



Administrative Data and Agricultural Statistics – What Strategy and Methods should we adopt?

Anders Wallgren and Britt Wallgren
BA Statistiksystem AB, formerly of Statistics Sweden
Norrhagevägen 30
SE-719 93 Vintrosa, Sweden
ba.statistik@telia.com

DOI: 10.1481/icasVII.2016.f33d

ABSTRACT

Agricultural surveys traditionally consist of censuses and sample surveys. The *sample surveys* commonly use stratified cluster sampling or stratified multiple-stage cluster sampling designs to select sampling units as farms or households from an area frame. The Farm Census is often based on a frame that is based on the agricultural households in the Population and Housing Census.

Many countries have started to use administrative registers with microdata to produce social, economic and agricultural statistics. The countries in northern Europe started the transition from a traditional, area frame-based statistical system into a completely register-based system in the 1960s. Since 1997 Eurostat has been in favor of the member countries' increasing use of administrative registers for statistical purposes.

In a register-based system, all sample surveys are based on frames that have been created with statistical registers. In addition, for a *census* such as the Farm Census, the Farm Register is used as a frame and questionnaires are sent to all farms to gain additional information that is not already in the register. A third kind of survey is based entirely on available microdata in the system of statistical registers. The traditional Population and Housing Census is replaced by a number of such *register surveys* that are based on statistical registers created by the national statistical institute. The statistical registers can be used to create an integrated register system, where microdata can be combined and used efficiently. Consistency and coherence are important quality dimensions that can be measured and improved through the systems approach.

Expert judgements or eye estimates are often used in Agricultural Routine Data Systems to produce agricultural statistics in developing countries. Unfortunately, this kind of data is often called *administrative data* as in FAO (2015e). The methods used for expert judgements have nothing in common with the methods used for censuses, sample surveys or register surveys. Administrative data originates from administration of separate units as persons or enterprises – expert judgements are not used in this way and should not be called administrative data. In contrast to administrative data, data collected in Agricultural Routine Data Systems do not represent something new in agricultural statistics.

The *Global Strategy to Improve Agricultural and Rural Statistics* by the World Bank, FAO and the UN (2011), was launched to modernize and improve agricultural statistics in developing countries. One important goal is that agricultural statistics should become an integrated part of the national statistical system. Another part of the strategy is for administrative data to play a more important role in the production of agricultural statistics than today. FAO has therefore published a number of working papers on *Improving the Methodology for Using Administrative Data in an Agricultural Statistics System* (FAO 2015a–e).

This paper discusses the long-term strategy for developing agricultural statistics to achieve integration with the national system and the more efficient use of administrative data. The ultimate goal is to create a statistical Farm Register that can be regularly updated with administrative sources. Note that administrative registers are also being used for non-agricultural statistics. Parallel to the work with improving agricultural statistics, national statistical systems are becoming increasingly register-based. When the Population and Household Census is replaced with statistical registers, the preconditions for a Farm Census will be completely different – the Population Census can no longer be used as a frame for agricultural surveys.

Keywords: Integrated statistical systems, Farm register, Coverage errors, Multiple frame approach

1. Agriculture and the rural population in a register-based system

This section describes how statistics on agriculture and the rural population are produced in a country where statistics production is based on registers. Statistics on agriculture are based on the Farm Register, but the Farm Register is only one of many registers used for statistics production. The Population Register and Business Register and the registers that have replaced the population and housing census constitute an integrated system.

1.1 The key features of a register-based system

In a traditional system based on area frames, the *location* of a statistical unit is the factor that determines if a unit is sampled and interviewed or not. A sampled person's name or identity is not important for the production of statistics. Consequently, information regarding the identities of units is not used in the production process. Countries with register-based production systems use *identity numbers* in all surveys: all censuses, sample surveys and register surveys. The Population Register, the Business Register and the Farm Register are used to create frames, and the statistical units are identified by unique numbers that are used in the production process. In sample surveys and censuses, the identities are used to contact the respondents by mail or telephone. In register surveys, microdata from different registers is combined to create the required data set for the specific survey. These combinations are created by record linkage using the identity numbers as matching keys.

In a national statistical system with a coordinated register system, all microdata in the system can be integrated. In a coordinated system, the populations in the different registers are consistent regarding coverage, and variables in different registers do not have contradicting values. This is what we mean by *an integrated statistical system*. Of course, perfect consistency is difficult to obtain. However, if coverage problems and variable inconsistencies are tolerable, we can say that the system is an integrated system. Integration of microdata will become possible without disturbing problems regarding lack of consistency between populations and variables.

Chart 1 illustrates a register-based production system. The basis is the register system consisting of four base registers (circles) defining the *populations* in the system and a number of statistical registers with important statistical *variables*. The lines in the chart illustrate the important *links* between units in different registers. We have added two sample surveys – the Labor Force Survey and the Investment Survey. These two sample surveys are only examples; all sample surveys can be linked to the register system.

Chart 1. A register-based production system – the register system and sample surveys

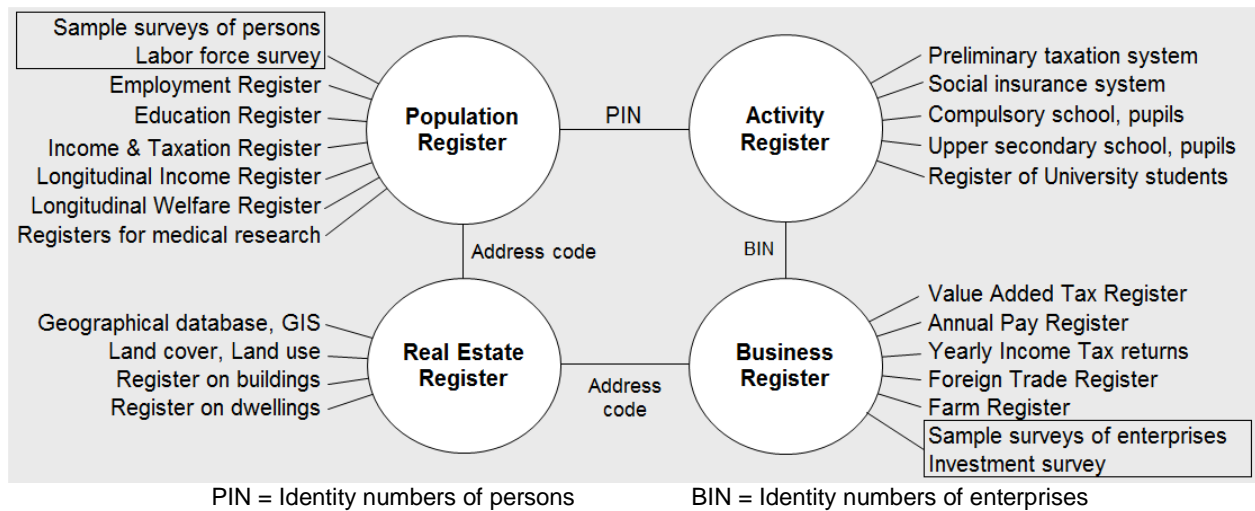
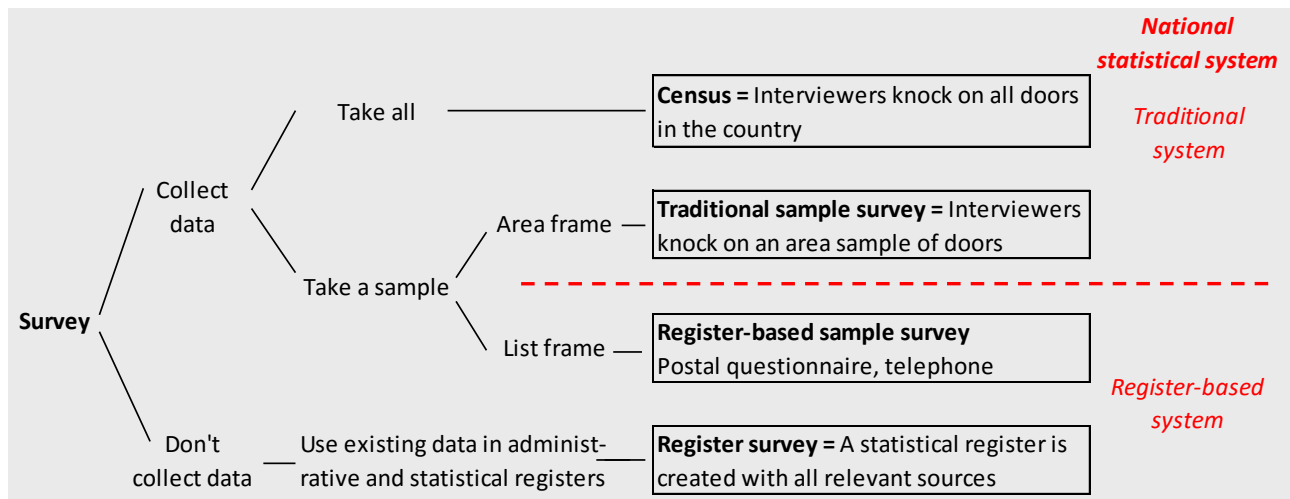


Chart 2 illustrates the different survey methods used in traditional and register-based production systems. In a register-based system the censuses are replaced with statistical registers, and sample surveys are not based on area frames. This explains why a register-based system is much more cost efficient than a traditional system.

Chart 2. Survey methods in traditional and register-based statistical production systems



1.2 The Farm Register in a register-based system

The population of the Farm Register is determined using one or more administrative registers. The Integrated Administrative and Control System (IACS) is a yearly source in countries that are members of the EU. Instead of conducting a census every tenth year with a census population that will soon be outdated, the IACS system generates a new register population every year.

The Farm Register and the Business Register should be consistent. All holdings in the Farm Register should be linked with enterprises in the Business Register. The holdings are local kind of activity units (LKAU) in the Business Register. The enterprises in the Business Register that are linked to the holdings should be active within NACE 1. There should be no other enterprises with NACE 1 as one of its economic activities in the Business Register.

1.3 Opportunities and problems with a register-based system

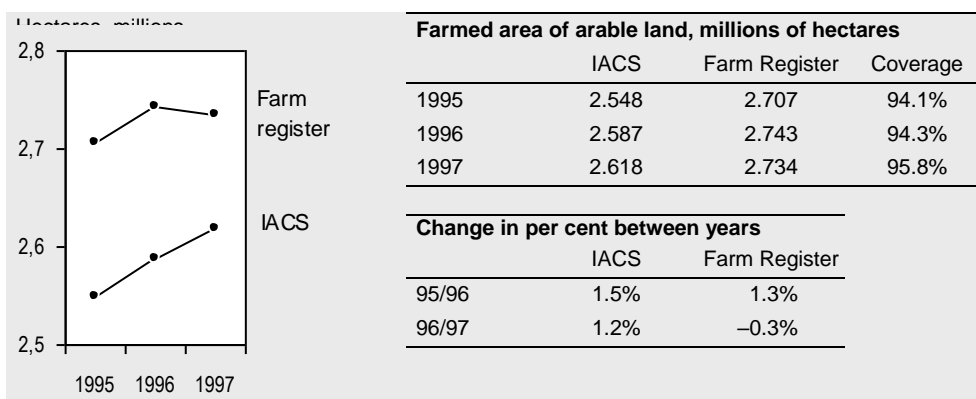
All registers in the system can be used to describe agriculture and the rural population, and there is no extra cost or response burden related to this way of gaining information. If the Farm and Business registers are consistent in the way described in Section 1.2, all administrative data linked with enterprises in the Business Register can be used to describe the agricultural sector.

All registers in Chart 1 contain information regarding agricultural enterprises and/or persons working for agricultural enterprises as self-employed or employed. The households related to these persons can be described as well as the households living in rural areas. With the GIS-system that is linked to the Real Estate Register, the rural regions can be defined freely without using administrative borders.

In Wallgren and Wallgren (2010) we discuss methodological issues and quality problems related to the Farm Register in a register-based statistical system. In a decentralized national system, agricultural statistics and economic statistics can be produced by different statistical institutes. If these institutes do not cooperate, the Farm and Business registers will be inconsistent and this will create problems for the National Accounts, for instance.

Selander et al. (1998) discuss the problem of undercoverage in the IACS register. If this undercoverage is not handled, the consequences can be misleading statistics. This is illustrated in the example in Chart 3, taken from Wallgren and Wallgren (2014). Even if coverage is rather high, variations in coverage will destroy time series patterns.

Chart 3. Undercoverage in an administrative register, the IACS-register



Selander et al. suggest that the population of holdings is divided into three parts, and coverage errors should be reduced by combining register data for the first part with statistical questionnaires for the other parts. This is actually a kind of multiple frame approach:

1. IACS 1 is the part of IACS that consists of the holdings that have reported the entire area of arable land divided among different crops.
2. IACS 2 is the part of IACS where there is no complete crop data in IACS, though there is information for register keeping.
3. The holdings that are outside the IACS system. This third part consists of the holdings that have not applied for subsidies. This part of the population can be monitored by using administrative data from registers with business data.

In Sweden in 1995, IACS 1 comprised some 69,000 holdings, IACS 2 comprised some 6,000 and the non-IACS population some 20,000. The lines of demarcation between the different parts of the population may be changed as a result of a change in agricultural policy.

Some general conclusions can be made based on this IACS example.

- A register-based system is dependent of the administrative systems that generate the administrative data. The coverage problems related with these administrative systems must always be analyzed and monitored over time. Changes in the administrative systems can produce changes regarding coverage.
- The example also illustrates that a multiple frame approach should be used to reduce coverage errors. Undercoverage can be reduced by combining many administrative registers ('frames'). This is an option in countries with many administrative registers available to the statistical institute.
- Coverage errors can be reduced by combining registers with list-based sample surveys or area samples. In the IACS example, list based samples or censuses covering the second and third part of the population will reduce undercoverage errors. In developing countries, combining registers with area sampling may be necessary to reduce the coverage errors related to the registers.

In Wallgren (2016b) we use the dual frame approach to discuss two problems:

- First, residential addresses are often not updated in administrative population registers in many countries. This means that it is difficult to use administrative population data for regional statistics. In countries with strong urbanization, this will lead to overcoverage of rural register populations and undercoverage of urban register populations.
- In addition, a difficult but important problem is the informal sectors that exist in many countries. Family farms in many countries can belong to this informal sector. Only the formal sector will be found in the administrative sources and these administrative registers must be combined with area samples that cover both the formal and informal sectors.

These two problems are common in developing countries. The problems must be solved so that these countries can modernize their traditional census-based statistical systems into register-based systems. The final solutions of these problems will require many years of applied methodological work.

As a rule, we assume that an area sample has no coverage errors. For the case when we work with flow data, such as production etc., the area sample must also cover all time. In the Swedish Labor Force Survey, we ask "*what did you do last week?*". To cover all time with the LFS, a new sample is interviewed every week during the calendar year. Thus, we must cover all weeks during the year, quarter or month to be able to assume that we have no coverage errors in the area sample. In contrast, administrative registers cover all time; births and deaths are registered continuously and, for instance, taxation systems measure income night and day during the entire year.

Our aim with this paper is to describe the long-term strategy that should be used when developing countries start to use administrative registers for statistical purposes. When we discuss how administrative registers can be used for agricultural statistics, we must at the same time understand that non-agricultural statistics are becoming increasingly register-based. This will change the preconditions for agricultural and rural statistics. Therefore, the transition processes that will change both agricultural and non-agricultural statistics should be considered together. Agricultural statistics must be an integrated part of the national system and should not be discussed and developed separately.

2. What strategy and methods should developing countries adopt?

This section illustrates that starting to use administrative registers for statistics production implies many new ‘non-statistical’ activities that are necessary for success. We discuss here the modernization of the national statistical system from a traditional census-based system with no statistical registers into a register-based system. This will be a general discussion describing how administrative registers can be used to create an integrated system of statistical registers. The discussion is based on our experiences from visits to a number of countries in Latin America and the Caribbean during 2011–2016.

If the system for agricultural statistics is to be fully integrated with the system for social and economic statistics, the Farm Register should be an integrated part of the register system. The statistical system should describe statistics on crops and livestock as well as rural regions and households that are dependent on agricultural activities.

2.1 Microdata with identities

All administrative registers contain *identities*. The registered transactions include the identity of the administrative unit responsible for the transaction and/or the identity of the unit that is administrated. These identities are the most important variables in a register-based statistical system. With the identities we can check quality issues as for example coverage and we can combine records for the same unit from different sources. The identities link the different parts of the system together and it becomes possible to create an integrated system and produce consistent statistics.

We also need access to *microdata* to enable work with administrative registers in a methodologically efficient way. In many countries the national statistical institute, NSI, receives only aggregate administrative data from the ministries due to a desire to protect the confidentiality of citizens and enterprises. However, quality assurance and quality improvements will then become difficult and in many cases impossible; and it will not be possible to combine data from different sources so that relations between variables can be studied.

The conclusion is that the statisticians at the NSI must gain access to administrative microdata with identities of good quality. The goal is to have identity numbers so that deterministic record linkage can be used in all statistics production. In addition, those who work with agricultural statistics should have access to microdata with identities. These identities can link together different sources with agricultural data and link agricultural registers with registers in the general system.

2.2 Survey design includes work with improving administrative systems

Many countries in Latin America have reformed their national registration systems so that almost all newborn receive their identity numbers soon after the birth. The situation is worse in other countries: there can be one old system of identity numbers, a new system for adults, one system for children and another system for persons born abroad. In addition, some persons may not have registered and consequently have no identity numbers.

Administrative systems that do not generate data of good administrative quality cannot be used for statistical purposes. When the administrative system is improved, both administration and statistics production will be improved – this is a win-win situation. The national registration system with identity numbers, registration of births, deaths and migration is the most important administrative system for statistics production. If this system is not robust, there will be no identities that can be used to identify farmers, family farms and households; and register-based rural statistics will not be possible.

In many of the countries we visited, the NSIs are actively working to improve the national registration system and the cadaster. All births and deaths should be registered and migration within the country must be registered to make regional statistics possible. Emigration must also be measured; otherwise, there will be overcoverage in the population register. These improvements require political support and sometimes new legislation. This is high-level survey design – the preconditions for all statistical surveys in the country are improved.

We have seen an example where an old, very bad system was replaced by a new system for criminal statistics. In the old system microdata with identity numbers were created by many persons working for two ministries, the police and the correctional authorities. The information could be recorded through hand-writing or with different IT-systems, and different authorities used different codes for classification. Registering information in the old system was difficult and retrieving information was almost impossible. The new system developed by the NSI is based on one form and will standardize and simplify reporting. Retrieving information is now easy. The administration of crime reporting, victims and criminals has been improved and microdata for the production of crime statistics is now available.

2.3 Protection of confidentiality

When a country changes the national statistical system from a traditional into a register-based system, new methods must be developed to protect confidentiality. Identities can be handled and should be handled in such a way that confidentiality is protected *at least as well* in the new register-based system as in the traditional system.

One unit at the NSI, consisting of a small group of persons, should be responsible for receiving all microdata that is stored in an input database with very restricted access. At this unit, names and addresses are removed and official identity numbers are replaced by anonymous numbers. After this work has been completed, the anonymized data is stored in the output database and those who need access to data for statistics production can use it for their purposes. Record linkage is carried out with the anonymous numbers if information in different statistical registers must be combined.

Providing microdata with identities to the NSI will be something new for ministries and other authorities. In many cases old attitudes must be changed. If the ministries understand how the NSI works to protect confidentiality, it will be easier for the NSI to gain access to the data.

2.4 Centralization, cooperation and legislation

Centralization is not as important in a traditional statistical system as it is in a register-based system. In the production system in Chart 1, we want the register populations to be consistent. The populations in different registers must be compared when checking for consistency. When the registers are created, other registers must be used in a systematic way. This work is simplified if all registers are stored together at the same statistical institute. A natural and important role for the NSI will be to create and maintain the system of statistical registers.

Thus, the transition from a traditional system into a register-based system will actualize a more centralized national system. We recommend that the NSI be responsible for all registers that replace the population and housing census and all registers that are used for the National Accounts.

Politicians, ministries, other administrative authorities and the NSI must work together to develop register-based statistics. This cooperation is necessary to improve the administrative systems. Cooperation between the NSI and the authorities that deliver administrative registers will be a permanent task in a register-based system. Old legislation must be replaced with new laws that define that the NSI should have access to microdata and that protect confidentiality.

Regional register-based statistics require that that all persons' residences are georeferenced and that all job activities are georeferenced. In all Nordic countries it is mandatory that a person's place of residence (dwelling) is registered and that all employers report at which local unit (establishment) each employee works. These two laws were passed by each Nordic parliament to make the register-based population and housing census possible.

2.5 Quality assessment

Problems related to field work, sampling errors and nonresponse are our main concerns when working with sample surveys. In register surveys we have no fieldwork, no sampling errors and we do not ask any questions and we do not have any nonresponse. This probably explains why many people feel bewildered when they start working with administrative registers – how shall we now work with quality assessment?

The understanding of field work and how it determines quality issues in a sample survey corresponds to the understanding of the administrative system that has generated the administrative register. The staff must develop this new competence by studying the administrative system, by having discussions with persons at the administrative authority and by spending time analyzing microdata from the source.

Instead of computing sampling errors, we should analyze coverage problems and other inconsistencies. We do this by comparing coverage at the micro level in different sources describing similar populations. If we combine two registers with record linkage, the mismatch gives us a measure of the differences in coverage (if we have good identity numbers, mismatch is caused by coverage differences). If we combine sources with similar variables, we can also study inconsistencies between variables in different sources. When we find inconsistencies regarding coverage and/or variables, the next step is to try to understand why we have these differences. This is difficult but necessary. Finally, we should correct the errors we have found and reduce the inconsistencies.

There is one important difference between sample surveys and register surveys. In sample surveys we cannot see the errors, we know that we have sampling errors and nonresponse errors but can only give an interval for the probable sampling error. In register surveys, however, many errors can be seen if microdata is analyzed. Here are two examples:

- We found that 16% of the persons between 20–24 years had different residences in the University Register and the Population Register. This is an example of inconsistent variable values. The Population Register can be corrected and improved with this information from the University Register.
- We found that the turnover of *Manufacture of pharmaceutical preparations* was underestimated by 17% due to undercoverage in the SBS, the Structural Business Statistics Survey. This is an example of inconsistent populations. The SBS can be corrected with the information we used.

Each administrative register should be evaluated for statistical purposes prior to use. However, administrative registers are often evaluated one by one and this may lead to wrong conclusions. This is because an administrative register can be used in many ways – for one use the register may be good, for another use the register may be bad. In addition, registers can be combined. Each register may appear to be bad when analyzed separately, but the register based on the combination can still be good. Systematic check of input data quality is discussed in Wallgren and Wallgren (2014, p 273ff). The main purpose is to find out *how* an administrative register should be used. Every new source should be considered as a methodological challenge.

3. Creating a system for agricultural statistics based on registers

FAO (2015e) describes the situation regarding agricultural statistics in some African countries. Three survey methods are used today:

- An agricultural *census* of agricultural holdings is interviewed perhaps every tenth year. This gives detailed statistics for small areas but the information is not timely.
- An *area sample* of agricultural holdings is interviewed. This gives timely information but not for small areas.
- A specific form of statistical survey where statistical data is collected regarding *eye estimates*¹ made by farmers and/or *expert judgements* made by regional officers. This gives timely information for small areas, but the content is not as detailed as in the census.

According to Ssekiboobo (2015), *administrative registers “are rarely used in developing countries because of quality and coverage issues.”* This means that register surveys are not used in African countries today, but if we want to modernize national statistical systems as many countries in Latin America do, then administrative registers must become available and be used for statistics production. If the present administrative systems are not sufficiently good, then the first step in the modernization process is to improve the registration systems as discussed in Section 2.2.

¹ This kind of survey is called *administrative data*, both by the FAO (2015a-e) and by Ssekiboobo (2015). It is a kind of statistical survey that only can be used for agricultural statistics – everyone can look at fields and can judge if the crop looks good or bad. These judgements can be collected and used for statistics based on eye estimates. For other kinds of statistics, it is impossible to make eye estimates – e.g. labor market statistics require the established kinds of sample surveys or register surveys.

It is confusing to call these *eye estimate surveys* for administrative data. We must be careful with our statistical terms – when we work with developing national statistical systems from traditional systems into register-based we must use the term *administrative register* in the correct manner. We ask our colleagues working with agricultural statistics: *Please, do not to destroy this term!* In Wallgren and Wallgren (2014, Chapter 2, particularly on page 28) we discuss the nature of administrative data.

3.1 The Global Strategy and the register system

The Global Strategy by the World Bank (2011) states that agricultural statistics should be an integrated part of the national statistical system. This is an important aspect that is not discussed in FAO (2015a-e). For us, the term *integrated system* is a statistical term with a specific meaning: *Populations and variables should be consistent and estimates should be coherent*. Chart 1 (and also Chart 4) illustrates the production system with base registers, other statistical registers and sample surveys. The following quotations come from the Global Strategy (our italics):

The integration and underlying methodology ... considers the quality dimensions, which include relevance and completeness, accuracy, timeliness, accessibility, coherence, and comparability.

The process of improving agricultural statistics will begin with the integration of agriculture into the national statistical system. This integration will be accomplished by the development of a master sample frame for agriculture to ensure relevance and completeness; its use *in implementing a coordinated data collection program to produce timely and accurate data that are coherent and comparable*;

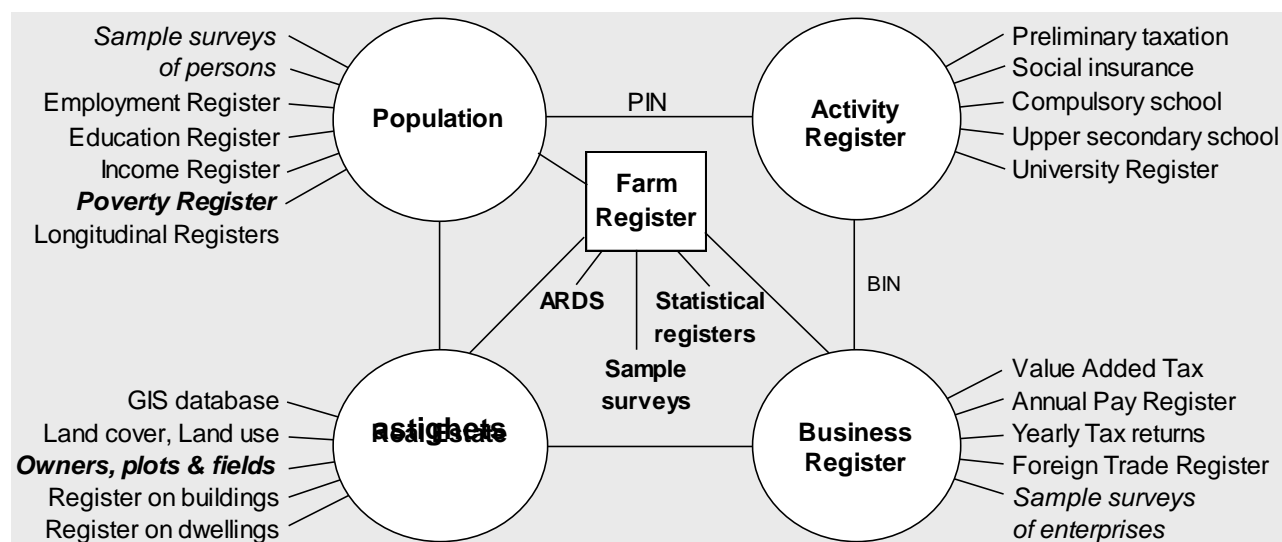
One of the shortcomings of current statistical systems in both industrialized and developing countries is that data are collected by sector, using different sampling frames and surveys. The division of data by sector leaves no opportunity to measure the impact of an action in one sector on another. In some cases, different organizations produce statistics for the same items, with different results, which confuse the data users.

Surveys are often conducted on an ad-hoc basis ... *It is therefore difficult to integrate data from various surveys for in-depth analysis with cross tabulation of variables*.

Integrated statistical systems can resolve many of these problems by avoiding duplications of effort, preventing the release of conflicting statistics, and ensuring the best use of resources. Concepts, definitions, and classifications become standardized, allowing more systematic data collection across sources.

The master sampling frame in the Global Strategy corresponds to the Farm Register in Chart 4. All sample surveys of holdings are based on the Farm Register and sources with crop areas, etc. are linked to the Farm Register. Thus, all data collection regarding agricultural variables is coordinated when the Farm Register is used in this way. As the Farm Register is linked to the Population and Business Registers, agricultural statistics can become an integrated and consistent part of the national system. Micro integration is possible, which makes “*in-depth analysis with cross tabulation of variables*” possible according to the quotation from the Global Strategy.

Chart 4. A register-based production system – the agricultural parts highlighted



The conclusion is: If you want an integrated statistical system that includes agricultural statistics, then the statistical method to achieve this goal is to create a system² of consistent registers according to the model in Chart 4.

² The eye estimate surveys denoted with ARDS (Agricultural Routine Data System) in Chart 4, can be included in this system if these surveys use data from agricultural holdings and contain identities of the holdings or holders.

3.2 Building the system step by step

The transition from a traditional statistical system based on censuses into a register-based system will take many years. The pioneer country was Denmark that managed to make this transition during 1968–1981. Austria managed to change into a completely register-based system during 2001–2011. The transition in Sweden was slower. The necessary legislative changes took many years, and the changes of the national registration of persons required much work with standardizing of addresses and the creation of a dwelling register. The Swedish transition took place between 1967 and 2011. This section provides an outline of the transition into a system for agricultural statistics that uses administrative registers.

The first step

In an integrated system, *identity numbers* play a very important role. The first step should therefore be to create a national registration system with good identity numbers. Newborns should be registered and receive their national registration number in connection with the birth. These personal identity numbers (PIN) should be used in all administrative systems. The identity numbers of holders can be used to identify holders or holdings in agricultural censuses and sample surveys and in administrative systems regarding agriculture.

The second step

The statistical Population Register is the most important part of a register-based statistical system. It is also important for agricultural statistics to have access to a Population Register with rural and agricultural households. The second step should therefore be to develop a *statistical Population Register*, which can be based on the administrative population register and supplemented by other sources to improve coverage and the quality of residential addresses. These addresses must have sufficient quality; otherwise, regional statistics will be impossible. Wallgren and Wallgren (2015) describes some of our experiences from Latin America.

The third step

The role of the base registers is to define statistical units and populations in the statistical system. The third step should therefore be to develop the Real Estate Register or Cadaster, the Business Register and the Farm Register. The Activity Register is also essential – all job activities are needed for the *Employment Register*, and all study activities are necessary for the *Education Register*. These two registers are important parts of the register-based population and housing census.

For register-based agricultural statistics, we need a Farm Register with all holdings and administrative registers with information regarding holdings. Within the European Union, three administrative sources can be used: the IACS-register, the Cattle Register and the Business register. In developing countries, a register can exist with big and/or commercial farms, but there is no register with small family farms. An inventory of sources with microdata regarding farms or farmers should be carried out. Cooperation between ministries can be important – other ministries than the Ministry of Agriculture may have data that includes farming households.

We can give one example from Latin America: The Ministry of Interior created a register to administrate support to poor households. The register covers between 60% – 95% of the population, with the highest coverage in poor rural regions. Here we suggest *integrated data collection* so that the Ministry of Interior will include questions when creating their register to identify agricultural households. This information can then be used by the Ministry of Agriculture and as a source for a Farm Register. Today, the Farm Census is often based on a frame based on the agricultural

households in the Population and Housing Census. In the future, this integrated data collection can replace this way of using the Population and Housing Census.

Agricultural statistics as an integrated part of the national system

When a country has established a system with national identity numbers for persons (PIN) and if there is a population register with information on where people live and together with whom, then the preconditions for statistics production have been greatly improved. It will now be possible to create statistical registers and produce register-based statistics on persons.

All administrative systems concerning persons and households should use identity numbers PIN. Then data from these systems can be used for statistics production. Also all surveys as population censuses, all area frame based sample surveys of households, all register-based (list based) sample surveys of persons and all register surveys³ with data on persons should include the identity numbers PIN. Then these surveys can be developed into an integrated statistical system that produces coherent statistics.

If we want to make agricultural statistics an integrated part of the national system, there must be a national system with registers that we can link to the agricultural registers. We assume therefore that the statistical Population Register exists and that family farms can be linked to this register with the identity numbers PIN of the holders. Commercial farms can be linked to the Business Register with business identity numbers BIN. If there exists a Real Estate Register, then the holders and holdings can be linked to the fields and plots they own via the identity numbers PIN and BIN. A Farm Register should be developed that can be linked to the registers in the national system. This is the necessary precondition that must be fulfilled if we want that agricultural statistics based on registers should be integrated with the national statistical system.

3.3 Agricultural surveys in an integrated system

In Chart 5 we give a short description of different sources with identities and agricultural data and how they can be used to create an integrated system of agricultural surveys. All relevant sources should be used to create a statistical Farm Register with as good coverage as possible. The Farm Register should also include basic classification variables as address, size and kind of farm.

Chart 5. Sources used to create a Farm Register and agricultural statistical registers

Source	Identity number	Survey
1. Population census with identity numbers PIN and a variable that indicates holders	PIN	Statistical Farm Register
2. Register to administrate support to poor households with identity numbers PIN and a variable that indicates holders	PIN	Statistical Farm Register
3. Agricultural Routine Data System, field staff is used to improve and update the Farm Register	PIN, BIN	Statistical Farm Register
4. Administrative Farm Register with commercial farms	BIN	Statistical Farm Register
5. Agricultural Routine Data System, field staff is used to report holding identities together with holding level data	PIN, BIN	Statistical Registers with crop areas, production etc.
6. Administrative registers with different content: Tax data for holders and holdings Land ownership Agricultural production and inputs Data from manufacturers and distributors Data from farm associations	PIN, BIN	Different statistical registers Can also be used to improve the Statistical Farm Register

³ A *register survey* is a survey where data from existing statistical registers and administrative registers are combined to create a new statistical register for specific statistical purposes.

Population census

If a country does a traditional population and housing census (perhaps for the last time), this census should be used not only to improve the Population Register but also to improve the Farm Register. Identity numbers, PIN, of all persons should be included in the census and also questions that identify holders and holding households. This information together with the residential address can be used to establish a first version of the family farm part of the Farm Register.

Register to administrate support to poor households

In Latin American countries, many ministries create their own registers with population data to administrate their different political programs. In one country we found six different registers with population data for different purposes. If some of these ministries include questions to find holders and holding households, then this kind data can be used to create the family farm part of a Farm Register.

Agricultural Routine Data System

FAO (2015e) describes the Agricultural Routine Data Systems, ARDS, that are common in many developing countries. There are many problems with these systems: the staff is not always motivated, there is lack of training and the reports are not standardized etc. However, the ARDS is an old and well established kind of statistical system that has field staff all over the rural parts of the country. We think that the ARDS gives opportunities that can be used in a register-based statistical system. But to become an integrated part of the new statistical system, the ARDS should include identities of the holders and collect microdata from holdings. Today the ARDS staff aggregates data from holders into village data. These aggregated data are then transferred to wards and regions etc. We suggest that gradually the ARDS changes from collecting the aggregated data that is used today into collection of microdata with identities of holders. This microdata can then be used to create registers with agricultural data.

The field staff working with the ARDS can also report changes in the stock of holdings. In this way the quality of the Farm register will be maintained. The knowledge that the field staff has regarding the farm population should be used to improve and update the Farm Register.

Administrative Farm Register with commercial farms

A register with commercial farms should be one of the sources that should be used to create the statistical Farm Register. There are as a rule also administrative sources with tax data and other economic information that are important for agricultural statistics. These sources can be combined with business identity numbers BIN for enterprises. One BIN identity can however be linked with more than one holding.

Administrative registers with different content

When the statistical Farm Register has been established, available administrative registers can be used to create new statistical registers with agricultural and rural information. The populations should be created with the Farm Register and the administrative registers will be used as sources of the variables in these new registers.

When the administrative system that generates the administrative data is old and bad, the data will have poor quality. The NSI should in such cases take the initiative so that the bad system is replaced with a modernized system. This requires cooperation with administrative authorities as we explained in section 2.2.

4. Conclusions

Administrative data and agricultural statistics – what strategy and methods should we adopt? The following two quotations come from Wallgren and Wallgren (2015) and are based on our experiences from Latin America.

“Administrative systems in the public sector generate vast amounts of data that is stored in administrative registers. These registers is a resource that can be used for statistics production. This resource is growing in volume and as the administrative systems gradually improve, the quality of the administrative data that is generated by these systems will become increasingly better.”

“If administrative registers are used for statistics production, costs can be saved and quality can be improved. Censuses can be replaced by register surveys, and new kinds of regional and longitudinal statistics can be developed. The national system can be changed into an integrated statistical system with consistent and coherent statistical registers and sample surveys. During the transition period, it may be necessary to combine the register-based statistics with estimates from area samples.”

According to the Global Strategy (World Bank 2011), agricultural and rural statistics in developing countries must be improved and become an integrated part of the national statistical system. When a country starts to use administrative registers for statistics production, more and more statistical registers are developed. With identity numbers, all registers and surveys can be linked and become parts of an integrated system that can be used to produce coherent statistics.

In this paper, we have tried to give an overview of the strategy and methods that should be used in this transition process. Many countries in Latin America are working in this direction and we think that within five years some of these countries will have improved their national systems considerably.

REFERENCES

- FAO (2015a) Technical Report 1: Reviewing the Relevant Literature and Studies on the Quality and Use of Administrative Sources for Agricultural Data. Food and Agriculture Organization of the United Nations.
- FAO (2015b) Technical Report 2: The Role of Administrative Data in Developed Countries: Experiences and Ongoing Research. Food and Agriculture Organization of the United Nations.
- FAO (2015c) Technical Report 3: Critical Analysis of Agricultural Administrative Sources Being Currently Used by Developing Countries. Food and Agriculture Organization of the United Nations.
- FAO (2015d) Technical Report 4: Analysis of Agricultural Administrative Data Gaps and Ways of Improving the Quality and Use of Administrative Data Sources for Agricultural Statistics. Food and Agriculture Organization of the United Nations.
- FAO (2015e) Technical Report 5: Strategy and Methodology for Improving the Use of Administrative Data – A Protocol for In-Country Testing. Food and Agriculture Organization of the United Nations.
- Selander, R., Svensson, J., Wallgren, A. and Wallgren, B. (1998) *Administrative Registers in an Efficient Statistical System – New Possibilities for Agricultural Statistics? How Should We Use IACS Data?* Statistics Sweden, Eurostat.

- Ssekiboobo, A. (2015) Methodology for Improving Administrative Data for Use in an Integrated Agricultural Statistics System. ISI 2015, Rio de Janeiro.
- Wallgren, A. and Wallgren, B. (2010) Using administrative registers for agricultural statistics, in R. Cenedetti, M. Bee, G. Espa and F. Piersimoni (eds), *Agricultural Survey Methods*, pp. 27–44. Chichester: John Wiley & Sons, Ltd.
- Wallgren, A., Wallgren, B. (2014) *Register-based Statistics – Statistical Methods for Administrative Data*. Second edition, John Wiley & Sons, Ltd, Chichester, UK.
- Wallgren, A., Wallgren, B. (2015) Towards an integrated statistical system based on registers. Paper presented at the Inter-American Bank’s conference *Driving the Data Revolution: Statistical Use of Administrative Registers*, Washington 16-18 September 2015.
- Wallgren, A., Wallgren, B. (2016a) Frames and Populations in a Register-based National Statistical system. *Journal of Mathematics and Statistical Science*, 2016, pp. 208–216.
- Wallgren, A., Wallgren, B. (2016b) Estimating frame coverage issues in agricultural surveys. GS Training and Outreach Workshop on Agricultural Surveys, FAO October 2016.
- World Bank (2011) *Global Strategy to Improve Agricultural and Rural Statistics*, the World Bank, FAO and the United Nations.