

Sources of Measurement Errors in Business Surveys

*Mojca Bavdaz*¹

Measurement errors are commonly ascribed to four sources: the respondent, the interviewer, the instrument (i.e., the survey questionnaire), and the mode of data collection. The unique characteristics of business populations and business surveys contribute to the occurrence of specific measurement errors. Although several authors have already exposed other sources of measurement errors in business surveys, in particular records, the information system, and the organization, some sources of measurement errors still lack exposure. This article proposes a typology to cover all known sources of measurement errors in business surveys based on previous research findings and an empirical study. It also elaborates on some implications for detecting, reducing, and preventing measurement errors in business surveys.

Key words: Business participant; records; respondent; survey characteristics; survey instrument; survey staff.

1. Introduction

Measurement error occurs at the time of data collection and is a type of nonsampling error, along with specification, frame, nonresponse, and processing errors (Biemer and Lyberg 2003). Measurement error can be defined as “a difference” (Hansen et al. 1951, p. 152) or “a discrepancy” (Sukhatme and Sukhatme 1970, p. 381) between the observed survey value and the true value, although the term *difference* might not be the most suitable description for variables that measure properties on nominal and ordinal scales (Lessler and Kalsbeek 1992). The definitions implicitly acknowledge the existence of a true value even though this is not always taken for granted (Kruskal 1991). The true value may be either treated as independent of the survey conditions or operationalized with regard to particular survey specifications (Lessler and Kalsbeek 1992). In practice, the true value is usually tied to particular survey conditions, so measurement error can also be defined as the observational gap between the ideal measurement and the response obtained (Groves et al. 2004). This article takes this latter perspective, which contrasts an actual survey response for an item with the best potential, because it takes into account what can be achieved for a particular item during data collection given the survey conditions.

In business surveys, measurement errors receive much more attention nowadays than they have in the past. On the one hand, the progress of sampling theory occupied most of the methodological development during the first half of the twentieth century

¹ University of Ljubljana, Faculty of Economics, Kardeljeva pl. 17, SI-1000 Ljubljana, Slovenia. Email: mojca.bavdaz@ef.uni-lj.si

Acknowledgments: I thank the staff of the Statistical Office of the Republic of Slovenia for their cooperation, as well as Lars Lyberg, Lea Bregar, and Jaak Billiet for their guidance and support in this research. I also thank the associate editor and referees for their helpful comments on an earlier version of this article.

(Lyberg 2003). On the other hand, fewer measurement errors have been expected in business surveys because respondents have been considered more competent and better equipped to perform the survey task (i.e., to provide accurate answers based on business records). Yet experience and sporadic research have shown that measurement errors in business surveys may be far from negligible. The new paradigm in survey methodology has shifted attention to causes of survey errors (Tourangeau 2003) and has produced several models of the response process applicable to business surveys (Edwards and Cantor 1991; Biemer and Fecso 1995; Sudman et al. 2000; Willimack and Nichols 2001; Lorenc 2006; Bavdaž in press). This article does not discuss those models, though it does refer to them. The focus here is on developing a typology of sources of measurement errors in business surveys along with relevant definitional issues.

Measurement errors are generally ascribed to four principal design features of the measurement process: (1) the interviewer, (2) the respondent, (3) the instrument (i.e., the survey questionnaire), and (4) the mode of data collection (e.g., Groves 1989). Sometimes errors arising from the information system and the interview setting are also considered (e.g., Biemer and Lyberg 2003). It is typical in business surveys that a respondent cannot answer survey questions by merely relying on memory and has to look up the business records to acquire relevant data about the organization; however, the recorded data may contain errors or require updating; be difficult to access; or deviate from the required definitions, unit of observation, time period, and so on. In business surveys, the records (Ponikowski and Meily 1989), the information system (Biemer and Fecso 1995; Biemer and Lyberg 2003), and the organization (O'Brien 2000) are therefore treated as sources of measurement errors. The organization being the broadest of these terms encompasses the information system as well as other important aspects that influence a respondent's behavior in the survey response process (e.g., organizational policy on surveys, internal organizational structure, interview setting). Another approach that applies only to establishment surveys conducted by mail defines the task, the information system, and the respondent as sources of measurement errors (Goldenberg et al. 1993).

As I discovered in an empirical study on the business survey response process, some sources of measurement errors in business surveys have not been treated previously. In addition, some findings from the literature on response models have not made their way into these lists. Listed sources of measurement errors would also benefit from a (different) systemization and clarification of relevant definitional issues. The aim of this article is therefore to develop a typology that addresses all known sources of measurement errors in business surveys based on previous and original research. The article first presents the empirical study, continues with the development of the typology, and concludes with a discussion of implications for detecting, reducing, and preventing measurement errors in business surveys.

2. Empirical Study

The purpose of the empirical study was to identify sources of measurement errors by examining in detail an actual response process in a business survey from start to finish. The Quarterly Survey of Trade (QST) conducted by the Statistical Office of the Republic of Slovenia (SORS) was an appropriate candidate given the importance and frequency of

mandatory recurring business surveys conducted by governmental organizations. Using an eight-page questionnaire accompanied by instruction and classification booklets, the QST mainly collected quantitative business data on sales, stocks, and employment to track changes in totals and structures. Data collection was self-administered and conducted by mail. Nonresponding businesses received up to three reminders. Telephone calls were typically made to nonresponding businesses and businesses that provided missing or highly inconsistent data.

I contemplated several methods from the qualitative research tradition, which enabled the collection of relevant data on the response process and sources of measurement errors in the QST. Taking into account the relative effectiveness and efficiency of the methods as well as the available resources and time frames, I selected a mix of expert evaluations, observations, and interviewing methods (for more details on selection, see Bavdaž 2009).

I gathered expert evaluations via interviews. Interviews with questionnaire administration experts from the SORS (subject-area specialists and people involved in activities after fielding the questionnaire) were based on a semistructured interview guide and aimed to acquire information on the conduct of the survey, its development over time, and problems encountered in the past. Interviews with two accounting experts consisted of think-aloud and semistructured questioning focused on comprehension of the survey request, conceptual issues, and data availability.

I collected the majority of data directly from businesses. Initially, 25 units from the existing QST sample were selected systematically across all business sizes to cover the heterogeneity of the response processes. Five units – namely, three businesses and two accounting firms acting on behalf of the businesses – eventually refused participation: three units referred to their work overload during the planned completion of the questionnaire, one unit claimed that the task consisted of writing down one figure, and one owner prohibited the respondent's participation. Participating units covered well the geographical location as well as the various combinations of trade activities and kinds of merchandise. However, none of the largest players in Slovenian trade was included. Therefore, I also included seven of those players in the sample, and all agreed to cooperate. Because several respondents referred to the difficulties and uncertainties involved in completing the questionnaire for the first time, I attempted to include some units that had not been in the QST panel. One unit out of four that were new in the QST panel agreed to participate, one refused, and two claimed ineligibility. Two units out of six that were not obliged to fill out the QST questionnaire agreed to participate; I also visited them but did not include them in further analyses because their hypothetical survey response process was performed without looking up the actual record. The findings are therefore based on 28 different-sized businesses: 13 small, 5 medium, and 10 large. I carried out on-site visits of the first 20 businesses in the first quarter of 2005 and on-site visits to the remaining eight businesses in the subsequent quarter.

The first contact with a business was established with the person indicated as a contact, if available, on an earlier questionnaire. I verified that the contact had actually filled out the questionnaire or identified such a person in a telephone conversation. The primary method of investigating in the businesses was the qualitative research interview. This interview largely relied on retrospective probing (Willis 2005) and ethnographic interviewing (Gerber 1999). I administered the interview with people who had filled out the QST

questionnaire. Most were already familiar with the survey, while only a few were new to it. In one case (and two omitted from further analyses) the interviewee came in contact with the survey instrument for the first time in my presence, so I applied the three-step test-interview (Hak et al. 2008). In the larger businesses additionally included I also conducted short, structured telephone interviews with people that respondents indicated had provided data for the questionnaire. As a complement to these interviews in eight cases, I observed the QST response process at least partially. The observation was mainly unstructured and its extent depended on the amount of data that respondents had retrieved before the visit.

The implementation of the research plan was challenging because of the necessary coordination with the SORS, the heavy burden imposed on selected units, and organization of the on-site visits during or immediately after the questionnaire was completed with me as the only available researcher. Thorough preparations preceded every research activity. Much effort was invested in building a relationship with the businesses to obtain their consent for participation (e.g., prenotification letters, telephone contacts for identification purposes, letters or e-mails with additional information, and several telephone contacts to arrange the on-site visits). Special attention was paid to minimizing the time that elapsed between the filling in of the QST questionnaire and the on-site visit to assure the validity of collected data. As a result, half of the units worked on the questionnaire on the day of the on-site visit, and only three units had a time gap longer than a week. In addition, three factors worked against the deterioration of recall when a short time lag occurred: business documentation, the recurrence of the survey response, and the advance announcement of the impending on-site visit. To minimize the bias due to a single researcher and increase the reliability of findings, I trained as a qualitative researcher and studied relevant methodological literature. I documented and justified all steps and decisions made in the research, and I recorded and transcribed all verbal communications with people who provided data on the response process, except for shorter interactions, which I documented immediately (for more details, see Bavdaž 2009).

The analysis has its roots in the literature review, which served as a starting point for identifying relevant themes (e.g., people involved in the response process) and aspects (e.g., respondent selection and communication between people involved in the response process). Segments of the transcribed text, notes and memos were attributed to the theme(s) they addressed and to the particular aspect they dealt with within a theme. Further structuring of aspects within the themes was used to generate initial codes. In the second round of text inspection, the coding became focused on and oriented toward the detection of patterns. At this point, all available material (e.g., interview transcriptions and, notes and memos from observations) was reexamined and findings revised. The patterns were pulled together into concepts and constructs with the help of various displays and with regard to previous study findings.

Although the study design and implementation were designed to ensure high validity and reliability of findings, the fact remains that this is still only one study with specific features that may limit the generalizability of its findings, particularly the specifics of the QST and the sample of units selected and included for examination. Caution and further research are therefore necessary when applying the findings to “pure” nontrade units, overloaded units, environments with different business philosophies and institutional contexts, and business

surveys with other characteristics. In light of these limitations, the findings should be considered mainly to help refine, systemize, and further develop existing research.

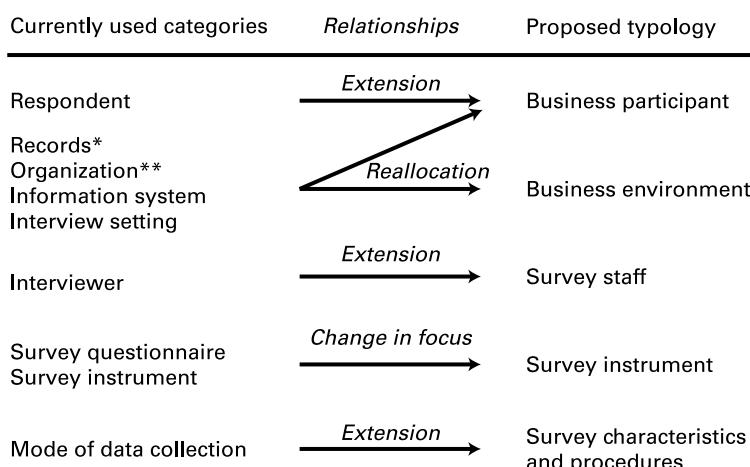
3. Developing a Typology of Sources of Measurement Errors in Business Surveys

The QST study showed that sources of measurement errors in business surveys are numerous and not completely covered by currently used categories. This section presents the results of the QST study and integrates them with previous findings to develop a typology of sources of measurement errors in business surveys. Figure 1 summarizes currently used categories and their relationships to the proposed typology.

The proposed typology first divides the sources of measurement errors in business-related sources that the survey organization can control for only indirectly, and survey-related sources that are under control of the survey organization even though such a classification is to some extent arbitrary, considering that the sources are often linked so that their interactions actually might cause measurement errors. This section first addresses the sources of measurement errors on the business side.

3.1. Business Participants

The literature on business surveys acknowledges that several people from the business are commonly involved in responding to such surveys. Edwards and Cantor (1991) pointed to multiple respondents and discussed gathering information “from or through someone else in the organization” (p. 227), delegating the survey task, and preparing or reviewing information. Tomaskovic-Devey et al. (1994) referred to boundary-spanning units while Groves et al. (1997) labeled them receptionists or gatekeepers. The latter also mentioned decision-makers and people who compile and update requested information. Sudman et al. (2000) exposed the frequent case of one person coordinating the data collection from



Source: Biemer and Lyberg 2003

*Ponikowski and Meily 1989

**O'Brien 2000

Fig. 1. Typology of Sources of Measurement Errors in Business Surveys

multiple (local) providers; they also mentioned company reporters and authorities. Willimack and Nichols (2001) actually labeled people who provide information about the organization “informants” or “reporters” because of their role as proxy respondents for the organization. Jones et al. (2005) distinguished between two types of respondent actors according to their activities: one type carries out comprehension, retrieval, and judgment, and the other communicates response and releases data. However, the current review identified only two definitions of a respondent and his or her different roles in business surveys, which is “to locate the source of the information and to provide it” (Goldenberg et al. 1993, p. 290) or “to supply the requested information, either by accessing the business’s information system or by relying on personal or other knowledge” (Biemer and Fecso 1995, p. 258).

In the QST study, several patterns with regard to the roles and number of people involved in the QST response process could be discerned. Most have already been described in the literature. A pattern that had been mentioned previously by O’Brien (2000) is the involvement of people on a contractual basis or from an accounting firm in the response process, in addition to those from the business itself. The pattern was noted in small- and medium-sized businesses. In the case of outsourced services the source of measurement error is located outside of the business and may cause correlation between measurement errors of different businesses within a single survey, as accounting firms often work for similar businesses.

Large businesses, on the other hand, generally had several people within the organization involved in the response process. Figure 2 presents the location and relationships between people involved in the QST response process in one business. The contact person was a head in the sales department. She copied the questionnaire upon receipt and sent the copy to the accounting department, where two colleagues shared it. She received the accounting and tax data from one of them, who had collected the tax data from the other. She requested employment items from a third colleague by phone. In addition, she used the data from her own and other reports to prepare turnover breakdowns. She regularly contacted the survey staff to arrange to postpone the deadline. She sought the signature of her supervisor for the sake of formality.

As this example indicates, the dispersion of information may require the involvement of more than one person in the response process, which raises questions as to the manner in which the survey instrument is transferred from one person to another and the communication issues associated with that. This scenario has already been mentioned in the literature with respect to the various participants and their roles in the response process. This study focused on the transfer of the survey instrument in more detail and revealed a variety of approaches (see Figure 3). Some people delivered both the questionnaire and the instruction booklet to the colleagues who provided data for some questions; however, they did so only for the first time or when they noticed changes. Others supplied colleagues only with a copy or a copied page of the questionnaire, and still others simply e-mailed excerpted questions. Some let their colleagues know only their own interpretations of the questions.

Considering the differences in roles and differential exposure to the survey instrument, it is impossible to label all the people involved in the response process as respondents. The two available definitions of respondents in business surveys focus on the provision of requested information and on the applied mechanism (i.e., the acts of locating and

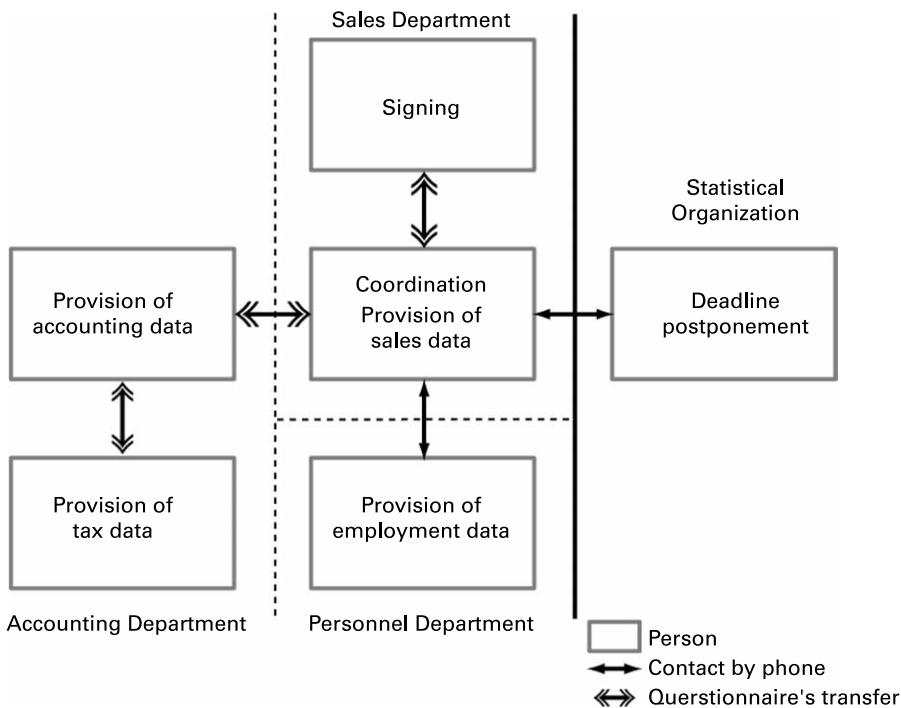


Fig. 2. Complexity of the QST Response Process in a Large Business

retrieving). None of them establishes the criteria that would separate those responding to a questionnaire from mere data providers. For instance, it is not unambiguous whether a person who receives a survey request through another person still qualifies as a respondent, or whether a person who delegates all retrieval tasks to colleagues but records the data on the questionnaire form qualifies as a respondent, or what exactly constitutes the act of providing or supplying information.

To tackle these issues, herein a respondent to a business survey is someone who provides data with the particular purpose of answering a survey question. Providing data is a necessary qualification for being a respondent, while contact with the survey instrument is a sufficient condition and differentiates a mere data creator or provider from a respondent. The provision of data means recalling data from one's memory, retrieving data from the business records, or collecting data from other people who recall the data from memory and/or retrieve them from the business records.

The contact with the survey instrument has to be sufficient to enable the person to autonomously identify the elements that are essential for comprehending the survey question and for judging the adequacy of the response. Exposure to a specific survey question (or label) as given in the questionnaire is the minimal contact that still qualifies a person as a survey respondent (see the dashed line in Figure 3). For example, in Figure 2 this means that the two people in the accounting department qualify as respondents because of their direct exposure to the survey questions, while the person in the personnel department may be treated only as a data provider depending on the interpretation of the contact person who is the respondent for employment data.

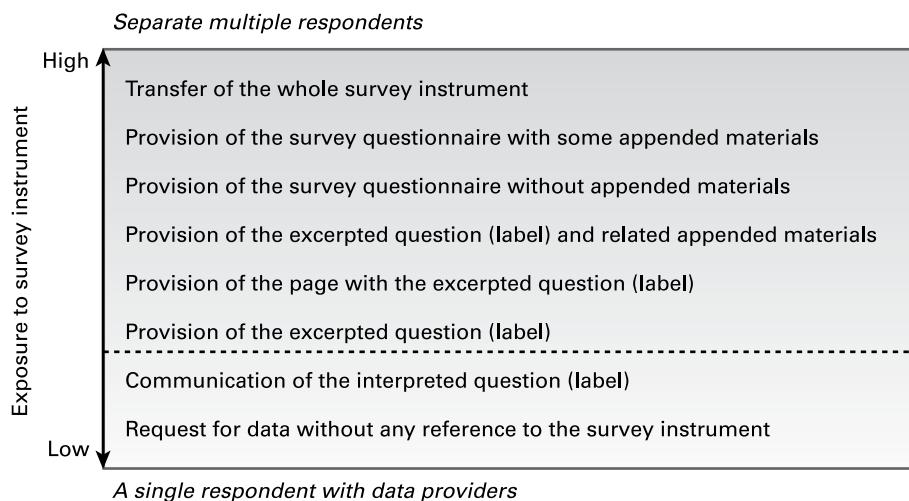


Fig. 3. Differential Transfer of the Survey Instrument among People Involved in the Response Process and the Resulting Differential Exposure to the Survey Instrument. (Note: Business surveys often use labels instead of real questions)

Several people may thus participate in the business response to the survey request. Among these business participants respondents are the most prominent – but still just one – group of people occupied with the survey task in the business. Other groups include response coordinators, data providers, authorities, and gatekeepers (e.g., boundary-spanning units, receptionists), and they may all be a source of measurement errors in business surveys, although their impact varies considerably.

3.2. Business Environment

As mentioned in the introduction, the literature recognizes that business participants operate with the records or the information system within a setting, organization, or environment, which may all be treated as sources of measurement errors. Edwards and Cantor (1991) and Lorenc (2006) dedicated their attention to the details of the record formation process, while Willimack and Nichols (2001) concentrated on the factors of record formation. Biemer and Fecso (1995) and Sudman et al. (2000) restricted themselves to enumerating problematic issues like multiple information systems. The QST study analyzed the record formation and resulting business records in accounting, as accounting data are among the essential ingredients of key economic indicators.

3.2.1. Business Records

The process of record formation in accounting includes book entries and their verification. Some larger businesses reported explicit deadlines for every phase, while smaller businesses often focused on administrative deadlines, apparently using data just to fulfill the legal requirements of bookkeeping, paying taxes and contributions, and so on. Quite a few businesses paced the process of record formation so as to comply with the submission deadline for the VAT tax return, which is set to the last day of the subsequent month for monthly returns. The QST deadline was usually set two weeks earlier (the 15th of the

month). Targeting a later deadline for final data is not problematic *per se*, but it does imply a greater inaccuracy of preliminary data, which may be then be used in survey questionnaires.

To illustrate, a respondent from an accounting firm explained that businesses made the book entries themselves and that she had to wait for the business to start with her work; she could begin verifying entries and preparing the VAT tax return only afterward, usually around the 15th or 20th of the month. In fact, a study of accounting firms in Slovenia showed that documentation disorder is among the biggest problems in client relationships (Horvat 2003). In contrast, a respondent in a company where accounts were checked daily and the books were closed one week after the end of month stated, “All these printouts will be the same down to the decimal point. They will be the same at any time.”

A consistent pattern was discovered in visits to all four foreign-owned businesses. All the businesses had tough requirements for internal reporting, which resulted in accounting records being up to date. Three were small businesses, nevertheless their monthly reports to the parent company were due in the first two weeks of the subsequent month, which was quite exceptional compared to the other units in the sample. One respondent was employed at an accounting firm that delivered services to a foreign-owned company and hence worked with several companies. She stated: “The first 15 days are the most stressful in this company. Everything has to be closed for the previous month, the sales report, exchange differences, depreciation, VAT, everything. . . Almost as if you closed the books.”

The architecture of records in businesses studied ranged from simple to complex. One small business did not even use profit and cost centers to allocate revenues and costs, although it was involved in three distinct business activities. Some businesses kept only the subsidiary books of accounts, while their accounting firms had access to the general ledger. A complex architecture of records was identified in large businesses; some respondents in these businesses mentioned that a detailed accounts chart was used for accounting records and/or that detailed data were available in a commercial database. A wealth of data could also be found in a small foreign-owned company; the respondent stated that the monthly report to the parent company required a lot of data. The degree of complexity influenced the problems encountered during the questionnaire’s completion; simpler architecture often resulted in aggregated data and the lack of necessary detail, while a more complex architecture resulted in problems of recategorization when detailed data had to be grouped to the requested survey categories.

Accounting records play a distinct role among business records because they translate the business activities into financial terms. A respondent described how sales activities were documented with invoices that were then entered into a commercial database. Afterward, the book entries were transmitted to the current book of accounts, to control invoice payments, and to the general ledger, to summarize the sales and their payments, which served as a basis for various reports.

These observations show that business records vary across companies with respect to three characteristics:

- Quantity of data representations
- Profusion of recorded data
- Variability in time.

The quantity of data representations denotes that the same phenomenon may appear in various forms and locations. It is closely related to the organization of work, which is usually more specialized in larger companies. In this study, for instance, monthly sales could be represented by: handshakes and merchandise deliveries; a pile of invoices; book entries in the commercial database, the current book of accounts, and the general ledger; internal reports to the sales director and director general; and external reports to the tax authority. The same item may thus be retrieved from several sources, which differ with respect to their access, reliability, detail, context, and so on. The profusion of recorded data refers to the amount of data about a business activity registered in business records, particularly the number of recorded variables and classification plans. In the study, for instance, some companies had no data about a business customer, while others collected certain information (e.g., name, address, VAT number, region, main activity, institutional sector, and size class). The profusion of recorded data was much higher in larger businesses; in smaller businesses people often kept much of the data in their heads. The variability in time indicates the changeability of a piece of data through time, which depends on the timing and accuracy of record formation. This can be expressed in terms of a deviation from the final value and the time lag necessary to achieve the final value. In this study, for instance, the monthly turnover in the general ledger sometimes amounted to zero until the first entries were made for that month with a weekly time lag. The best identified approach consisted of daily book entries and many promptly executed checks. Unfortunately, many businesses were lagging behind with entries and verification and did not perform many checks.

3.2.2. Beyond Business Records

Because the information is dispersed across the organization as a result of the division of labor, establishment of branch plants, subsidiaries, and so on (Tomaskovic-Devey et al. 1994), several people may need to be involved in the response process (see Figure 2). However, people faced with the survey task not only look up the records but also sometimes turn to other people for information. For instance, a respondent in an accounting firm sought additional information from the client (the business included in the QST sample) before deciding which data to report in the QST questionnaire. Another respondent in a large business had to match turnover data classified by its own commercial groups with the requested commodity groups. However, headings provided in the business records were not informative enough for those who were unacquainted with the details of the company's product range. The respondent decided to turn to the sales department for explanations of some major headings but did not feel comfortable requesting so much information to thoroughly understand all relevant headings. In contrast, the same respondent was completely satisfied with the monthly personnel report and did not seek additional information from colleagues. Observation of the response process detected data verification ("The shop has 60 square meters, doesn't it?"), short exchanges of opinion, and questions about the retrieval ("Is 21 or 24 their cost center?") with colleagues who did not otherwise participate in the process of completing the QST questionnaire.

These examples show that business participants are not the only people in the organization who may have an effect on the response process and the occurrence of measurement errors. There is a thin line between those who participate in the process and those who are not engaged in it. The criteria for distinguishing these two groups are

activeness and participation. Business participants are active in the response process, in contrast with the passive others whose activities remain unaffected by the survey request. Although the activities of both groups may influence the survey response process and its outcome, the effect of the survey response process on these groups is differential: it is noticeable in the case of business participants, and negligible, at least in the short term – if it has any influence at all – in the case of others.

The previously mentioned examples also show that business records may represent only a part of the information support available to participants in the business survey response process: other people may be another source of support. The business information system is thus broader than business records, as it encompasses data in business records and from other people. The system typically reflects the organizational structure that supports business activities, including statistical reporting (for an example of the organizational structure that contributed to the complexity of the QST response process in a large business, see Figure 2). In the QST study, the complexity also increased when an organization outsourced at least part of its accounting function, which meant that the business's information system had to be linked to another system in the accounting firm. This link may represent another source of measurement errors, as a result of asymmetrical information and a lesser motivation to carefully complete the questionnaire (there is a fee for accounting services, and businesses were unwilling to pay extra for statistical reporting, which is an organizational source of measurement error). The QST study also pointed to other well-known organizational matters, such as the existence of (informal) policies on survey participation and concern with data confidentiality, as well as the less frequently discussed issue of evidence-based decision-making. As already mentioned in the previous section, foreign-owned businesses had tough requirements for internal reporting. As a result, the businesses appeared to show a particular respect for data and their quality.

To capture all relevant sources of measurement errors, it may be necessary to go beyond the organization to the business environment. For instance, the QST study pointed to the importance of the institutional environment, which sets the regulatory framework and various standards for the functioning of the economy, including those concerning statistical reporting, standards in accounting, and so on. Figure 4 summarizes the sources of measurement errors that arise in a business environment.

Answering questions in business surveys is a typical business task (O'Brien 2000; Willimack and Nichols 2001). Biemer and Lyberg (2003) explicitly mention interview setting as a source of measurement error in surveys and refer to the immediate environment in which the response process takes place. In business surveys such as the QST, the physical setting can be an office with file cabinets; shelves full of binders, books, and journals; a desk with a computer, a telephone, and an adding machine; and so on. However, it may be more important to consider what is behind this physical appearance (e.g., retrieval options in the general ledger, access to the accounting information system, availability of personnel reports) as well as other cultural factors (e.g., business policies and priorities, respect for standards). The setting is therefore inseparably linked to the response process and may encompass various sources of measurement error in the business environment that go well beyond business records (see Figure 4). The sources of measurement error, however, are not only found on the business side; some of them are also related to the survey. Survey-related sources are presented in the following sections.

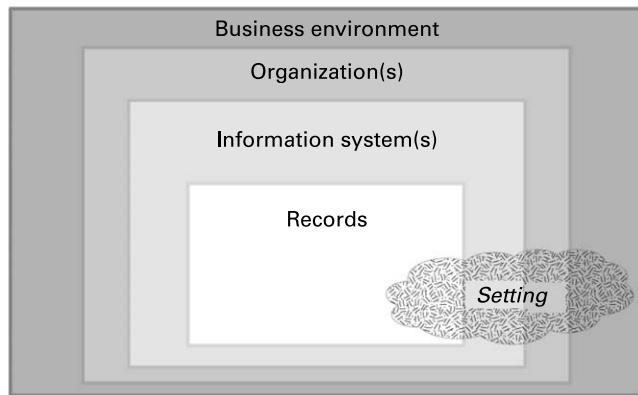


Fig. 4. Sources of Measurement Errors in Business Environment

3.3. Survey Staff

The literature on measurement errors in surveys often points to interviewers as sources of measurement errors. Data collection in business surveys is prevalently self-administered, so there is no typical interviewer's task of asking questions and recording answers. Nevertheless, survey staff occasionally come in contact with people in businesses (Sudman et al. 2000). In the QST study, some questionnaire administration staff carried out nonresponse follow-up, others tried to resolve editing failures by contacting businesses, subject specialists were available for consultation, and so on. These contacts usually contributed to the prevention or reduction of measurement errors as well as unit and item nonresponse. For instance, a respondent called the survey organization while first completing the questionnaire to verify the appropriateness of the approach that the previous respondent had used (i.e., whether to include own products in the merchandise category). The survey staff explained that the approach was not correct and helped set up the correct one. Nonetheless, the study provided some evidence that the survey staff involved in the response process could also become a source of measurement error.

Such evidence, for instance, came from those businesses that contacted survey staff for advice on a survey question that asked for the breakdown of turnover by buyer. Many businesses claimed that their business records did not have readily available data to answer the question and so used rough estimates based on familiarity with the business activity or prepared approximations or good estimates for one period of time and then used the same percentage breakdown for subsequent periods. A few respondents stated that the survey staff agreed with such an approach when asked for advice, which survey staff acknowledged in interviews. This suggests that survey staff may have caused measurement error rather than item nonresponse. Another example refers to a survey question on turnover in foreign markets, which had to be classified as turnover of either services or merchandise. One respondent stated that she agreed to do as the survey staff wanted and classify the business's activities as turnover of merchandise and retained this approach for continuity but was not convinced that it was correct (the approach was indeed wrong). In the previously mentioned cases, the survey staff had a direct impact on the occurrence (or nonoccurrence) of measurement error. Therefore, survey staff are a potential source of measurement error.

3.4. Survey Instrument

The QST study confirmed the importance of a good survey instrument in self-administered surveys, particularly given the complexity of measuring economic aspects, for which many concepts may be elusive. The study revealed many known issues with business surveys, such as use of jargon, the mixing of professional and colloquial language, questions separated from instructions, lack of instructions, deficiencies in the questionnaire design, unclear question intent, and so on. In this sense, it is possible to list specialists for questionnaire content and design as sources of measurement error as far as they are involved in the questionnaire design and evaluation processes (Esposito 2003). As expected on the basis of previous findings (e.g., O'Brien et al. 2001), the study also showed that the use of two booklets (i.e., for instruction and classification) appended to the questionnaire was infrequent, which proved detrimental in some cases because the booklets contained information essential to the provision of accurate answers.

Moreover, other materials appended with or linked to the questionnaire can be sources of measurement error. Too often, however, only the questionnaire receives attention while the other parts that support data collection are rarely discussed. Because the aim of the typology is to cover all known sources of measurement error, it seems more appropriate to use the holistic approach and focus on the survey instrument as a whole.

3.5. Survey Characteristics and Procedures

Among all survey characteristics and procedures, the mode of administration is usually discussed as a source of measurement error. The mode of administration refers to the medium used to contact respondents and collect their answers to survey questions (Biemer and Lyberg 2003). The usual concern in business surveys relates to measurement errors that arise from self-administration of data collection, which accommodates data retrieval but places greater weight on the performance of the survey instrument and a greater burden on respondents. This was also the case with the QST, a typical mail survey using a paper questionnaire. When problems occurred in the response process, respondents could contact QST staff by phone or e-mail, but few actually did this. Many reported that they had been perplexed when completing the questionnaire while the survey instrument had not provided the necessary information and explanations for all uncertainties. This contributed to the occurrence of measurement error and the negative experience of completing the questionnaire that some respondents felt.

The usual procedure of contacting only key businesses and those units with missing data and major inconsistencies in data meant that many, presumably smaller, measurement errors were deliberately left in the data. An exceptional procedure was set up when the value-added tax was introduced to the tax system, and the survey staff contacted businesses in advance to clarify the upcoming changes in statistical reporting, which likely helped reduce measurement error.

Other survey characteristics influenced the response process of the QST and the resulting measurement errors. The recurrent administrations of the QST to the same business shaped this process considerably (Bavdaž in press). In some cases the routine completion of the questionnaire resulted in the perseverance of measurement errors, which is consistent with previous findings (Sudman et al. 2000). However, this also creates

opportunity for new measurement errors. For instance, a respondent had problems shifting from reporting trade items to reporting production items when the business's activity slightly changed from purchasing, repackaging, and selling roasted coffee to purchasing raw beans, roasting, repackaging, and selling coffee. From this point of view it is not impossible (though it is rather improbable) that business activities change in such a way to make measurement errors become smaller or even disappear.

It should be noted that the effect of recurrence may vary with the frequency of recurrence (i.e., the periodicity of the survey). Although the empirical study did not foresee comparisons with surveys of other periodicities, there were some indications of the effect of periodicity on a recurrent response process. For instance, a respondent from an accounting firm who had already participated in many surveys commented that quarterly repetitions were distant enough to forget the details and thus require substantial cognitive effort.

Business surveys can be mandatory or voluntary. Participation in the QST was required by law, which means that nonresponse, a late response, or an inaccurate response could all have been sanctioned. The SORS was only partially successful in using its legal mandate in the QST: fewer than half of the businesses returned the QST questionnaire before the first mail reminder and one-third of the QST questionnaires required some sort of treatment or intervention, mainly because of incomplete and inconsistent data. Nevertheless, the final response rates were high (i.e., constantly higher than 90%). Enforcing a response was therefore less problematic than enforcing a timely and accurate response, which is consistent with previous findings (Willimack et al. 2002). Some reduction of nonresponse error may have resulted in increased measurement error. For instance, some respondents described their reluctance to participate in the survey as well as their abatement after SORS had contacted them intensively; they returned the questionnaire but took the path of least resistance when filling it out (e.g., providing very rough estimates, ignoring some questions or breakdowns even if applicable). In addition, there was some evidence that the sponsor could have an effect on respondents' perceptions. For instance, the central bank is an authorized producer of official statistics, as is SORS. One respondent, however, completed a survey sponsored by the central bank right after finishing tax reports and left the SORS-sponsored survey for later completion.

4. Discussion of Typology Implications

The proposed typology has several implications for detecting, reducing, and preventing measurement errors in business surveys. It first distinguishes different people involved in the survey response process and labels them according to their roles in this process and their relationships to the survey instrument. Distinction between a respondent and a data provider thus reflects the potential effect of the survey instrument. The implication of this distinction is that the survey instrument must perform proficiently for those who are exposed so that they are adequately equipped for further data collection from their colleagues. A survey instrument that does not communicate all the necessary elements of a question to respondents is likely to produce a measurement error in the collected items. When the people who provide data to respondents are not exposed to at least some part of the survey instrument, they also might lose the chance of detecting critical information themselves and thus influencing the adequacy of the data.

In the operational context this means that statistical organizations should investigate how to adapt instruments for inappropriate respondents and how to convince the recipient of a survey instrument to identify (better) respondents and deliver a sufficient portion of the instrument to those persons. This seems to be particularly important when respondents use similar but not identical concepts (e.g., employment based on payroll versus employment based on personnel records) or lack adequate knowledge (e.g., of accounting concepts and terminology). Although respondents are the most prominent group of business participants, it may be convenient for statistical organizations to target other business participants, particularly gatekeepers and authorities, each group with a different strategy because of their different roles in the survey and in the business.

The typology also puts forward several aspects of the setting in which the business response process occurs, from narrowly determined business records to the broadly defined business environment. Characteristics of business records provide important indications for response outcome: the quantity of data representations shows the richness of data sources and potential respondents; the profusion of recorded data suggests what researchers cannot ask without risking measurement error; and the time variability implies how much measurement error researchers incur at different points in time. The reference to the business environment as a source of measurement errors, on the other hand, indicates that some sources simply cannot be dealt with in the framework of a single business survey. Such sources require long-term, well-chosen strategic endeavors on the part of survey organizations if they are to be influenced to their advantage. In the short run, it seems more effective to act on those sources of measurement errors that are under direct control of the survey organization: survey staff, survey instrument, and survey characteristics and procedures.

Previous research has used survey staff in business surveys as a source of information on potential measurement errors (Rowlands et al. 2002; Giesen and Hak 2005). The typology suggests that they should also be viewed as a potential source of these errors – despite the sensitivity of this issue. As a consequence, appropriate training should be developed and organized for all survey staff, not only for interviewers.

The typology stresses the importance of a holistic approach to the survey instrument, as in business surveys the survey instrument rarely consists of a survey questionnaire only. Not using of the entire instrument when its use is expected may be a source of measurement errors, as is the use of the instrument when the instrument fails to communicate the necessary or correct information. This implies that efforts should be directed to improving the whole survey instrument and its usage.

The mode of administration is a recognized source of measurement errors in surveys. The typology suggests that other survey characteristics (e.g., recurrence and periodicity, legal mandate, sponsor) can contribute to the occurrence of measurement error. The first participation of a business or a person in a self-administered recurring business survey calls for intensive and extensive support and control of the statistical organization, because it is then that the questions about survey completion emerge and require answers. This first attempt to complete the survey task sets the standards and thus determines the quality of all subsequent completions of the questionnaire. This also seems important for the enforcement of any existing legal requirements. The question of the sponsor effect requires action in the long term.

There is one controversy with regard to legal requirements and their effect on measurement error. On the one hand, businesses are required by law to provide accurate data to statistical organizations. On the other hand, the statistical organizations usually concede the use of estimates when precise data are unavailable. This concession is legitimate, as good estimates at the business level usually suffice for good population estimates. However, this calls for some rules or guidance on what constitutes a good estimate and how much measurement error is still acceptable. The lack of such information may then serve as an excuse for making measurement errors and for providing bad data to the statistical organization.

Last, the typology includes survey procedures as a source of measurement error. The usual self-administration of data collection in business surveys contributes to the fact that most procedures are reactive (i.e., established to react to unwanted situations, such as lack of adequate response). To reduce measurement error more effectively it seems necessary to move to a proactive approach and use every opportunity to enrich the survey experience with positive and instructive messages.

The typology is applicable to businesses of all sizes even though the sources of measurement error may manifest in different ways (e.g., through unavailability of qualified staff and adequate records in smaller businesses, through multilayered relationships and sophisticated information systems in larger ones).

5. Conclusion

This article proposed a typology that aims to cover and systemize all known sources of measurement error in business surveys. The typology extended the currently used category of respondents to include all business participants who are involved in the response process and who need to be distinguished from the business environment, which remains unaffected by the survey task. It also extended the category of interviewers to include all survey staff and the category of mode to encompass various survey characteristics and procedures. It also calls for more focus on the survey instrument as a whole. Such a typology may prove a useful tool for addressing all potential sources of measurement error in a particular business survey and for synthesizing research findings across surveys. It must be noted, though, that because the research underlying the typology is mainly limited to governmental business surveys, it is important that it is further evaluated with regard to business surveys that have other characteristics. More research is also warranted with regard to the kind of effect that specific sources and their interactions have on measurement error in business surveys.

6. References

- Bavdaž, M. (2009). Conducting Research on the Response Process in Business Surveys. *Statistical Journal of the IAOS*, 26, 1–14.
- Bavdaž, M. (2010). The Multidimensional Integral Business Survey Response Model. *Survey Methodology* (in press).
- Biemer, P.P. and Fecso, R.S. (1995). Evaluating and Controlling Measurement Error in Business Surveys. In *Business Survey Methods*, B.G. Cox, D.A. Binder, B.N.

- Chinnappa, A. Christianson, M.J. Colledge, and P.S. Kott (eds). New York: Wiley-Interscience.
- Biemer, P.P. and Lyberg, L.E. (2003). Introduction to Survey Quality. Hoboken, NJ: Wiley-Interscience.
- Edwards, W.S. and Cantor, D. (1991). Toward a Response Model in Establishment Surveys. *Measurement Errors in Surveys*, P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (eds). New York: Wiley-Interscience.
- Esposito, J.L. (2003). A Framework Relating Questionnaire Design and Evaluation Processes to Sources of Measurement Error. 2003 FCSM Conference Papers. <http://www.fcsm.gov/03papers/Esposito2.pdf>
- Gerber, E.R. (1999). The View from Anthropology: Ethnography and the Cognitive Interview. In *Cognition and Survey Research*, M.G. Sirken, D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur, and R. Tourangeau (eds). New York: Wiley-Interscience.
- Giesen, D. and Hak, T. (2005). The Response Process Model in Business Surveys: Lessons Learned by Using a Multi-Method Approach. 2005 FCSM Conference Papers. http://www.fcsm.gov/05papers/Giesen_Hak_AXB.pdf
- Goldenberg, K.L., Butani, S., and Phipps, P.A. (1993). Response Analysis Surveys for Assessing Response Errors in Establishment Surveys. Proceedings of the International Conference on Establishment Surveys ICES, American Statistical Association. <http://www.bls.gov/ore/pdf/st930240.pdf>
- Groves, R.M. (1989). Survey Errors and Survey Costs. New York: Wiley-Interscience.
- Groves, R.M., Cantor, D., Couper, M., Levin, K., McGonagle, K., Singer, E., and Van Hoewyk, J. (1997). Research Investigations in Gaining Participation from Sample Firms in the Current Employment Statistics Program. Proceedings of the American Statistical Association, Surveys Research Methods Section, http://www.amstat.org/sections/srms/Proceedings/papers/1997_047.pdf
- Groves, R.M., Fowler, Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E., and Tourangeau, R. (2004). Survey Methodology. Hoboken, NJ: Wiley-Interscience.
- Hak, T., van der Veer, K., and Jansen, H. (2008). The Three-Step Test-Interview (TSTI): An Observation-based Method for Pretesting Self-completion Questionnaires. *Survey Research Methods*, 2, 143–150.
- Hansen, M.H., Hurwitz, W.N., Marks, E.S., and Mauldin, W.P. (1951). Response Errors in Surveys. *Journal of the American Statistical Association*, 46, 147–190.
- Horvat, R. (2003). Računovodski servisi v Sloveniji (Accounting firms in Slovenia). IKS, 30, 10, 21–33. [In Slovenian].
- Jones, J., Rushbrooke, J., Haraldsen, G., Dale, T., and Hedlin, D. (2005). Conceptualising Total Business Survey Burden. *Survey Methodology Bulletin*, 55, 1–10.
- Kruskal, W. (1991). Introduction. In *Measurement Errors in Surveys*, P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (eds). New York: Wiley-Interscience.
- Lessler, J.T. and Kalsbeek, W.D. (1992). Nonsampling Error in Surveys. New York: John Wiley.
- Lorenc, B. (2006). 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.
- Lyberg, L. (2003). Quality Improvement in European National Statistical Institutes. Washington Statistical Society, 23 November.
- O'Brien, E.M. (2000). Respondent Role as a Factor in Establishment Survey Response. Proceedings of the Second International Conference on Establishment Surveys ICES-II. Alexandria, VA: American Statistical Association.
- O'Brien, E., Fisher, S.K., Goldenberg, K., and Rosen, R. (2001). Application of Cognitive Methods to an Establishment Survey: A Demonstration Using the Current Employment Statistics Survey. Proceedings of the Annual Meeting of the American Statistical Association [CD-ROM]. Alexandria, VA: American Statistical Association.
- Ponikowski, C.H. and Meily, S.A. (1989). Controlling Response Error in an Establishment Survey. Proceedings of the American Statistical Association, Survey Research Methods Section, http://www.amstat.org/sections/srms/Proceedings/papers/1989_043.pdf
- Rowlands, O., Eldridge, J., and Williams, S. (2002). Expert Review Followed by Interviews with Editing Staff: Effective First Steps in the Testing Process for Business Surveys. International Conference on Questionnaire Development, Evaluation and Testing Methods QDET, 14–17 November, Charleston, SC. http://www.jpsm.umd.edu/qdet/final_pdf_papers/rowlands.pdf
- Sudman, S., Willimack, D.K., Nichols, E., and Mesenbourg, T.L. (2000). Exploratory Research at the U.S. Census Bureau on the Survey Response Process in Large Companies. Proceedings of the Second International Conference on Establishment Surveys ICES-II. Alexandria, VA: American Statistical Association.
- Sukhatme, P.V. and Sukhatme, B.V. (1970). Sampling Theory of Surveys, (Second Edition). Ames: Iowa State University Press.
- Tomaskovic-Devey, D., Leiter, J., and Thompson, S. (1994). Organizational Survey Nonresponse. *Administrative Science Quarterly*, 39, 439–457.
- Tourangeau, R. (2003). Cognitive Aspects of Survey Measurement and Mismeasurement. *International Journal of Public Opinion Research*, 15, 3–7.
- Willimack, D.K. and Nichols, E. (2001). Building an Alternative Response Process Model for Business Surveys. Proceedings of the Annual Meeting of the American Statistical Association, 5–9 August [CD-ROM]. Alexandria, VA: American Statistical Association.
- Willimack, D.K., Nichols, E., and Sudman, S. (2002). Understanding Unit and Item Nonresponse in Business Surveys. In *Survey Nonresponse*, R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little (eds). New York: Wiley.
- Willis, G.B. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage Publications.

Received October 2007

Revised September 2009