The World Bank's Framework for Statistical Capacity Measurement: Strengths, Weaknesses, and Options for Improvement

Floribert Ngaruko[1]

Abstract

Using the results-chain approach to capacity building and the PARIS21 framework, this paper discusses the use of statistical activities and outputs by the World Bank to measure statistical capacity. The paper focuses on the strengths and weaknesses of the World Bank's approach, and explores options for improving the indicator that results from it. While international comparability and cost effectiveness are the main strengths of the approach, the overemphasis that it places on statistical activities and outputs to the detriment of characteristics of statistical systems and data-producing agencies represents its main weakness, which causes the indicator to capture performance instead of actual capacity. To improve the World Bank approach would require refining the method of aggregation of the ratings of various aspects of statistical capacity, activities and outputs, and more critically, to take due account of statistical capacity utilization, which is the missing link in the World Bank approach between statistical activities/outputs and statistical capacity.

Key Words: *Statistical capacity indicator, Capacity building, Capacity utilization, PARIS21, Capacity building results-chain framework*

Résumé

En recourant au cadre conceptuel de PARIS21 ainsi qu'à l'approche par la chaîne des résultats aujourd'hui utilisée dans le renforcement des capacités, cet article discute de l'utilisation par la Banque mondiale, des activités et de la production statistiques pour mesurer les capacités des systèmes statistiques des pays. L'article identifie les forces et les faiblesses de cette approche, et explore les options pour améliorer l'indicateur qui en résulte. Si la comparabilité au plan internationale et le faible coût sont les principaux atouts de l'approche, sa focalisation sur les activités et la production statistiques au détriment des caractéristiques des systèmes statistiques et des agences de production des données reste sa principale faiblesse, qui fait que l'indicateur reflète la performance

[1] Coordinator of the ACBF Technical Advisory Panel and Network on National Statistics and Statistical Systems (STATNET); The African Capacity Building Foundation (ACBF); ZB Life Towers; Cnr. Jason Moyo Ave / Sam Nujoma St.; Harare; Zimbabwe (F.ngaruko@acbf-pact.org); (00-263) 4 702 931/2; 790 398/9; 700 208/10/14; 799 783/87; 799 810/12.

This paper does not necessarily reflect the views of the ACBF or STATNET.

davantage que les capacités à proprement parler. Pour améliorer l'approche, le papier suggère l'affinement de la méthode d'agrégation des indicateurs des différentes composantes des systèmes statistiques, des activités et de la production statistiques, et surtout la prise en compte du degré d'utilisation des capacités existantes qui est le chaînon manquant dans l'approche de la Banque mondiale, entre activités/production et capacités statistiques.

Mots clés : *Indicateur de capacités statistiques, Renforcement des capacités, Utilisation des capacités, PARIS21, Approche du renforcement des capacités par la chaîne des résultats.*

1. INTRODUCTION

The World Bank (henceforth WB) has been rating countries around the world for their statistical capacity since 2004. The WB's indicator of statistical capacity uses information publicly available to assess three aspects, namely statistical practices, data collection activities, and statistics availability. Recently, the WB has indicated its intention to improve its framework for statistical capacity measurement by shifting toward a new approach focusing on four dimensions: institutional framework, statistical methodology, data sources, and data dissemination.[2]

The objective of this paper is to contribute to this improvement effort, notably by proposing an analysis of the strengths and weaknesses of the WB's existing approach to statistical capacity measurement, and by exploring some options for its enhancement. Specifically, the paper seeks to explore the rationale and modalities of the WB's use of statistical activities and outputs to measure statistical capacity, to determine the problems that this poses, and to devise some options as to how the WB's indicator could be improved to better capture actual statistical capacity.

The paper hypothesizes that the WB's focus on statistical activities and outputs to measure statistical capacity, to the detriment of aspects related to statistical systems and data-producing agencies, results in the indicator poorly reflecting actual statistical capacity. To support this hypothesis, the paper uses a two-pronged approach. First, the paper builds on the new understanding of, and approach to capacity and capacity enhancement. Indeed, the development of the WB's framework for measuring statistical capacity occurred when international efforts were underway to revamp

[2] Fantom and Watanabe (2008).

the approach to capacity building and align it to the paradigm of results-based management. These efforts resulted in the framework known as the "capacity building results-chain approach." The paper refers to this framework in its attempt to capture the very meaning of statistical capacity.

Second, the paper relies on recent developments in terms of the measurement of statistical capacity at the international level. The formulation of the WB's framework also took place at a time when there was considerable momentum to improve statistics, statistical capacity, and statistical capacity measurement in developing countries. In November 1999, this resulted in the launch of the PARIS21 Consortium, a global forum and network aimed at promoting and facilitating statistical capacity building and better use of statistics. It also led to the establishment of the PARIS21 Task Team, which in May 2001 was tasked with devising an approach to statistical capacity building measurement. For the purposes of this paper, the PARIS21 framework provides benchmarks against which the WB's approach to statistical capacity measurement can be assessed.

The paper's reliance on the capacity building results chain and on the PARIS21 approaches resonates not only with the specific subject matter but also with the instrumental role that the WB played in the development of these approaches. Indeed, not only did it spearhead the promotion of the capacity building results chain approach but, in concert with the UN, OECD, IMF, and the EC, the WB also co-founded the PARIS21 Consortium. Moreover, it was also part of the PARIS21 Task Team, and assumed the PARIS21 Task Team's secretariat.[3]

Using both the PARIS21 framework and the capacity building results chain approach, this paper explores the WB's use of statistical activities and outputs to measure statistical capacity. It identifies the problems that this poses, and provides insights into how the WB's indicator could be improved so as to better reflect statistical capacity.

[3] The membership of the PARIS21 Task Team was as follows: the IMF, Chair (Ms. L. Laliberté, Chairperson, Mr. T. Morrison, Mr. J. Bové and Mr. S. Khawaja), the World Bank, Secretariat (Mrs. M. Harrison, Secretariat, Mr. M. Belkindas, and Mr. G. Eele), the UN Statistics Division, UNSD (Mr. W. de Vries), the UN Economic Commission for Latin America and the Caribbean, UN ECLAC (Ms. B. Carlson), and the UN Economic Commission of Europe, UNECE (Mr. J-E. Chapron), and AFRISTAT (Mr. Lamine Diop). The Consultants to the PARIS21 Task Team were Mr. D. Allen, Mr. T. Holt and Mr. J. van Tongeren.

The rest of the paper proceeds as follows. Following this introductory section, Section 2 focuses on the characteristics of the WB's indicator of statistical capacity, analyzes some of its advantages and shortcomings, and presents some of the questions that it leaves open. Section 3 identifies the aspects of the statistical capacity results chain that this indicator captures, and discusses an improvement in data aggregation as a requirement for the WB's indicator to better reflect statistical capacity. Section 4 focuses on capacity utilization as the missing link in the WB's approach between statistical capacity and statistical activities, and as a central factor that has to be taken into account for statistical activities and outputs to consistently reflect statistical capacity. Section 5 concludes the paper.

2. THE FRAMEWORK

The WB's framework for measuring statistical capacity uses three broad components to derive a composite indicator of statistical capacity: (i) statistical practice, (ii) data collection, and (iii) statistics availability (see Annex 1). Statistical practices are captured by 10 indicators that refer to the base year of national accounts; the use (or not) of the balance of payments manual; the status of external debt reporting; the base year of the Consumer Price Index; the index of industrial production; the availability of IMF's import/export prices; the government finance accounting concept; the frequency of the enrollment reporting to UNESCO; the frequency of vaccine reporting to WHO; and the subscription to the IMF's Special Data Dissemination Standard. As regards data collection, five indicators are used. These include: the periodicity of population censuses; the periodicity of agricultural censuses; the periodicity of poverty-related surveys (IES, LSMS, etc.); the periodicity of health-related surveys (DHS, MICS, priority survey, etc.); and the completeness of vital registration systems. As for the cluster of statistical indicator availability, this comprises 10 indicators: income poverty, child malnutrition, child mortality, immunization, HIV/AIDS, maternal health, gender equality, primary school completion, access to water, and per capita GDP growth.

Within each of the three clusters cited above, the various items are scored on the same scale and given equal weight. To obtain the overall score of statistical capacity, the three components' scores are given equal weight (see Annex 1). Such overall scores are then used for comparisons among countries and over time.

Over the past years, the indicator has revealed substantial contrasts among the highest and lowest performers. For example, Egypt achieved a score of 89 on a zero-to-100 scale in 2007, against a score of only 17 for Liberia. The WB's indicator also shows important changes over time. For example, the scores for Nigeria and Rwanda rose from 51 and 58 in 2006 to 62 and 71 in 2007, respectively, suggesting that these countries' statistical capacities had improved by 22% over a single year. Over the same period, the scores for Libya, Guinea-Bissau, and the Central African Republic fell from 41, 37, and 42 in 2006 down to 31, 29, and 33 in 2007, pointing to a decline in statistical capacity by 24%, 22%, and 21% in these countries, respectively.

A closer look at these scores for the three dimensions of statistical capacity shows that variations over time are mostly due to changes in data collection in the Central African Republic, Guinea Bissau, Libya, and Nigeria. The figures suggest that in these countries, from 2006 to 2007 the magnitude of change in data collection ranged between 40% and 50%.

As these figures illustrate, the WB's indicator can result in huge variations over short periods. Given the time needed to build statistical systems, the question arises whether the indicator adequately captures statistical capacity. Indeed, the criticism has been made that the WB indicator obliges countries to continue to carry out statistical activities and produce outputs at regular intervals in order to maintain high scores, as to do otherwise would result in significant declines in their scores of statistical capacity.[4]

The relevance of the indicator has more specifically been questioned by national studies for its inability to reflect actual statistical capacity changes over time. For example, in Niger since 2004, there has been the enactment of a new law on statistics; the establishment of a National Council of Statistics; the transformation of the former Directorate of Statistics and National Accounts (DSCN) into the new and more dynamic National Institute of Statistics; and the transformation of statistical units in core ministries into fully-fledged Directorates. Other changes include improved coordination among data-producing agencies; the formulation of the National Statistical Development Strategy based on the PARIS21 principles; the provision of a substantially increased budget; better-trained and motivated staff; and improved work conditions.[5]

---

[4] ACBF (2007).
[5] Gafishi et al. (2008).

In view of all the above changes, it is apparent that the capacity of Niger's statistical system has been significantly strengthened. Gafishi et al. (2008) argue that these changes have resulted in improved data collection and enhanced indicator availability to some extent, as well as in the National Institute of Statistics enjoying a reputation as an emerging center of excellence. However, the authors forcefully stress the contrast with the WB indicator scores, which have barely changed over the period despite these improvements. Accordingly, the authors call for further investigation into the reasons for this apparent paradox.[6]

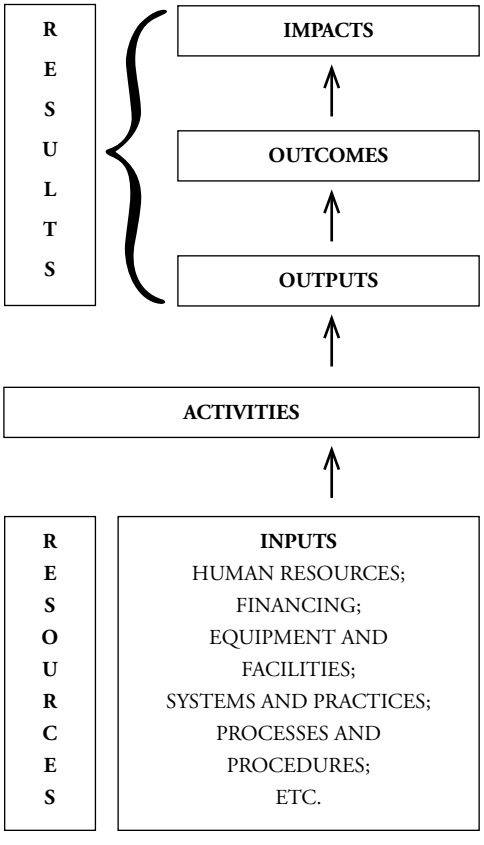
To fully comprehend the strengths of the WB's indicator and the problems that it poses, it may be useful to look at its underlying framework and the context of its development. The WB began to use the approach at a time when the PARIS21 initiative, the first systematic attempt at the international level to develop statistical capacity building indicators, had already produced its framework for measuring statistical capacity. This framework provides for comprehensive reviews of national statistical systems, requiring detailed country visits. As has been pointed out by various commentators, the problem with the PARIS21 approach is that it is expensive and time-consuming. Moreover, it imposes an additional burden on the often-limited capacity of the statistical systems being evaluated, especially in low-income countries.[7] It is also likely to result in idiosyncratic descriptions that render international comparisons difficult.

The WB opted for a different, more limited approach to mitigate these shortcomings. This approach uses a smaller set of indicators (Annex 1), for which the data are publicly available, and which allow for the capture of differences in statistical capacity among countries and over time. In this context, the WB's success is undeniable, as it has cost-effectively been rating statistical capacity for 144 countries around the world since 2004.

The development of the WB's indicator of statistical capacity also occurred at a time when the popularization of the results-based management approach lent support to the emergence of another new framework, known as the "capacity building results chain framework" (Figure 1). This promotes analysis of cause and effect; clarifying the relationships among long-term goals, mid-term outcomes, and immediate objectives, and the resources, strategy, and actions required to achieve them in a results-oriented manner. The capacity building results chain framework also aims to identify

[6] Gafishi et al. (2008: 13).
[7] ACBF (2007).

Figure 1: Capacity Building in the Results-chain Framework



key assumptions and performance indicators for each stage of the process; and to visualize the process in context, by considering the external factors that might influence outcomes.

In such a framework, statistical capacity and statistical capacity enhancement are viewed as part of a chain going from statistical resources to activities, to outputs, to outcomes (as statistical indicators come to influence policymaking and development monitoring and evaluation, for example) and finally to impact (in terms of poverty reduction and welfare improvement, for instance).

The PARIS21 initiative followed this approach. Its stated objective was to develop demand-driven statistical capacity indicators as tools to be applied to specific goals. The approach aimed to measure the statistical capacity to meet those goals, as these would be geared toward statistical results to meet users' needs.[8]

Hence, in the light of the capacity results-chain approach and the PARIS21 framework, and in view of the aforementioned variations in the scores of statistical capacity, the WB's approach poses some questions. These include the following:

- Can actual statistical capacity be so volatile, as illustrated by the cases of Nigeria, Rwanda, Libya, Guinea-Bissau, and the Central African Republic?

[8] PARIS21 Task Team (2002a).

- Does the WB's statistical capacity indicator capture the relevant dimensions of statistical capacity?
- What – if any – corrections are required for the WB's indicator to reflect statistical capacity more consistently?

The next sections attempt to address these issues. However, first it may be useful to define statistical capacity, and to look at the implications of this definition for the WB's approach. This is the objective of the next section.

3. CAPACITY VERSUS PERFORMANCE: WHAT DOES THE WB'S INDICATOR MEASURE?

The concept of capacity has frequently been given different definitions, as disparate motivations and objectives have led development actors to choose to address different aspects of capacity. For the purpose of this paper, capacity is simply and generally viewed as the *resources* of a society to achieve societal goals. At the level of an institution or agency, capacity represents the resources of that institution to deliver its mandate.

In the specific case of statistical systems, the experience of countries that have successfully enhanced statistical capacity shows clearly that statistical capacity building requires focusing on a wide spectrum of factors.[9] These include human resources and infrastructure (such as buildings and power, among others), and other material resources. Resources also include data-producing agency staff and human resources management practices, such as hiring and firing, promotion, rotation, training, and career development.[10]

Resources also include financing and its characteristics – such as its level, sources, and stability – which largely determine the flexibility and independence of the statistical system, as well as related processes and procedures that determine the efficiency of the use of funds. Resources should also include computing facilities, including the availability, maintenance, and updating of information technology infrastructure such as servers, communication network, and computers. Also falling under the rubric "resources" are transportation and communication systems equipment,

[9] For example, see Gafishi et al. (2008) and Kiregyera (2006).
[10] PARIS21 Task Team (2002b).

inclusive of operational support, printing equipment, office supplies and sundries, and other aspects of the office environment.[11]

Last but not least, statistical practices and the regulatory framework of statistics are key resources of a statistical system. Statistical processes and procedures are resources as well, to the extent that they determine access to information, preserve confidentiality of individual data, ensure adequate coordination, planning, monitoring and evaluation of statistical operations, guarantee the independence of statistical activities, and ensure the accuracy, reliability, and accessibility of statistical data and metadata such as information on underlying concepts, definitions, classifications, methodology, data sources, accuracy, etc.

All this serves to confirm that the dynamics of statistical capacity – that is, the changes in the characteristics of statistical resources – are multifaceted. Statistical capacity enhancement may consist in increasing the number of staff and their skills and/or altering human resources management practices so that a core contingent of highly trained staff can be retained and maintained through regular recruitment and training. Other changes that may be construed as statistical capacity building include:

- improvements in the statistical regulatory framework, systems and practices, and processes and procedures;
- the provision and upgrading of infrastructure such as buildings, information technology resources;
- improvements in other material resources such as transportation and communication systems equipment and other aspects of the office environment; and
- the improvement of financing and its characteristics.

In light of the capacity building results chain framework discussed earlier, it appears that the WB's indicator mostly captures statistical activities and outputs that are *outcomes* of statistical resources rather that statistical resources per se. For instance, the 10 items that determine the score of statistical indicator availability clearly capture outputs. As for statistical activities, they are scattered in the statistical practice and data collection clusters. They include population censuses, agricultural censuses, poverty surveys, and health surveys. They also include UNESCO reporting, and setting a new base for national accounts, the consumer price index, import and export price indexes, and industrial production indexes.

[11] PARIS21 Task Team (2002b).

In fact, of the total of 25 items that the WB's indicator encompasses, as many as 19 items (accounting for nearly 77% of the overall score) relate to statistical activities and outputs. Only six items (accounting for about 23% of the overall score) relate to statistical capacity aspects. These are indicators of whether (or not) the UN considers the vital registration system coverage as complete; whether the country has subscribed to the Special Data Dissemination Standard; national and WHO/UNICEF estimates of the national immunization coverage are consistent; government finance accounts are consolidated; the data used for external debt reporting are actual or preliminary; and the latest edition of the Balance of Payments Manual is in use (see Annex 1). This demonstrates that the WB's indicator of statistical capacity is mostly driven by statistical activities and outputs.

The fact that the WB's indicator is largely driven by statistical activities and outputs and only marginally by statistical resources means that it mostly captures statistical performance rather than statistical capacity. Such performance is specifically captured quantitatively, particularly in terms of frequency. As Annex 1 illustrates, the higher the frequency of statistical activities and outputs, the higher the score. As regards the few items that focus on actual capacity, they mostly reflect the ability of the statistical system to produce quality statistical outputs and are at the core of the Data Quality Assessment Framework (DQAF) that has been developed by the IMF.

The WB indicator's focus on performance rather than capacity poses a conceptual problem. While capacity and performance are related – in the sense that the latter is an outcome of the former – yet they are different, and performance indicators cannot be substituted for capacity enhancement indicators. As Mizrahi argues, the failure to distinguish between the two can lead to inappropriate policy. According to the author, unlike performance indicators, indicators of capacity and capacity enhancement provide information about sustainability by revealing information on the extent of institutionalization or routinization of improvements to a system.[12]

This principle holds true not only for those cases where performance is reflected in quantitative terms; it is valid also in those cases where performance is captured in *qualitative* terms. In this context, statistical capacity measurement has to be approached in a different way to mere statistical quality measurement. As the PARIS21 Task Team forcefully argued, a statistical output entirely financed and executed through external sources

[12] Mizrahi (2003: 5).

may be of high quality, yet be a poor measure of statistical capacity in terms of domestic expertise and sustainability. Accordingly, the PARIS21 Task Team emphasized the need for a sound statistical capacity measurement framework to take account of indicators applicable to agencies (e.g. central statistical agencies, the statistical units of central banks and line ministries) that produce statistical outputs.[13]

In some ways, the weakness of the WB's indicator of statistical capacity is partly due to its method of aggregating the various items it considers to be relevant dimensions of statistical capacity. Indeed, minimizing the relative weights of statistical activities and outputs would improve the indicator. The most straightforward and radical option would be to remove the items related to activities and outputs from the indicator (or to apply a zero weight to them), and to keep only the system- and agency-related aspects. Such an option is particularly appealing since the WB's approach, by indiscriminately using inputs and outputs to measure statistical capacity, seems to double-count a number of aspects and to mix up flows (periodic activities and outputs) and stocks (improvements that occur once and for all) or performance- and capacity-related aspects.

However, despite its attraction, the option of totally eliminating indicators of activities and outputs from the overall WB indicator raises a serious problem, which relates to the complexity of the relationship between inputs and outputs in statistical processes. Indeed, this relationship is more complex than suggested by the earlier discussed capacity building results chain framework.

Statistical activities and outputs cannot be regarded simply as an outcome of statistical resources, as some may also serve as inputs for downstream activities and outputs. For instance, a number of resources that form the capacity base of a statistical system are required to carry out household surveys to obtain survey data. These data are necessary to undertake further work to produce statistical indicators, such as a consumer price index, inequality index, and employment statistics, among others. In such a process, household surveys appear to be the first-level activity, while the calculation of the indicators forms the second-level activity. Between the two are the survey data as a first-level output. The statistical indicators constitute second-level outputs.

[13] PARIS21 Task Team (2002a: 7).

Therefore in actual statistical processes, a number of "intermediate" activities and products can be viewed as both outputs and inputs, depending on the stage considered. The PARIS21 Task Team forcefully acknowledged the complexity of this relationship between inputs and outputs in statistical processes. Not only did they view survey data as inputs for the production of final statistical outputs. They also recognized that statistical activities contribute to the maintenance and development of basic resources of statistical systems such as skills, as they allow for such skills to be continuously in use.[14] Thus, not only can statistical activities and outputs serve as resources or inputs for downstream activities and outputs; there is also a retroactive loop, whereby statistical activities and outputs impact on basic statistical capacity. Such a loop makes it even more difficult to isolate resources from results, and inputs from outputs in statistical processes.

Another compelling reason to retain statistical activities and outputs in the WB's indicator is because they are needed as proxies for statistical basic resources, which would otherwise be difficult to quantify. The PARIS21 Task Team suggests, for example, that crude measures of volume activities could be considered as a proxy for the mass of general statistical expertise of a data-producing agency.[15] In other words, although some statistical activities are by no means relevant dimensions of statistical capacity, they may serve as proxies. In light of the relationship between statistical resources and statistical activities/outputs, the use of the latter may be appropriate, especially as a number of statistical aspects for which comparable data exist internationally relate to statistical activities and outputs, more than to statistical systems and data-producing agencies.

In conclusion, it may be useful to highlight the serious challenge that the use of statistical activities and outputs represents to the WB approach. As argued in this section, relying on statistical activities and outputs to measure statistical capacity causes the indicator to capture performance rather than capacity. Yet, statistical activities and outputs need to be used as proxies of genuine dimensions of statistical capacity that would otherwise not be accounted for. The greater the weight of such dimensions in statistical capacity, the greater the weight that needs to be given to statistical activities and outputs, hence the greater the distortion in the overall indicator. The next section argues that this problem may be mitigated by, first, a better understanding of the relationship between statistical capacity

[14] PARIS21 Task Team (2002b).
[15] PARIS21 Task Team (2002b: 12).

and statistical activities and, second, by adequately accounting for capacity utilization.

4. TAKING CAPACITY UTILIZATION INTO ACCOUNT

In recent years, the development community has increasingly recognized that in most African countries, capacity is often underutilized as a result of the misallocation of available skills and talents. Some authors have even found that, in a number of circumstances, capacity is prevented from being put to work.[16] By and large, the literature points to incentives provided by the prevailing public sector management system and governance factors to explain the underutilization of capacity. As Obadan argues, when talents are misallocated, and recruitment and promotions based on personal connections and loyalties rather than on merit, existing capacity is likely to be underutilized.[17]

Capacity utilization may vary across time and space, and be understood in quantitative and qualitative terms. A society may utilize different amounts of its capacity at different times, just as different societies may utilize variable amounts of their capacities at a given time. On the other hand, a society may utilize the same amount of its capacity but with different efficiency at different times; whereas different societies may utilize the same amount of their capacity but with different efficiency at a given time. Thus, not only is the relationship between a statistical system's activities/outputs and its capacity very complex, as earlier discussed, but it is by no means linear. This relationship is affected by many factors, which result in different rates of capacity utilization in different circumstances.

There is a growing consensus to consider that a society's capacity to utilize effectively and efficiently its capacity is a key aspect of its capacity per se, and that the effort to increase the rate of utilization of its capacity should be considered as capacity enhancement in its own right.[18] This means that capacity utilization could be considered as a capacity resource. Yet, capacity utilization is a *particular* resource. This particularity resides in the fact that in specific circumstances, capacity utilization and overall capacity may evolve independently, or even in opposition.

---

[16] Obadan (2007).
[17] Ibid.
[18] Obadan (2005).

To illustrate, let C, A, and U be statistical capacity, statistical activity, and the rate of utilization of the existing statistical capacity C in period t, respectively. One then can represent the relationship among these variables as follows:

$$A_t = \alpha C_t U_t \qquad (1)$$

α being a constant that captures the average "productivity" of statistical resources. Equation (1) simply means that at any point of time t, statistical activity is undertaken using all or part of existing statistical resources. Taking the logarithm and then the difference of equation (1) yields:

$$\Delta A = \Delta C + \Delta U \qquad (2)$$

with Δ indicating the percentage change in the amount of the concerned variable during the period. Equation (2) can be rearranged as follows:

$$\Delta C = \Delta A - \Delta U \qquad (3)$$

As Equation (3) shows, a given change in the amount of statistical resources is equal to the difference between the change in statistical activity and the change in utilization of this amount of statistical capacity, all changes expressed in terms of percentage points. Thus, Equation (3) implies that after a period of low statistical activity, during which only part of existing statistical resources were in effective use, additional statistical activity and outputs could be taken up by mobilizing latent capacity rather than by acquiring fresh capacity. In such a situation, the increase in statistical activity would be absorbed by the increase in the rate of utilization of existing capacity without any impact on the overall mass of resources that the statistical system has at its disposal.

Equation (3) also means that if during a period, the change in the rate of capacity utilization were higher than the change in the amount of statistical activities undertaken in that period, then statistical capacity would have declined, not increased, during that period, even if the change in statistical activity were positive. In fact, a given statistical activity growth rate implies a commensurate change in statistical capacity only if the rate of capacity utilization remains constant during that period. This is what the WB implicitly assumes, causing its approach to rely on a highly restrictive and largely unrealistic assumption.

Indeed, this condition is likely to be met in two particular circumstances, namely situations where existing capacity is in full use and where there is no possibility to further increase it, or where latent capacity exists but no actionable policy is available or possible to mobilize such unused capacity. In such situations, a given percentage point increase in the amount of statistical activities undertaken would mean a commensurate increase in statistical capacity, since one would have, for example, to hire and train new staff, and equip them with new equipment to take up these additional activities. If existing capacity were not fully used, and actionable policy available or possible to mobilize (at least partially) the latent capacity, then additional activity would translate into the use of all or part of the unused capacity, resulting in the new tasks being taken up without a commensurate increase in statistical capacity.

While this suggests that latent capacity allows for the absorption of the impact of changes in statistical activity/outputs on statistical capacity, it is noteworthy that in some circumstances it could amplify such an impact. For instance, if the volume of statistical activities and outputs increases at a given rate during a given period, while for some reason the rate of capacity utilization is decreasing over the same period, then statistical capacity would have increased at a rate higher than that of statistical activity and outputs.

Equation (3) suggests that the relationship between statistical capacity and statistical activities/outputs spills over a third variable, that is, capacity utilization. The above model suggests that there is an identity linking changes in the level of statistical capacity in any given country with changes in the level of statistical activities and outputs, and changes in statistical capacity utilization. The relationship among these three variables may be compared to that among the angles of a triangle, as the interdependency among them means that to derive reliably the variation in any of the three variables, a sufficient condition is to know variations in the other two variables.

Thus, the "Capacity–Activity–Capacity Utilization Triangle" model illustrates a relationship between statistical capacity and statistical activity/outputs that is more sophisticated than the one assumed by the WB, and which implies a mechanistic and univocal impact of changes in statistical activities/outputs on statistical capacity. The model rather suggests that taking capacity utilization into account every time statistical activities and outputs are used to measure statistical capacity is crucial if the WB's indicator is to reliably capture actual statistical capacity. One interesting question that this conclusion raises is how, as an improvement to the WB's

approach to statistical capacity measurement, accounting for capacity utilization would compare with the earlier-mentioned need to improve on indicator aggregation.

One consequence of an improved method of score aggregation in the WB's approach would be a reduction in the distortion created by the overemphasis on statistical activities and outputs to the detriment of statistical capacity. An improved aggregation method would mean the WB's approach giving greater weight to statistical resources than to statistical activities and outputs, leading to a reduction in the distortion and to less volatility in the overall score of statistical capacity. Furthermore, it would mean less pressure on countries to continue to carry out statistical activities and to produce statistical outputs at regular intervals in order to maintain high scores, as is the case under the current measurement system.

However, none of these problems would totally disappear solely as a result of an improved aggregation method. Indeed, these problems cannot disappear as long as statistical activities and outputs have to be used as proxies for those aspects of statistical capacity that would not be accounted for otherwise. In circumstances where such aspects represent a significant share of statistical capacity, the problem of volatility would be particularly acute, since scores of statistical activities and outputs would have to be given relatively greater weight to account for these aspects adequately.

In contrast, taking capacity utilization into account would likely improve the WB's measurement system considerably. This is borne out by the perspective that the "Capacity–Activity–Capacity Utilization Triangle" model offers to address the problem of volatility. Specifically, the apparent volatility of statistical capacity, as it results from the WB's current approach, is largely a consequence of the neglect of capacity utilization, since this approach assumes that variations in statistical activity and outputs imply commensurate variations in statistical capacity.

A proportion of the changes in statistical activities may simply be absorbed by changes in statistical capacity utilization, implying that variations in statistical activity/outputs do not mean commensurate variations in statistical capacity. By the same token, countries would not have to continue to carry out statistical activities and to produce statistical outputs at regular intervals to maintain high scores, since part or the totality of these activities and outputs could be undertaken using existing capacity.

Thus, the "Capacity–Activity–Capacity Utilization Triangle" model provides a plausible explanation for the abnormally high volatility of statistical capacity as measured by the WB's indicator. Furthermore, it allows for a more accurate envisioning of the relationship between statistical capacity and statistical activities and outputs. In this model, capacity utilization is like a filter placed between statistical capacity and statistical activities/outputs to separate those variations in statistical activities/outputs that are associated with changes in latent capacity, from those changes that are genuinely associated with variations in actual statistical capacity.

Hence, whereas the WB's indicator wrongly attributes the bulk of statistical activities and output variations to the change in statistical capacity systematically, the "Capacity–Activity–Capacity Utilization Triangle" model makes up for this weakness. The model makes it clear that capacity utilization is the missing link in the WB's approach to statistical capacity measurement, between statistical capacity and statistical activities and outputs.

5. CONCLUSION

The objectives of this paper were to explore the rationale and modalities of the WB's use of statistical activities and outputs to measure statistical capacity, to identify the problems that this poses, and to devise some options on how the WB's indicator could be improved so as to better capture actual statistical capacity. To address these issues, the paper referred to the PARIS21 framework and the capacity building results chain approach. Choosing to use these frameworks was particularly apposite, given the paper's subject matter and the central role that the WB played in their emergence and promotion.

The paper identified two major shortcomings in the WB's approach, which have resulted in its failure to adequately capture statistical capacity. These are (i) its overreliance on statistical activities and outputs and (ii) its neglect of statistical capacity utilization. These compel a country to carry out statistical activities at regular intervals in order to maintain a high score; they also explain why statistical capacity, as measured by the WB's indicator, appears to be highly volatile. The interactions between these two factors tend to aggravate the bias of the WB's indicator: the greater the reliance of the indicator on statistical activities and outputs, the greater the impact of its neglect of capacity utilization. Conversely, if the WB's approach properly took capacity utilization into account, then the impact of its emphasis on statistical activities and outputs would be moderated.

In this paper, an effort has been made to identify options for an improved approach. It was argued that such an approach should be consistent with the PARIS21 framework and the capacity building results chain approach as much as possible, while not conflicting with cost-effectiveness and international comparability. The latter two factors were identified as the bedrock of the WB indicator's success to date, in spite of its highlighted weaknesses.

Although removing activity- and output-related items from the WB indicator might initially appear to be the most appropriate solution to the problem, this option has been rejected for a number of reasons. One is the difficulty in clearly distinguishing between inputs and outputs in statistical processes. Indeed, not only do some activities and outputs serve as inputs for downstream activities and outputs, but there also exists a retroactive loop whereby downstream activities and outputs in turn affect basic statistical capacity. Moreover, the data available to measure statistical capacity in a cost-effective manner and to allow for international comparability, mostly relate to statistical activities and outputs, and only marginally to actual statistical capacity.

Against such a background, it was argued that improvements in the aggregation of various items in the WB's indicator methodology might be a better option for an improved measurement system. Also, the paper argued that when using scores of statistical activity and outputs to measure statistical capacity, taking due account of statistical capacity utilization would be helpful.

These improvements would have differing impacts, however. An improved aggregation method would reduce the distortion caused by the use of the volume of statistical activities and outputs to measure statistical capacity. However, taking into account statistical capacity utilization would allow differentiation in statistical activities and outputs between, on the one hand, those associated with changes in actual statistical capacity and, on the other, those likely to be absorbed by changes in mere capacity utilization. It is therefore recommended that the WB should work on improving the aggregation method, and more importantly, take into account capacity utilization in its approach, as the missing link between statistical activities and outputs and statistical capacity.

REFERENCES

ACBF (2007), *Towards Reforming National Statistical Agencies and Systems: A Survey of Best-Practice Countries with Effective National Statistical Systems in Africa*, ACBF Best-Practice Study Series BPS 01/2007. Harare: The African Capacity Building Foundation.

Fantom, N. and N. Watanabe (2008), "Improving the World Bank's Database of Statistical Capacity," *African Statistical Newsletter*, vol. 2, no. 3, pp. 21-22.

Gafishi, P., I. MacAuslan, and C. Spanneut (2008), *Evaluation de l'Appui au Renforcement des Capacités Statistiques. Etude de Cas du Niger*. Mimeo. Niamey, Niger: Institut National de Statistique (INS).

Kiregyera, B. (2006), *Strengthening National Statistical Systems in Sub-Saharan Africa –Some Lessons from Ugandan Experience*. ACBF Lessons Note, Harare.

Mizrahi, E. (2003), *Capacity Enhancement Indicators: Review of the Literature*. World Bank Institute Evaluation Studies No. EG03-72. Washington, DC: The World Bank Institute.

Obadan, M. (2005), "Challenges in the Building of Public Service Capacity in Africa," ACBF Working Paper No. 5. Harare: ACBF.

Obadan, M. (2007), "Capacity Utilization, Retention, and the Use of African Diasporan Communities as Development Actors: Challenges and Opportunities." Paper presented at the Second Forum on Capacity Building, Maputo (Mozambique), 1-3 August.

PARIS21 Task Team (2002a), *The Framework for Determining Statistical Capacity Building Indicators*, April.

PARIS21 Task Team (2002b), *Statistical Capacity Building Indicators Final Report*. September.

Wingfield-Digby, P. (2008), "Africa's STATS League – The Movers and Shakers 2006-2007," *African Statistical Newsletter*, vol. 2, no. 1, pp. 26-28.

World Bank (2006), *Statistical Capacity Improvement in IDA Countries – Progress Report*. Washington DC: The World Bank, May 16.

ANNEX 1: COMPONENTS OF THE WORLD BANK'S INDEX OF STATISTICAL CAPACITY

I. STATISTICAL PRACTICE				
Indicators	**1**	**0**	**Max**	**Weight**
1. National accounts base year	Within last 10 years or annual chain linking	Otherwise	1	10
2. Balance of payments manual in use*	Balance of Payments Manual, the fifth edition	Otherwise	1	10
3. External debt reporting status*	Actual or preliminary	Otherwise	1	10
4. Consumer Price Index base year	Within last 10 years or annual chain linking	Otherwise	1	10
5. Industrial production index	Produced and available from IMF	Otherwise	1	10
6. Import/export prices	Produced and available from IMF	Otherwise	1	10
7. Government finance accounting concept*	Consolidated central government accounts	Otherwise	1	10
8. Enrollment reporting to UNESCO	Annual or missed reporting only once in the last 4 years	Otherwise	1	10
9. Vaccine reporting to WHO*	Nationally reported data on measles vaccine cover-age consistent with WHO estimates	Otherwise	1	10
10. IMF's Special Data Dis-semination Standard*	Subscribed	Otherwise	1	10
				Maximum total score: 100

II. DATA COLLECTION					
Indicators	**2**	**1**	**0**	**Max**	**Weight**
1. Periodicity of population census	≤ 10 years		Otherwise	2	10
2. Periodicity of agricultural census	≤ 10 years		Otherwise	2	10
3. Periodicity of poverty related surveys (IES, LSMS, etc.)	≤ 3 years	≤ 5 years	Otherwise	2	10
4. Periodicity of health related surveys (DHS, MICS, Priority survey, etc)	≤ 3 years	≤ 5 years	Otherwise	2	10
5. Completeness of vital registration system*	Complete		Otherwise	2	10
					Maximum total score: 100

/cont…

ANNEX 1: COMPONENTS OF THE WORLD BANK'S INDEX OF STATISTICAL CAPACITY (Cont.)

III. STATISTICS AVAILABILITY						
Indicators	**3**	**2**	**1**	**0**	**Max**	**Weight**
1. Periodicity of income poverty indicator	≤ 3 years	≤ 5 years	> 5 years	Otherwise	3	5
2. Periodicity of child malnutrition indicator	≤ 3 years	≤ 5 years	> 5 years	Otherwise	3	5
3. Periodicity of child mortality indicator			National or international estimates available	Otherwise	1	5
4. Immunization indicator			Annual	Otherwise	1	5
5. HIV/AIDS indicator			National or international estimates available	Otherwise	1	5
6. Periodicity of maternal health indicator	≤ 3 years	≤ 5 years	> 5 years	Otherwise	3	5
7. Periodicity of gender equality in education indicator	≤ 3 years	≤ 5 years	> 5 years	Otherwise	3	5
8. Primary completion indicator			At least one observation in the last 5 years	Otherwise	1	5
9. Access to water indicator			National or international estimates available	Otherwise	1	5
10. Periodicity of GDP growth indicator	Annual	≤ 1.5 years	> 1.5 years	Otherwise	3	5
						Maximum total score: 100

Source: World Bank.
* Components not related to statistical activities and outputs.