

Data editing by reporting enterprises

National Statistical Institute, Spain

Pedro Revilla (previlla@ine.es)

Manuel González (mgdavila@ine.es)

Margarita González (mgonzalez@ine.es)

José Quesada (pquesada@ine.es)

Summary

- **Prerequisites to get high quality incoming data**
- **Auditing by reporting enterprises**
- **Web survey opportunities**

**Data editing “good practice”:
moving editing closer to respondents**

**Going a step further: integrating the
respondents in editing**

TQM approach

The essential ideas

- **The suppliers (reporting units) and the customers (data users) are part of the productive system**
- **The suppliers and the customers determine the definition of quality to be used by our organisation**

The basic principles

- **Customer satisfaction**
- **Continuous improvement**
- **Facts-based management**
- **People-based management**

TQM approach

**the suppliers and the customers are part
of the productive system**

β

**integrating reporting enterprises in our
productive system**

How?

Prerequisites to quality

- **high quality statistics cannot be produced from poor microdata**
- **no data editing method would be able to amend low quality raw data**
- **the statistical process begins with the production and transmission of microdata by the reporting enterprises**

Strategies to get high quality data

- data collection have to be adapted to respondent conditions and possibilities
- statistical agencies should implement corporate strategies to encourage the respondent to fill in questionnaires
- data editing should improve the response quality (Granquist, 1997)

**“By learning more about
the product, the process
and the customer, we can
do a better job...”**

**BOX, G.E.P (1994) “STATISTICS AND QUALITY
IMPROVEMENT”. JOURNAL OF THE ROYAL
STATISTICAL SOCIETY.**

Achieving high quality incoming data

key success factors

- adapting questionnaires to the accounting practices**
- improving relationship with enterprises**

Adapting questionnaires to the accounting practices of the enterprises

- **Requested variables and valuation rules adapted to those of the reporting enterprises**
- **Observation units adapted to those units for which enterprises have available information**
- **Using different models of questionnaires adapted to the branch and size of the enterprises**
- **Personalised questionnaires**
- **Transmission formats (paper, fax, diskette, e-mail, web) adapted to the choice of the enterprises**



Underlying principle: enterprises provide data in the same way they produce them for their own use, and the statistical agency re-elaborates them for analytical purposes, if necessary

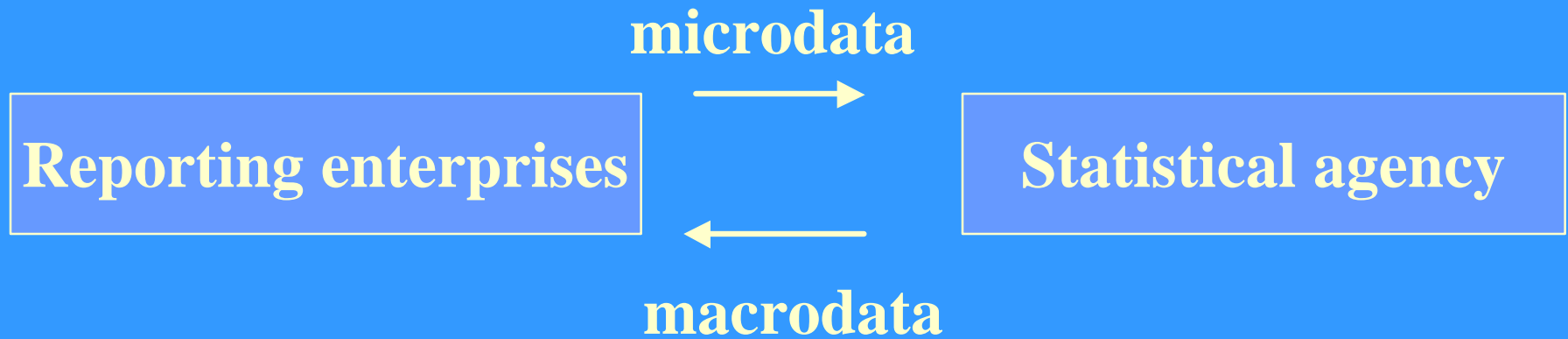
Adapting questionnaires to enterprises



- **Simplicity of response**
- **Speed in answering questionnaires**
- **Less errors to edit**



quality improvement & reducing burden



Key success factor:

**Improving the twofold relationship
between the reporting enterprises and
the statistical agency**

**Reporting enterprises position about
questionnaires:**

**...so, what use is
answering this to me?**

Hard contradiction:

- Our customers are (more or less) satisfied
- Our suppliers are strongly unsatisfied

solution?

Let's try to change our suppliers into customers!

Offering tailored data in exchange for the questionnaires

We provide the reporting enterprises with answers to the following questions:

- What is my market share in my business activity?**
- How many enterprises have bigger market share than mine?**
- What is the overall share of those enterprises with bigger market share than mine?**

“The industrial surveys carried out by the INE are the best example of a joint venture between industrial enterprises and public statistical offices. In other words, they are an excellent illustration of the advantages that may be achieved by means of an efficient collaboration between public and private sector”

Mr. BALTASAR ERRASTI

President of the Industry Commission of the Spanish Confederation of Enterprise Organisations.

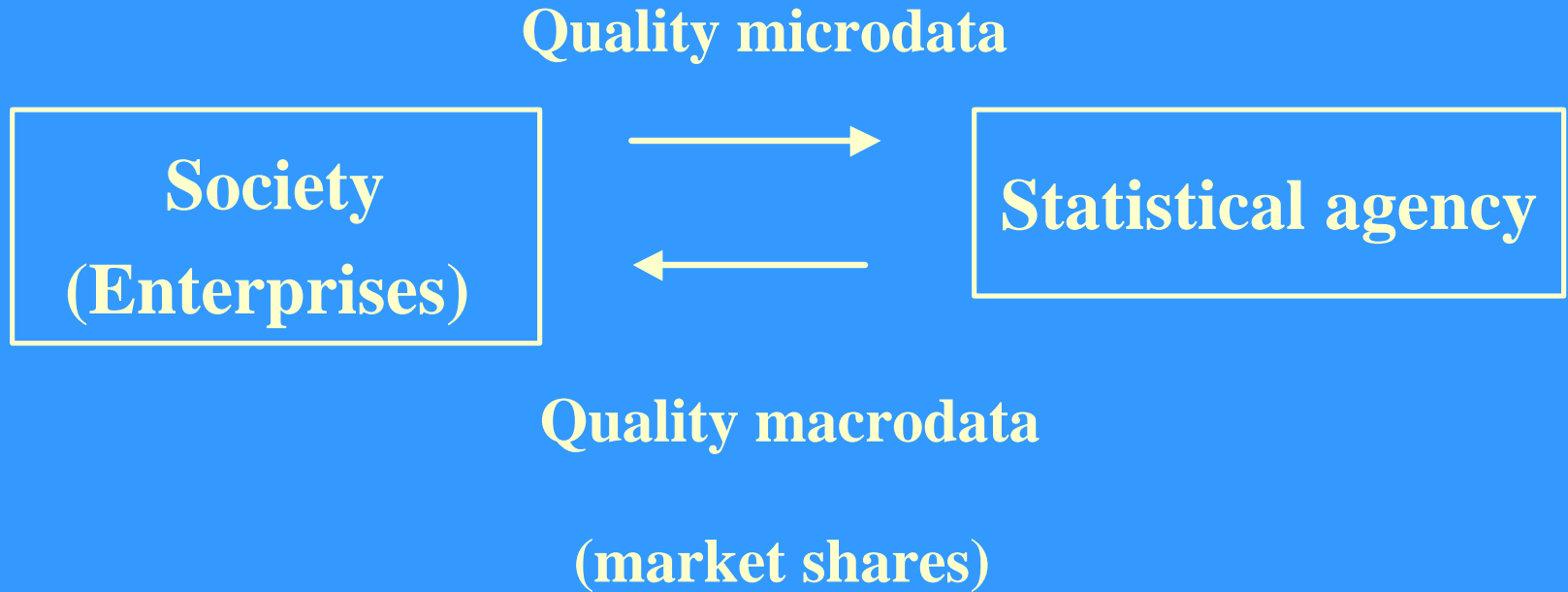
**Now, sometimes, reporting
enterprises say:**

**...the Statistic gives
something in change!**

A new model of relationship with enterprises:

“the joint venture model”

Fellegi's virtuous circle



action to improve our relation with enterprises

β

**enterprises call us when data do not match
expectations**

β

**errors detected and corrected
questions and definitions improved**

Auditing by reporting enterprises

- reporting enterprises may have an active role on editing and quality evaluation**
- prior knowledge about their sectors and markets**

Current situation

- **11000 enterprises market studies**
- **70000 enterprises sample**

β

- **11000 “external auditors”**
- **only 1 in 6 have participated**

We have to make a lot of efforts to inform enterprises’ management of our data offer

EDR and Web surveys opportunities

- changes in technology have broadened the possibilities for data collection**
- offering an electronic reporting option**

Data editing by reporting enterprises

- **Paper questionnaires ⊆ few opportunities**
- **Electronic questionnaires ⊆ many opportunities (built-in edits)**

EDR enable more editing to take place at data collection by the reporting enterprises

Web surveys advantages over other EDR methods

- **mature technology**
- **widespread acceptance**
- **few prerequisites**

The Web makes it simple to put electronic forms at the disposal of almost every enterprise

Web advantages

- **more opportunities to introduce more edits (edits that use respondents previous data)**
- **more security (https)**

Web advantages

- **improving accuracy**
- **improving timeliness**
- **reducing cost**
- **reducing enterprise burden**

Web advantages/improving accuracy

- **elimination of data keying at statistical agency**
- **built-in edits**

Web advantages/ improving timeliness

- elimination of data keying**
- data transfer on the web much faster than postal system**
- some electronic devises (automatic data fills and calculations, skips of no applicable questions, etc.) could help to fill in the questionnaire faster**

Web advantages/ reducing cost

- **reducing storage, packing, and postal charges**
- **eliminating data keying and data verification**
- **some of the editing task could be reduced from built-in edits**

A lot of expectation about the Web surveys in the years to come...

...the use of Web surveys has often been lower than expected

Take-up of Internet surveys

Branson (2002)

generally less than 10%, often less than 5%, Australia

Granjean (2002)

18% Index of Industrial Production, France

Mayda (2002)

5% -25% quarterly surveys on business and agriculture,
Canada

Spanish Population Census 2001

less than 1%

Why the rate of using Internet surveys is quite low, while technical conditions are available for many of the respondents ?

Advantages for reporting enterprises

- **The tasks that take the most time are looking for the required information and computing answers**
- **no time difference keying data on a screen to fill in a questionnaire on paper**
- **the reporting enterprises benefits depend on the way metadata support filling in questionnaires (help texts, auto-fill rules, pre-filled data, etc.)**

Encouraging the use of Web questionnaires by enterprises

- **explaining respondent benefits**
- **consider Web questionnaires in a wider context of all administrative duties and all EDR (e-commerce, e-administration, etc.)**
- **incentives (temporary access to information, free deliveries of tailored data, etc)**

EDR cannot be the only way of data collection

β

mixed mode of data collection

β

global editing strategies

Crucial questions

- **What kind of edits?**
- **How many?**
- **Mandatory?**
- **Automatic corrections?**

How many?

- too few \mathbb{P} losing an essential advantage
- too many \mathbb{P} respondent giving up

Mandatory edits

- **non plausible data have to be corrected or footnoted**
- **appropriate balance**
 - **a lot of assures quality**
 - **too many increase enterprises burden**

INE Web-based projects

- **Population Census 2001**
- **General project giving respondents Internet option**
- **Turnover and New Orders Survey**

Turnover and New Orders Survey

- very simple form
- avoiding problems of different components, versions, speeds, ...

β

no need to install any software on the respondent computer

Turnover and New Orders Survey

- **few edits are mandatory**
- **the rest are include as a message**

mandatory edits

- at least some data should have been entered
- alphabetic characters are not allowed in numeric fields
- 3 accounting equalities

3 accounting equalities

- **turnover =S turnover by markets**
- **new orders =S new orders by markets**
- **stock of orders at the end =**
= stock of orders at the beginning +
+ new orders - cancellations - orders invoiced

edits include as a message

- some fields are missing
- historic edits
- stock or orders at the end $t-1$ =
= stock of orders at the beginning t

Offering tailored data from the Web

- when an enterprise sends a valid form, it immediately receives tailored data from the server
- the Web has some advantages over paper

Editing strategy

- after Web edits no traditional microediting is needed
- selective editing based on time series modelling

β

all fatal errors and the most important query errors can be corrected before the index is disseminated for the first time

Final remarks

- **still far from the “joint venture model”**
- **the number of electronic questionnaires users too small**
- **more research needed on selective editing**

but

the combination of TQM, EDR and selective editing, may be in the near future a key success factor in producing high quality data, at a lower cost and with a lower respondent burden