

# The experience of a National Statistical Institute after a law change: Estonia

First Regional Workshop

Microdata Access in European Countries — Cooperation  
Between National Statistical Institutes and Social  
Science Data Archives

Tuulikki Sillajõe

24.04.2013



# Contents

- Current legal framework and changes that the new law has brought
- Results of the researchers' satisfaction survey
- Further challenges

## Legal basis for microdata access

- The Official Statistics Act is in force since Aug 2010.
- Procedure for access to microdata for scientific purposes established by Director General of Statistics Estonia (SE).
- Microdata are allowed to be used for research purposes only by legal persons or agencies, but not by freelance natural persons.
- Students pursuing a Master's or Doctor's degree are also considered researchers.
- **The same rules apply to domestic and foreign researchers.**

## Decision process

- A legal person or agency in need of microdata for research purposes must submit a written application to the producer of official statistics.
- Applications are considered by the Confidentiality Council (an internal body of SE).
- In case the application was satisfied, SE shall enter into an agreement with the legal person having submitted the application.
- Confidentiality commitments are signed by each researcher mentioned in the agreement.

## Rules for making the decision

- In case of every requested data set, the Confidentiality Council shall consider two aspects:
  - risk of identifying a statistical unit;
  - the impact that the identification of a statistical unit may have (sensitivity of data).
- If the identification risk and sensitivity are rated as high, confidential data can be used only in a safe centre, by remote access or remote execution
- If both aspects are rated as low, the data applied for can also be delivered on removable devices.

## Modes of access

- Safe centre
- Remote execution (LISSY)
- Remote access
- Scientific use file on CD-ROM or via FTP-server
- Public use files on the web
- Order for information
  
- SAS, SPSS, STATA, R

# All data categories are available

- Social surveys
- Business data
- Censuses
- Also, data derived from administrative records and other databases may be disseminated for scientific purposes.
- Data that permit direct or indirect identification of a statistical unit can be equally disseminated for scientific purposes.
- **No consent or permission of data supplier needed.**

## Microdata dissemination till 2010

- Business microdata were extremely rarely accessed and not available at all for some years.
- Microdata allowing the identification of a natural person could be accessed for scientific purposes without the person's consent.
- In practice, there were few agreements with some major universities.
- Practices and communication about them were very diverse, but accepted and appreciated by researchers who happened to know about them and were thus able to use the service.



## Major changes since 2010

- Switch from institution-based agreements to project-based agreements
- Introduction of confidentiality commitments
- Equal access and transparent procedure for researchers
- Direct link between microdata used and research published (<http://www.stat.ee/teadustood>)
- SE stores all microdata disseminated for 25 years in the form they were made available for a particular research
- More categories of microdata are available
- A slight rise in the use of microdata by researchers

# Progress made, 19.10.2010–23.04.2013

Year	Applications	rejected	cancelled	Agreements made	remote access or use of safe centre	dissemination on removable devices
2010	12	0	2	7	7	0
2011	15	1	2	13	8	5
2012	37	3	2	27	19	8
2013	10	0	1	12	10	2
Total	74	4	7	59	44	15

# User satisfaction survey 04.–15.04.2013

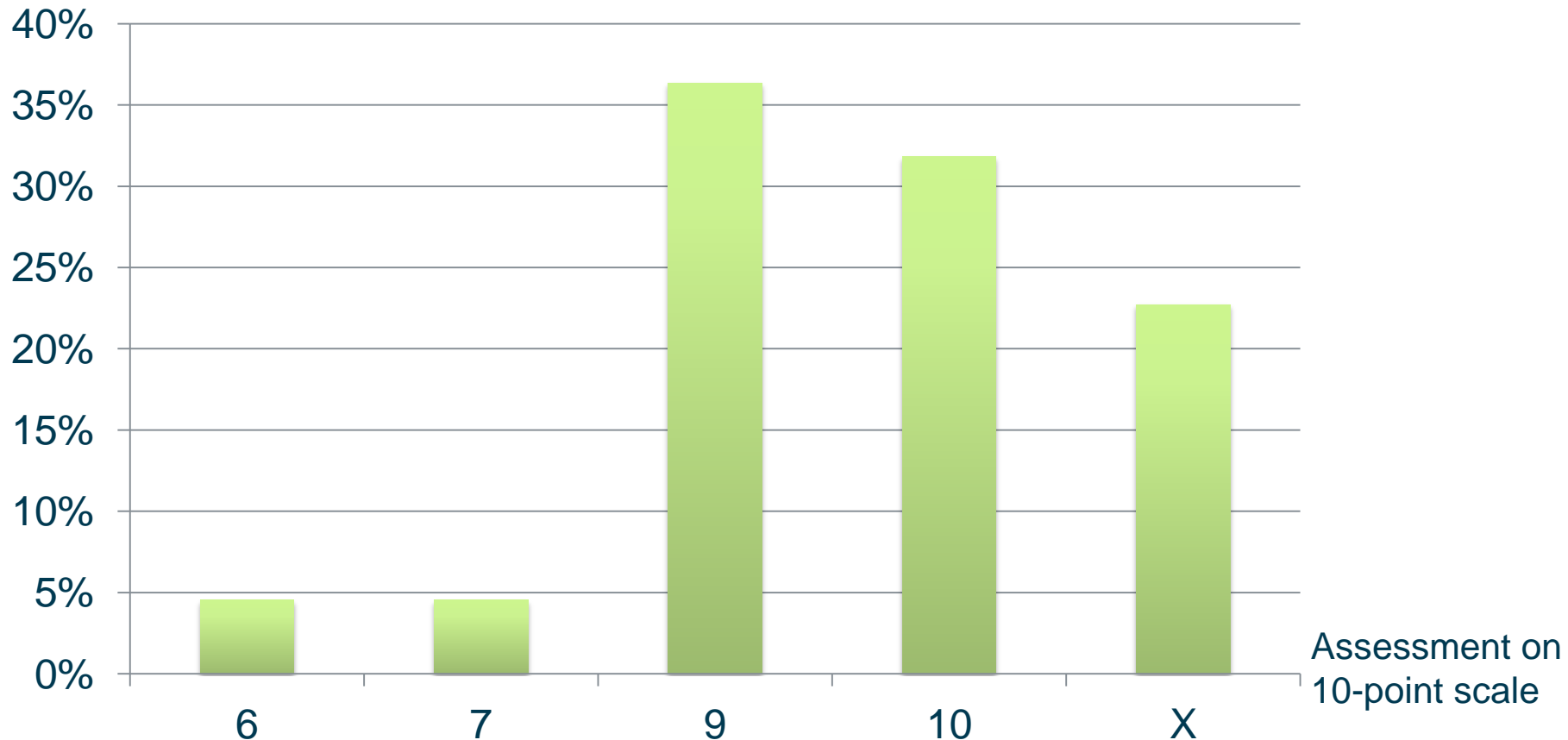
- Target group — researchers mentioned in agreements for remote access or use of safe centre (both: finished and valid agreements), 51% of them had used the service
- Data collection method — CAWI
- Response rate — 39%, whereas 50% of these had used the service
- 33% of respondents had not used the service, mainly because actual data processing was done by other members of the project team (one respondent considered it unhandy)

## Users of remote access or safe centre

- 67% of respondents had used remote access or safe centre.
- How did you get information about the availability of the service?
  - 55% from colleagues
  - 32% from elsewhere
  - 18% from SE
- The majority had used remote access (86%).
- The following slides concern only these 67% of respondents.



# Application procedure

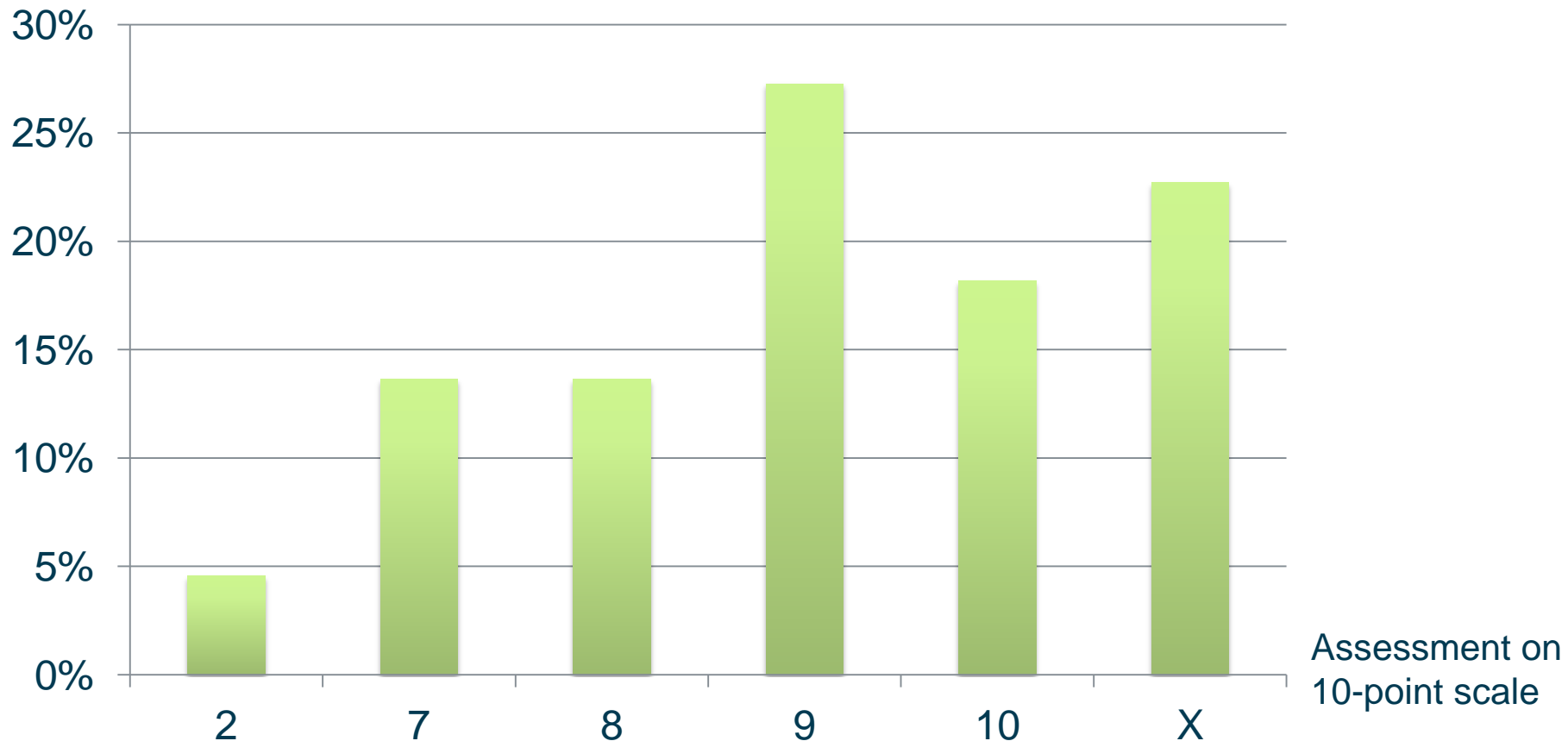


**Average rating — 9.1**

X — not personally involved in the process, because it was done by another member of the research team (23% of respondents)



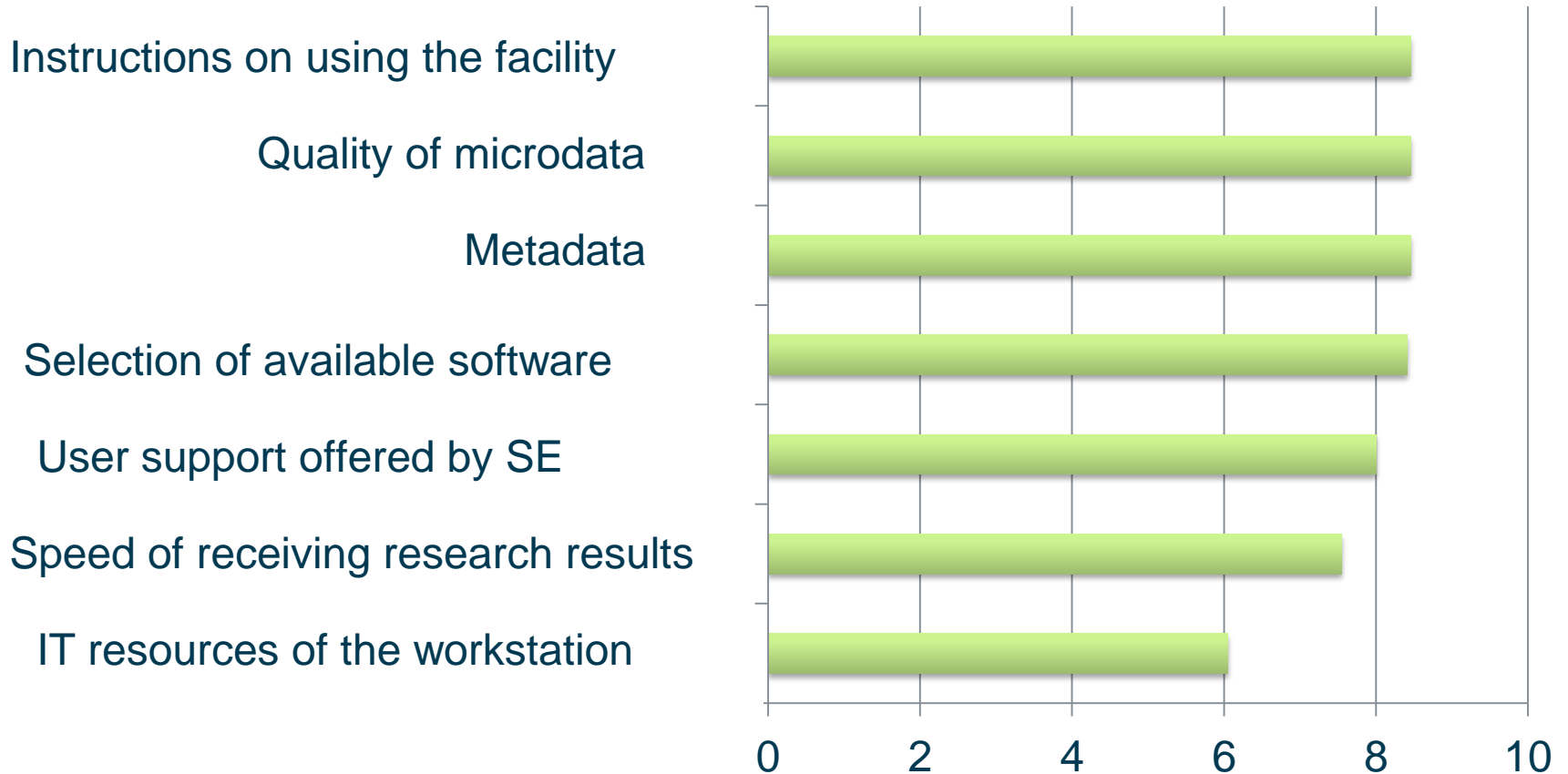
# Speed of access to microdata



**Average rating — 8.3**

X — not personally involved in the process, because it was done by another member of the research team (23% of respondents)

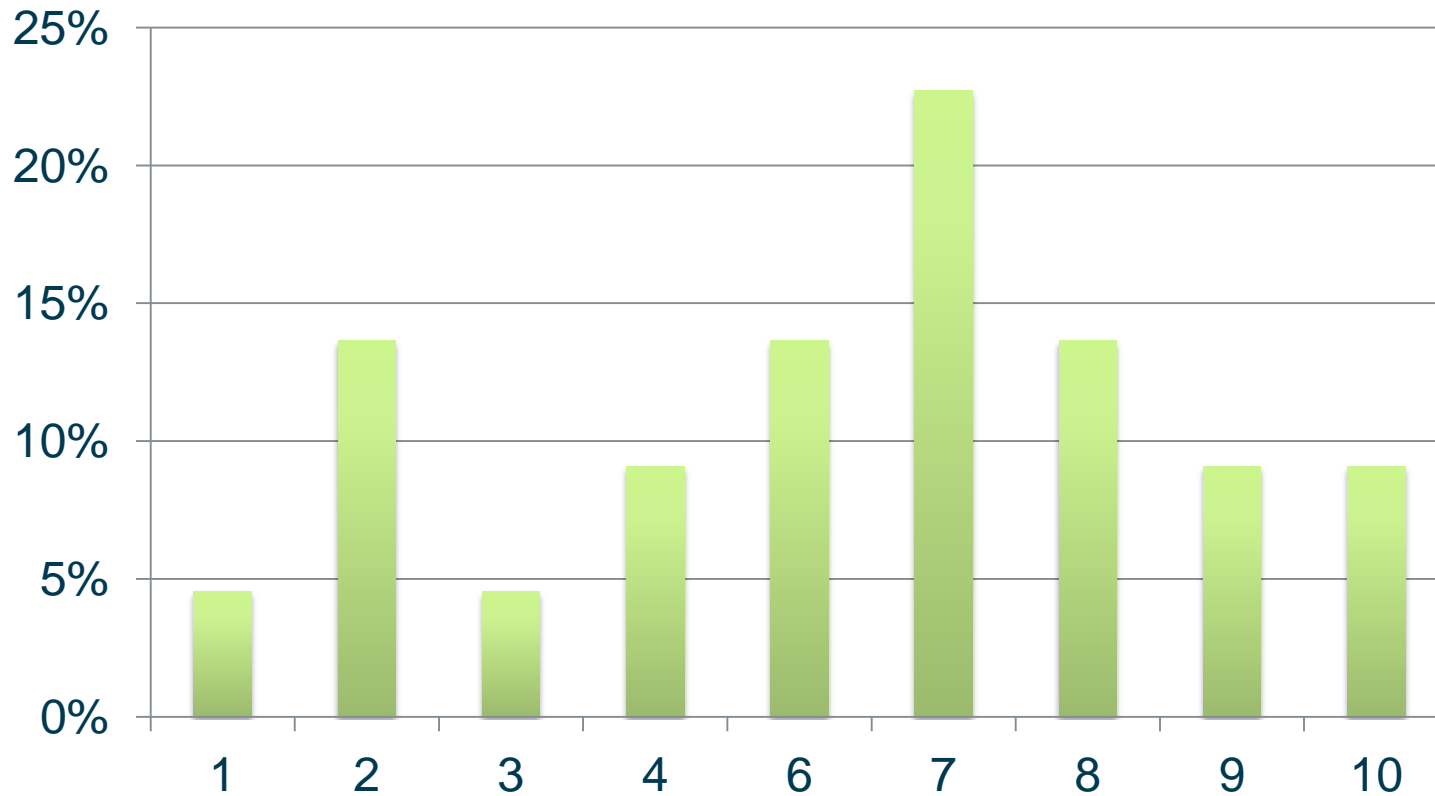
# Satisfaction with ...



**Average rating — 7.9**



# IT resources of the workstation



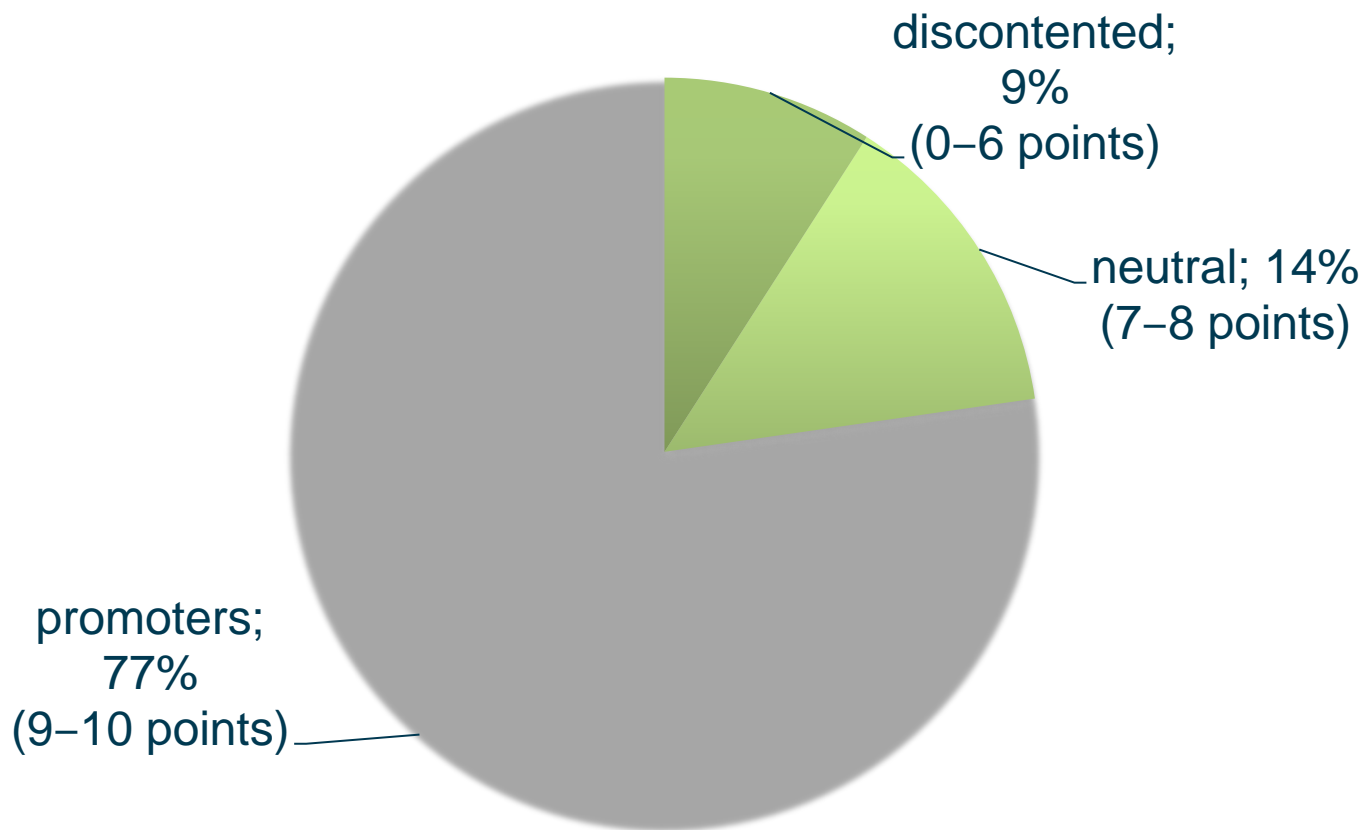
Assessment on  
10-point scale

**Average rating — 6.1**





# Promoter index



## Net Promoter Score — 68

(An NPS that is higher than zero is considered good and an NPS of 50+ is considered excellent)

## Lessons learned

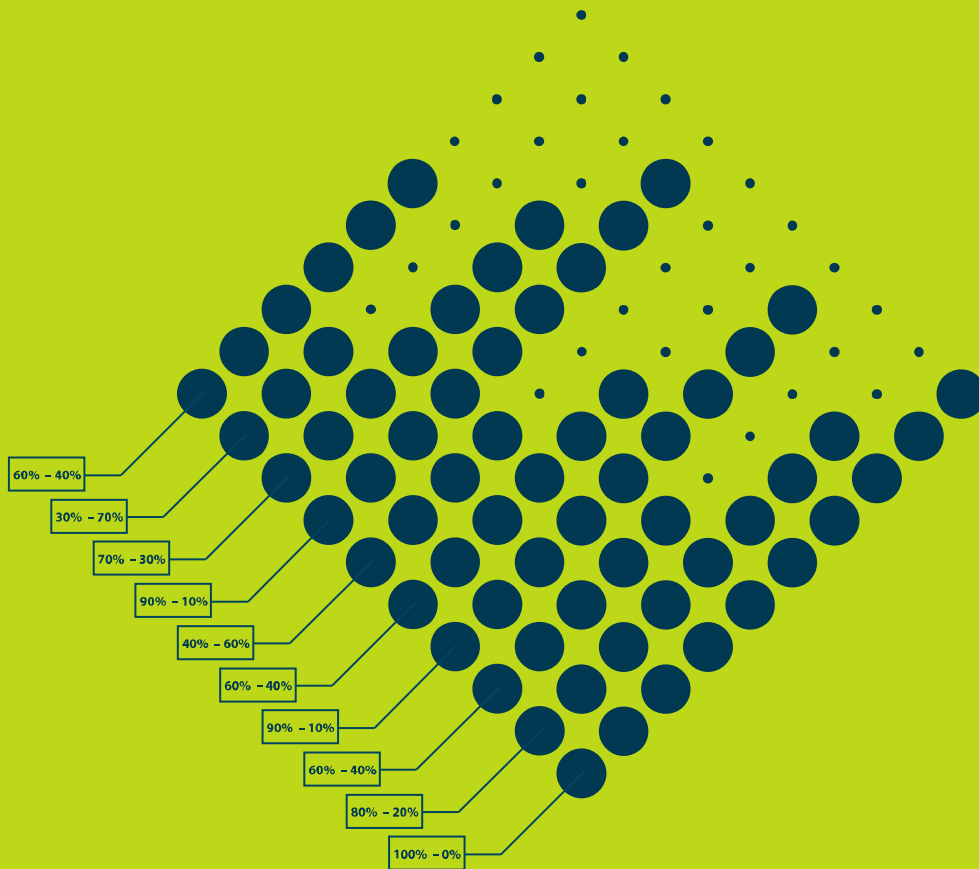
- There is no point in reinventing the wheel.
- Supply creates demand — a clear and quick procedure enhances access.
- Offer as good a service as you can.
- The aim of SE was to widen the reuse of the microdata collected for production of official statistics. But, at the beginning, SE still gained a kind of negative image as an impeder of research.
- Communication with stakeholders and interest groups is essential.
- Collect facts.

## Researchers' satisfaction could be increased by...

- upgrading Office 2003 software to allow faster processing of large amounts of data in Excel and Access;
- upgrading VPN (virtual private network) software which does not work with Windows 7 and Windows 8 now;
- improving the written instructions on how to install VPN and use the workstation;
- informing about the possibility to have additional packages of R installed by the IT-staff of SE;
- informing researchers if the checking of the results of their research takes up to three days (i.e. not available on the same day);
- improving metadata (adding information on sampling).

**Thank you for your attention!**

[tuulikki.sillajoe@stat.ee](mailto:tuulikki.sillajoe@stat.ee)



Microdata Access in Estonia