

Business data collection methodology: Current state and future outlook

Mojca Bavdaž^{a,*}, Ger Snijkers^b, Joseph W. Sakshaug^c, Türknur Brand^d, Gustav Haraldsen^e,
Bilal Kurban^f, Paulo Saraiva^g and Diane K. Willimack^h

^a*School of Economics and Business, University of Ljubljana, Ljubljana, Slovenia*

^b*Statistics Netherlands, Heerlen, The Netherlands*

^c*Institute for Employment Research, Ludwig Maximilian University of Munich and University of Mannheim, Germany*

^d*Central Bank of the Republic of Turkey, Ankara, Turkey*

^e*Statistics Norway, Kongsvinger, Norway*

^f*Turkish Statistical Institute, Ankara, Turkey*

^g*Statistics Portugal, Lisbon, Portugal*

^h*U.S. Census Bureau, Washington, DC, USA*

Abstract. Collecting data from businesses faces ever-larger challenges, some of them calling for an overhaul of underlying methodology, e.g. motivation for participating is low; technology is shaping data collection processes; response processes within businesses are imperfectly understood while alternative data sources originating from digitalization processes push the response process (thus also response quality) further out of our sight. The paper reviews these challenges, discusses them in light of new developments in the field, and proposes directions for future research. This review may help those that collect data from businesses (e.g. national statistical institutes, academia, and private statistical agencies) to reconsider their current approaches in light of what promises to work (or not) in today's environment and to build their toolkit of business data collection methods.

Keywords: Business surveys, establishment surveys, data collection, administrative data, registers, digitalization, alternative data sources

1. Introduction

Data on businesses has an enormous impact on our societies. These data serve governments for designing policies and monitoring economic progress; academia for testing theoretical models; organizations representing businesses and businesses for benchmarking, among others [1,2]. Collecting data from businesses on a regular basis is largely in the domain of official statistics with some notable exceptions being, for example, research institutes and non-profit organizations. These institutions are often at the leading edge of applying

emerging methodologies in the field of business data collection but they are scattered around the world. Many other organizations share the interest in business data collection methodology but no single outlet exists for keeping abreast of the latest developments. Two types of meetings were initiated to fill this gap and help advance the field: large-scale meetings known as the *International Conference on Establishment Surveys* (ICES) and small-scale meetings known as the *European Establishment Statistics Workshop* (EESW) and the *Business Data Collection Methodology Workshop* (BDCM). This latter series of workshops offers an interactive setting for exchange of information among methodologists and practitioners, discussing findings and research in progress, hands-on experience, mistakes and pitfalls as well as emerging issues, thus contributing to the increase of expertise of all participants.

*Corresponding author: Mojca Bavdaž, School of Economics and Business, University of Ljubljana, Kardeljeva pl. 17, 1000 Ljubljana, Slovenia. Tel.: +386 15892630; Fax: +386 15892698, E-mail: mojca.bavdaz@ef.uni-lj.si.

This paper provides a snapshot of what is currently on the agenda in business data collection methodology, exposes new developments as they are happening in the field or influencing the field, and discusses the potential of novel ideas and insights, inspired by presentations and discussions at the 2018 BDCM Workshop.¹ The paper is organized around three big topics: the response burden and process in business surveys (Section 2), re-engineering business data collection (Section 3), and contact, communication, and motivation strategies (Section 4). The paper concludes with a discussion of what was learned, where the field is headed, and how this impacts future research (Section 5).

2. Response burden and process in business surveys

2.1. Qualitative research on the response burden

Compared to social surveys, business surveys have some unique characteristics that call for special attention because of their impact on accuracy of collected data [1,3–5]. Among these characteristics is the business response process where: (i) the business management, not the surveyor, decides who should respond; (ii) respondents often need to carry out an internal data collection; (iii) survey questions typically require knowledge of work-related concepts; and (iv) data usually have to be retrieved from the business information system [6,7]. The response process thus evolves at two levels: people go through their own cognitive processes of comprehension, retrieval, judgement, and communication [8,9] while taking part in organizational processes [7,10–15]. Qualitative research (like questionnaire pretesting) sheds light on the response process and helps understand what and why errors happen in this process (see Section 2.2). Technology adds to the complexity of the response process in web business surveys but also opens up new possibilities for studying, supporting and extending this process, like the analysis of paradata (see Section 2.3).

As business survey response occupies people's paid time and is often mandatory, burden imposed on businesses is considered one of the key constraints in sur-

veys, next to internal costs, time, and quality considerations [16–18]. Sources of response burden are numerous [19] as are the activities that national statistical institutes (NSIs) have developed to deal with burden [20,21].

Measuring burden is not easy and reported implementations are rare [19]. Dale and Haraldsen [22] proposed an operationalization of both actual and perceived response burden along with perceived causes of burden and motivational factors with nine questions. Statistics Finland implemented a reduced set of only four questions as part of a voluntary response burden questionnaire accessible through a click-through link on the last page of every business web survey in 2018: two questions collecting actual burden in time, one assessing perceived burdensomeness, and one open question for feedback [23]. However, due to unacceptably low response rates, the click-through link was removed and instead respondents were routed automatically to the response burden questionnaire after logging off the main survey. This procedure achieved an average response rate of 15 percent (range: 3–37 percent). More work is thus needed to address how best to activate respondents to participate in a response burden questionnaire and whether burden measures suffer from non-ignorable nonresponse. Detecting burden hotspots that may include not only large but also small establishments in certain sectors (e.g. [24]) also warrants further attention.

2.2. Qualitative research on the response process

Questionnaire pretesting has been considered indispensable for designing or revising data collection instruments in official statistics [16,25]. For social survey questionnaires, guidelines on testing methods have been discussed since the 1980s. For business surveys, the guidelines on implementation of such qualitative inquiry are still developing [26] with only [27–29] providing an overview of methods for questionnaire development and testing tailored to the business context.

Snijkers [30] reported their experience in using pre-field qualitative research in response to a mandate for combining two existing business surveys: Statistics Netherlands' Quarterly and Annual Survey of Finances of Enterprises and the Dutch Central Bank's Monthly Balance of Payments Survey. Before drafting the questionnaire, the researchers sought to study the response processes of businesses by launching a qualitative feasibility study including on-site visits to large non-financial businesses. The results led to the conclusion that combining both surveys would yield no

¹BDCM 2018 (<https://www.ine.pt/scripts/bdcm/index.html>) was held at Statistics Portugal in Lisbon on 19–21 September 2018. Fifty participants came from 24 countries and represented mostly national statistical institutes and a central bank, but also four non-profit organizations and two academic institutions.

new benefits to the businesses, even though the proposal seemed logical to the two sponsors. Specifically, the researchers concluded that data quality improvements and response burden reduction may only be achieved with a combined questionnaire tailored to the business context.² Businesses' internal structures, organizing the response process, and giving businesses enough time to get prepared should also be taken into consideration. Based on this experience, the researchers argued that starting questionnaire pre-testing at the stage when a draft questionnaire has been developed is too late. More preferable is to study the business context earlier in order to obtain a fuller picture of how information is distributed between people, between sources, between departments, and over time, and use these insights when drafting a questionnaire.

Irastorza et al. [31] applied pre-field qualitative research in the context of extending the survey universe to include micro businesses of 5–9 employees in the European Survey of Enterprises on New and Emerging Risks (ESENER), a survey which focuses on how workplaces manage occupational safety and health risks (OSH) in practice. In-depth interviews with these micro businesses revealed that most OSH-related responsibilities were outsourced to external service providers who handled the relevant arrangements. Thus, the investigators identified a potential mismatch between paper compliance and actual workplace practice. The study concluded that comprehension of survey questions may vary considerably with business characteristics, in particular size, sector, and the way OSH obligations are met (in-house vs. external service provider), and that regular qualitative inquiries are needed to reconcile survey answers with the realities of workplace practices.

Brand et al. [32] consider the use of post-field qualitative interviews to investigate a peculiar response pattern observed in the Business Tendency Survey (BTS), a monthly panel survey conducted by the Central Bank of Turkey to track business conditions and sentiment. One of the key questions is, what is your impression of the general economic conditions in your industry, compared to the previous month? More optimistic, remain unchanged, or more pessimistic. In the period 2007–2018, between 60–70 percent of respondents replied “remain unchanged” with the exception of some months in 2008 until mid-2009 [33]. A concern was raised regarding question interpretation as the literature suggests that a neutral response can be a hidden “don't

know” [34], a face saving “don't know” [35], a result of hesitancy or lack of information [36], or an outcome of satisficing [37]. In-depth interviews with respondents that choose the “remain unchanged” answer for a long time revealed that most of these respondents interpret “remain unchanged” as the same increase or decrease in optimism as before. Only if the monthly reported increase or decrease in optimism rose or fell, these respondents would choose the increase or decrease option. The study thus revealed that once a positive/negative expectation was reported, all subsequent positive/negative expectations of the same magnitude were reported as “remain unchanged”. Therefore, the study quashed initial concerns about satisficing and raised concerns about unintended interpretation that should be further investigated because of its consequences for the interpretation and relevance of derived indicators (calculated as a difference between positive and negative answers).

2.3. Response process in web surveys

In web surveys, a useful quantitative complement to qualitative research for studying the response process is the collection of *paradata*, that is automatically generated process data [38] or any data about the process by which the data was collected. Once the registration system is set up, paradata is typically created in real time and inexpensively, and has a wide range of uses from managing fieldwork to evaluating survey errors [39,40]. Paradata in business surveys has received increasing attention [2]. In particular, audit trails (a time-event schedule which records date and time of each mouse click and keystroke in the questionnaire) are being used for identifying different response profiles and highlighting problems in comprehension, which can be addressed using other means (e.g. cognitive interviewing, usability testing). NSIs, such as Statistics Netherlands [41] and, more recently, the UK Office of National Statistics [42] have reported their experiences. The raw logs contain massive amounts of noisy data in terabyte-size form, so a key issue is to extract relevant information in an efficient manner. Statistics Netherlands, for instance, reported problems with processing huge amounts of paradata which the UK Office of National Statistics largely avoided by analyzing micro-level data from a subsample and by relying on aggregate data from Google Analytics for the full sample.

Statistics Canada has developed an automated procedure to clean large amounts of data by parsing out the most important information and creating a report in some hours [43]. Their paradata is generated by

²For more information on the combined business survey, see <http://www.cbs.nl/balanceofpayments>.

Microsoft Internet Information Services logs, which records HTTP transactions, showing actions taken in the web questionnaire (e.g. time per page, type of browser and device used, help button usage, edits triggered, path through questionnaire etc.). The logs contain neither the respondent's answers nor the data on web pages accessed outside of the questionnaire. Analysis of these data makes it possible to identify critical pages where the respondent spent a lot of time, saved the data, or abandoned the questionnaire. It also reveals scanty use of help pages (e.g. used by only one percent of respondents in the Monthly Survey of Manufacturing) and modest corrections of data upon triggered edits (e.g. done by as few as eight percent of respondents in the Monthly Survey of Manufacturing; only irregular reporters changed the data more than half of the time). Paradata can be further used to improve data collection strategies. Statistics Canada is in the process of developing models to predict the time of a response using web logs and call history data. Units predicted to respond soon might, for example, be placed lower on the prioritization list for follow-up calls to allocate resources more efficiently.

Another priority research area among business surveys is cutting down on the amount of manual editing and follow-up calls. In business surveys, editing collected data may take up to 40 percent of the budget [44,45]. Asking respondents to correct or validate data when they are entered in the questionnaire may reduce the need for verification on the survey organization's side [46], but it also presents challenges. For example, the Australian Bureau of Statistics noted issues with *embedded validation* (also referred to as in-form edits or error messages) in relation to the response process [47]. Embedded validation should be neither a replacement for good questionnaire design nor an impediment to the majority of respondents with no need for assistance to complete the questionnaire. The respondent should always have a way out and never get stuck on a question; so hard edits should be used with extreme care and mainly limited to routing questions. The MoSCoW prioritization technique (Must have, Should have, Could have, and Won't have [48]); may help decision making about which validation rules to embed. However, to support response and validation, a list of all survey questions or sections and a summary or reconciliation page should be available [42,49]. Other recommendations for embedded validation include providing a clear diagnosis of the problem and the required action, or at least provide an opportunity to explain an unusual combination [50]. As proposed by Loren-

et al. [46], validation messages should be constructed and evaluated in a questionnaire design, evaluation, and testing (QDET) framework, but Price [47] notes that cognitive and usability testing is challenging because validation messages are rarely triggered, thus testing requires careful purposive sampling of businesses with problematic response behavior.

A related area of research is the validation of external data within the survey. For example, Statistics Norway has implemented an implicit validation check of government subsidy amounts so that the amount is provided in brackets as part of the question text. The respondent either re-enters the same number or provides a different one [49]. However, it remains unclear whether such an implicit dependent interviewing approach actually improves the data quality, or results in a confirmation bias [50].

3. Re-engineering business data collection

Official statistics continuously evolves its ways of obtaining business data. Many NSIs use multiple data sources and some have started to use the collected data for multiple purposes. The use of alternative, secondary data sources (such as administrative sources, registers and big data) that originated in the 1970s is increasing while primary data collection is shifting to more online and system-to-system data collections [51]. Many challenges in modernizing current processes give rise to new opportunities and innovations.

3.1. Integrating alternative (non-survey) data sources

In Europe, for example, the general recommendation is to use "administrative and other data sources whenever possible to avoid duplicating requests for data" [16]. Interestingly, countries are still at very different stages regarding the use of administrative data. Usage of these data varies from supporting survey data collection and replacing some survey data to estimating errors and using the administrative data as the main source that could be improved by survey data.

Some NSIs are in the early stages of exploiting administrative data by ensuring legal bases and verifying the source quality [52]. Recently, the Turkish Statistical System (TurkStat) has overcome legal milestones to acquire on-demand and scheduled administrative data transfers of tax records and other information from the Revenue Administration and Social Security Institution for the production of short-term indicators, annual

business statistics, and national accounts since March 2018 [53]. The Hellenic Statistical Authority (ELSTAT) has recently put in place a new production system for the compilation of Structural Business Statistics based on a combination of multiple data sources: tax data, social insurance data, commercial data, and a small-scale supplemental sample survey for the estimation of variables not available in the administrative sources [54].

However, even when a legal basis for the use of administrative data has been formed, its content, format, and scope might not be satisfactory to be used for official statistics. The Statistical Office of the Republic of Serbia (SORS) found out that some important variables were missing in the Central Register of Compulsory Social Insurance for the calculation of registered employment and had to develop new methodologies to augment this administrative data with the Statistical Business Register data [55]. The Statistical Business Register itself relies on survey and administrative data for updating [56]. On the other hand, administrative data may help estimate errors, among them particularly problematic coverage errors [57,58].

Statistics Canada plans to eliminate (almost) all of its 40 agricultural surveys and a census that collects data directly from farm operators every five years [59]. Two types of data sources are considered essential to reach this goal: a) about 300 separate administrative datasets, e.g. tax data, data from supply-management sectors (including dairy, chicken, eggs and turkey), data from insurance agencies on planted and insured crops, land valuations, honey bee permits, winery establishment grants, etc.; and b) satellite imagery with remote-sensing and geospatial information. Some data will come from other harmonized business surveys and model-based predictions. A key planned innovation is to evaluate the quality of a single item in the alternative data source (micro approach) and use it as a replacement for a respondent's answer if a sufficient quality threshold is reached, thus, smartly personalizing the questionnaire at the farm-unit level whenever reliable data is available for that unit. This idea will be tested in the 2021 agricultural census which will run a direct data collection (featuring also an interactive web mapping application for geo-coding of reported data) in parallel to predicting values with alternative data sources.

3.2. Automation and technical standards

A typical characteristic of business surveys is that businesses may receive several surveys and often receive the same survey several times [5]. Poor survey

coordination and administration may cause irritation among businesses and are often consequences of the traditional non-tailoring and stove-pipe (or silo) statistical production. An increasing number of NSIs are moving towards a centralized and integrated coordination system.

One such infrastructure model, proposed by the Centre for the Development of Vocational Training (Cedefop), includes several components, including a relational database that is linked to sample sources (e.g. official registers, panels, etc.), the survey software (supporting questionnaire design, translation, fieldwork, etc.), statistical software (for data processing, analysis, output creation, etc.), data storage, and other documentation (e.g. metadata) [60]. Such models support data quality management from questionnaire development through fieldwork to data dissemination (e.g. [61]).

A prominent example of an NSI changing to a centralized coordination and integration system is Statistics Portugal. Portugal has moved from paper to electronic questionnaires, embracing a multi-source and mixed-mode approach, and implementing more automated data collection [51,62]. Over 98 percent of all questionnaires were completed online in 2017. This effort was the result of an extensive centralization of data collection and the creation of an integrated infrastructure to support the statistical production (a new central data collection department and an Integrated Survey Management System, denoted as SIGINQ). One of its components is WebInq, an internet data collection service which offers a flexible way of linking various respondents, companies, and surveys. It also allows respondents to find help on electronic questionnaires and the products and services prepared with the collected data. Another important milestone was the creation of the Simplified Business Information System (IES) in 2006 [63]. This system facilitates several dispersed but identical legal obligations of businesses, e.g. annual accounts, tax statements, statistical information, etc. Statistics Portugal achieved complete coverage of the business universe (from 50,000 to 400,000); reduction of information availability from 12 to 6.5 months; automatic data collection by electronic means; a significant increase in detail; and the complete elimination of one of the costliest surveys. Plans to further simplify the IES include pre-filling questionnaires with previously-submitted and already-validated tax data. As a complement to web questionnaires, Statistics Portugal's Automated Data Transmission (TAD) service, which also involves the upload of XML files, as well as web services, already reached 24 percent of collected questionnaires in 2017.

By integration, standardization and harmonization of data collection systems in Portugal, uniform processes across the various production units could be created, eliminating differences between common processes. By economies of scale, productivity increased, the risk of failures decreased, and resources were freed up to invest in innovative techniques and develop new capabilities. On the other hand, a harmonized system resulted in loss of flexibility. Critical success factors include strong top management sponsorship, the involvement of a multi-disciplinary team, extensive design requirements, and internal change management. To the best of our knowledge, Portugal is the only country that has implemented such an advanced and integrated system. Another country also having implemented a coordinated system is Belgium.

Statistics Belgium adopted XBRL (eXtensible Business Reporting Language) as a standard for business data collection [64]. XBRL is a freely available global standard for exchanging business (especially financial) information, such as financial statements. In 2008 Statistics Belgium started to develop an XBRL-based web questionnaire as an e-government tool and as the standard for data collection amongst businesses and finalized it in 2017. Since then all surveys adopting the XBRL format have been monitored in one single system (StatData), directly connected with the Business Register. All the surveys follow the same standardized process in terms of loading the sample, loading data to pre-filled forms, creating user-ids and web forms, and exporting data. The XBRL adoption makes it possible to re-use existing concepts and create new specific concepts for each survey. Apart from the cost reduction, the standardized steps allow for more flexible human resources, as input, output, and data-processing of different surveys have similar characteristics. For instance, a survey 'organizer' can make changes and add validations directly to the form, a task that used to be carried out by a computer scientist. Further, since all surveys can be organized in a similar way, the daily data transfers can be similarly processed and stocked in a library. Since all of these libraries contain similar tables with identical variables and similar structure, the process of accessing data of a variety of statistics is simplified.

3.3. Use of mobile devices

Nowadays, offering respondents the option to complete surveys online is standard practice in household and establishment surveys alike [6,65]. However, with the proliferation of mobile devices (e.g. smartphones,

tablets), survey designers have to ensure their web surveys are compatible, and that the use of a mobile device doesn't harm response rates or data quality. Previous research has found that unit and item nonresponse rates tend to be higher on smartphones compared to PCs and tablets, and that measurement error tends to be higher due to visibility issues related to small screens [66]. However, other research finds that differences in measurement error tend to be small and respondents using smartphones tend to provide similar responses to those responding via a tablet or PC [67].

Business surveys have not faced the same urgency as household surveys to optimize their web surveys for mobile devices, generally because the response process tends to a) rely on records, b) involve multiple individuals in completing the survey request, and c) be lengthy and complicated to complete. Businesses are often assumed to use desktop and laptop machines, though the empirical evidence supporting this assumption is lacking. Morrison et al. [68] shed light on this issue by reporting device usage in five establishment surveys across three U.S. government agencies. Using paradata, the researchers find that a very small percentage of establishments who respond to a web survey do so through a smartphone. The highest rates of smartphone usage among web survey respondents were observed for the National Agricultural Statistics Service (NASS)'s 2017 Census of Agriculture (COA; 6 percent) and the 2018 June Crops Agricultural Production Survey (2 percent), respectively. This was followed by the Bureau of Labor Statistics (BLS)' Annual Refiling Survey (1.4 percent) and Job Openings and Labor Turnover Survey (0.5 percent), and the National Center for Science and Engineering Statistics (NCSES)' Higher Education Research & Development Survey (0.2 percent). Interestingly, the BLS and NCSES surveys have higher percentages of web responses, but lower percentages of mobile responses, compared to the NASS surveys. The finding suggests that the prevalence of mobile response depends on the population; for example, NCSES respondents, located at colleges and universities, are contacted at their offices and likely to be at their desks when they receive the survey request. Additional findings from the NASS COA survey show a lower rate of breakoffs for phone and tablet (both 0.2 percent) compared to desktop machines (1.6 percent), lower average session durations for phone and tablet devices compared to desktops, and higher rates of device switching from mobile devices to desktops than the reverse sequence among respondents with two or more sessions.

Mobile devices are also used by survey organizations for collecting responses and other administrative in-

formation in the field. Cadena [69] reports a novel use of mobile devices for updating the Mexican Statistical Business Register and georeferencing establishments. One of the key derivatives of the Business Register is the National Statistical Directory of Economic Units (DENUE), a highly demanded data product which consists of a complete listing of all businesses in the country, uploaded in a geographic information system (GIS). The DENUE contains detailed data on every business unit concerning identification, precise location, contact, establishment size, and activity they carry out. With these characteristics, users are able to filter the subuniverse of interest and augment it to additional layers of geographical information using digital cartography.

Since 2004, the Mexican Economic Census, conducted every five years, has experimented with the use of mobile devices to locate and collect information from businesses. Interviewers were equipped with a Personal Digital Assistant (PDA) for 10 percent of geographical areas in 2004 and for all in 2009. In 2014, PDAs were replaced with a subnotebook (changeable to a tablet), which included the most recent directory containing over five million economic units, questionnaires, digital cartography, and satellite images, among other tools. These devices allowed for updating in real-time the status and location of every business unit in the digital cartography and DENUE. The use of mobile devices led to an estimated cost savings of about 20 percent compared to previous data collections on paper because of multiple efficiencies and savings in manual labor costs [69].

3.4. Transformation: Content simplification and harmonization

Setting up an integrated system requires standardization and harmonization [51]. Harmonizing units and content is at the core of a “respondent-centric approach” at the U.S. Census Bureau [70]. The unit problem has attracted more research attention only recently [5,71]. The U.S. Census Bureau first concluded that for the 52 largest, high impact businesses, one reporting unit does not meet the needs of all businesses researched. In the next step, a “complexity score” was created for each multi-unit business based on the number of establishments, industries (unique 8-digit NAICS), states, and tax reporting entities. The research concluded that the top one-percent with the highest complexity score represent 20 percent of total burden and requires an account manager. A kind-of-activity unit will be implemented as the harmonized unit by late-2020 for all

other multi-unit businesses because 60 percent of these businesses have a single kind-of-activity unit.

Another important aspect is content harmonization, namely harmonizing definitions, survey questions, and instructions. The U.S. Census Bureau has begun with concepts that are common across surveys and most critical for economy-wide statistics (inventory, payroll, sales, control data, expenses) but encountered obstacles, including the failure to fully understand user needs, resistance from subject-matter experts, and lack of funding and resources for testing harmonized content [70].

Transformations that aim at coordination and integration within NSIs are complex [72–74]. As a case example of what transformation entails, the UK Office of National Statistics is proposing to move away from publishing short-term statistical outputs separate for the retail, motor trade, and wholesale industries, and merge these industries together into a single Distributive Trade output [75]. Qualitative feasibility research conducted with businesses across the three sectors studied how amenable businesses were to harmonized reporting, for instance, whether retail can provide turnover data exclusive of VAT like wholesale; or whether businesses can provide “total retail/wholesale/motor trade turnover” instead of “total turnover”, which could result in missing or misclassified turnover data should businesses have activities outside of their industry classification. Transformation should ensure consistent data at the point of collection, thus reducing the need for later adjustments.

4. Contact, communication and motivation strategies

4.1. Developments in communicating with businesses

Communicating with businesses about the survey request is a necessary process for delivering questionnaires to sampled units and obtaining timely, complete, and accurate responses. In follow-up to Cox and Chinnappa [3] Snijkers and Jones [6] identified three stages of business survey communication: a) pre-field stage (e.g. establishing contact with selected businesses, informing businesses about the survey and the intended respondent); b) field stage (e.g. sending advance letters, administering reminders, implementing nonresponse follow-up procedures, providing businesses with additional resources such as FAQs, websites, and help desks); and c) post-field stage (e.g. feedback). According to Snijkers and Jones [6] an effective communication strategy does not only provide the necessary in-

structions about the survey procedure, it also provides a tailored and persuasive means of inducing participation.

Any communication strategy has to primarily convince businesses that the benefits of taking part in the survey outweigh the time, resources, and costs required to complete the task. Although this challenge is somewhat different for mandatory versus voluntary business surveys, the same question applies: what is an effective communication strategy?

Recently, research by several statistical agencies shed some light on modestly researched effectiveness of response enhancement measures and improved communication strategies. For example, in response to the majority of their agricultural surveys becoming mandatory, Statistics Netherlands conducted a pilot study introducing new targeted communication products to improve the image of their agency in the eyes of sampled agricultural businesses [76]. These products included instructional videos, factsheets, and pre-due date reminder cards, and were implemented in the pilot in the same way for the newly mandatory surveys and a reference survey that remained voluntary. For the mandatory surveys the response rate improved significantly (from 50 to 75 percent), while for the voluntary reference agricultural survey the response rate remained at the initial low level (of about 50 percent). Even though this was not a rigorously designed experiment, Houben et al. [76] concluded that the communication products failed to increase the response rate for the voluntary survey, and the legal obligation must have had an effect. A second pilot study investigated the effectiveness of including additional information about the approaching enforcement procedure in a reminder letter for the mandatory 2017 Research and Development Survey. This contributed to a 6-percentage point higher response rate compared to previous years. Another pilot study, implemented experimentally in the non-mandatory Survey on Arts and Culture Education, showed that sending an attractively designed pre-due date reminder card tailored to the target group, was less effective than the standard pre-due date letter: 7.7 percent response rate increase compared to 10.7 percent; though each reminder was effective in increasing the response rate and sending additional reminders was more effective than sending fewer reminders. A general conclusion from these pilot studies is that direct communication measures (like advance, pre-due date, and reminder letters) work best to get response. Additional analyses at Statistics Netherlands showed that these direct measures are indeed effective response enhancement measures [77]. These measures may be seen as a nudge

for businesses to complete the questionnaire [78]. The more indirect materials (like factsheets and cards) may be nice to have, but their effect on response rates seems low or negligible, though they could be important for maintaining a good relationship with businesses.

Many NSIs are centralizing and standardizing their data collection [6,21,74], while at the same time tailoring their communication to targeted groups of businesses [79,80]. Statistics Estonia developed an automated notification system and training to facilitate cooperation in all mandatory business surveys to avoid the need to impose penalties for nonresponse [81]. The automated system provides online help resources and notifies establishments about obligations in the following calendar year and approaching deadlines. Statistics Estonia also invites previous nonrespondents and new sample units to attend face-to-face training with survey staff but results so far are disappointing: attendance is very low, and businesses do not follow-up with these training procedures [81].

A move towards improved facilitation for businesses was also implemented by the Italian National Institute of Statistics (ISTAT), which centralized and enhanced their inbound and outbound contact center services in 2016–2017 [82]. The enhanced inbound service provides assistance to responding units in the access and navigation of the data submission portal through a toll-free telephone number and dedicated email address. Telephone is the preferred channel among businesses, representing 75 percent of all inbound contacts. Three levels of assistance are provided with increasing levels of specialization and complexity. The outbound service initiates telephone contacts with the most relevant non-responding units who are registered in the business statistical portal, but who have not yet responded some time before, or a few days after, the punctual deadline. The service provides several functions, including reminding non-responding units to comply before the approaching deadline, offering support for the questionnaire/interpretation of questions, data entry assistance, and insertion of register changes in the appropriate section of the statistical portal. Although costly, the first results seem very promising as the response rates increased considerably and the number of days to get response decreased substantially.

The Polish Reporting Portal is another example of a centralized system that facilitates business response [83]. The system automatically sends three types of announcements (advance pre-notification to all units about the upcoming survey, a pre-due date reminder, and a final reminder to non-responding units), but also

offers other functions, including addressing break-offs by outbound messaging to respondents who started but didn't finish the task, and offering questionnaire support through a centralized telephone service or direct contact with specialized staff responsible for the survey. Every business has to designate one employee as a person in charge of all statistical reporting. The system allows this person to delegate parts of their mandate to other people (e.g. filling in specific parts of a questionnaire). Statistics Poland started developing this portal in 2007 and, as a result of legal changes, electronic submission of data became obligatory in 2009 for all except the smallest businesses. The number of user accounts increased from about 34,000 in 2008 to 871,000 in July 2018; the number of collected questionnaires increased from 320,000 to about 3 million.

4.2. Experiments in communication and contact strategies

One of the challenges in assessing the effectiveness of a new or enhanced communication strategy is that they are rarely implemented in a controlled setting. Production usually trumps experimentation and, thus, isolating the effect of a single or multiple design changes can be difficult if the implementation is altered for production reasons. However, recent efforts suggest that the frequency of controlled experiments in establishment surveys is increasing, implementing the idea of embedding research in recurrent business surveys propagated by Willimack et al. [27].

One active area addresses the impact of invitation/reminder modes on the response rate, nonresponse bias, and costs. The U.S. Bureau of Labor Statistics (BLS) conducted a series of experiments to evaluate the email as an invitation mode compared to more expensive paper based invitations by postal mail [84]. In one experiment email and postal mail invitations yielded equivalent overall response rates, although the email group responded at a slower rate than the postal mail group. The email group elicited a higher yield of web responses (about 30-percentage points higher) and led to a 21-percent reduction in costs overall compared to the postal mail.

A second experiment examined the effect of different mode sequences for non-response follow-up in a mixed-mode web/paper survey. Businesses with email addresses on the sampling frame were randomized into three groups, each receiving an email invitation followed by three additional email follow-ups. The email follow-ups were augmented with postal mail for each

of the three follow-ups in the first group, for the second and third follow-ups in the second group, and only for the third follow-up in the third group. The first group (fewest email-only/most combined email-paper contacts) yielded the highest response rate overall and was more likely to respond online compared to the other groups. Thus, the results suggest that email should be used as a supplement to postal mail, and not as a replacement.

Sakshaug et al. [85] also found email-only contacts to be ineffective, this time in a web-only survey of businesses conducted at the German Institute for Employment Research (IAB). In their experiment, the invitation and reminder modes were fully crossed with email and postal contacts for businesses with both contact information on the sampling frame. Email addresses were collected in a forerunner survey. The costliest paper invitation-paper reminder sequence yielded the highest response rate and lowest nonresponse bias overall. The next-best sequence was email-paper followed by paper-email and, lastly, email-email, which yielded the lowest response rate and largest nonresponse bias overall. The study also found that administering supplementary paper contacts (invitation and reminder) to the units with invalid email addresses was effective for increasing response rates and reducing nonresponse bias. For businesses without an available email address, the authors tested a strategy of requesting email addresses through a prenotification letter for the purpose of emailing the survey invitation. However, the return rate was very low (8 percent) and only about 40 percent of those businesses who provided an email address actually participated in the web survey. Thus, compared to a reference paper-only contact without prenotification letter, the email address request was ineffective from a response rate, nonresponse bias, and cost perspective.

Providing incentives to survey units to activate the social exchange process and increase the perceived trust of the survey organization [86,87], is another design feature which is often embedded within a communication strategy to induce cooperation, particularly in household surveys. Although incentives do not need to be monetary in nature to be effective, they should be viewed as having some value and interest to the business. A case example is Statistics Portugal, which delivers regular Personalized Feedback Reports to all responding establishments [51,88]. These reports contain a mix of individual and aggregate statistical data to benchmark individual companies against the sector or economy based on the collected data. The report achieves a high level of value recognition from busi-

nesses. Whether such incentives produce higher participation rates is an open question. Statistics Netherlands experimented with providing an unconditional incentive in the form of a small folder with the main results of the previous year's survey enclosed in the invitation letter [76], but the results indicated that the folder did not improve response rates.

Willimack [89] reports the results of several experiments conducted at the U.S. Census Bureau to aid the development of an effective communication plan for the mandatory, web-only 2017 Economic Census. The experiments, conducted in several annual and sub-annual surveys, examined variations in the type, timing, and sequence of contacts (advance notice, certified vs non-certified mail, earlier due date reminder, accelerated nonresponse follow-up letter; combination of earlier due date reminder and accelerated follow-up), optimal targeting of escalation techniques under adaptive design scenarios (targeted allocation of industries with low response rates), envelope appearance and labeling (ink color, envelope size, pressure vs. non-pressure sealed), and alternative motivation strategies. The separate experiments generally found support for the use of pre-due date reminders, red ink on the envelope to emphasize due dates and past due notices, pressure-sealed envelopes, accelerated nonresponse follow-ups, certified mail for later follow-up mailings, targeted selection of units for certified follow-up mailings. Little (or no) support was found for advance notices, half-page (vs. standard) size envelopes, flyers inserted in mailings with motivational messages, and subsampling without targeting. An open question, which the authors acknowledge, is whether the effectiveness of these individual (and separately tested) design features is likely to translate if strung together in a holistic communication strategy.

Instead of testing individual design features across separate surveys, the UK Office of National Statistics (ONS) took an enhanced holistic approach by applying several behavioral “nudges” to help businesses understand the purpose and expectation of the requested survey participation [90]: reciprocity (a prenotification letter and FAQ), messenger effect (a helpful and informal voice in the prenotification letter and an official and monitoring voice in the invitation letter), hassle factor (a simplified message and an illustrative 3-point checklist diagram), head start (pre-checking the first box of the checklist diagram as “already done!”), “make it timely” (highlighting what information is required and when), and “getting something back” (an infographic based on previous survey results relevant to the business and a big thank you note).

The intervention was implemented in the Monthly Business Survey for a random sample of businesses in the construction sector with fewer than 20 employees for newly selected sample units with the “business-as-usual” control group. The results were promising as this communication plan yielded a higher rate of responses before and after the punctual deadline compared to the control group. As with any holistic approach, however, it is unclear whether the improved response rate is driven by a small set of features or all of them.

The U.S. Bureau of Labor Statistics (BLS) tested whether an existing data collection structure of the Annual Refiling Survey, a web survey that is very fast at collecting data on every site's main business activity(ies) and location from a large sample, could be leveraged to conduct a second “piggyback” survey, especially to reach small target populations [91]. After completing the first survey, respondents were shown a transition page and asked to complete a few additional survey questions that took about two minutes to complete. These questions asked whether the respondent could (or not) report data on several topics (e.g. the current number of job openings, total sales, the top three revenue producing products or services) and about their relationship to the sampled company and department. In one test, a proportional random sample of single-site companies was drawn from the existing survey, mirrored the standard procedure of two email invitations whenever an email was available and two postal mails to all units, and yielded an overall 43 percent response rate with only a slight difference between emailed and mailed invitations. An additional sample was drawn from the population not included in the initial survey, which yielded a 19 percent response rate, to evaluate the possibility of getting a sample representative of the whole US economy. The results of the piggyback survey are thus promising but more research is needed to understand the types of questions that could be asked and the impact on subsequent response.

5. Discussion and conclusions: Surveys and beyond

Today, in the digital era with internet as a data source, digitalization of the business information chain, sensing technology and the Internet of Things, business data in principle is no longer scarce but might be more difficult to collect, process, or interpret [92]. To quote Robert Groves, based on his keynote speech ‘Official Statistics and Big Data’ at the 2013 conference on New

Techniques and Technologies for Statistics (Brussels, Belgium): “*We live in interesting times, in which new developments bring new challenges for statistical agencies. This requires new methods and new skills. ‘We have work to do.’*”

Previous sections exposed many research gaps where indeed “we have work to do”. Innovations that call for further methodological consideration include web questionnaire design on various devices, automated data collection, and more tailored and intelligent questionnaire communication. Future research should address how to make questionnaires more intelligent, for instance, how to embed validation rules (what validation would be motivating and not discouraging, how to communicate and solve “error messages” that involve several cells, etc.) and how to ensure a good overview. Future research should also investigate how effective the respondent’s support/help is and what to do next (if effective, how to promote the use of support/help; if ineffective, look for alternatives and repurpose currently available material, e.g. towards users and new survey staff). Some of these research questions require first more insights in the business response process. Paradata could provide some insights but analyses of these data are still in their infancy. More insight is expected from future qualitative research in the business response process, for instance on the interplay between question intent and comprehension issues to shed light on specification and measurement errors; on comprehension and data availability by business size to evaluate the need of tailoring questions; on retrieval processes in larger businesses to guide questionnaire design, etc. Outcomes of this research should also be better documented.

Although business web surveys are typically completed behind a PC, optimization for other devices is likely to become relevant at least in some surveys when (event) data need to be entered on the spot in diary-like questionnaires/forms (e.g. the use of pesticides on farms), to exploit their embedded functionalities (e.g. GPS in transportation surveys) or to offer applications instead of a questionnaire. Establishing contact and motivating businesses for response remains an important area of research, with questions such as what adequate sources of email addresses are and how to update them, whether and how to use piggyback surveys, what is the impact of invitation and reminder modes on response rates and bias. Furthermore, system-to-system data collection opens up many new questions: what the unit of observation is and who the respondent or the contact person is in this case, how to deal with changes in survey items, whether sampling is still efficient given that

the most time-consuming part is done only once, how much businesses want to control data flows, etc.

Since the use of administrative and register data keeps penetrating most of the statistical domains in NSIs and is expected to expand further, NSIs should continue their efforts to keep a sustainable cooperation and build partnerships with source owners to prevent surprises, maintain continuity of data delivery, and aim at solutions that better suit statistics. With some more recent alternative data sources (e.g. satellite imagery), data collection is similarly out of NSI control and calls for cooperation with source owners while with some other alternative data sources (e.g. web scraping email addresses), the burden of data collection and processing falls entirely on the shoulders of NSIs.

Studies into the internet as a data source, and the application of sensor-generated data are another promising avenue to explore. An example of a study that uses internet as a data source is an OECD study on MNEs (Multi-National Enterprises) [93] and the work on the Global Group Register by the United Nations [94]. This work also shows the limited scope of traditional methods and individual NSIs (because of national borders). Since NSIs are limited to measure MNE activities on a country level, the United Nations and OECD set up databases using an international and MNE centric approach. Data for this database is collected by combining traditional data sources (like annual enterprise reports) and data from MNE websites, using innovative data collection methods (XBRL, web-scraping and text analytics), among others. In the context of measuring globalization, sourcing activities, and the Sustainable Development Goals, these are relevant new data sources and methods.

An example of the use of sensing technology and the Internet of Things would include the collection of sensor-generated data and data communication via the internet, as is used e.g. in smart farming (see e.g. [95]) as well as in other industries (e.g. in the transportation industry using chipped containers). These technological innovations generate immense numbers of data [96], but are still at their very beginning, as is the exploration to collect and use these data for official statistics purposes [97,98]. Statistics Canada has touched on this development by combining available administrative datasets and remote sensing data [59] (see Section 3.1). In the Netherlands a project started to study locally generated farm sensor data in relation to data reporting using questionnaires [99]. The idea, like with the Canadian project, is to reduce response burden, as well as expand the statistical output. This development

could be an answer to ever existing demands for more detailed and timely statistics, as well as statistics on broader phenomena, moving towards smart statistics.

Like with survey-generated data, also these new data sources and collection methods involve quality considerations. At the data level, these include e.g. validity and reliability, coverage issues and representativeness, and unit issues (i.e. to what unit are the data related?). Here, the Total Survey Error Framework may serve as a quality framework [5,100,101]. The big difference of the use of existing data as compared to 'designed' survey data is that an NSI is not in control of these issues. For example metadata (like data definitions and data formats) may not be (readily) available. With regard to collecting these data, issues to consider are: data availability, access, sharing, and data harmonization (data definitions, data formats), especially when the same kind of data from large numbers of individual providers are to be collected and combined. This was also discussed with regard to automated data collection (see Section 3.2). Time-related issues include: stability of data definitions and stability of delivery. In addition, user quality considerations like relevance and timeliness need to be considered, as well as cost considerations both from the producer perspective and the data provider perspective. Taking all quality perspectives and considerations together, the quality diamond could be applied [101], as well as quality considerations for register data [102]. The Total Survey Error Framework may be turned into a Total Data Error Framework [51].

Even with these developments in mind, it is our strong conviction that also in the future, surveys as collections of designed data [103], will remain an important method to collect data about businesses for various reasons: to collect data that are not available in other sources, to supplement data from other sources, or as a way to validate data coming from other sources. However, to keep fulfilling this role of a trusted data collection method, with the use of all kinds of alternative data sources and the emergence of new technologies, the surveys that will be conducted in the future need to be of high professional standard and need to take the above mentioned innovations into account. That is why methodological research into these innovations is important. Determining quality of alternative sources at the individual (business) level is an interesting idea that deserves further exploration because it opens up the possibility of using a larger number of sources. The idea of having dynamic questionnaires where questions covered in alternative sources would not be asked or would have pre-filled answers promises reductions in

response burden, but also calls for evaluation of unintended consequences (e.g. context effects, confirmation biases).

Modern methodologists and official statisticians cannot forget about the established business survey methods (2); they have to constantly improve survey methods and address research gaps in business survey methodology. However, they also have to be (come) knowledgeable about new developments in businesses (e.g. with regard to data availability), economy (e.g. digital economy) and society (e.g. sustainable development), about other sources of data on businesses and methods for their collection, processing and interpretation, and integrate this knowledge in an updated business data collection methodology. Issues that continue to be at the heart of business data collection and are also key to data integration include: (a) data quality from a methodological perspective, assessment based on knowledge of the data generation process and using the Total Survey/Data Error Framework as a starting point [5,98] and from user and producer perspectives [5,101]; (b) the business unit [71,101] and; (c) metadata [104].

Acknowledgments

We would like to take the opportunity to thank all participants of the 2018 BDCM workshop for their presentations and discussions, and Statistics Portugal for being an excellent host. We would also like to thank Leanne Houben for reviewing an earlier version of this manuscript and two anonymous reviewers for helpful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the policies of their employers.

References

- [1] Snijkers G, Bavdaž M. Business Surveys. In: Lovric M, editor. *International Encyclopedia of Statistical Science* [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011 [cited 2019 Mar 31]. pp. 191-4. Available from: 10.1007/978-3-642-04898-2_625.
- [2] Snijkers G, Haraldsen G, Jones J, Willimack D. *Designing and Conducting Business Surveys*. John Wiley & Sons; 2013; 489 p.
- [3] Cox BG, Chinnappa BN. Unique Features of Business Surveys. In: Cox BG, Binder DA, Chinnappa BN, Christianson A, Colledge MJ, Kott PS, editors. *Business Survey Methods* [Internet]. John Wiley & Sons, Ltd; 1995 [cited 2019 Mar 31]. pp. 1-17. Available from: 10.1002/9781118150504.ch1.
- [4] Rivière P. What Makes Business Statistics Special? *International Statistical Review*. 2002; 70(1): 145-59.

- [5] Snijkers G. Achieving Quality in Organizational Surveys: A Holistic Approach. In: Liebig S, Matiaske W, editors. *Methodische Probleme in der empirischen Organisationsforschung* [Internet]. Wiesbaden: Springer Fachmedien Wiesbaden; 2016 [cited 2019 Mar 31]. pp. 33-59. Available from: [10.1007/978-3-658-08713-5_3](https://doi.org/10.1007/978-3-658-08713-5_3).
- [6] Snijkers G, Jones J. Business Survey Communication. In: Snijkers G, Haraldsen G, Jones J, Willimack DK, editors. *Designing and Conducting Business Surveys*. Wiley, 2013, pp. 359-430.
- [7] Willimack DK, Snijkers G. The Business Context and Its Implications for the Survey Response Process. In: Snijkers G, Haraldsen G, Jones J, Willimack DK, editors. *Designing and Conducting Business Surveys* [Internet]. John Wiley & Sons, Ltd; 2013 [cited 2019 Mar 31]. pp. 39-82. Available from: [10.1002/9781118447895.ch02](https://doi.org/10.1002/9781118447895.ch02).
- [8] Tourangeau R. Cognitive Sciences and Survey Methods. In: Jabine T, Straf M, Tanur J, Tourangeau R, editors. *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington (DC): National Academy Press; 1984, pp. 73-100.
- [9] Tourangeau R, Rips LJ, Rasinski K. *The Psychology of Survey Response*. Cambridge University Press, 2000.
- [10] Bavdaž M. The Multidimensional Integral Business Survey Response Model. *Survey Methodology*. 2010; 36(1): 81-93.
- [11] Edwards WS, Cantor D. Toward a Response Model in Establishment Surveys. In: Biemer PP, Groves RM, Lyberg LE, Mathiowetz NA, Sudman S, editors. *Measurement Errors in Surveys* [Internet]. John Wiley & Sons, Ltd; 2011, pp. 211-33. Available from: [10.1002/9781118150382.ch12](https://doi.org/10.1002/9781118150382.ch12).
- [12] Haraldsen G. Response Processes and Data Quality in Business Surveys. In: Lorenc B, Smith PA, Bavdaž M, Haraldsen G, Nedyalkova D, Zhang L-Ch, et al., editors. *The Unit Problem and Other Current Topics in Business Survey Methodology*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing, Lady Stephenson Library; 2018, pp. 155-76.
- [13] Lorenc B. Two Topics in Survey Methodology: Modelling the Response Process in Establishment Surveys; Inference from Nonprobability Samples Using the Double Samples Setup. Doctoral Dissertation. Stockholm: Stockholm University, Department of Statistics, 2006.
- [14] Sudman S, Willimack DK, Nichols E, Mesenbourg TL. Exploratory research at the US Census Bureau on the survey response process in large companies. In: *Proceedings of the Second International Conference on Establishment Surveys*, American Statistical Association. 2000, pp. 327-337.
- [15] Willimack DK, Nichols E. A Hybrid Response Process Model for Business Surveys. *Journal of Official Statistics*. 2010; 26(1): 3-24.
- [16] Eurostat. *European Statistics Code of Practice – Revised Edition 2017* [Internet]. Eurostat, Luxembourg, 2017, Available from: <https://ec.europa.eu/eurostat/web/ess/-/european-statistics-code-of-practice-revised-version-2017>.
- [17] Snijkers G, Haraldsen G, Jones J. Planning the Survey. In: *Designing and Conducting Business Surveys* [Internet]. John Wiley & Sons, Ltd; 2013 [cited 2019 Apr 17]. pp. 127-63. Available from: [10.1002/9781118447895.ch04](https://doi.org/10.1002/9781118447895.ch04).
- [18] Willeboordse A, editor. *Handbook on the Design and Implementation of Business Surveys*. Eurostat, Luxembourg, 1998.
- [19] Haraldsen G, Jones J, Giesen D, Zhang L-C. Understanding and Coping with Response Burden. In: Snijkers G, Haraldsen G, Jones J, Willimack D, editors. *Designing and Conducting Business Surveys* [Internet]. Hoboken, New Jersey: John Wiley & Sons, Ltd; 2013 [cited 2019 Mar 31]. pp. 219-52. Available from: [10.1002/9781118447895.ch06](https://doi.org/10.1002/9781118447895.ch06).
- [20] Bavdaž M, Giesen D, Černe SK, Löfgren T, Raymond-Blaess V. Response Burden in Official Business Surveys: Measurement and Reduction Practices of National Statistical Institutes. *Journal of Official Statistics*. 2015; 31(4): 559-88.
- [21] Giesen D, Vella M, Brady CF, Brown P, Ravindra D, Vaasen-Otten A. Response Burden Management for Establishment Surveys at Four National Statistical Institutes. *Journal of Official Statistics*. 2018; 34(2): 397-418.
- [22] Dale T, Haraldsen G. *Handbook for Monitoring and Evaluating Business Survey Response Burdens* [Internet]. European Commission, Eurostat; 2007 [cited 2019 Mar 31]. Available from: <https://ec.europa.eu/eurostat/documents/64157/4374310/12-HANDBOOK-FOR-MONITORING-AND-EVALUATING-BUSINESS-SURVEY-RESPONSE-BURDEN.pdf/600e3c6d-8e8d-44f7-a8f5-0931c71d9920>.
- [23] Niemelä A. Response Burden Measurement Project in Statistics Finland [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_10_43_BDCMLisbon2018_StatisticsFinland_AnnaNiemel%C3%A4_PPT.pdf.
- [24] Šegan V. Should We Approach Differently to Data Collection from Large Businesses? [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_11_13_BDCMLisbon2018_Statistical%20Office%20of%20the%20Republic%20of%20Slovenia_Vojko%20C5%A0egan_PPT.pdf.
- [25] Office of Management and Budget. *Standards and Guidelines for Statistical Surveys* [Internet]. 2006. Available from: https://unstats.un.org/unsd/dnss/docs-nqaf/USA_standards_stat_surveys.pdf.
- [26] Bavdaž M, Giesen D, Moore DL, Smith PA, Jones J. Qualitative Testing for Official Establishment Survey Questionnaires. *Survey Research Methods*. 2019; 13(3): 267-88.
- [27] Willimack DK, Lyberg L, Martin J, Japac L, Whitridge P. Evolution and Adaptation of Questionnaire Development, Evaluation, and Testing Methods for Establishment Surveys. In: Presser S, Rothgeb JM, Couper MP, Lessler JT, Martin E, Martin J, et al., editors. *Methods for Testing and Evaluating Survey Questionnaires* [Internet]. Hoboken, New Jersey: John Wiley & Sons, Ltd; 2004 [cited 2019 Mar 31]. pp. 385-407. Available from: [10.1002/0471654728.ch19](https://doi.org/10.1002/0471654728.ch19).
- [28] Willimack DK. Methods for the Development, Testing, and Evaluation of Data Collection Instruments. In: Snijkers G, Haraldsen G, Jones J, Willimack DK, editors. *Designing and Conducting Business Surveys* [Internet]. John Wiley & Sons, Ltd; 2013 [cited 2019 Mar 31]. pp. 253-301. Available from: [10.1002/9781118447895.ch07](https://doi.org/10.1002/9781118447895.ch07).
- [29] Snijkers G, Willimack DK. The Missing Link: From Concepts to Questions in Economic Surveys. Paper presented at the European Establishment Statistics Workshop, Neuchatel, Switzerland, 2011.
- [30] Snijkers G. The Response Process in Large Businesses [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_01_40_BDCMLisbon2018_StatisticsNetherlands_Ger%20Snijkers_PPT.pdf.
- [31] Irastorza X, Wadsworth W, Cleary A. An International Establishment Survey and the Coverage of Micro Enterprises – Does One-Size Fit All? [Internet]. Presented at the Fifth

- International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_02_02_BDCM%20Lisbon%202018_EUOSHA_Xabier%20Iraistorza%20et%20al_PPT.pdf.
- [32] Brand T, Aysoy C, Genç B. Exploring Responding Behavior Behind “Remain Unchanged” Answers: Case Study for the Business Tendency of Turkey [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_03_29_BDCMLisbon2018_Central%20Bank%20of%20Turkey_Turknur%20Brand%20et%20al_PPT.pdf.
- [33] Aysoy C, Brand T. Investigations on Responding Behaviour: Case Study for the Business Tendency Survey of Turkey [Internet]. Research and Monetary Policy Department, Central Bank of the Republic of Turkey; 2017 [cited 2019 Mar 31]. (CBT Research Notes in Economics). Report No.: 1705. Available from: <https://ideas.repec.org/p/tcb/econot/1705.html>.
- [34] Bovi M. Economic versus psychological forecasting. Evidence from consumer confidence surveys. *Journal of Economic Psychology*. 2009; 30(4): 563-74.
- [35] Sturgis P, Roberts C, Smith P. Middle Alternatives Revisited: How the neither/nor Response Acts as a Way of Saying “I Don’t Know”? *Sociological Methods & Research*. 2014; 43(1): 15-38.
- [36] Giovannini E, Uysal A. Statistics, Knowledge and Policy: What Do We Know About What People Know? OECD Workshop on Business and Consumer Tendency Surveys, Rome, 2006.
- [37] Krosnick JA. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*. 1991; 5(3): 213-36.
- [38] Couper MP. Measuring Survey Quality in a CASIC Environment. In: Proceedings of the Section on Survey Research Methods of the American Statistical Association; 1998. pp. 41-9. Available from: <http://www.websm.org/uploadi/editor/measuring%20survey%20quality%20in%20a%20casic%20environment.pdf>.
- [39] Durrant G, Kreuter F. Editorial: The use of paradata in social survey research. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2013; 176(1): 1-3.
- [40] Kreuter F. Improving Surveys with Paradata: Analytic Uses of Process Information. John Wiley & Sons, 2013.
- [41] Snijkers G, Morren M. Improving Web and Electronic Questionnaires: The Case of Audit Trails. Presented at the European Conference on Quality in Official Statistics, Helsinki, May, 2010.
- [42] Stewart J, Sidney I, Timm T. Paradata as an Aide to Questionnaire Design. In: Lorenc B, Smith PA, Bavdaž M, Haraldsen G, Nedyalkova D, Zhang L-C, et al., editors. *The Unit Problem and Other Current Topics in Business Survey Methodology*. 2018, pp. 177-97.
- [43] Doiron R. Exploring Web Survey Paradata to Improve Survey Design [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_04_21_BDCMLisbon2018_Statistics%20Canada_R%C3%A9jean%20Doiron_PPT.pdf.
- [44] National Academies of Sciences, Engineering, and Medicine. Reengineering the Census Bureau’s Annual Economic Surveys [Internet]. Washington (DC): The National Academies Press; 2018. Available from: 10.17226/25098.
- [45] Norberg A. Editing at Statistics Sweden – Yesterday, today and tomorrow [Internet]. Proceedings of the MSP2009; Modernisation of Statistics Production; 2009. Available from: https://www.scb.se/contentassets/0c872ed44afa41cda0f5f47ffb982c6b/b_3_qa_systems_norberg_final.pdf.
- [46] Lorenc B, Norberg A, Ohlsson M. Studying the impact of embedded validation on response burden, data quality and costs. In: Lorenc B, Smith PA, Bavdaž M, Haraldsen G, Nedyalkova D, Zhang L-C, et al., editors. *The Unit Problem and Other Current Topics in Business Survey Methodology*. 2018. pp. 199-211.
- [47] Price T. Subject Matter Edits in Online Forms: Problem Definition, Exploration, Development and Alternatives [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_05_42_BDCMLisbon2018_Australian%20Bureau%20of%20Statistics_Tanya%20Price_PPT.pdf.
- [48] Clegg D, Barker R. *Case Method Fast-Track: A Rad Approach*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1994.
- [49] Hole B. Designing an Instrument for Collecting Data from Political Organisations [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_06_32_BDCMLisbon2018_Statistics%20Norway_Bente%20Hole_PPT.pdf.
- [50] Haraldsen G. Questionnaire Communication in Business Surveys. In: *Designing and Conducting Business Surveys* [Internet]. John Wiley & Sons, Ltd; 2013 [cited 2019 Apr 17]. pp. 303-57. Available from: 10.1002/9781118447895.ch08.
- [51] Buiten G, Snijkers G, Saraiva P, Erikson J, Erikson A-G, Born A. Business Data Collection: Toward Electronic Data Interchange. Experiences in Portugal, Canada, Sweden, and the Netherlands with EDI. *Journal of Official Statistics*. 2018; 34(2): 419-43.
- [52] Daas P, Ossen S, Tennekes M, Zhang L-C, Hendriks C, Haugen KF, et al. List of Quality Groups and Indicators Identified for Administrative Data Sources. Eurostat: Luxembourg, 2011.
- [53] Kurban B, Eskici HB, Öztürk M, Arslanoğlu S, Tüzen MF. Extending the Use of Administrative Data in the Production of Business Statistics [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_08_24_BDCMLisbon2018_Turkish%20Statistical%20Institute_Bilal%20Kurban%20et%20al_PPT.pdf.
- [54] Benaki V. Introduction of a New Production System for the Compilation of Business Statistics [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_09_37_BDCMLisbon2018_Hellenic%20Statistical%20Authority_%20BENAKI%20Vassiliki_PPT.pdf.
- [55] Gavrić L. Combining Data From Administrative and Statistical Sources in Producing Labour Market Statistics [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D1_07_11_BDCMLisbon2018_Statistical%20Office%20of%20the%20Republic%20of%20Serbia_Ljiljana%20_Gavric_PPT.pdf.
- [56] Cimbaljević S. Statistical Business Register Survey on the

- Local Units of Large and Medium Sized Enterprises [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D2_13_12_BDCMLisbon2018_Statistical%20Office%20of%20the%20Republic%20of%20Serbia_Sini%C5%A1a%20Cimbaljevi%C4%87_PPT.pdf.
- [57] Elshani H. Frame Error Impact on Structural Business Statistics Surveys [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/bp/D3_26_07_BDCMLisbon2018_Kosovo%20Agency%20of%20Statistics_Hysni%20Elshani_BP.pdf.
- [58] Meyer BD, Mittag N. An Empirical Total Survey Error Decomposition Using Data Combination. National Bureau of Economic Research, 2019.
- [59] Thomassin M. The Migration of the Canadian Census of Agriculture to an Integrated Business Program Without Contact with Respondents [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D2_22_01_BDCMLisbon2018_StatisticsCanada_Mathieu%20Thomassin_PPT.pdf.
- [60] Scheuregger D. Integrating Survey Design and Data Quality Management [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_24_15_BDCMLisbon2018_CEDDEFOP_Daniel%20Scheuregger_PPT.pdf.
- [61] MacGoris S. Quality Assurance for the European Company Survey 2019 [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_25_35_BDCMLisbon2018_Eurofound_Sophia%20MacGoris_PPT.pdf.
- [62] Saraiva P. Surveys Only As a Last Resort: Automated Business Data Interchange. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September, 2018.
- [63] Rodrigues S, Chumbau A. Simplified Business Information: Today and Tomorrow [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D2_19_36_BDCMLisbon2018_Statistics%20Portugal_Sofia%20Rodrigues%20et%20al_PPT.pdf.
- [64] Vanhoucke S. Standardization of the Data Collection of Business Statistics in Belgium [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D2_18_30_BDCMLisbon2018_Statistics%20Belgium_Sem%20Vanhoucke_PPT.pdf.
- [65] Callegaro M, Manfreda KL, Vehovar V. Web Survey Methodology. SAGE; 2015; 345 p.
- [66] Antoun C, Katz J, Argueta J, Wang L. Design Heuristics for Effective Smartphone Questionnaires. *Social Science Computer Review*. 2018; 36(5): 557-74.
- [67] Tourangeau R, Sun H, Yan T, Maitland A, Rivero G, Williams D. Web Surveys by Smartphones and Tablets: Effects on Data Quality. *Social Science Computer Review*. 2018; 36(5): 542-56.
- [68] Morrison RL, Ridolfo H, Sirkis R, Edgar J, Kaplan R. Smartphone Usage in Establishment Surveys: Case Studies from Three US. Federal Statistical Agencies [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: [https://www.ine.pt/scripts/bdcm/doc/ppt/D2_20_20_BDCMLisbon2018_U.S.%20Federal%20Statistical%20Agencies%20\(NSF,%20NASS%20and%20BLS\)_Rebecca%20L.%20Morrison%20et%20al_PPT.pdf](https://www.ine.pt/scripts/bdcm/doc/ppt/D2_20_20_BDCMLisbon2018_U.S.%20Federal%20Statistical%20Agencies%20(NSF,%20NASS%20and%20BLS)_Rebecca%20L.%20Morrison%20et%20al_PPT.pdf).
- [69] Cadena SP. Use of Computing Mobile Devices in the Economic Censuses for Updating the Mexican Statistical Business Register and Georeferencing Establishments [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: [https://www.ine.pt/scripts/bdcm/doc/ppt/D2_21_27_BDCMLisbon2018_INEGI%20\(Mexico\)_Susana%20P%C3%A9rez%20Cadena_PPT.pdf](https://www.ine.pt/scripts/bdcm/doc/ppt/D2_21_27_BDCMLisbon2018_INEGI%20(Mexico)_Susana%20P%C3%A9rez%20Cadena_PPT.pdf).
- [70] Wellwood J, Marquette E. Harmonizing Economic Surveys [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: [https://www.ine.pt/scripts/bdcm/doc/ppt/D1_12_39_BDCMLisbon2018_Census%20Bureau%20\(USA\)_Jessica%20Wellwood%20et%20al_PPT.pdf](https://www.ine.pt/scripts/bdcm/doc/ppt/D1_12_39_BDCMLisbon2018_Census%20Bureau%20(USA)_Jessica%20Wellwood%20et%20al_PPT.pdf).
- [71] Delden A van, Lorenc B, Struijs P, Zhang L-C. Letter to the Editor. *Journal of Official Statistics*. 2018; 34(2): 573-80.
- [72] Saint-Pierre E. Redefining roles and responsibilities in a new harmonized statistical production process: opportunities and challenges. Presented at the 2015 Conference of European Statisticians (CES), 14–16 September 2015, Budapest, Hungary.
- [73] Saint-Pierre E, Couture F. Obtaining wide support for Statistics Canada's Integrated Business Statistics Program: a key task in the project plan. Presented at the 2014 Conference of European Statisticians (CES), 28–30 April 2014, Paris, France.
- [74] Snijkers G, Goettgens R, Hermans H. Data collection and data sharing at Statistics Netherlands: yesterday, today, and tomorrow. Presented at the 2011 Conference of European Statisticians (CES), 14–16 June 2011, Geneva, Switzerland.
- [75] Thorsteinsson K, Huxley B, Edmunds R. Short-Term Statistics Transformation: The Business Perspective [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: [https://www.ine.pt/scripts/bdcm/doc/ppt/D2_17_31_BDCMLisbon2018_Office%20for%20National%20Statistics%20\(UK\)_Kate%20Thorsteinsson%20et%20al_PPT.pdf](https://www.ine.pt/scripts/bdcm/doc/ppt/D2_17_31_BDCMLisbon2018_Office%20for%20National%20Statistics%20(UK)_Kate%20Thorsteinsson%20et%20al_PPT.pdf).
- [76] Houben L, Groot W, Debie D, Goris G, Opreij A, Geers X, et al. A Continuous Search to Improve Targeted Communication in Business Surveys. In: Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September [Internet]. 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_31_38_BDCMLisbon2018_Statistics%20Netherlands_Houben%20A%20et%20al_PPT.pdf.
- [77] Snijkers G, Geurden-Slis M, Goris G, Burger J, van den Hombergh L. The Effect of Response Measures in Business Surveys [Internet]. Presented at the Workshop on Statistical Data Collection, UNECE Conference of European Statisticians, Geneva, October; 2018. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.58/2018/mtg7/DC2018_S2_Snijkers_et_al_Netherlands_AD.pdf.
- [78] Thaler RH, Sunstein CR. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press; 2008, 295 p.

- [79] Geurden-Slis M, Giesen D, Houben L. Using the Respondents' Perspective in Redesigning the Survey on International Trade in Goods Statistics [Internet]. Paper Presented at the Fifth International Conference on Establishment Surveys ICES-V, June 21-23; Geneva, Switzerland; 2016. Available from: https://www2.amstat.org/meetings/ices/2016/proceedings/065_ices15Final00102.pdf.
- [80] Vennix K. The Treatment of Large Enterprise Groups Within Statistics Netherlands [Internet]. In Proceedings of the Fourth International Conference on Establishment Surveys (ICES-IV), Montreal, June; 2012.; Available from: <https://www2.amstat.org/meetings/ices/2012/papers/301992.pdf>.
- [81] Pellmas H. Notifying and Training Procedures for Respondents in Statistics Estonia [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_34_14_BDCMLisbon2018_Statistics%20Estonia_HeidiPellmas_PPT.pdf.
- [82] Bellini G, Bosso P, Binci S, Curatolo S, Monetti F. Centralized Inbound and Outbound Contact Center Service as a New Strategy in Data Collection [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_33_19_BDCMLisbon2018_Italian%20National%20Institute%20of%20Statistics_Giampaola%20Bellini%20et%20al_PPT.pdf.
- [83] Szymankiewicz P. Data Collection and Information Flow Management in Statistical Surveys Conducted With the Use of Reporting Portal [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D2_16_06_BDCMLisbon2018_StatisticsPoland_Pawe%C5%82%20Szymankiewicz_PPT.pdf.
- [84] Langeland J. Evaluating Mode Sequence When Email is used as the Initial Contact in Establishment Surveys [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_27_10_BDCMLisbon2018_U.S.%20Bureau%20of%20Labor%20Statistics_Josh%20Langeland_PPT.pdf.
- [85] Sakshaug JW, Vicari B, Couper MP. Paper, E-mail, or Both? Effects of Contact Mode on Participation in a Web Survey of Establishments. *Social Science Computer Review*. 2019 Dec 1; 37(6): 750-65.
- [86] Dillman DA. *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley, 1978.
- [87] Dillman DA, Phelps G, Tortora R, Swift K, Kohrell J, Berck J, et al. Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research*. 2009; 38(1): 1-18.
- [88] Moreira A, Saraiva P. Motivating Respondents in Business Surveys [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D3_32_34_BDCMLisbon2018_Statistics%20Portugal_Almiro%20Moreira%20et%20al_PPT.pdf.
- [89] Willimack DK. From Experimentation to Implementation: Putting the Pieces Together to Form a Cohesive Contact Strategy for the US. Economic Census. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September, 2018.
- [90] Tosi A, Moore H, Smith P, Best B. Leveraging Behavioural Insights to Improve Construction Businesses' Survey Response [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: [https://www.ine.pt/scripts/bdcm/doc/ppt/D3_29_23_BDCMLisbon2018_Office%20for%20National%20Statistics%20\(United%20Kingdom\)_Alessia%20Tosi%20et%20al_PPT.pdf](https://www.ine.pt/scripts/bdcm/doc/ppt/D3_29_23_BDCMLisbon2018_Office%20for%20National%20Statistics%20(United%20Kingdom)_Alessia%20Tosi%20et%20al_PPT.pdf).
- [91] Edgar J, Dalton M, Thomas E. Wait! Before You Go, Just a Few More Questions: Pilot Test of a Piggyback Survey [Internet]. Presented at the Fifth International Workshop on Business Data Collection Methodology, Lisbon, September; 2018. Available from: https://www.ine.pt/scripts/bdcm/doc/ppt/D2_15_03_BDCMLisbon2018_USBLS_Emily%20Thomas%20et%20al_PPT.pdf.
- [92] Snijkers G. Innovations in Business Statistics Data Collection. In 2018. Available from: <http://konference.ef.uni-lj.si/emos/innovations-in-business-statistics-data-collection/>.
- [93] Ahmed N, Doyle D, Pilgrim G. Measuring MNEs using Big Data: The OECD Analytical Database on Individual Multinationals and their Affiliates (ADIMA). In 2019.
- [94] Demolin Hermans sH. Advancing New Collaborative Mechanisms for the Profiling of MNEs in National, Regional and Global Group Registers. *Statistical Journal of the IAOS*. 2020.
- [95] Wolfert S, Ge L, Verbouw C, Bogaardt M-J. Big Datas in Smart Farming. *Agricultural Systems*. 2017; 153: 69-80.
- [96] Srinivasan V. *The intelligent enterprise in the era of big data*. Hoboken: Wiley, 2017.
- [97] Kowarik A. TSS from Smart Farming Systems. [Internet]. Presented at the ESSnet Meeting on Big Data, June 3-5 2019, London. Available from: https://webgate.ec.europa.eu/fp-fis/mwikis/essnetbigdata/images/b/bf/WPL_Meeting_2019_06_03-05_London_Presentation_Task_1_AT.pdf.
- [98] Snijkers G, De Broe S. Smart Business Statistics: How to Integrate Technology and Official Statistics [Internet]. Presented at the European Conference on Quality in Official Statistics, Krakow, June; 2018. Available from: https://www.q2018.pl/wp-content/uploads/Sessions/Session%2025/Sofie%20De%20Broe/Session%2025_Sofie%20De%20Broe.docx.
- [99] Punt T, Snijkers G. Exploring sensor-generated data in precision farming; towards official statistics using business sensor data. Heerlen: Statistics Netherlands, 2019.
- [100] Groves RM, Fowler F, Couper MP, Singer E, Tourangeau R. *Survey Methodology*. 2nd ed. New York: Wiley, 2009.
- [101] Haraldsen G. Quality Issues in Business Surveys. In: Snijkers G, Haraldsen G, Jones J, Willimack DK, editors. *Designing and Conducting Business Surveys* [Internet]. John Wiley & Sons, Ltd; 2013 [cited 2019 Mar 31]. pp. 83-125. Available from: 10.1002/9781118447895.ch03.
- [102] Daas P, Ossen S, Tennekens M, Zhang L-Ch, Hendriks C, Foldal Haugen K, et al. Report on Methods Preferred for the Quality Indicators of Administrative Data Sources. Luxembourg: Eurostat, 2011.
- [103] Groves RM. Official Statistics and "Big Data" [Internet]. Keynote presentation at the New Techniques and Technologies for Statistics (NTTS) Conference, Brussels, March; 2013. Available from: https://ec.europa.eu/eurostat/cros/system/files/1A03_Groves.pdf.
- [104] Zhang L-C. Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica*. 2012; 66(1): 41-63.