III, but still

Chart 1: Infrastructure indicator for 114 developing countries, 2003.



Source: Author's calculations.

Table 1. LDC Classification according to K-means cluster analysis, 2003

Clusters	Countries	Elec-tricity	Phone-density	Pv-Roads	Road-density	GDP pc at 2000 $	Populat density (per km)	Area (sq km)	Urbaniza-tion (% of total)
Group I: LDCs with best infra-structure (11 LDCs)	Benin Bhutan Gambia Haiti Laos Lesotho Mauritania Senegal Sudan Togo Yemen	21.5	64.8	29.4	2.4	443.6	73.4	421,129	35.0
Group II: LDCs with average infrastructure (14 LDCs)	Bangladesh Burkina Faso Cambodia Eritrea Ethiopia Malawi Mali Mozambique Nepal Niger Rwanda Sierra Leone Tanzania Uganda	8.6	18.7	13.7	1.5	235.9	157.7[a]	454,739	22.5
Group III: LDCs with poor infra-structure (6 LDCs)	Angola Central African Republic Chad Guinea Guinea-Bissau Madagascar	5.7	12.3	8.8	4.3	343.4	23.0	664,043	33.0

Source: Author's calculations based on Estache and Guinoechea (2005) and World Bank, World Development Indicators, 2005.

Note: The data reported in the table are not normalized and should be read with the usual reference unit measures.

[a] Excluding Bangladesh, the population density falls to 88.3.

far below that of Group I. The average electrification rate of Group II is less than half that of Group I. Furthermore, their phone-density and share of paved roads are a third of the average of the group with best infrastructure. Out of 31 LDCs, only 11 belong to Group I. Although the four variable averages for Group I are still below the averages for the other developing countries, the Gambia, Mauritania and Senegal seem to have better infrastructure than some other developing countries. Furthermore, there seems to be a clear geographical difference in the infrastructure stock available to the LDCs: Asian LDCs have better infrastructure than the African LDCs. Asian LDCs are included in Groups I and II, while the cluster with poor infrastructure, or Group III, is only composed of African LDCs.

As expected, the 11 LDCs of Group I have the highest real GDP per capita. Group III has a higher real per capita GDP than Group II, due to the presence of Angola and Chad, two oil-exporting countries. If they were excluded from the computation, the average real GDP per capita for the LDCs with poor infrastructure would have been 257.11, which is not drastically different from the mean of group II.

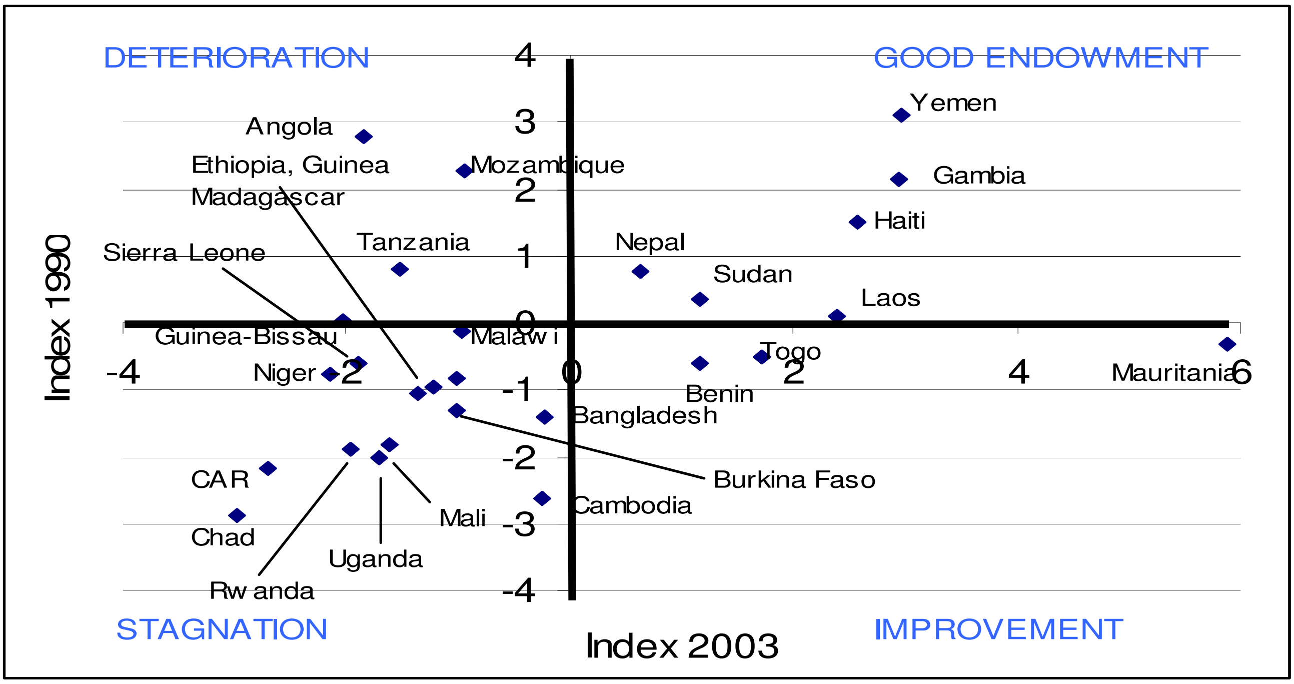
The group of LDCs with the best infrastructure is composed of relatively small countries, in terms of land area, with a high population density as well as the high shares of urbanization. Mauritania and Senegal, which are amongst the LDCs with the best infrastructure, have the two highest urbanization rates (above 49.6%) of the countries considered. The group of LDCs with the worst infrastructure is composed of large countries with a low population density. The size of the LDCs included in group III is a third bigger than that for the LDCs of Group I. Furthermore, the population density for the poor infrastructure cluster is only a third of that for the LDCs with the best infrastructure.

Angola, Central African Republic, Chad, Guinea, Guinea-Bissau and Madagascar have the poorest infrastructure stock of the LDCs considered. It is rather surprising to find two oil exporting countries with poor infrastructure. This finding raises some policy concerns over the future development-oriented strategies that could make for better use of the oil revenues. Investing in infrastructure-building and/or infrastructure-upgrading projects has large poverty-reducing effects (see for example, ADB et al., 2005), as well as high benefit/cost ratios (see Fan et al., 2004 and 2005 for country studies on the benefit/cost ratios for different types of public investments, one of which covers road building).

4. Have the infrastructure stocks of the LDCs improved over time?

Using the same methodology and variables employed in section 2, two infrastructure indices have been calculated to assess and compare the stock of infrastructure available in both 1990 and 2003. Chart 2 scatters the infrastructure indicators calculated for 28 LDCs for the years 1990 and 2003, and table 2 summarises the main results by listing the countries according to the evolution of their infrastructure stock.

Chart 2: Evolution of the infrastructure index for 28 LDCs, 1990 vs 2003



Source: Author's calculations

The **I** quadrant contains the countries that had good infrastructure in both 1990 and 2003 compared to the remaining LDCs. The **II** quadrant contains the countries that have experienced a worsening of their infrastructure levels from 1990 to 2003. The **III** quadrant contains those LDCs whose infrastructure stock has stagnated over time, while the **IV** quadrant contains the LDCs that have experienced an improvement in their infrastructure stock from 1990 to 2003.

Out of the 28 LDCs analysed, only three have improved their infrastructure stocks, while the infrastructure of 14 LDCs has stagnated, and that of the other four has deteriorated over time. Amongst the seven LDCs that have had a good infrastructure level in 1990 and 2003, only Nepal was clustered in Group II (see table 1 for reference). The remaining LDCs belong to the cluster with the best infrastructure (Group I).

The infrastructure stock of Angola, Guinea-Bissau, Mozambique and Tanzania had shown a marked deterioration in 2003 compared to 1990, due to conflicts, switch in external aid from physical to social infrastructure, and the failure of the private sector to replace the Government in the provision of these services (UNCTAD, 2006). The three countries that have improved their infrastructure stock over time, namely Benin, Mauritania and Togo, have been clustered in Group I.

This analysis has shed light on a worrying trend: The vast majority of the LDCs has not managed to upgrade their infrastructure stock over time. More attention should be devoted to this issue.

5. Conclusion

In spite of having the lowest level of infrastructure in the world, there are large differences amongst the levels of infrastructure within the LDCs.

This paper has used a non-hierarchical cluster method, the K-means, to cluster the LDCs according to whether their infrastructure stock was good, average or poor. The analysis has been carried out relying almost exclusively on the internationally-comparable infrastructure database compiled by Estache and Goicoechea (2005). Due to data unavailability, the cluster analysis was carried out for only 31 LDCs and relied on four main variables. The analysis reflects the infrastructure situation of 2003, the latest available year, as well as the change of infrastructure stock that occurred with respect to 1990. It showed that the vast majority of the LDCs considered have not successfully upgraded its infrastructure stock over time. This is an issue which clearly deserves further attention.

Gambia, Mauritania and Senegal have been found to have the best infrastructure level of the LDCs analysed, while Angola, Central African Republic, Chad, Guinea, Guinea-Bissau and Madagascar have the worst. The presence of two oil-exporting LDCs among those countries with the worst infrastructure calls for some targeted Government initiatives aimed at devoting a share of oil revenues into infrastructure-building or upgrading.

Table 2:. The evolution of infrastructure stocks in the LDCs between 1990 and 2003.

Quadrants	Description	Countries
I	Good infrastructure stock (7 LDCs)	Bhutan, Gambia, Haiti, Laos, Nepal, Sudan, Yemen
II	Deterioration of infrastructure stock (4 LDCs)	Angola, Guinea-Bissau, Mozambique, Tanzania
III	Stagnation of infrastructure stock (14 LDCs)	Bangladesh, Burkina Faso, Cambodia, Central African Republic, Chad, Ethiopia, Guinea, Madagascar, Mali, Malawi, Niger, Rwanda, Sierra Leone, Uganda
IV	Improvement of infrastructure stock (3 LDCs)	Benin, Mauritania, Togo

Source: Author's calculations based on chart 2

About the author:

Lisa Borgatti has a Ph.D in international economics from the Graduate institute of international Studies of Geneva, Switzerland. She has been working for the Division of Africa, LDCs and Special Programmes, UNCTAD, since 2002. She has contributed to the UNCTAD Least Developed Countries Reports 2004, 2006, and 2007.

Endnotes

i. The latest year available.

ii. Angola, Bangladesh, Benin, Bhutan, Burkina Faso, Cambodia, Central African Republic, Chad, Eritrea, Ethiopia, Gambia, Guinea, Guinea-Bissau, Haiti, Laos, Lesotho, Madagascar, Malawi, Mali, Mauritania, Mozambique, Nepal, Niger, Rwanda, Senegal, Sierra Leone, Sudan, Tanzania, Togo, Uganda, and Yemen.

iii. This assumption implies that missing observations take the same average value as the non-missing observations.

iv. The number of clusters that best minimizes the distance between each observation and the cluster mean, and maximizes the distance between each cluster, is three. A sensitivity analysis, carried out using the Calinski and Harabasz Pseudo F-index, shows that the index is highest when the countries are grouped in 3 clusters (12.47), followed by 8 clusters (12.01).

v. For a stability check, a second cluster analysis was carried out to verify whether the country groupings were steady or not. Three other variables were added to the initial specification, the choice of which was made dependent only on the best data coverage, namely phone density outside the largest city, the cost of a local phone call (in US dollars), and the share of population with access to improved sanitation facilities. The results of this analysis show that out of 31 countries only 3 have changed groups. Benin moved to Group II, while Lesotho and Tanzania moved to Group III.

vi. Eritrea, Lesotho and Senegal have been excluded for lack of data in 1990.

References

1. Asian Development Bank (ADB), Japan Bank for International Reconstruction and World Bank (2005) Connecting East Asia: A new Framework for Infrastructure. http://www.adb.org/Projects/Infrastructure-Development/Infrastructrue-study.pdf
2. Berlage, L. and Terweduwe, D. (1988) "The Classification of Countries by Cluster and by Factor Analysis", World Development, 16(12), 1527-1545.
3. Estache, A. and Goicoechea, D. (2005) "A Research Database on Infrastructure Economic Performance", World Bank Policy Research Working Paper No. 3643.
4. Fan, S., Zhang, X., Rao, N. (2004) "Public Expenditure, Growth and Poverty Reduction in Rural Uganda", DSGD Discussion Paper No. 4, International Food Policy Research Institute, Washington, DC.
5. Fan, S., Nyange, D., Rao, N. (2005) "Public Investment and Poverty Reduction in Tanzania: Evidence from Household Survey Data". DSGD Discussion Paper No. 18, International Food Policy Research Institute, Washington, DC.
6. Limao, N. and A. Venables (2001) 'Infrastructure, Geographical Disadvantage, Transport Costs and Trade', The World Bank Economic Review, 15(3): 451-479.
7. United Nations Conference on Trade and Development (UNCTAD) (2006) The Least Developed Countries Report. Develop Productive Capacities. UN Sales No. F.06.II.D.9

Annex 1: A Brief Review of the Methodology

A cluster analysis is a statistical technique that allows for the creation of homogenous groups of variables without prior information on the classification of the data. "The objective is to sort observations into groups called clusters so that the degree of statistical association is high among members of the same group and low between members of different groups" (Berlage and Terweduwe, 1988:1529). Each cluster is composed of elements that have a small distance from each other and a relatively large distance from the elements of another cluster. In other words, all available variables for *n* countries are classified in a given number of clusters *c* characterised by (i) a small variability within the cluster and (ii) a large variability across different clusters.

Although there are various ways to calculate the distance or proximity between two observations, this paper uses the most commonly used distance function, i.e. the Euclidean distance function, $d_{i,j}$, which is calculated as follows.

$$d_{i,j} = \left[\sum_{K=1}^{p} (X_{ik} - X_{jk})^2 \right]^{1/2}$$

where $d_{i,j}$ represents the distance between observations *i* and *j*, X_{ik} is the value of the *ith* observation of the variable *k*, and *i= 1,...,n.*

This paper uses the K-means clustering method, which is one of the most used non-hierarchical methods. A hierarchical procedure (i.e. average linkage method, not shown here) has however been used to calculate the initial number of clusters used as starting point for the K-means clustering analysis.

The K-means clustering method allocates the observations to a specified number of clusters in an iterative way in order to minimize the distance between each observation and the cluster means. The error component of the K-means can be defined as

$$E[P(n,c)] = \sum_{i=1}^{c} \sum_{j=1}^{n} \delta_{ji} d_{j,i}^{2}$$

where $P(n,c)$ stands for the partition of *n* observations into *c* clusters and δ_{ji} is an indicator function that takes the value 1 if the *jth* observation is in cluster *i* and 0 otherwise. The error component is calculated for each observation until no improvement in the within-cluster variance can be reached resulting in an optimal allocation of the *n* observations into the *c* clusters.

Furthermore, the Calinski & Harabasz pseudo-F index was used to identify the number of clusters that best maximise the distance function. This index measures the separation between clusters and is calculated as follows:

$$\frac{S_b/(k-1)}{S_w/(n-k)}$$

where S_b is the sum of squares between the clusters, S_w is the sum of squares within the clusters, *k* is the number of clusters and *n* is the number of observations. The higher the Calinski & Harabasz pseudo F-index, the greater the divergence between the clusters and, therefore, the best the country groupings that result from the analysis.