

The Thin Line Between Statistical and Public Information in Scandinavia

Olav Ljones
Statistics Norway
POB 8131 Dep
N-0033 Oslo Norway
olj@ssb.no

Otto Andersen
Statistics Denmark
Sejrogade 11
DK 2100 Kobenhavn O
oan@dst.dk

Jon Epland
Statistics Norway
POB 8131 Dep
N-0033 Oslo Norway
jep@ssb.no

1. Introduction

This paper intends to discuss two concerns that may be in conflict with each other. One aspect is the human right for privacy and integrity. Even though privacy is a common value, it seems to be some differences of opinions in what kind of information that is valued to be most sensitive. The other perspective is that individuals, corporations and government need information and demand transparency. Rational decisions both by individuals and the public should be based upon solid information. There are however examples of both positive use and misuse of information.

The National Statistical Institutes and Official Statistics can play an active role by submitting official statistics as information for improved understanding of the society. Official statistics should - if it follows the basic principles for official statistics - at the same time give transparency and protect privacy.

The recent evolution of modern IT techniques has in a rather dramatic way changed the setting for privacy, data protection and transparency. The possibilities for those who want to use modern IT techniques to offend integrity and confidentiality have increased significantly. Electronic traces are left more or less all over. Enormous databases with potentials for both positive use and misuse are one result. For official statistics it is important to capture data from all available sources. Official statistics are dependent on broad public confidence, in order to get access to confidential data.

2. Some challenges

To balance public information in an optimal mixture of transparency and confidentiality seems difficult. The importance of integrity and confidentiality should be accepted as a basic need and an essential part of human rights. One of the results of the IT revolution is that the media by which the information is distributed are crucial for how fast the information is spread and accordingly to whom the information in practice is accessible to. This wide spread of information will have a negative effect on privacy.

Confidentiality is not only an emotional question. A totalitarian regime or others may and will misuse private information for economic or political discrimination, if the relevant information is available and relevant.

Transparency between individuals, institutions and government may also - if it is not misused - be a value by itself. Society can give civil rights (voting, legal protection, social insurance, health care etc) based on identification. For businesses it is important that potential investors, customers etc. have access to information about products, prices etc. The land register, the motor vehicle register etc, even patent registers may be seen, as examples of the common need for transparency.

Some characteristics are regarded more sensitive than others. In Norway handling of personal data has to be done in accordance with the law on data-protection. Strict rules have to be followed for data sets that include sensitive data. Sensitive data are (legally) defined as: a) Racial and ethnic origin, or political opinions, philosophical or religious belief, b) The fact that a person has been suspected of, charged with, indicted for or convicted of a criminal act, c) Health, d) Sex-life, e) Trade-union membership.

Normally the sensitivity of data will increase when different sets of data are merged, even when each variable by itself is non-sensitive. It seems also that what type of information that is regarded as sensitive, varies between individuals and societies. For the business sector a lot of variables that are collected, can be sensitive. For businesses the sensitivity of some information is very dependent on the actuality. Some data and statistics may provide "inside information" that are of value for those who are active in trading stocks or currency.

Data are collected from private sector by the government for various purposes. In many countries there are a common political concern to reduce the overall response burden that government put on businesses. The reuse of information is one instrument that can reduce the response burden. By opening for exchange of information between different parts of public administration, a government can help to reduce the burden for the respondents.

In a NSI or other administrative bodies, a lot of data from different sources are collected and stored. To protect data from misuse and guarantee confidentiality the NSIs and others have to follow strict rules and regulations. There is however always to consider some probability for an accident or some criminal or other planned breaking of security rules, with data leakage as the result.

In the fight against crime, the use of merged data can be an effective tool. But the question is - can we accept that there are no limitations in this direction, without breaking essential human right principles? It seems essential to make a distinction between the rights for the NSIs and other bodies as regards the merging of administrative data. If the NSI follows the basic principles merged data should never be given away for other purposes than statistics. In some cases the NSIs have been approached by investigators to help with certain kind of information (e.g. illegal foreign trade with countries that are under boycott). The NSIs have, however, (in the examples from Denmark and Norway) got legal support for a policy to never give information for such purposes.

In most countries there are - especially from the research community - a strong demand for easy access to micro data. To serve the demand from the research community for access to micro-data, data may be anonymized; the researchers can work within the data-network of the NSI for a limited time period, they have to sign a pledge of secrecy etc.

When official statistics are used directly or indirectly for welfare policy, different groups of the population can gain or loose from changes in welfare policies. There are potential dilemmas for a NSI in this aspect. The normal situation is that the NSI will claim to be politically independent, without responsibilities for the welfare effect of a change in policy based on official statistics. The ideal is for a NSI to use best practice methods, but within this concept, it may in many cases be arbitrary choices. A situation where the NSIs are confronted with the political consequences of such "arbitrary" choices can be foreseen. Another problem is linked to a situation when official statistics can be used to identify the welfare situation for particular groups of the population, for example described by race or ethnic background. In Norway there is no strong recent tradition for the use of ethnic groups as a class-variable. After some years with relatively high immigration (from countries outside Scandinavia), there is however a request, especially from right-wing political parties, for more statistics to show the "costs" of immigration - the concept of immigration account has been

used. We believe that most NSIs will feel close to having an ethical dilemma in such cases. This means that some variables/characteristics may be sensitive, even when they are from an individual point of view treated with full confidentiality.

3. Experiences

In Norway a collection of administrative registers (Bronnoysund Registers) give public access to information. These registers cover many parts of economic and business life, moveables, enterprises, annual reports and accounts, bankruptcy, marriage settlements, environment and eco-management, public questionnaires and response burden etc. The purpose is to serve both those who register and those who seek information from the register. Those who seek information have to pay a fee.

In Norway there is also open access to some main income and tax items for each individual. The law says that when taxation is finished the tax office (a local office) shall make public a protocol of all that have been taxed. The argument was that this was necessary to give the individual a possibility to do his or her own checking of the tax assessment made by the tax office. Still the list is in general only available on paper. For the press there is however a special access to a CD ROM version.

The NSIs are by the Statistics Act given the rights to either use data from administrative sources included administrative registers, for statistical purposes. Strict rules for confidentiality have to be followed. Data can be merged from different administrative sources. The other main method is to collect data directly from individuals (persons/household or firms).

There is a common concern about the total response burden that is put on private corporations by the government. As a policy measure (in Norway) it is established a central register where all government questionnaires have to register. With the intention to reduce the response burden, information that are collected for administrative purposes should only be collected once and accordingly government bodies are obliged to exchange information - with one important exception; Statistics Norway shall not give individual information that are collected by them to others.

To guarantee privacy and integrity for individuals in several countries there is a special law on data protection and personal registers. There was in 1980, furthermore, established a special administrative body called The Data Inspectorate. The purpose of the act and the Inspectorate is to safeguard the protection of individual and social privacy. Personal data shall not be registered without proper justification. Sensitive information should not be registered unless it is necessitated by the purpose of the register.

There is a heavy demand for access to micro-data from researchers. The NSIs seems to try to meet this demand by submitting anonymized data and by putting different legal restrictions on the use. In Norway there is a long tradition for a co-operation between the NSI and a Data Archive owned by the Norwegian Research Foundation. In Denmark the basic principle has been that researchers are given the opportunity to use micro-data from a special service-office within the NSI

4. Public debate and Public opinion.

The general impression is that there has been no severe or major break of confidentiality from the NSIs. The positive effect from the activities by bodies such as the Data Inspectorate should not be neglected. There seems also to be an ongoing debate about the response burden and bureaucracy. In Norway people until recently had yearly to fill in their tax return form. As a result of the wide access the tax offices have to the income receipts for each wage earner and to bank and insurance accounts for each individual, the tax return form is now pre-filled with all or most items by the tax authorities. This is for most people a very practical improvement, but it also is a proof of to what detail the tax authorities have access to information. Since there is an obvious gain to the public from this pre-filled tax report, it seems like the new practice has not created critics with reference to data protection and integrity.

There is in Norway a debate about the practice that some main items showing income, assets and tax-paid are made public for all taxpayers. This practice has been under debate in the parliament in 1997 and 2000, but both times the majority voted in favour of the present practice. Those who belong to the minority argue that such open lists (registers) are an offence of data protection etc. The practice can also indirectly be of help for those who want to abuse privacy by re-identifying anonymized statistical micro files. Those who are in favour of the practice, argue that such openness is crucial for the political acceptance of the tax system and it is important to have this detailed list to identify weak points in the tax rules. The press argues that this practice is important for the freedom of the press.

In 1997 Statistics Norway did a statistical enquiry about what kind of information people consider to be sensitive. 57 % feel that information about health is very sensitive while only 15 % feel that information about political opinions is very sensitive. Information about own economic situation is felt to be very sensitive by 35 % and ability to pay (solvency) is felt to be very sensitive by 39 %. The proportion that regards information as very sensitive seems to increase for most type of information with the level of education of the respondent.

5. Concluding remarks

It is a common understanding that it is important to balance confidentiality and transparency. Confidentiality is important as a human right and as a measure to protect individuals against crimes etc. Transparency can also help to secure human rights and to give important information that is a prerequisite for a free and well functioning market economy and democracy. It is important to have an open debate about all parts of the structure of the common information systems, both the administrative part and the statistical part. One should also bear in mind that it is a possibility for misuse and criminal attacks towards the systems. "Firewalls" of all kinds have to be built and maintained. It is also necessary to know the history of misuse of data, performed by politicians, bureaucrats and regimes.

Official statistics can at the same time give information and create transparency without offending the individual right for privacy. The rights that the NSIs have to use administrative data for official statistics seem not to be an important threat to data-protection and misuse of data. One effect of the use of administrative data by the NSIs is the reduced response burden.

The NSIs seem to regard their role as public bodies that are neutral guardians and suppliers of information to citizens and corporations. To protect this role by managing both privacy and transparency is a duty.