Report on Updating

of the

Statistical Business Register

Kenya National Bureau of Statistics
Nairobi, Kenya

2018

Table of Contents

Foreword

Kenya National Bureau of Statistics is the main producer of economic statistics and also a coordinator and supervisor of the National Statistical System (NSS). To undertake economic based surveys, the Bureau maintains a Central Business Register (CBR) which is used as a sampling frame for all Establishment Based Surveys (EBS) in the formal sector of the economy. Relevant economic statistics are important from a more general aspect since the Vision 2030 is anchored on macroeconomic stability and reforms. In this perspective, a comprehensive Statistical Business Register (SBR) is a fundamental part of fulfilling the vision.

To produce quality statistics that meet, international standards, there is need to maintain a comprehensive business register. With this realization, KNBS identified upgrading and updating of the SBR as one of the core activities under the Kenya Statistics Program for Results (KSPforR), a World Bank facility aimed at improving quality of statistics. With a new register, KNBS seeks to improve its institutional capacity to generate quality economic statistics.

This report presents the status of the revamped and updated SBR which was supported by the Statistics Sweden through four Technical Assistance (TA) missions. For this, and for peer reviewing this report, we remain grateful.

I also take this opportunity to thank the World Bank which extended financial support through the Kenya Statistics Program for Results (KSPforR).

I extend my gratitude to all persons who were involved in the revamping of the SBR and in particular to the three consultants Staffan Blomquist, Christian Ekström and Bjorn Berglund from Statistics Sweden who provided technical support. The support from Rasmus Larsson, the long term chief technical advisor from Statistics Sweden is also highly appreciated.

Finally, I also thank James Gatungu, Robert Nderitu, Benson Karugu, Newton Amugune and Francis Mwandembo for their contribution and efforts during the revamping process.

Zachary Mwangi
Director General

The Revamped Statistical Business Register (SBR)

1 Background to the SBR

Kenya's long-term development strategy is guided by the Kenya vision 2030 which outlines a wide range of milestones that the country strives to achieve by 2030. The vision is implemented in five-year Medium-Term Plans. The Plans recognize the importance of statistics in monitoring and evaluating the government programmes within the framework of Vision 2030. Relevant economic statistics are important from a more general aspect since all pillars of the vision are anchored on macroeconomic stability and reforms.

As the principal agency of the Government for collecting, analyzing and disseminating statistical data, KNBS therefore has a critical role in the process for Kenya to realize its vision. The Bureau maintains a Statistical Business Register (SBR) which is used as a sampling frame for all establishment based surveys in the formal sector of the economy.

2 Justification to revamp/update the SBR

The Business Register had shortcomings regarding comprehensiveness and quality resulting to being used in a limited way as a sampling frame for establishment based surveys. This necessitated the need for improvement, and thus identified as one of the prioritized activities under the World Bank's Kenya Statistics Program for Results (KSPforR). With an updated register, KNBS seeks to improve its institutional capacity to produce statistics in line with international standards.

3 Methodological approach

The process of upgrading the Business register was supported by Statistics Sweden through four Technical Assistance (TA) missions. This was in addition to consultations through Skype. The main task for the TA missions was to analyse the existing, as well as future needed variables in the Business Register (SBR). The analysis was to be based on user needs within the Kenyan National Bureau of Statistics (KNBS). Based upon these analyses, a detailed plan and time schedule was presented for future activities. Further, the mission provided a draft design regarding the system of the new updated/upgraded BR at the Kenyan National Bureau of Statistics (KNBS).

4 Technical Assistance Missions

The first mission on the subcomponent Statistical Business Register took place in October 2015. A second two-week TA mission was conducted in October 2016. In the same year there was a study visit by KNBS staff to Statistics Sweden (SCB) in December. In July 2017,

the consultants at SCB performed a one-week preparatory work for the final mission in October 2017.

5   Expected Outputs from the Technical Assistance Missions

The following were the Expected outputs from the TA missions:
- Documented Status of the business register
- Meetings with key stakeholders (KRA, NSSF and County Governments among others)
- Five (5) officers trained in Business Register management
- Comprehensive and up to date sampling frame
- Documented revised design (Business Architecture) for the Statistical Business Register
- Support system for SBR validation, maintenance, sampling and estimation
- Developed and comprehensive Business Register

6   Results of the Technical Assistance Missions

This section gives a summary of the tasks agreed upon to be undertaken during the review missions as set out under the terms of reference and their completion status.

6.1   SBR Review

6.1.1   Completed tasks

The following table depicts status of the expected outputs.

Status of Expected Outputs

Output	Completed (Yes/No)
Analysis of data from the KRA and other administrative source	Yes
Analysis of links between LEU and LOU	Yes
Analysis of the structure of the business register data base and consideration of new requirements	Yes
Preparation of a Detailed plan for and time schedule for future activities	Yes
Analysis of existing and future needed variables in the SBR	Yes
Assessment of user demands for the SBR	Yes
Drafting design of updated/upgraded SBR	Yes
Analysis of data from the Census of Establishment	Yes
Use of data collected in the Census of establishment	Yes
Setting up a business register with possibilities to view, edit and evaluate business data	Yes
Setting up a database	Yes
Restarting the sample module in the business register application	Yes

Notes:

- The terms legal unit (LEU) and firm have been used interchangeably.
- Terms establishment and local unit (LOU).
- The terms Statistical Business Register (SBR), Master file and Master frame all denote a list or table of statistical units and variables in a database.

6.1.2 General finding of the SBR review

The previous SBR could not fulfill the role of serving as a backbone and a source for a comprehensive frame for business statistics. There were shortcomings regarding; coverage, timing, process for updating, rules for maintenance, establishment of statistical units and general quality of information. It was further observed that, the updating of the SBR was being done once a year, and mostly in the beginning of the year. The SBR was mainly constructed using administrative information received from KRA (Kenyan Revenue Authority). In addition, other administrative sources used included National Hospital Insurance Fund (NHIF) and National Social Security Fund (NSSF).

The main source of information for updating the SBR was the Annual Survey of the employed and self-employed (LE). All firms in the SBR received the survey questionnaire and were required to give the following information about the establishment; name, address, demography and ownership structure, status, economic activity   and main

products/services produced or sold. The survey also required information about key variables including annual turnover, number of persons employed and their earnings. It was from this survey that KNBS received information on the level of the establishment.

The establishment was linked to the firm utilizing information received from KRA. It was a time-consuming process since information about relations is sparse. To facilitate the process LE respondents were requested to give information about the Personal Identification Number (PIN), which is a unique identity of the firm established by KRA. Respondents were generally reluctant to give information about the PIN and therefore non-response was high. KNBS regarded the coverage of the financial sector as high since information was made available from the Central Bank and the Insurance Regulatory Authority.

According to the file received from KRA, there were approximately 140,000 firms in Kenya. The SBR contained 72,000 of these firms, one reason being that many of the firms in the tax list were "Briefcase operators" with no information on physical location.

Information regarding the informal sector was of high priority and KNBS expressed a wish to create a specific register covering units active in the informal sector.

6.1.3 Main findings during the study visit

- A census of firms and establishments had been planned to be carried out by the KNBS in 2017.
- The population for the census was to be based on the most recent population information of firms from the KRA. The population would then be the taxpaying firms at a given point in time, which in turn would make up for the largest share of the formal sector in the economy[1].
- Generally, a cut-off was used; firms with legal form equal to sole proprietor were included on the census.
- There were eventually three main data sources used for the 2017 KNBS - SBR as suggested by KNBS: KRA, Census of establishments and the county councils business registers.
- The census aimed at collecting information on both firms and establishments.
- Data was collected using tablets and paper forms. The results were loaded to a server and then transferred to the SBR database.
- The census exercise was finalised by the KNBS in the beginning of 2017.
- One data base table for firms was created.
- Relations and links between firms and establishments were discussed.

---

[1] The BR-team suggested that the formal sector be made up of all registered firms formally registered at the Registrar of Companies. However, having a lot of dormant firms in the Registrar records, the criteria for firm status "active" was created by an additional registration with either the KRA or the NSSF.

- There was only one industry code assigned to each establishment. The unit establishment is characterised as Local Kind of Activity Unit (LKAU)
- Some adjustments in the data base table for establishments (LKAUs) were necessary depending on what variables were to be included.
- The links between firm and establishments (LKAUs) needed to be further elaborated by the KNBS.
- The front end system for sampling would be, as suggested by the KNBS, be part of the SBR in July 2017.
- Data base language Dbase was used.
- The ICT-expert from SCB had an advisory role in the project. This means no code writing in neither Dbase nor Fox Pro.
- Some basic arrangements were required for the ICT-environment at KNBS regarding the SBR, such as a back-up of the register.
- All development work was made in a copy of the current KNBS SBR data base.
- The KNBS SBR data base required referential integrity to avoid duplicates and inconsistencies.
- A test environment was set up regarding the data base and the front-end application.
- Parts of the collected COE-variables were to be tested before being loaded into the new BR data base.
- Final loading of data from the COE to the SBR data base was undertaken in June 2017.

6.1.4   User demands on the BR

- The mission included seven interviews with representatives from the following statistical domains and potential BR-users; Labour statistics, Agricultural statistics, ICT statistics, National accounts, Industrial statistics, Construction statistics and financial statistics.
- Target units for these domains were firms and establishments. A firm was deemed necessary in all domains, whereas establishment was pointed out by four. KNBS has an obligation to produce regional statistics if possible. This meant establishment units were also necessary in the SBR.
- Yearly updates of the SBR would do for all domains but one. In this case quarterly updates were required.
- Important variables for the users were identifiers, classifications, contact variables and stratifications variables.
- Four interviewees also wanted subject matter variables included in the SBR.
- Moreover, a general view was that the SBR must be produced and made available for all enterprise based surveys at the KNBS.

6.1.5   Continued assessment of status of the SBR

- A deeper assessment of the BR was made based on information loaded in the database table with information on establishments.
- A number of identities were used for firms in Kenya. Administrative data providers usually assigns identities of their own, and the KNBS also creates one firm identity (Firm-ID). For establishments the KNBS creates an identity denoted as EIC. Establishment level data was mainly provided by in house surveys or originates from County Governments. For the latter, each record has a Business Permit ID for firms. There are no links between the ID-numbers provided by the following administrative data sources; KRA, National Social Security Fund (NSSF), Registrar of Companies (ROC), National Hospital Insurance Fund (NHIF) and the County Governments.
- There were blanks in the relations between firm and establishment units in the BR database. This made it difficult to keep a register with consistency between statistical units.
- There were also several blanks in variables.
- The linking between firms and establishments in the database was based on matching procedures for addresses and names of the firm or local unit.
- One delivery from the KRA had been made since the 2015 mission. It contained about 40,000 records of the largest firms and this information was being used to update the SBR
- Manual updates had already been made to the SBR with data originating from County Governments on establishment level.
- Batch updates were not implemented.
- Variables central to a SBR such as turnover and activity code (section level) existed and were also available on legal unit level via administrative sources.
- Some other variables such as status, activity code (class level) and number of employees were not found in administrative records. However, information to derive the variable "status" could be collected from administrative data records.
- It was not possible to evaluate data from KRA in detail or to investigate variables potentially available from the KRA. The BR-team suggested that a population based on tax records would include about 80 per cent of the formal sector. The rest would be found in other administrative sources such as the NSSF and the ROC. It was however generally agreed that KRA was to be the main supplier of data to the BR.

6.2   Database, sampling function, ICT-environment

- An overview of the data model was presented. The database was constructed in FoxPro language and linked to Dbase.
- The data model contained data, documents, and codes.

- Data tables included one separate table for firms and one for establishments. This means the two objects could be separated in the database. Also, there were separate tables for codes and classifications.
- The database seems well adapted to the BR needs.
- Support from previous consultants for FoxPro and Dbase had expired.
- Some adaptions of the database were found to be necessary depending on if any new variables were to be included.
- Any adaptions of the database would be handled by an external consultant.
- There was a front-end system for maintenance, reports and sampling.
- It was not possible to use the front-end sampling function. Information from the BR-team suggested that depending on deficiencies of how data was organised in the database, i.e. a mix of units and insufficient links between them, as well as empty cells for many variables.
- Any adaptions of the database also required programming upgrades of the front-end functions.
- As regards to the ICT-environment, it was found that the entire SBR was stored on one or more PCs, and not on a server.

6.2.1 Final Database

- The database is constructed in FoxPro language and linked to Dbase.
- Data tables include one that constitutes the SBR and a number of reference tables.
- The database was adapted to the existing functionalities in the front-end system.
- Any adaption of the database in terms of adding further variables also requires programming upgrades of the front-end functions.
- The SBR is stored on a   PC and not on a server.

6.3 Documented revised design (Business Architecture) and systematic approach for the SBR management/maintenance

The revised design describes how a fully and comprehensive SBR could be achieved at KNBS. It defines the processes and rules for regularly updating of the structured list of business units engaged in the production of goods and services. The design describes functionality and services required for maintenance of the SBR, e.g. how to link units, identify and construct statistical units. The new design includes procedures and information required to derive coding (ISIC, Institutional sector), unit characteristics (turnover, employment) and   identification of activity of special interest. For efficient allocation of samples it is, besides the   methodology, important that different characteristics of the establishment are of high quality. These issues have been identified as particularly important at KNBS.

The design of the SBR is developed based on the insight on how the statistical business register could fulfil user needs and assist the compilation of statistical data as a tool for the preparation of surveys. The SBR design was to be developed in close cooperation with the statistical domains, as users of the SBR. At KNBS, was the ambition that the SBR, not only was to have the capacity to assist the statistical domains regarding sampling, but also regarding estimation, e.g. treatment of non-response and grossing-up. Preferably the new SBR design was to elaborate the processes for issues related to sampling and estimation methodology. Capacity building within the area was also highlighted as a key measure to reduce the degree of dependence on external support.

The design took into consideration how the SBR could serve as a source of information for statistical analysis of the business population and its demography.

Another important feature of the design was that it could define the scope of the register. Although recognized that information about the informal sector is of high priority it was recommended that the scope of the SBR component was to be limited to the formal sector, at least in the first phase of the new design process. Regardless, it was important to decide on how to define the formal and the informal sector. It was recommended that international references be applied.

"The informal sector is broadly characterised as consisting of units engaged in the production of goods or services with the primary objective of generating employment and incomes to the persons concerned. These units typically operate at a low level of organisation, with little or no division between labour and capital as factors of production and on a small scale. Labour relations – where they exist – are based mostly on casual employment, kinship or personal and social relations rather than contractual arrangements with formal guarantees. The informal sector thus defined excludes households producing exclusively for own final use."

It is worth mentioning that KNBS was to present statistics on the informal sector at the end of the period of cooperation. KNBS had plans to carry out a business module within Kenya Integrated Household Budget Survey (KIHBS), the Micro, Small and Medium Enterprise (MSME) survey, collect information from counties and conduct a Census of Establishments (COE) which would contribute to extensive information about the informal sector. This was also a reason to why it could be preferred to delay the introduction of the informal sector into the SBR (even as a separate part). Maintenance of the register with an informal component was expected to be very difficult and expensive.

6.4 Support system for SBR validation, maintenance, sampling and estimation

Depending on the results from the design phase and the identified SBR functions and services required, there was need to improve existing support systems/applications or to

develop new systems/applications. Statistics Sweden limited their TA to the requirements of the system.

An expressed wish from KNBS was that an automated process could support estimation and sampling. There was already an IT-system in place but it had so far not been possible to implement in statistical production. Another issue pointed out was the need for common services in production, e.g. common software and a platform for validation and analysis of data. This was outside the scope of the component being implemented.

6.4.1 Sampling Module

The front-end system was restarted. This was to be used for some maintenance, producing some reports and parts of the sampling function. The sampling function that was restarted regards the LE-survey. This had been identified as a priority for KNBS. There was also a sampling function called ad hoc sampling. This was not restarted. As far as both teams could perceive, the output of LE-survey sampling function seemed reasonable. However, there was no guarantee that the samples produced are accurate; a thorough assessment of the back-end programming was subsequently deemed necessary.

6.4.2 Statistical Business Register

- The main unit in the SBR is the local kind of activity unit. All firm information was disaggregated if more than one economic activity was recorded. Then, another unit was created with a specific activity and in most cases a specific regional locality.
- Both firm level ID and establishment level ID have been created. This means that data could be aggregated at both national and regional levels.
- There were blanks in some variables. This was as a result of use of multiple sources that do not cover the same information.
- Maintenance could be performed by using either the maintenance function in the front-end application or by manipulating data by programming.

6.5   Personnel trained in Business Register management

Currently, three people are involved in the work with the business register, but far from three full-time staff. In principle, the register is updated with information from administrative sources during one month at the beginning of the year. Four to five months after updating, the statistical register is established through profiling and identification of relations between establishments and firms, a work that lasts for no more than two months.

To achieve the ambitious change proposed in the KNBS strategy and in this plan, the team recommended that staff be increased in the section and the staff be released to work on full time basis. Much of the competence would be gained from on-the-job training, experience setting up the design and the practical BR development process.

A study visit to Statistics Sweden was recommended and subsequently undertaken. It

should be noted that the Bureau's intention to strengthen the competence concerning sampling and business statistics methodology, which would require specific training in sampling and estimation methodology.

6.6   Census of establishments and secondary sources

The LE and MSME surveys and the census of Establishments are important sources to SBR update and therefore closely related to the SBR component. It was recommended that one of the tasks of the assessment was to identify and describe the relations between this surveys/census and their contribution to the quality of the SBR. Since the LE is a main source for statistics regarding the labour market, general quality aspects of economic statistics, e.g. coverage and reliability should be reflected. The MSME was carried out in 2015/16.

KNBS undertook a Census of Establishment in 2017. A number of variables were collected, but not all were included in the SBR because, both the database and the application were not adapted to all new variables. (For a complete list of variables in the SBR, see appendix 3). Results from the census made up for about 45,000 of the nearly 145,000 records in the BR. More than two thirds instead came from secondary sources. The majority of these records were from KRA. Below is a list of secondary sources and estimated reference years.

Source	Type	Reference years
Census of establishments	KNBS Census	2016
KRA	Administrative	2014, 2015, 2016
MSME	KNBS Survey	2015
ICT	KNBS Survey	2015
Census of Schools	Census/Administrative	2015
FIS	KNBS Survey	2015

7   Milestones of achievements from the Technical Assistance Missions

The outcomes from the missions were an updated register of establishments pointing at reference year 2016, an application with a sample module and a database. There is also a read-only database for back-up purposes. The register contains BR core variables such as ISIC, sector, status, region, number of employed and turnover. The SCB and KNBS teams also exchanged knowledge and experience on issues related to register methodology and commenced work on future steps for developing the Bureau's BR. The cooperation will continue with the support of introducing a new and modern business register system as a joint TA from Statistics Sweden and Statistics Norway (planned for 2018-19).

8 The Way forward

- The KNBS and SCB jointly made a draft plan for the period ending July 2017. For further details regarding the plan forward, see appendix 5. In short, the main features are:

8.1 Data model and application (front end system)

- Objects, identities and variables must be made clear and defined
- The KNBS will suggest a design
- Rules for back-up and restore is needed
- Documentation is needed

8.2 Business rules

- These will regard the most essential variables such as number employed, turnover, ISIC, production value.
- Rules for imputation etc.
- Data sources to be used

8.3 Plan for loading census data into the BR data base

- Rules for batch updates
- Data cleaning
- Mapping variables
- Tests in different steps

8.4 Future plans

- The KNBS plans to be conducting a census of firms every 3 years. This will be useful for updating the business register.

9 Revamped SBR Status as at 2016

9.1 SBR Information System

The system is organized around a series of menus. Each menu lists choices from which ones makes a choice of the action to take. The system helps to organize and present the transaction data on masterfile and LE surveys. The system consists of the following elements:

MASTERFILE
- o MASTERFILE MAINTENANCE
  - Establishment data
- o REPORTS
  - MASTERFILE TABULATIONS
    - Number of establishments by industry and employment group
    - Number of establishments by County and Activity
    - Employment by industry and employment group
    - Employment by sector for the last 5 years
    - Employment by sector and gender
    - Employment by Industry and gender
    - Employment by occupation
    - Checklist by id
    - Checklist by name
    - Checklist by county
    - Checklist by county and district
    - Checklist by specific activity (ISIC)
    - Checklist by specific activity (Range) (ISIC)
    - Checklist by main activity
    - Checklist by county, district and activity
    - Number of establishments by county, district and activity
    - County Employment Statistics
    - List of added establishments in a period
  - MAILING LABELS
    - Mailing labels by county/district
    - Mailing labels by activity/ISIC

SURVEY PROCESSES
- o LE SURVEYS
  - Create sample with weights
  - Re-Calculate weights after survey
  - LE sample Listing
  - LE Sample Management
  - DATA ENTRY
    - Large Form (Parts II, III &IV)
    - Small Form
  - REPORTS
    - LE TABULATIONS
- o Ad-Hoc Sample
- o Mailing Labels for ad-hoc sample

REFERENCE TABLES
- o Counties
- o Districts
- o Locations
- o Sub-locations
- o Towns
- o Constituencies
- o Economic activities
- o ISIC
- o Status
- o Occupation
- o Type of Establishments

ADMIN
- o User Maintenance
- o Roles Maintenance
- o Update Masterfile from excel
- o Update ISIC table
- o Backup Data Files

9.2 Number of Establishments as at 2017

Activity	Previous Masterfile	Current Masterfile
A - Agriculture, forestry and fishing .. ……………………………………	748	4,189
B - Mining and quarrying………………………………………………	141	297
C - Manufacturing…………………………………………………	3,384	6,038
D - Electricity, gas, steam and air conditioning supply…………………………	74	84
E - Water supply; sewerage, waste management and remediation activities…	156	598
F - Construction……………………………………………………	1,155	11,846
G - Wholesale and retail trade; repair of motor vehicles and motorcycles……	12,070	44,020
H - Transportation and storage…………………………………………	1,310	2,241
I - Accommodation and food service activities…………………………………	3,359	8,362
J - Information and communication ………………………………………	1,399	3,480
K - Financial and insurance activities…………………………………………	2,799	4,682
L - Real estate activities………………………………………………	522	3,680
M - Professional, scientific and technical activities……………………………	2,445	3,667
N - Administrative and support service activities……………………………	1,212	3,621
O - Public administration and defence; compulsory social security…………	1,427	1,734
P - Education………………………………………………………	17,034	29,581
Q - Human health and social work activities…………………………………	8,870	11,296
R - Arts, entertainment and recreation………………………………………	392	575
S - Other service activities……………………………………………	2,629	8,440
U - Activities of extraterritorial organizations and bodies……………………	85	131
Not elsewhere classified ………………………………………………	10,954	–
Grand Total	**72,165**	**148,562**

10 Recommendations for improvement of the SBR

10.1 Sampling module

- The front-end system was restarted. This can be used for some maintenance, producing some reports and parts of the sampling function.
- The sampling function that was restarted regards the LE-survey. This was the priority of the KNBS. There is also a sampling function called ad hoc sampling. This was not restarted and therefore will require to be restarted.
- As far as both teams could perceive, the output of LE-survey sampling function seemed reasonable. However, this is no guarantee that the samples produced are actually accurate; a thorough assessment of the back-end programming is deemed necessary for that.

10.2 Future updates

10.2.1 Increase cooperation with KRA. Collect comprehensive data based on tax records.

- Set up an agreement between KNBS and KRA regarding scheduled data deliveries. Specify grade of detail, format and time for delivery. Special attention should be given to establish secure data transmissions.
- The dataset from KRA should be considered to form the population for the proposed census on firms and establishments.
- Data from KRA should be updated annually. It should also be investigated if data can be updated from KRA more frequently, for instance on a quarterly basis.
- For each record with a KRA PIN-number, a KNBS Firm-ID could be created. Blanks in either of the two identities cannot be allowed.

10.2.2 Distinguish statistical units and variables in the SBR, establish links

- Create a set of Legal Units (Firms) and Local Units (Establishments). It is important to create and maintain primary keys (identifiers) to each one of them. It is also important to establish links (foreign keys) between the two of them.
- The primary ID for firms should be Firm-ID. Blanks in Firm-ID cannot be allowed. Each Firm-ID must be related to another ID, such as the KRA PIN-number.

10.2.3 Status of units

- Create variables that confirm the current status of a unit. Suggestion: active, dormant or closed. The status could be derived from information in the administrative data records on firms.

10.2.4 Update the BR with Census of Establishments data

- Design the census form strictly from a BR-perspective. This means to include BR-core variables only and to have a clear division between questions related to firms and establishments.
- Strongly consider using a population of firms from the KRA.
- Update information on firms and establishments for the firms included in the population of firms from KRA only.
- Assign all planned full-time resources for the SBR (or preferably more) to edit census data collected and to maintain the links between firms and establishments.
- Based on information from the tax records and the census, some basic events for multi located firms regarding the secondary establishment could be as presented below:

Multi located firms

Event	Firm Source	Secondary Establishment Source
Registration (birth into SBR)	KRA	Different sources i.e. census, SBR-survey and maintenance
Deregistration (death in SBR)	KRA	Different sources i.e. census, SBR-survey and maintenance

- When creating a new, secondary establishment it should always be linked to a firm with an existing PIN code.

10.2.5 Other administrative data: Collect more administrative data.

- Expand data collection of administrative data to the National Social Security Fund (NSSF) and the Registrar of Companies (ROC).

  If inconsistencies appear between two data sources – which most certainly will be the case – it must be clear how to prioritize between the data sources. Data will be linked using other characteristics different from ID-numbers, such as matching procedures for names and addresses. From a maintenance perspective this is time consuming.

- New entries will include at least one firm ID-number from a source approved by the KNBS BR.
- Relations must be established from the identifiers from these sources to the KNBS Firm-ID in the database. The link is Firm-ID.
- Firm-ID cannot be blank. At least one relation to one of the external ID must be established.
- Prepare for updates of records with another reference year. Handle start-ups and close downs in the database, as well as changed values in variables. One option will be to create one total frame annually.
- Update data from administrative sources more frequently in an year.

10.3 Other recommendations

10.3.1 One SBR for the KNBS

- KNBS to take measures in order to make the BR more available to all users in the bureau.

10.3.2 Estimate total firm and establishment population by using other sources

- A mission on Methodology should be planned to estimate the size of the informal sector and the SBR under coverage. By using results from KIHBS and other surveys, under coverage of firms and local units could be calculated. This could also provide an estimate of the magnitude of the informal sector. This is not only requested by the KNBS, but also by the KRA[2].

10.3.3 Evaluate the system for sampling

- The proposed mission on methodology should focus on evaluating the methods for sampling implemented in the sampling system. The results should be input for the work on restarting the sampling system.

10.3.4 Secure the ICT environment at KNBS

- It is very important that the BR operates in a secure ICT environment with functioning servers, automatic back-ups, etc. KNBS will have to promote actions to secure the ICT environment. In addition, since the FOXPRO software is no longer being supported, the KNBS should acquire supported software.

Appendices

Appendix 1: Excerpt from the First Mission Report

KENSTAT 2015:10
Business activities

Economic Statistics

Report from a mission to Kenya National Bureau of Statistics, KNBS Nairobi, Kenya

4 – 12, May, 2015

by

Mats Haglund & Jens Olin

Report from a short-term mission to Kenya National Bureau of Statistics

Project:	Statistics development
Project funding:	Sida
Project component:	Economic Statistics
Project subcomponent:	Inception mission
Cooperation project partners:	Statistics Sweden (SCB), Kenya National Bureau of Statistics (KNBS)
Mission period:	4 – 12 May, 2015
Place:	Nairobi, Kenya

4.1.1 Business Register

For the sub-component 3.2 Business Register there are three outputs specified in the RBM matrix:

- Documented current state of the business register
- Meetings with key stakeholders (KRA, NSSF, County Governments)
- 5 officers trained in Business Register management

In this context it is relevant to also consider 3.1 sub-component output

- "Comprehensive and up to date sampling frame".

According to the RBM is the component objective to achieve a developed and comprehensive Business Register. The possibility to extract a sampling frame of high quality is a logic output of the objective.

Other outputs to consider in the component are:

- A documented revised design (Business Architecture) for the Business Register and a systematic approach for Business Register management / maintenance
- Support system for BR validation, maintenance, sampling and estimation

General finding

The review of the situation regarding the Business Register (BR) confirms the findings from PIM January 2014. In this report is the concepts master file and business register used interchangeably. The present BR cannot fulfil the role as a backbone and serve as a source for a comprehensive frame for business statistics. There are shortcomings regarding: coverage, timing, process for updating, rules for maintenance, establish statistical units and general quality of information. Updating of the BR is made on an ad-hoc basis. However, normally updating is done once a year, in the beginning of the year. The BR is constructed using administrative information received from KRA (Kenyan Revenue Authority). There are also several other administrative sources used, e.g. National Hospital Fund and National Social Security Fund (NSSF). The main source of information for the BR is the Annual Employment and Earning Survey (EAE). All firms in the BR receive the survey and are required to give information about the establishment: name, address, demography and ownership structure, status, economic activity and main products/services produced or sold. The survey also requires information about the key variables annual turnover, number of persons employed and their earnings. It is from the survey KNBS receives information on the level of the establishment. The establishment is linked to the firm utilizing information received from KRA. It is a time consuming process since information about relations is sparse. To facilitate the process EaE respondents are requested to give information about the PIN, which is a unique identity of the firm established by KRA. Respondents are generally reluctant to give information about the PIN and the non-response is high. KNBS regards the coverage of the financial sector as high, while administrative information is available from the Central Bank and the Insurance Regulatory Authority.

According to the file received from KRA there are approximately 140,000 firms in Kenya. The BR contains 72,000 of these firms, one reason being that many of the firms in the tax file are inactive. Another reason is that many of the firms are "Briefcase operators" where there exists no information on physical location. The activity to improve the coverage of the BR is prioritized.

KRA has information on firms that pay tax and have more than 5 million KSH turnover. Consequently has the BR no information of small and medium sized establishments. In this context is the large informal sector in Kenya a particularly challenging issue. In 2011, 9.27

million people[3] were working in the informal sector, which is approximately 80 per cent of total employment. The Government goal is that the ratio of informal to total employment should be reduced to 0.6 in 2017. Information about the informal sector is an area of high priority and KNBS has expressed a wish to create a specific register covering units active in the informal sector.

Documented current state of the BR
It is proposed that the first activity within the component will be to carry out a comprehensive assessment of the BR. An assessment will give insights into how to redesign and develop the business register with the aim to increase consistency, continuity, coverage and timeliness of statistics.

The EaE survey, the MSME survey and the census of Establishments are important sources to BR and therefor closely related to the BR component. It is recommended that one of the tasks of the assessment is to identify and describe the relations between these surveys and their contribution to the quality of the BR. Since EaE is a main source for statistics regarding the labour market general quality aspects of economic statistics, e.g. coverage and reliability should be reflected. The assessment of EAE is preferably coordinated with the assessment of statistics regarding industrial production carried out in the component "general quality". According to KNBS is the MSME survey already in place but lack of resources has delayed the deployment, originally planned to the finance year 14/15. If financing is solved, the MSME is planned to be carried out in 15/16. In addition, the census of establishments is planned to be launched in 15/16. However, the survey has not yet been designed in detail and KNBS have concerns regarding the survey methodology. The last census of establishment was a disappointment with approximately 20-30 thousand respondents. It is believed that the main reason for the result was the contact strategy used. Nevertheless, any technical assistance from Statistics Sweden is not foreseen regarding the census of establishments. KNBS do not plan to review the questionnaire used in earlier censuses of establishments.

Meetings with key stakeholders (KRA, NSSF, County Governments)
This activity can preferably be integrated in the process already from start. The relations with key stakeholders, in particularly data providers, are crucial for the success of the component. To increase the awareness and promote the importance of the BR it is important to involve key stakeholders in the process. A first meeting is proposed to take place already in the assessment phase.

Documented revised design (Business Architecture) and systematic approach for BR management/maintenance
A revised design describe how a fully and comprehensive statistical business register can be achieved at KNBS. It should define the processes and rules for regularly updating of the structured list of business units engaged in the production of goods and services. The design will describe functionality and services required for maintenance of the BR, e.g. how to link units, identify and construct statistical units. In this context are procedures and information required to derive coding (ISIC, Institutional sector), unit characteristics (turnover, employment) and identification of activity of special interest. For efficient allocation of samples it is, beside the methodology, important that different characteristics of the establishment are of high quality. These issues have been identified as particularly important at KNBS.

---

[3] KNBS, Kenya Facts and Figures 2012, Employment by Industry, 2008-2011, page 32

It is important that the design of the BR is developed based on the insight on how the statistical business register can fulfil user needs and assist the compilation of statistical data as a tool for the preparation of surveys. The BR design should be developed in close cooperation with the statistical domains, as users of the BR. At KNBS is the ambition that the BR, not only should have the capacity to assist the statistical domains regarding sampling, but also regarding estimation, e.g. treatment of non-response and grossing-up. Preferably will the BR design elaborate the processes for issues related to sampling and estimation methodology. Capacity building within the area was also highlighted as a key measure to reduce the degree of dependence on external support.

The design should take into consideration how the BR can serve as a source of information for statistical analysis of the business population and its demography.

Another important feature of the design is that it should define the scope of the register. Although recognized that information about the informal sector is of high priority it is recommended that the scope of the BR component is limited to the formal sector, at least in the first phase of the redesign process. Regardless, it is important to decide on how to define the formal and informal sector. If is recommended that international references are applied.

"The informal sector is broadly characterised as consisting of units engaged in the production of goods or services with the primary objective of generating employment and incomes to the persons concerned. These units typically operate at a low level of organisation, with little or no division between labour and capital as factors of production and on a small scale. Labour relations – where they exist – are based mostly on casual employment, kinship or personal and social relations rather than contractual arrangements with formal guarantees. The informal sector thus defined excludes households producing exclusively for own final use."[4]

It is worth mentioning that KNBS will be in a better position to present statistics on the informal sector at the end of the period of cooperation. KNBS plans to carry out a business module within KHIBS, the MSME survey, collecting information from counties and the establishment survey will all contribute to extensive information about the informal sector. This is also a reason to why it might be preferred to delay the introduction of the informal sector into the BR (even as a separate part). It is also a matter of resources, to keep a register on informal establishments updated can be very difficult and expensive.

Support system for BR validation, maintenance, sampling and estimation
Depending on the results from the design phase and the identified BR functions and services required, there might be a need to improve existing support systems/applications or develop new systems/applications. In this phase, it is crucial that KNBS can devote appropriate IT knowledge and resources to the component. Statistics Sweden will limit the TA to the requirements of the system.
An expressed wish from KNBS is that an automated process should support estimation and sampling. There is already an IT-system in place but it has so far not been possible to implement in statistical production. Another issue pointed out was the need for common services in production, e.g. common software and a platform for validation and analysis of data. This is outside the scope of the component.

---

[4] Measuring the Non-Observed Economy: A Handbook, OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, 2002.

5 officers trained in Business Register management
Currently, three people are involved in the work with the business register, but far from three full-time staff. In principle, the register is updated with information from administrative sources during one month at the beginning of the year. Four to five months after updating is the statistical register established through profiling and identification of relations between establishments and firms, a work that last for no more than two months.

To achieve the ambitious change proposed in the KNBS strategy and in this plan, there must be staff, able to work continuously with the business register during the whole year available. With the goal of five persons to be well versed in the work of the business register, these five persons have to work, in principle, fulltime with the business register. Much of the competence will be gained from on-the-job training, experience setting up the design and the practical BR development progress. A study visit to Statistics Sweden is recommended in this component.

It should be noted that KNBS intention to strengthen the competence concerning sampling and business statistics methodology, would require specific training in sampling and estimation methodology.

Comprehensive and up to date sampling frame
This is the main output of the Business Register and frame methodology applied should be documented in the design phase.

KENSTAT 2016:xx
Economic statistics

Business Register

Report from a mission to Kenya National Bureau of Statistics
Nairobi, Kenya

19 – 30 September, 2016

by

Staffan Blomquist
Christian Ekström

Report from a short-term mission to Kenya National Bureau of Statistics

Project:	Cooperation project KNBS/SCB
Project funding:	Sida
Project component:	Economic statistics
Project subcomponent:	Business register
Cooperation project partners:	Statistics Sweden, KNBS
Mission period:	19 – 30 September 2016
Place:	Nairobi, Kenya

1. Summary of the report

The main task for the mission was to analyse the existing, as well as future needed, variables in the Business Register (BR). The analysis was to be based on user needs within the Kenyan National Bureau of Statistics (KNBS). Based upon these analyses a detailed plan and time schedule should be presented for the future activities. Further, the mission should provide a draft design regarding the system of the new updated/upgraded BR.

The KNBS BR consists of legal and local units. Data is captured from both administrative sources and from in house surveys. A census of firms and establishments is scheduled for end 2016. The situation for the BR is more or less the same as in October 2015 with a mix of objects in the database and insufficient links between them. Usage of the BR is still very low.

It seems like the KNBS has a good opportunity within a period of one year to create a BR that serves many user needs. The main recommendations in order to achieve this are to continue cooperation with the KRA and to update the BR with a comprehensive dataset based on tax records. The census on firms and establishments could improve information on establishment units in the BR. It is however important to increase the effort to establish links between the firm and establishment units in order to develop two separate and distinct objects in the BR.

2. Background

The overall project objective is that KNBS will improve its institutional capacity to produce statistics in line with European Union and international standards. The project document for technical support to KNBS identifies six main components; management capacity, statistical methodology (including dissemination and analyses), economic statistics (including labour statistics), agriculture statistics, environment statistics, and gender statistics.

3. Project assessment

Kenya's long term development strategy is guided by the Kenya vision 2030 which outlines a wide range of Milestones that the country will strive to achieve before the year 2030. The vision is implemented in five year Medium Term Plans. The current Medium Term Plan (2013-17) recognizes the importance of a National Statistical System (NSS) supporting planning, monitoring and evaluation of government programmes within the framework of Vision 2030. This is certainly true for economic statistics as one of the three pillars in the vision addresses economic development. The Bureau is the principal custodian of the central business register which is usually used as a sampling frame for all establishment based surveys in the modern sector of the economy. Therefore, as a main producer of economic statistics and as coordinator of the NSS, KNBS has a critical role in the process for Kenya to realize its vision. Relevant economic statistics are important from a more general aspect since all pillars of the vision are anchored on macroeconomic stability and reforms. In this perspective a comprehensive Business Register is a fundamental part of fulfilling the vision.

However, the current Business Register has shortcomings when it comes to comprehensiveness and quality. As a consequence, the Business register is used in a limited way as a sampling frame for establishment based surveys. There is therefore great need for improvement, something that also the World Bank has recognized in their Kenya Statistics Program for Results (KSPforR) were an upgraded and updated Business Register is one of the major goals.

I)   Project component

Component objective
Improved quality and coverage of economic statistics.
Component
Economic statistics.
Subcomponent
Business Register.

II)   Previous results

A first mission on the subcomponent Business Register was carried out in October 2015.

4.   Results of the mission

In the following the terms legal unit and firm have been used interchangeably and the same goes for the terms establishment and local unit. The terms Business Register (BR), master file and master frame all denote a list or table of statistical units and variables in a database.

I)   Completed tasks

The following table depicts the expected outputs, if they were achieved or not and the reasons why not.

Output	Completed (Yes/No)	Comment
Analysis of existing and future needed variables in the BR	Yes	
Assess user demands of the BR	Yes	
Detailed plan for and time schedule for future activities	Yes	
Draft design of updated/upgraded BR	Yes	
Analysis of data from the KRA	No	This activity was also planned for the mission. It could not be fully accomplished due to lack of comprehensive KRA data.

II) Continued assessment of the current status of the BR

- A deeper assessment of the BR was made based on information loaded in the database table with information on establishments. Some basic characteristics are shown in the table below.

	Active units		Closed units	
LE-Year[5]	**Units**	**Employees**	**Units**	**Employees**
0	861	0	26,246	0
2010	270	0	40	0
2011	54,002	1,570,417	3,003	78,524
2012	3,706	130,108	3,707	231,150
2013	144	10,779	19	310
Total	58,983	1,711,304	33,015	309,984

- A number of identities are used for firms in Kenya. Administrative data providers assign identities of their own and the KNBS also creates one firm identity (Firm-ID). For establishments the KNBS creates an identity denoted EIC. Establishment level data is mainly provided by in house surveys or originates from County Governments. For the latter, each record has a Business Permit ID for firms. There are no links between the ID-numbers provided by the following administrative data sources: KRA, National Social Security Fund (NSSF), Registrar of Companies (ROC), National Hospital Insurance Fund (NHIF) and the County Governments.
- There are blanks in the relations between firm and establishment units in the BR database. This makes it difficult to keep a register with consistency between statistical units.
- There are also several blanks in variables.
- The linking between firms and establishments in the database is based on matching procedures for addresses and names of the firm or local unit.
- One delivery from the KRA had been made since the 2015 mission. It contained about 40 000 records of the largest firms. These data was overviewed briefly. The dataset included number of employees, but it was uncertain if this information originated from the KRA. This file has not updated the BR.
- Manual updates have however been made to the BR with a data originating from County Governments on establishment level.
- Batch updates are not implemented.
- Variables central to a BR such as turnover and activity code (section level) exist and are also available on legal unit level via administrative sources.
- Other variables such as status, activity code (class level) and number of employees may not be found in administrative records. However, information to derive the variable "status" could be collected from administrative data records.
- There has been no possibility to evaluate data from KRA in detail or to investigate variables potentially available from the KRA. The BR-team suggests that a population based on tax records will include about 80 per cent of the formal sector. The rest would be found in other administrative sources such as the NSSF and the ROC. It is however our view that the KRA should be the main supplier of data to the BR.

---

[5] Denotes the year of the LE-survey

III) Database, sampling function, ICT-environment

- An overview of the data model was presented. The database is constructed is FoxPro language and linked to Dbase.
- The data model contains data, documents, and codes.
- Data tables include one separate table for firms and one for establishments. This means the two objects can be separated in the database. Also, there are separate tables for codes and classifications.
- The database seems well adapted to the BR needs.
- Support for FoxPro and Dbase has expired.
- Some adaptions of the database may be necessary depending on if any new variables are to be included.
- Any adaptions of the database will be handled by an external consultant.
- There is a front end system for maintenance, reports and sampling.
- It is not possible to use the front end sampling function. Information from the BR-team suggests this depends on deficiencies of how data is organised in the database, i.e. a mix of units and insufficient links between them, as well as empty cells for many variables.
- Any adaptions of the database also require programming upgrades of the front end functions.
- As regards the ICT-environment, we found that the situation remains similar to the previous mission: no regular back-ups are made for the BR and the entire BR is stored on one or more PC, and not on a server.

IV) Census of firms and establishments

- A census of firms and establishments is planned to be carried out by the KNBS in late 2016.
- It was not clear what population will be the base for the census. One alternative was to use the existing, but somewhat outdated, information in the BR. Another suggestion presented by the BR-team was to use a population of firms from the KRA. The population would then be the taxpaying firms at a given point in time, which in turn would make up for the largest share of the formal sector in the economy[6].
- The census aims at collecting information on both firms and establishments.
- Data will be collected using tablets and paper forms. The results will be loaded to a server and then transferred to the BR database.
- A draft form for the census was presented. There was a large resemblance with other survey forms that were presented in the previous mission, such as the LE-survey and the BR-survey. The similarities regarded both questions and layout.
- A number of questions deemed out of scope for a BR, according to our view, were included on the form.
- The census form was to be finalised by the KNBS during a workshop which was planned after the BR-mission.

V) User demands on the BR

---

[6] The BR-team suggests that the formal sector be made up of all registered firms. All firms must be formally registered at the Registrar of Companies. However, having a lot of dormant firms in the Registrar records, the criteria for firm status "active" could be created by an additional registration with either the KRA or the NSSF.

- The mission included seven interviews with representatives from the following statistical domains and potential BR-users: Labour statistics, Agricultural statistics, ICT statistics, National accounts, Industrial statistics, Construction statistics and Financial statistics.
- Target units for these domains were firms and establishments. Firm was deemed necessary in all domains, whereas establishment was pointed out by four. KNBS has an obligation to produce regional statistics if possible. This means establishment units is also necessary in the BR.
- Yearly updates of the BR would do for all domains but one. In this case quarterly updates were required.
- Important variables for the users were identifiers, classifications, contact variables and stratifications variables.
- Four interviewees also wanted subject matter variables included in the BR.
- Moreover, a general view was that the BR must be produced and made available for all enterprise based surveys at the KNBS.

VI) Future plans

- The KNBS plans for the census on firms and establishments. From our experience a census is a challenging task that requires an additional input of resources. At the time of the mission however, we just got a little insight into the planning.
- A coming deadline for WB-funding is August 1, 2017. For this to be realised the KNBS should accomplish developments in the BR.

VII) Detailed plan for future activities and missions

See recommendations, chapter 5.2 and appendix 5.

5. Recommendations
I) Recommendations to reach component output(s)

Below are our recommendations for reaching the outputs of the Business register component. A plan for future missions is found in annex 5.

The main recommendation is to consider a stepwise development of the BR having the WB-funding deadline in mind. It is our opinion that the KNBS has a good opportunity to create a BR that serves many user needs. An update with a comprehensive dataset with administrative data on firm level would make up for a start. We include 4.1.1, 4.1.2, 4.1.3 and 4.1.4 as outlined below in the first step recommendations.

II) Step 1, administrative data: Increase cooperation with KRA. Collect comprehensive data based on tax records.

- Make a request to the KRA in order to investigate what variables KRA may provide, as well as the number of records.
- Set up an agreement between KNBS and KRA regarding scheduled data deliveries. Specify grade of detail, format and time for delivery. Special attention should be given to establish secure data transmissions.
- The dataset from KRA should be considered to form the population for the proposed census on firms and establishments.

- Data from KRA should be updated annually. It should also be investigated if data can be updated from KRA more frequently, for instance on a quarterly basis.
- For each record with a KRA PIN-number, a KNBS Firm-ID must be created. Blanks in either of the two identities cannot be allowed.
- Based upon information from the tax records, some basic events for single located firms could be as follows:

Single located firms

Event	Firm Source	First Establishment Source
Registration (birth into BR)	KRA	KRA
Deregistration (death in BR)	KRA	KRA
Active**	KRA/KNBS method	KRA/KNBS method
Not active***	KRA/KNBS method	KRA/KNBS method

** turnover or employees
*** no turnover and no employees

III)   Step 1: Distinguish statistical units and variables in the BR, establish links

- Create a set of Legal Units (Firms) and Local Units (Establishments). It is important to create and maintain primary keys (identifiers) to each one of them. It is also important to establish links (foreign keys) between the two of them.
- The primary ID for firms should be Firm-ID. Blanks in Firm-ID cannot be allowed. Each Firm-ID must be related to another ID, such as the KRA PIN-number.

IV)   Step 1: Status of units

- Create variables that confirm the current status of a unit. Suggestion: active, dormant or closed. We believe the status could be derived from information in the administrative data records on firms. Questions on the status of firms have also been suggested to the census form.

V)   Step 1, census: Update the BR with census data on firms and establishments

- Design the census form strictly from a BR-perspective. This means to include BR-core variables only and to have a clear division between questions related to firms and establishments.
- Strongly consider using a population of firms from the KRA.
- Update information on firms and establishments for the firms included in the population of firms from KRA only.

- Assign all planned full time resources for the BR (or preferably more) to edit census data collected and to maintain the links between firms and establishments.
- Based on information from the tax records and the census, some basic events for multi located firms regarding the secondary establishment could be as presented below:

Multi located firms

Event	Firm Source	Secondary establishments Source
Registration (birth into BR)	KRA	Different sources i.e. census, BR-survey and maintenance
Deregistration (death in BR)	KRA	Different sources i.e. census, BR-survey and maintenance

- When creating a new, secondary establishment it should always be linked to a firm with an existing PIN code.

VI) One BR for the KNBS

- We suggest the KNBS to take measures in order to make the BR more available to all users in the bureau. Whether it may be of an organisational nature or not, we leave to the KNBS to investigate.

VII) Step 2, administrative data: Collect more administrative data.

- Expand data collection of administrative data to the National Social Security Fund (NSSF) and the Registrar of Companies (ROC).

  This step requires strict rules for maintenance. If inconsistencies appear between two data sources – which most certainly will be the case – it must be clear how to prioritize between the data sources. This step also involves linking records by using other characteristics than ID-numbers, such as matching procedures for names and addresses. From a maintenance perspective this is time consuming.

- New entries must include at least one firm ID-number from a source approved by the KNBS BR.
- Relations must be established from the identifiers from these sources to the KNBS Firm-ID in the database. The link is Firm-ID.
- Firm-ID cannot be blank. At least one relation to one of the external ID must be established.
- Prepare for updates of records with another reference year. Handle start-ups and close downs in the database, as well as changed values in variables. One option is to create one total frame annually.
- Update data from administrative sources more frequently than once a year.

VIII) Remaining recommendations from 2015

Estimate total firm and establishment population by using other sources

- A mission on Methodology should try to estimate the size of the informal sector and the BR under coverage. By using results from KIHBS and other surveys, under coverage of firms and local units could be calculated. This could also provide an

estimate of the magnitude of the informal sector. This is not only requested by the KNBS, but also by the KRA[7].

Evaluate the system for sampling

- A mission on methodology should focus on evaluating the methods for sampling implemented in the sampling system. The results should be input for the work on restarting the sampling system.

Secure the ICT environment at KNBS

- It is very important that the BR operates in a secure ICT environment with functioning servers, automatic back-ups, etc. No mission is suggested since no ICT-component is included in the overall cooperation project. Instead, the KNBS will have to promote actions to secure the ICT environment. In addition, since the FOXPRO software is no longer being supported, the KNBS requires to acquire a supported software.

I) Recommendations to reach the component outcome(s)

No recommendations for the component outcome are required at this point.

II) Risks and risk management

Resources

It is very important that KNBS can provide the planned resources. The planned output for 2016 is 5 trained staff at KNBS. Although training is essential, it is vital that staff is also assigned to the BR in order to work with maintenance of the register. Furthermore, it would be valuable to also train/assign ICT-competence to the BR-team. Then resources assigned for subject matter duties would not be burdened by ICT-related issues.

III) Time schedule

See 5.2 and appendix 5.

6. Conducted activities

I) Meetings and workshops

- Assessment of the KNBS BR continued from last mission regarding
  - Data model
  - Brainstorming: draft vision of register contents (to be state)
  - Populations, statistical units and variables
- Discussion, administrative data input from KRA
- Discussion, census of firms and establishments
- Interviews with potential BR-users from seven statistical domains within the KNBS.

II) Input from short term consultants

- Documented comments on questions and layout of the census form
- Documented suggestion on variables to be checked with the KRA that KRA may supply to the KNBS BR

---

[7] Kenstat 2015:11

- Comments on data model with two main objects and the links between them
- Comments on current use of Dbase and FoxPro for KNBS Business Register
- Presentation and discussion on statistical objects legal unit, local unit, enterprise unit, kind of activity unit and enterprise group
- Presentation and discussion characteristics of the Albanian SBR and in particular the following variables: status of the unit, activity codes, time stamps, information on sources, BR-team maintenance signature, and time stamp for BR-team maintenance signature
- Plan for the way forward

A list of participants and conducted meetings is included in appendices 2 and 4.

III) Actions needed for preparing the next mission

Action	Deadline (MM YYYY)	Responsible person
Request what variables the KRA can provide and the total number of records	Prerequisite for study visit to Sweden in December	KNBS
Collect the above mentioned data from the KRA	Prerequisite for study visit to Sweden in December	KNBS
Provide a sample of census data for analysis (if the census has started)	Prerequisite for study visit to Sweden in December	KNBS
Census data ready to be inserted into the BR database	Prerequisite for BR mission to KNBS in February	KNBS

Business Register

Report from a study visit to Statistics Sweden
Örebro, Sweden

12 – 16 December, 2016
by

Björn Berglund
Staffan Blomquist
Christian Ekström

Report from a short-term mission to Kenya National Bureau of Statistics

Project:	Cooperation project KNBS/SCB
Project funding:	Sida
Project component:	Economic statistics
Project subcomponent:	Business register
Cooperation project partners:	Statistics Sweden, KNBS
Mission period:	12 – 16 December 2016
Place:	Örebro, Sweden

1. Summary of the Report

The planned outputs from this study visit were an analysis of KRA and census data in order to create variables on objects LEU and LOU. Further the analysis should include how to crate links between LEU and LOU. In addition to this the data base structure of the business register should be analyzed and changed accordingly to new requirements. An updated plan for the forward work on the development of the BR should also be presented.

During the workshops KNBS and SCB jointly made a draft plan for the period ending in July 2017 including:

- Data model and application
- Business rules
- Plan for loading census data into the BR data base

SCB also was to suggest a draft project plan for the work to come until July 2017. The plan is attached to this report.

2. Background

The overall project objective is that KNBS will improve its institutional capacity to produce statistics in line with European Union and international standards. The project document for technical support to KNBS identifies six main components; management capacity, statistical methodology (including dissemination and analyses), economic statistics (including labour statistics), agriculture statistics, environment statistics, and gender statistics.

I) Project assessment

Kenya's long term development strategy is guided by the Kenya vision 2030 which outlines a wide range of Milestones that the country will strive to achieve before the year 2030. The vision is implemented in five year Medium Term Plans. The current Medium Term Plan (2013-17) recognizes the importance of a National Statistical System (NSS) supporting planning, monitoring and evaluation of government programmes within the framework of Vision 2030. This is certainly true for economic statistics as one of the three pillars in the vision addresses economic development. The Bureau is the principal custodian of the central business register which is usually used as a sampling frame for all establishment based surveys in the modern sector of the economy. Therefore, as a main producer of economic statistics and as coordinator of the NSS, KNBS has a critical role in the process for Kenya to realize its vision. Relevant economic statistics are important from a more general aspect since all pillars of the vision are anchored on macroeconomic stability and reforms. In this perspective a comprehensive Business Register is a fundamental part of fulfilling the vision.

However, the current Business Register has shortcomings when it comes to comprehensiveness and quality. As a consequence, the Business register is used in a limited way as a sampling frame for establishment based surveys. There is therefore great need for improvement, something that also the World Bank has recognized in their Kenya Statistics Program for Results (KSPforR) were an upgraded and updated Business Register is one of the major goals.

II) Project component

Component objective
Improved quality and coverage of economic statistics.

Component
Economic statistics.
Subcomponent
Business Register.

III) Previous results

A first mission on the subcomponent Business Register was carried out in October 2015, and a second, two-week mission, was carried out in October 2016.

3. Results of the mission

In the following the terms legal unit and firm have been used interchangeably and the same goes for the terms establishment and local unit. The terms Business Register (BR), master file and master frame all denote a list or table of statistical units and variables in a database.

I)   Completed tasks

The following table depicts the expected outputs, if they were achieved or not and the reasons why not.

Output	Completed (Yes/No)	Comment
Analysis of data from the KRA and census	Yes	
Analysis of links between LEU and LOU	Yes	
Analyse the structure of the business register data base and consider new requirements	Yes	
Detailed plan for and time schedule for future activities	Yes	Draft version

II)   Preparations by SCB

- Analysis of the current KNBS BR database
- Introducing SCB ICT-expert Björn Berglund to the KNBS BR database
- Sending questions to KNBS to be prepared for the study visit

III)   Preparations by the KNBS

- Collecting a data file from the KRA with information on firms.
- Draft version of census questionnaire

IV)   Main findings during study visit

- A census of firms and establishments is planned to be carried out by the KNBS in spring 2017.

- The population for the census will be the based on the newly collected population of firms from the KRA. The population would then be the taxpaying firms at a given point in time, which in turn would make up for the largest share of the formal sector in the economy[8].
- A cut-off will be used; firms with legal form equal to sole proprietor and with 5 number of employees or larger will be included on the census. There may however be exceptions from this general rule.
- There will eventually be three main data sources in use for the 2017 KNBS BR as suggested by the KNBS: KRA, Census of establishments and the county councils.
- The census aims at collecting information on both firms and establishments.
- Data will be collected using tablets and paper forms. The results will be loaded to a server and then transferred to the BR database.
- The census form is to be finalised by the KNBS in the beginning of 2017.
- One data base table for firms needs to be created.
- Relations and links between firms and establishments were discussed.
- There will be only one industry code assigned to each establishment. The unit establishment may therefore be better characterised as local kind of activity unit (LKAU)
- Some adjustments in the data base table for establishments (LKAUs) may be necessary depending on what variables will be included.
- The links between firm and establishments (LKAUs) need to be further elaborated by the KNBS.
- The front end system for sampling should, as suggested by the KNBS, be part of the BR in July 2017.
- Data base language Dbase will be used.
- The ICT-expert from SCB will have an advisory role in the project. This means no code writing in neither Dbase nor Fox Pro.
- Some basic arrangements will be required for the ICT-environment at KNBS regarding the BR, such as a back-up of the register.
- All development work will be made in a copy of the current KNBS BR data base.
- The KNBS BR data base requires referential integrity to avoid duplicates and inconsistencies.
- A test environment will have to be set up regarding the data base and the front end application.
- Parts of the collected COE-variables must be tested before loaded into the new BR data base.
- Final loading of data from the COE to the BR data base is scheduled to take place in June.
- The follow-up by the World Bank will be in the form of an audit performed by consultants. A demonstration of the BR, including the front end application, during a visit to the KNBS is expected.

---

[8] The BR-team suggests that the formal sector be made up of all registered firms. All firms must be formally registered at the Registrar of Companies. However, having a lot of dormant firms in the Registrar records, the criteria for firm status "active" could be created by an additional registration with either the KRA or the NSSF.

V) What will be in place for the World Bank audit

A KNBS data base with newly collected data from 2016 and 2017 and a front end application. Further specifications needs to be elaborated by the KNBS.

VI) The role of SCB

SCB will assist the KNBS in its work to complete a BR in time for the World Bank audit scheduled to July 2017. The assistance will be advisory.

VII) The plan forward

- The KNBS and SCB jointly made a draft plan for the period ending in July 2017. For further details regarding the plan forward, see appendix 5. In short, the main features are:

1. Data model and application (front end system)
   - Objects, identities and variables must be made clear and defined
   - The KNBS will suggest a design
   - Rules for back-up and restore is needed
   - Documentation is needed

2. Business rules
   - These will regard the most essential variables such as number of employed, turnover, ISIC, production value.
   - Rules for imputation etc.
   - Data sources to be used

3. Plan for loading census data into the BR data base
   - Rules for batch updates
   - Data cleaning
   - Mapping variables
   - Tests in different steps

VIII) Detailed plan for future activities

See chapter 5.2 and appendix 5.

4. Recommendations

I) Recommendations to reach component output(s)

The main recommendation is to enhance and elaborate the draft plan forward as presented briefly in section 3.1.7 and the future activities as presented in section 5.2 and appendix 5.

We also want to stress remaining recommendations from the previous report regarding estimate total firm and establishment population by using other sources, evaluate the system for sampling, secure the ICT environment at the KNBS, and to address the risk of lack of resources.

However, during the study visit we learned that one person with ICT-competence at KNBS could be assigned to work with the BR database and make necessary changes in order to have a BR running in July 2017.

II) Time schedule

See 5.2 and appendix 5.

5. Conducted activities

I) Meetings and workshops

- Issues of the KNBS BR regarding:
  - datamodel
  - data from KRA
  - future data input from census of establishments
  - register contents (to be state)
  - populations, statistical units and variables
- Workshop on SBR as a sampling frame
- Data collection from Individuals and Households
- Strategy, Co-operation
- Swedish system of statistical registers

II) Input from short term consultants/SCB

- Power point presentations linked to the above mentioned meetings and workshops
- Wiesbaden conference paper on new tools for data collection and maintenance in the Swedish Statistical Business Register
- Introduction of SCB ICT-expert Björn Berglund to the project
- Draft version of plan for the way forward

A list of participants and conducted meetings is included in appendices 2 and 4.

III) Actions needed for preparing the next mission

Action	Deadline (MM YYYY)	Responsible person
Prepare a draft project plan for January – July 2017	13-01-2017	SCB
Final decision on COE-variables/questionnaire	January	KNBS
Provide a first version of database structure of the KNBS BR	January	KNBS

KENSTAT 2017:14
Economic statistics

Business Register

Report from a mission to Kenya National Bureau of Statistics
Nairobi, Kenya

2 October - 13 October, 2017

by

Bjorn Berglund
Staffan Blomquist
Christian Ekstrom

Report from a short-term mission to Kenya National Bureau of Statistics

Project:	Cooperation project KNBS/SCB
Project funding:	Sida
Project component:	Economic statistics
Project subcomponent:	Business register
Cooperation project partners:	Statistics Sweden, KNBS
Mission period:	2 - 13 October 2017
Place:	Nairobi, Kenya and Orebro, Sweden

1. Summary of the report

The main task for the mission was to finalise the setup of a Business Register (BR) at the Kenyan National Bureau of Statistics (KNBS).

The outcome from the mission was an updated register of establishments pointing at reference year 2016, an application with a sample module and a database. There is also a read-only database for back-up purposes. The register contains BR core variables such as ISIC, sector, status, region, number of employed and turnover. The SCB and KNBS teams also exchanged knowledge and experience on issues related to register methodology and commenced work on future steps for developing the KNBS BR.

2. Background

The overall project objective is that KNBS will improve its institutional capacity to produce statistics in line with European Union and international standards. The project document for technical support to KNBS identifies six main components; management capacity, statistical methodology (including dissemination and analyses), economic statistics (including labour statistics), agriculture statistics, environment statistics, and gender statistics.

2.1   Project assessment

Kenya's long term development strategy is guided by the Kenya vision 2030 which outlines a wide range of Milestones that the country will strive to achieve before the year 2030. The vision is implemented in five year Medium Term Plans. The current Medium Term Plan (2013-17) recognizes the importance of a National Statistical System (NSS) supporting planning, monitoring and evaluation of government programmes within the framework of Vision 2030. This is certainly true for economic statistics as one of the three pillars in the vision addresses economic development. The Bureau is the principal custodian of the central business register which is usually used as a sampling frame for all establishment based surveys in the modern sector of the economy. Therefore, as a main producer of economic statistics and as coordinator of the NSS, KNBS has a critical role in the process for Kenya to realize its vision. Relevant economic statistics are important from a more general aspect since all pillars of the vision are anchored on macroeconomic stability and reforms. In this perspective a comprehensive Business Register is a fundamental part of fulfilling the vision.

However, the current Business Register has shortcomings when it comes to comprehensiveness and quality. As a consequence, the Business register is used in a limited way as a sampling frame for establishment based surveys. There is therefore great need for improvement, something that also the World Bank has recognized in their Kenya Statistics Program for Results (KSPforR) were an upgraded and updated Business Register is one of the major goals.

2.2   Project component

Component objective
Improved quality and coverage of economic statistics. *Component*
Economic statistics. *Subcomponent* Business Register.

2.3 Previous results

A first mission on the subcomponent Business Register took place in October 2015, and a second, two-week mission, in October 2016. The same year there was a study visit to SCB in December. In July 2017, the consultants at SCB performed a one-week preparatory work for the final mission.

3. Results of the mission

In the following the terms legal unit and firm have been used interchangeably and the same goes for the terms establishment and local unit. The terms Business Register, master file and master frame all denote a list or table of statistical units and variables in a database.

3.1.1 Completed tasks

The following table depicts the expected outputs, if they were achieved or not and the reasons why not.

Output	Completed (Yes/No)	Comment
Use data collected in the Census of establishment	Yes	-
Set up a business register with possibilities to view, edit and evaluate business data	Yes	-
Set up a database	Yes	-
Restart the sample module in the business register application	Yes	-

3.1.2   Data: Census of establishments and secondary sources

- KNBS undertook a census during 2017 based on a population of firms from the Kenya Revenue Authority (KRA). A number of variables were collected, but not all is included in the BR. The reason is that the database and application is not adapted for all new variables and this was not the task for the mission. For a complete list of variables in the BR, see appendix 3.
- Results from the census make up for about 45 000 of the nearly 145 000 records in the BR. More than two thirds instead come from secondary sources. The majority of these records come from the KRA. For a list of secondary sources and estimated reference years, see appendix 4.

3.1.3   Business Register

- The main unit in the BR is the local kind of activity unit. All firm information was disaggregated if more than one economic activity was recorded. Then, another unit was created with a specific activity and in most cases a specific regional locality.
- Both firm level ID and establishment level ID have been created. This means there is a possibility to aggregate data on both national and regional level.
- There are blanks in some variables. This is due to the use of multiple sources that do not cover the same information.
- Maintenance could be performed by using either the maintenance function in the front end application, or by manipulating data by programming.

3.1.4   Database

- The database is constructed is FoxPro language and linked to Dbase.
- Data tables include one table that constitutes the BR and a number of reference tables.
- The database was adapted to the existing functionalities in the front end system.

- Any adaptions of the database in terms of adding further variables also require programming upgrades of the front end functions.
- The BR is stored on one PC and not on a server.

3.1.5 Sampling module

- The front end system was restarted. This can be used for some maintenance, producing some reports and parts of the sampling function.
- The sampling function that was restarted regards the LE-survey. This was the priority of the KNBS. There is also a sampling function called ad hoc sampling. This was not restarted.
- As far as both teams could perceive, the output of LE-survey sampling function seemed reasonable. However, this is no guarantee that the samples produced are actually accurate; a thorough assessment of the back end programming is deemed necessary for that.

4. Recommendations

4.1 Recommendations

- Manual updates, if any, should be concentrated to improve the quality in variables, such as the ISIC codes.
- Insertion of additional units should be avoided since there is a high risk of creating duplicates which in turn cause inconsistencies in the BR
- A mission on methodology should focus on evaluating the methods for sampling implemented in the sampling system.
- A mission on Methodology should also try to estimate the size of the informal sector and the BR undercoverage. By using results from KIHBS and other surveys, undercoverage of firms and local units could be calculated. This could also provide an estimate of the magnitude of the informal sector.
-

5. Conducted activities

5.1.1 Meetings and workshops

- Setting up the KNBS BR
  - Programming
  - Testing
  - Validating
- Clarifying details regarding data contents
- Clarifying details regarding contents in the reference tables

5.1.2 Input from short term consultants

- A read only data base for back-up purposes
- Documentation in English of the Swedish BR
- Workshop to create reports from the BR by running scripts outside of the application
- Presentation and discussion on the role of the BR in a statistical bureau in Europe