



A report on Swedish equality data

The role of statistics in fighting discrimination: the need for strategies and credibility.

English executive summary

Introduction

In December 2011 the Swedish government commissioned a preliminary study concerning the development of national equality data. The task to conduct the study was given to the Swedish Equality Ombudsman (EO). The study was to examine those methods that might be suitable for collecting data in relation to five of the seven¹ discrimination grounds in the Discrimination Act (age and sex were excluded from the study) as well as in relation to the five recognized national minorities.

The government articulated two main reasons for such a study. The first reason is previous international criticism concerning the lack of disaggregated data that could shed light on the living conditions of different protected groups. The second reason is that the lack of disaggregated data according to different discrimination grounds or belonging to a recognized national minority constitutes an obstacle to formulating and following up the national equality policy.

The main objective of the study is to examine if it would be possible to use the Survey on Living Conditions (ULF/SILC) for the production of equality data, thus including in the Survey new variables that would correspond to five discrimination grounds as well as to belonging to a national minority and use of a national minority language.

Terms of reference

According to the terms of reference, the government needed a preliminary study that would show what kind of equality data it would be possible to produce in a large-scale national survey.

¹ The seven discrimination grounds in the Discrimination Act are: sex, transgender identity or expression, ethnic origin, religion or other belief, disability, sexual orientation and age. Age and sex were excluded from the study as they are already included in the Official Statistics of Sweden.

During the study, the EO was expected to consult Statistics Sweden and the Swedish Data Inspection Board as well as organizations from civil society that may be assumed to represent different protected groups.

Covered variables

Discrimination grounds:

- ethnic origin,
- religion or other belief,
- disability,
- sexual orientation
- transgender identity or expression.

Recognized national minorities and national minority languages:

- Jews: language – Yiddish
- Roma: language - Romany Chib (all varieties),
- Sami (also as indigenous people): language - Sami (all varieties)
- Swedish Finns: language - Finnish
- Tornedalers: language - Meänkieli (Tornedal Finnish).

The Survey on Living Conditions (ULF/SILC)

As Sweden does not carry out a population census, the official statistics of Sweden are based on administrative registers. Statistics Sweden, which is the national statistics agency, conducts an official survey on living conditions on a yearly basis - “Survey on Living Conditions” (abbreviated as ULF in Swedish). During recent years, parts of the Survey were integrated with the European Union Statistics on Income and Living Conditions (EU-SILC). This Survey is often referred to as ULF/SILC.

The Survey is sample-based, it is conducted every year through telephone interviews. The Survey covers the following areas: housing, income, health, leisure, civic activities, social relationships, employment and security.

Currently, ULF/SILC includes the following variables that are of relevance for the equality data study: age, gender and nationality. Nationality is a complex variable that includes the place of birth of a person combined with the place of birth of the person’s parents.

The report

This 160-page report was published electronically on the website of the Ombudsman² in November 2012 and printed in December 2012. The report

² <http://www.do.se/sv/Material/Statistikens-roll-i-arbetet-mot-diskriminering/>

is currently available only in Swedish. This document is an English executive summary of the report.

Conclusions

High policy ambitions require a broad knowledge-base

The Swedish national policy on anti-discrimination and the policy on national minorities have a broad approach and articulate an ambition to achieve equality in rights, equality in opportunities and equality in outcome. The report introduces a new terminology in Swedish: “jämlikhetspolitik” for “equality policy”, “jämlikhetsdata” for equality data and “jämlikhetsstatistik” for equality statistics. The word “jämlikhet” (Swe. “equality”) in the report does not cover equality in terms of social class which is the more traditional use of the term in Swedish. The concept concerns equality in rights, opportunities and outcomes to the extent that it relates to anti-discrimination and in relation to the grounds in the Discrimination Act and to the recognized national minorities.

A first conclusion is that following-up an ambitious national equality policy requires the development of a comprehensive knowledge-base of equality data. Relevant literature in the field defines four main sources for equality data – official statistics, research, complaints data and workplace and service delivery monitoring. Different data sources deliver different kinds of knowledge and highlight different aspects of discrimination and inequality. These sources must be seen as complementary to each other. It is therefore not possible to replace statistical data with data from non-statistical data sources and vice versa.

The EO made an attempt to identify the use of the above mentioned data sources in order to place equality statistics in a broader perspective. The attempt indicated that equality statistics are available only in relation to equality between women and men. Thus, it was possible to identify a gap in the national equality knowledge-base – the lack of equality statistics in relation to the other discrimination grounds as well as the recognized national minorities. The lack of equality statistics confirms what was stated by the government in the terms of reference, namely that there are shortcomings in the knowledge-base that underlies the formulation and the follow-up of the equality policy. The EO concluded that production of equality statistics would be an important and necessary measure that is needed to complete the national knowledge-base.

Legal requirements for processing sensitive personal data

A first conclusion is that processing sensitive personal data for the purpose of combating discrimination is not forbidden in the international legal instruments concerned with data protection, such as the EU Data Protection Directive (Directive 95/46/E) and the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS 108). The relevant international instruments only provide a legal framework

for processing personal data and formulate a number of general principles that need to be respected in order for the processing to be legal.

From a Swedish perspective, it is the Personal Data Act (1998:204) that regulates the processing of (sensitive) personal data. According to Section 13 of the Act, the processing of sensitive personal data is prohibited. This term refers to data that reveal race, ethnic origin, political opinions, religious or philosophical beliefs, membership in a trade union or health and sex life. There are a number of exceptions to this prohibition in Sections 14-19 of the Act. At the same time, the Personal Data Act is also subsidiary. The Official Statistics of Sweden are regulated mainly through the Official Statistics Act (2001:99) and the Official Statistics Ordinance (2001:100). For instance, the processing of data on health – which is classified as sensitive - in ULF/SILC is explicitly regulated through the Official Statistics Ordinance.

Currently, ULF/SILC includes two variables that have relevance for the equality policy – age and sex. For new variables to be introduced in the ULF/SILC, the regulations on the Official Statistics of Sweden need to be amended since these variables fall within the scope of the legal definition of “sensitive personal data”. In case other methods than ULF/SILC are used for processing such kinds of data, it is necessary to define the purpose, the scope and the collection method before an assessment of the necessary legal amendments can be made. In this respect, Section 20 of the Personal Data Act allows the government to introduce further exceptions from the prohibition on processing sensitive personal data.

Challenges to introducing new variables in ULF/SILC

The technical challenges to developing ULF/SILC as a source for equality statistics concern the formulation of the questions and the alternative pre-defined answers that would be added to the Survey. In this respect, it is crucial that a consistency in relation to the notions of definitions, classification and categorizations is fully respected.

The importance of respecting consistency is underlined by the fundamental differences between the variables that are subject to the study – the five discrimination grounds and the belonging to a recognized national minority and use of a national minority language. The variables under discussion are not always comparable and concern heterogeneous parts of the population. The legal definitions, when available, are constructed in different ways.

EO's conclusion is that it is not possible to use the same method for producing equality statistics in relation to all the protected groups that are covered by the study. Certain protected groups are too small in relation to total population and would probably not generate data in a sample-based survey, as in the case of the Roma community. Other parts of the population express a lack of confidence in the state and would tend not to participate in a survey when asked about their ethnic origin or religious affiliation, as in the case of the Jewish community. For these parts of the population there is a need to adopt alternative data collection methods.

The main issue from a technical point of view is designing alternative data collection methods that would both allow longitudinal comparability and comparability with the results from ULF/SILC. These aspects of comparability are crucial for the possibility to follow-up the equality policy measures.

Those parts of the population that feel a lack of confidence in the state need to feel that they have control over the data collected and processed. A hypothetical possibility is to use the same questionnaire as in the ULF/SILC, by allowing a researcher that the group has confidence in to head the collection and handling of the data. A number of members of the group could be specially trained to handle the collection of the data. The above mentioned approach is merely an illustration of how the design issue needs to be approached. At the end of the day, the alternative data collection methods need to be designed in consultation with those parts of the population that are concerned by these collection methods.

For an assessment of different methodological approaches there is a need to involve both researchers and statisticians. The issues of representativeness and sampling will be crucial even when using the same questionnaire in respect to different data collection methods. Sampling problems have been highlighted in both national and international studies. Past experiences indicate that it is not possible to construct identical and ultimate samples for the alternative data collection methods. However, consultations with the concerned parts of the population as well as with researchers and statisticians would make it possible to reach reasonable compromises.

Sampling issues are crucial for the usefulness of data. Statistical data with certain shortcomings in their representativeness can still be useful for formulating and following-up concrete policy measures as they can be used as indicators for identifying the areas that need to be targeted in terms of further studies or specific measures. However, data with shortcomings in representativeness may not always be useful in strategic litigation in the anti-discrimination area as their validity can be questioned.

Still, it must be stressed that alternative data collection methods should not be replaced by qualitative studies, even if they won't produce fully comparable outputs. Substituting statistical data with non-statistical data would maintain the lack of:

- tools to evaluate and follow-up equality policy measures,
- tools to track changes in the living conditions of certain groups longitudinally,
- statistical data that may be used in strategic litigation in the anti-discrimination area.

The point is, no matter what compromises may be reached through consultations with the involved protected groups, it is crucial that all parties involved are aware of the consequences of the compromise. In case a given protected group completely opposes the production of statistical data concerning the group, this group should be thoroughly informed of the

consequences of such a choice. Of course, this does not mean that the state's obligations towards such a group would diminish. But the state's possibilities to live up to its obligations towards that group will be limited. There is no doubt that every group's free will must be respected – no one can be or should be forced to declare their sensitive personal data. But everyone has also the right to be correctly informed on the state's possibilities to evaluate and follow-up policy measures with and without statistical data. This is crucial for reaching compromises on the future policy measures concerning every involved group.

Recommendations for a large-scale study

Currently available data

According to Statistics Sweden, ULF/SILC includes currently following data of relevance:

- Data on health that is relevant for the discrimination ground disability, namely: sight impairments, hearing impairments, moving impairments, (severe) mental stress, low physical activity mainly due to health problems.
- Data on "nationality" i.e. data on the country of birth combined with data on the country of birth of the parents.

One conclusion is that it is not possible to identify all data available in the national statistical system, it takes a special mapping study to be able to do that. A tentative conclusion, is that it is not unfair to say that the statistical system is only used for disaggregated data on sex and does not deliver disaggregated data on the variables covered by this study. A proper mapping study to identify all available data in the statistical system may be a first suitable measure before conducting a large-scale study on living conditions.

Data collection methods

- No single data collection method is suitable for all the variables covered by the study. The choice of method depends on two main factors: the type of the variable involved and the sensitivity of the involved population towards collecting and processing personal data. It is therefore recommended that a choice of method is made in collaboration with the involved groups and that the consistency in terms of definitions, classifications and categorizations is respected.
- ULF/SILC may be a suitable data collection method for the following discriminations grounds: ethnic origin, religion or other belief, disability and sexual orientation. It may also be suitable for the following two national minorities: the Swedish Finns and the Sami.
- ULF/SILC is not suitable for the discrimination ground transgender identity or expression as well as for the following national minorities: Jews, Roma and the Tornedalers. Equality statistics on these groups need to be produced through alternative data collection methods.
- Alternative data collection methods need to make it possible to compare data from other population groups as well as data with the general population. Which alternative data collection method is

suitable for a given population group needs to be designed after consultation with the group in question as well as involving researchers and statisticians.

Guaranteeing anonymity and non-violation of personal integrity is a complex issue. To guarantee anonymity, data needs to be made anonymous in a reliable way. To guarantee non-violation of personal integrity, data needs to be collected on the principle of unconditioned self-identification. The notion of safeguards in this respect has three aspects – a technical aspect, a legal aspect as well as a credibility aspect. Formal safeguards in terms of a secured technical solution and a well-formulated legal framework are credible only to the extent that the state's ambitions concerning the data collection are credible in the eyes of the general population and the protected groups. It is therefore recommended to involve the concerned protected groups in a consultation to ensure the credibility aspect prior to a large-scale study.

Legal requirements

- International legal instruments concerning data protection do not prohibit processing sensitive personal data for the purpose of combating discrimination.
- Which specific legal requirements need to be fulfilled depend on the method used to collect data as well as the scope and purpose of the data processing.
- For ULF/SILC, the legal framework regulating Official Statistics of Sweden needs to be amended.
- Defining a legal framework allowing the processing of sensitive personal data in relation to alternative data collection methods is only possible when the methods are identified and the purpose and the scope of the processing is defined. In this respect, Section 20 of the Personal Data Act allows the government to introduce further exceptions from the prohibition on processing sensitive personal data. There is a need to consider how the output of alternative data collection methods can be incorporated in the Official Statistics of Sweden.

Further recommendations for a large-scale study

- There is a need to develop a comprehensive national strategy on equality data and equality statistics for the sake of formulating and following-up the equality policy. It is recommended that the government adopts a cross-sectoral approach in working out such a strategy, involving equality groups, researchers and statisticians. The Finnish experience in adopting such an approach may provide some important lessons. When ULF/SILC is involved, some lessons might also be learned from the Irish experiences.
- A large-scale study involving sensitive personal data would imply a paradigm shift in the Swedish state's position on statistics. A number of the variables covered by the study concern the whole population. It is recommended that the government should as a first step, introduce

the concepts of equality data and equality statistics in the public debate to give the general public a possibility to understand and accept the need to collect and process sensitive personal data.

- The variables covered by the study have different characteristics. Different population groups concerned by these variables express different levels of sensitivity towards the issues of collecting and processing sensitive data. Thus, there will be a need to use different data collection methods in relation to different variables and groups. The cost-effectiveness of a large-scale study depends on the comparability of the results from different data collections methods. It is therefore recommended that a large-scale study is initially discussed with the concerned protected groups in order to establish the foundation for a common ground. EO's experience in consulting with different protected groups during this study indicates that it is more valuable to conduct a consultation with all of the concerned groups "under the same roof" rather than with each group separately.
- Different protected groups have different positions on equality statistics as well as different expectations. The government should be ready to handle a number of situations that, in this respect, we chose to refer to as paradoxes. One such paradox is that among the groups that are skeptical to processing personal data, one would find those who would probably benefit the most from a tool for following-up the equality policy measures. Another such paradox relates to the data on ethnic origin and data on language. The Sweden Finnish Delegation expressed an interest in collecting data on language in the first place and regarded data on ethnic origin as complementary to data on language. On the other hand, The Cooperation Group for Ethnic Associations in Sweden (SIOS) has the opposite position, expecting data on mother tongue for ethnic minorities (even officially non-recognized minorities) to be complementary to data on ethnic origin. At the same time, both the Swedish Finns and the Sami representatives made it clear that they expect to see the explicit categories "Swedish Finns" and "Sami" if data on ethnic origin is to be collected. These paradoxes underpin the need to consult with the protected groups "under the same roof" rather than with each group separately.
- ULF/SILC is a powerful and cost-effective tool for the production of statistics. However, Statistics Sweden has several times pointed out that the response rate has been declining in recent years. If the general public does not get a clear picture of the rationale for introducing several sensitive variables in the Survey, there is a risk that the response rate might come down to even lower levels. This would have negative consequences on the quality of the statistics produced by ULF/SILC. In this respect, it should be recalled that the consulted protected groups have emphasized the importance of credibility when collecting and processing sensitive data. However it is not the credibility of Statistics Sweden that was questioned but the credibility of the state (or government) in general terms. On the contrary, the

consulted protected groups made it clear that Statistics Sweden enjoys a very good reputation. There is therefore a risk that the reputation of Statistics Sweden might get damaged, unless the government ensures that the concepts are introduced into the public debate in a manner that wins the support of the general public, before a large-scale study is introduced. It is recommended that the government should move forward carefully and start by introducing those variables that concern those protected groups that are positive to the use and development of equality data.

- Collecting data on living conditions is mostly meaningful when it is done on a longitudinal basis. Equality statistics are not an end in themselves but rather a means to reach other goals. Differences in living conditions between different groups that can be highlighted by surveys of the ULF/SILC-type may depend on a number of factors, discrimination is one of them. Data from ULF/SILC needs to be complemented by a number of qualitative studies that can help understanding which differences are relevant and to what extent the differences relate to discrimination. It is recommended that the government at an early stage of a large-scale study ensure that such studies are adequately funded. It is only when the results of ULF/SILC are further analyzed in terms of discriminatory factors that an equality policy may be formulated and followed-up in a better way.

Further information on the report and citation

Contact for further information on the report:

Research Officer Yamam Al-Zubaidi (yamam.al-zubaidi@do.se).

Suggested citation:

Al-Zubaidi, Yamam (2012). Statistikens roll i arbetet mot diskriminering: en fråga om strategi och trovärdighet. [” The role of statistics in fighting discrimination: the need for strategies and credibility”]. The Swedish Equality Ombudsman: Stockholm.