

Response Burden and Data Quality in Business Surveys

Marco Bottone¹, Lucia Modugno¹, and Andrea Neri¹

Response burden has long been a concern for data producers. In this article, we investigate the relationship between some measures of actual and perceived burden and we provide empirical evidence of their association with data quality. We draw on two business surveys conducted by Banca d'Italia since 1970, which provide a very rich and unique source of information. We find evidence that the perceived burden is affected by actual burden but the latter is not the only driver. Our results also show a clear link between a respondent's perceived effort and the probability of not answering some important questions (such as those relating to expectations of future investments and turnover) or of dropping out of the survey. On the contrary, we do not find significant effects on the quality of answers to quantitative questions such as business turnover and investments. Overall, these findings have implications for data producers that should target the perceived burden, besides the actual burden, to increase data quality.

Key words: Response burden; data quality; business surveys.

1. Introduction

Policymakers need high quality and detailed information on firms' decisions and performances to monitor the state of the economy and to assess the effectiveness of their policies. On the other hand, participating in a survey is a cost for businesses. It does not lead to any obvious financial return and it takes time away from profitable activities. If they receive frequent inquiries or if the information to be provided is burdensome, businesses are likely to give a low priority to these requests, especially if they are not mandatory. This means that they may refuse to collaborate or, even if they agree to participate, they may provide low-quality data (with low timeliness, with a high number of missing items or with measurement errors). It is therefore crucial to measure and monitor business response burden. This principle is clearly stated in the Quality Assurance Frameworks produced by several international organizations. The European Statistical Associations states that "response burden must be proportionate to the needs of the users and must not be excessive for respondents. The statistical authorities should monitor the response burden and set targets for its reduction over time." Along the same line, the United Nations require statistical agencies to "choose data sources taking into account accuracy and reliability, timeliness, cost and the burden on respondents." The literature generally distinguishes between actual and perceived burden (see [Willeboordse 1997](#) and [Haraldsen et al. 2013](#) for a comprehensive overview and discussion of this topic). While actual burden refers to objective measures of the complexity of the survey, perceived

¹ Bank of Italy, Via Nazionale, 91, Rome, 00184, Italy. Emails: marco.bottone@bancaditalia.it, lucia.modugno@bancaditalia.it and andrea.neri@bancaditalia.it

Acknowledgments: The views expressed are not necessarily shared by the Bank of Italy.

burden is essentially a subjective measure provided by the respondents. Another relevant distinction made by the literature is between gross and net burden. While the former only considers the costs of survey participation, the latter takes account the benefits enjoyed by respondents for their contribution (such as the feedback of survey results). To the best of our knowledge, in all the empirical studies actual burden is considered in gross terms, probably because the benefits are difficult to measure. There are a few different approaches for measuring such burden. Ideally, it should be calculated as the product of three factors: the number of people involved in the survey, the time they spent and their average hourly salary. Yet, this measure is rarely used, first because it depends upon a very noisy estimate of the cost/hour for the respondent, second because it is burdensome to collect. Generally, it is calculated by multiplying the number of completed questionnaires by an estimate of the average time required for completing and submitting the response, multiplied by the number of survey cycles during the year. Sometimes the burden is computed using the total sample size rather than the number of completed responses. In a minority of cases, it is measured in financial terms by multiplying the hours spent by the average hourly cost of respondent's time (Snijkers et al. 2013; Bavdaž et al. 2015). The main limit of actual burden is that the positive effects of survey participation can hardly be measured in objective terms. Perceived burden refers to the respondents' assessment of how burdensome they find it to comply with the data request. The importance of perceptions was initially stressed by Bradburn (1978). Respondents' assessment is likely to also include other aspects that may affect the burden that merely time measurement does not take into account. For example, Fisher and Kydoniefs (2001) suggest that respondents' perceptions can be affected by factors such as their motivation and belief in the utility of surveys or the method of data collection, or the item sensitivity. Yan et al. (2019) find that low motivation, difficult recall tasks, challenging survey effort, and negative perception of the survey all directly contribute to respondents' perception of burden. Dale et al. (2007) suggest to use two qualitative and simple questions to measure perceptions. The first relates to the perception of time taken while the second relates to the perception of the overall burden. They also recommend asking questions about which conditions mainly make the survey burdensome. Summing up, the measurement of perceptions is important for two main reasons. First, it enables us to capture more factors relating to the burden other than those relating to time and money. Second, it is an easy way to also consider the positive effects of survey participation. The final rate provided by the respondents is the balance between burdens and gratifications (Haraldsen 2004). Measurement of burden can be undertaken using different approaches. It can be collected directly from business survey participants including a small set of questions at the end of the questionnaire, or subsequently by follow-up contact with a sub-sample of the surveyed population. It can also be estimated indirectly, as a product of recordkeeping studies, tests or experts' valuations. Practices relating to the response burden measurement are likely to vary a lot across institutions. Dale et al. (2007) provide an overview of the questions used by several statistical agencies to collect information on the response burden. Bavdaž et al. (2015) conducted a survey of 41 national statistical institutes (NSIs) from 39 countries. They find large heterogeneity in measurement practices not only between institutes but also within them. Most institutes do some kind of measurement of actual burden, while perceived burden is less frequently assessed. This heterogeneity raises several issues. Probably, the

main one is which measures of response burden should be collected for monitoring data quality.

The current study contributes to the literature by addressing two research questions. The first one is to what extent actual burden is associated with perceived burden. There are not many studies dealing with both issues simultaneously, probably because of data limitations. Yet, for data producers it is useful to know to what extent such measures are associated: in the extreme case in which actual and perceived burden are perfectly correlated, it would not be necessary to collect both, reducing the overall burden on respondents. The second and main research question is whether perceived burden has a direct impact on data quality once many other potential associates are controlled for (including actual burden). Even if respondent burden is widely recognized to have an impact on quality, to the best of our knowledge, only a few articles attempt to empirically validate this statement and the evidence provided is often indirect. For example, [Haraldsen and Jones \(2007\)](#) use the number of values corrected during editing; whereas [Giesen \(2012\)](#) takes the timeliness of the response to evaluate the effects of perceived burden on response behavior. For data producers, it is important to know which factors affect data quality in order to identify strategies to improve it.

We draw on a unique and rich dataset made by two main business surveys conducted by the Bank of Italy, which collect information on both actual and perceived burden. Actual burden is measured by five indicators: number of pages of the questionnaire, number of questions, number of people involved in the survey, use of external consultancies and the average completion time. Perceived burden is measured by a qualitative question asking respondents to rate their effort to complete the questionnaire. It is also worth mentioning that these surveys are not mandatory. Because of that, respondents' burden may have very negative effects on data quality. Our research plan consists of first studying the relationship between perceived and actual burden and their dynamics over time. We then provide evidence of the relationship between perceived burden and data quality while controlling for actual burden and other factors. The structure of the article is as follows. Section 2 describes the conceptual framework we use in our analysis, while Section 3 presents our data. In Section 4, we provide an empirical analysis of the association between our measures of actual and perceived burden. In Section 5 we investigate the relationship between perceived burden and several measures of data quality. Section 6 concludes.

2. A Conceptual Framework for Response Burden and Data Quality

A first model of survey burden was developed in 1978 by Bradburn. He suggests a definition of response burden consisting of four elements: interview length, required respondent effort, frequency of being interviewed and the stress of psychologically disturbing questions, which may be asked ([Bradburn 1978](#)). He also pointed out that it is necessary to carefully focus on the respondent's subjective perceptions of the time and effort required to fill in the questionnaire. [Fisher and Kydoniefs \(2001\)](#) propose a model in which response burden is a combination of 32 elements that can be grouped in three categories: respondent burden (personality traits and attitudes of the respondents), design burden (characteristics of the survey such as duration of the interview and the wording of

the questions) and interaction burden (what happens when respondents with certain characteristics are confronted with a survey that has certain properties). Haraldsen (2004) and Haraldsen et al. (2013) further elaborate on the previous models by highlighting that it is crucial to assess whether the perception of the burden outweighs the positive aspects of the survey. The key variable to monitor for data quality is the final perceived balance between negative and positive aspects. In this article, we adopt the Total Business Survey Burden Model (TBSB) discussed by Jones et al. (2005) and Dale et al. (2007). Within this framework, perceived burden is described as an intermediate variable between data quality and three elements: survey design components, respondent characteristics and contextual business factors. Survey design refers to aspects such as the data collection method, the communication strategy, the length and content of the questionnaire. A respondent's characteristics mainly relate (her/his) cognitive ability, motivation and capacity to collect the necessary information. Contextual factors are linked to the organization of the businesses (which may facilitate or prevent the collection of information) and to its strategies. Two important points can be drawn from the total business survey burden model. First, perceived burden is the only factor that directly affects data quality. All the other factors relating to the survey design, respondent characteristics and the general context only contribute to shaping this burden. Second, perceived burden is a more general concept than actual burden since it may originate from many sources other than those related to the time and/or money it takes to comply with a survey request.

This framework can be applied in order to gain a better understanding of the effects of perceived burden in two important phases of the survey: the recruitment of respondents and data collection.

In the initial phase of recruitment, the decision of businesses to participate depends on their perceptions of anticipated burden. These perceptions are based on the information they are provided. In web surveys, for instance, the advertised interview length is negatively associated with response rates (Galesic and Bosnjak 2009; Yan et al. 2010a) and it is positively associated with break-offs (Galesic 2006). A similar result is found for mail surveys (Edwards et al. 2002). In the case of longitudinal surveys, the expected burden is also determined by respondents' prior experience of the survey. For instance, Bergman and Brage (2008) find that respondents with a negative experience are less willing to accept new survey requests. Research has shown that other factors may mediate or worsen this initial perception. For instance, interest in the survey topic is found to be an important contributor to the decision to participate (Groves et al. 2006). Moreover, a positive opinion the sponsor also generally acts as a mediator of the estimated burden and increases the probability of participation. Surveys sponsored by government agencies have higher response rates than surveys sponsored by non-government agencies (Presser et al. 1992). Tomaskovic-Devey et al. (1994) suggest that the expected burden is also related to the overall context. For example, the organizational practices and divisions of labor and information may facilitate or inhibit the assembly of relevant knowledge to reply adequately to survey requests. Establishments with greater performances and financial resources are more likely to have the organizational slack to complete a survey.

When respondents start filling in the questionnaire, their initial perceptions of burden may change because the questions prove to be more (or less) complex than expected or because of changes in other factors relating to the overall context. Even if a business has

decided to participate in the survey, an increase in perceived burden may result in decisions to break off, to skip a question or to provide poor quality data (Yan et al. 2014). One factor that may contribute to this situation is the complexity of the questionnaire. Long questions or a high number of response options taking longer to answer are more prone to response order effects (Holbrook et al. 2007; Galesic and Bosnjak 2009) and are more likely to induce break-offs (Peytchev 2011) or item nonresponse (Yan et al. 2010b) and less reliable responses (Tourangeau et al. 2019). Other articles focusing on web surveys find that higher dropout rates are associated both with the formal characteristics of questions, for example, the position of questions, the number of questions on a page and poor visual design, and with the characteristics of the respondents (Crawford et al. 2001; Heerwegh and Loosveldt 2002; Galesic 2006). Bavdaž (2010) suggests a model (later expanded by Haraldsen 2013) that links the difficulty to retrieve the requested information to the quality of the responses. The likely outcomes range from exact information to item nonresponse. The alternatives in between are approximations, solid or rough estimates and blunders. Contextual aspects may play a role too. Organizations insulated from their environment and in unregulated environments may have little interest in disclosing information (Tomaskovic-Devey et al. 1994). On the contrary, businesses such as publicly traded firms that are dependent on their environment for resources generally have a higher motive to respond. The current study tests two hypotheses, drawn from this conceptual framework. The first one is that the measurement of perceived difficulty is necessary for data producers since it enables the capture of different information from the one contained in the measures of actual burden (relating time and money). The second assumption is that perceived burden has a direct effect on several dimensions of data quality (such as the propensity to participate in the survey, to provide all the information requested and to give accurate answers) even after controlling for actual burden.

3. Data

The Bank of Italy has a long-standing tradition of conducting business surveys aimed to monitor the economic outlook, to study firms' behavior and to assess the effectiveness of economic policy measures.

The main one is the Survey of Industrial and Service Firms (INVIND hereafter), carried out between the end of January and mid-May, which gathers information on investments, gross sales, the workforce, expectations and other economic variables relating to Italian industrial and service firms with at least 20 employees. It began in 1972 (although the microdata available are those from the wave conducted in 1985) and only covered industrial processing firms with at least 50 workers. From 2002 onward, the sample has consisted of about 4,000 firms, of which around 3,000 belong to the industrial sector and the remaining firms belong to the service sector.

The questionnaire is usually composed of two parts: a core part that collects quantitative information on actual and expected structural characteristics (such as turnover, investments and number of employees) and a monographic section dealing with special topics aimed at a more conjunctural analysis. The core questions are compulsory, in the sense that without a response the whole questionnaire is not considered complete and therefore corresponding data are treated as a unit nonresponse.

It is worth noting that we have access to firm's fiscal identifiers and therefore we have been able to link survey data with other administrative records and in particular with register data on firms' balance sheets. This data linkage is very useful both in order to enrich the survey with some variables not directly measured and to perform ex-post comparison between the same quantities measured from the survey and available from the administrative sources. We use this data linkage in the analysis of measurement error.

A second survey that we use is the Business Outlook Survey of Industrial and Service Firms (BOS), carried out between September and October. It has been conducted since 1993 to respond to short-term economic analysis needs. It mainly collects qualitative information on firms' performance and their future expectations. The target sample size is 4,000 units.

Both surveys are conducted by the local branches of the Bank of Italy. The surveys use a mixed mode approach of self-enumeration (web and interactive pdfs), telephone interviews and face-to-face interviews. In the 2019 wave, the INVIND and BOS surveys achieved the following percentages of questionnaires by mode (INVIND response rate is shown first): web collection (6% and 11%); interactive pdfs (74% and 45%); telephone interviews (13% and 45%); and face-to-face (7% and 1%). Moreover, the participation is voluntary in both surveys.

The two surveys are conducted on approximately the same sample of firms. Those who have participated in past waves (of either survey) and are still in the target population are always contacted for a new data collection wave. A firm can drop out of the sample either because of its choice, because of bankruptcy or because it has fallen below the surveyed size threshold. Larger businesses (with more than 5,000 employees) are always contacted in the next wave. The response rates are around 70% for INVIND and 75% for BOS. A refreshment sample is selected to compensate for attrition to reach the minimum target of 4,000 firms in each wave. As a consequence of this design, the samples used for the two surveys are made of almost the same units: in the period 2015–2019, most firms participated in both surveys in the same year (Table 1).

Our analyses are mainly based on the INVIND survey. We exploit the lower and more stable complexity of the BOS survey for two main purposes. First, we use it for a preliminary investigation on the existence of a 'hardening survey climate' phenomenon in our data. This issue refers to the fact that it is now more difficult to conduct a survey involving people as respondents than it used to be (either in business or household surveys). If this is the case, many results that we observed in our study could be driven by such a phenomenon, rather than by the complexity of the survey. Second, we use the BOS survey to study whether the association between perceived and actual burden across time is different from the one resulting from the INVIND survey.

Table 1. Number of firms participated in INVIND, BOS and in both in the period 2015–2019.

Survey	Year of data collection					Total
	2015	2016	2017	2018	2019	
INVIND only	490	597	491	502	404	2,484
BOS only	552	473	673	504	647	2,849
Both	3,770	3,798	3,717	3,889	3,807	18,981
Total	4,812	4,868	4,881	4,895	4,858	24,314

4. The Measurement of Actual and Perceived Response Burden

Our data enables us to construct five measures of actual burden. The first one is the number of pages of the questionnaire. This measure reflects not only the number of questions and their length but also the need to provide instructions or clarifications to help respondents. The second one is the number of fields to fill in (the number of variables in each questionnaire). In case of multiple choice questions, each response option is considered as a different field, since respondents have to read it, think about it and possibly retrieve the necessary information. Both measures have been reconstructed for both surveys since 1985.

Figure 1 shows the evolution of these two measures over time. The complexity of the INVIND questionnaire has increased overall from 1985 to 2019. In particular, the number of variables more than tripled during the period, with greater growth occurring between the late 1990s and the early 2000s. Since then, the total number of variables in each questionnaire has remained stable, while the number of pages has increased considerably: it has grown more than two times since 2009 and seven times since the beginning of the survey. The marked increase in 2010 was due to the positioning of the instructions for the respondents below each question instead of in a separate document as previously done. This change increased the length of the questionnaire by about five pages in that year alone. The BOS questionnaire has also undergone similar changes, albeit more gradually, particularly in terms of the number of variables (Figure 1).

Starting from the 2017 wave, the INVIND survey also collects information on the number of people who have contributed to the completion of the survey, whether or not external consultancy was required and the time spent filling in the questionnaire. Unlike the previous indicators, these are individual-specific and capture the difficulty personally faced by each respondent.

Our measure of perceived burden is based on a simple question that can be translated as follows: “How would you rate the level of complexity of the survey?”. Response options range from 1 (“low”) to 4 (“excessive”). This question has been present since 2004 for the INVIND survey and since 2010 for the BOS survey. Starting from the 2017 wave, respondents are also asked to rate, with a score from one to ten, the contribution of five

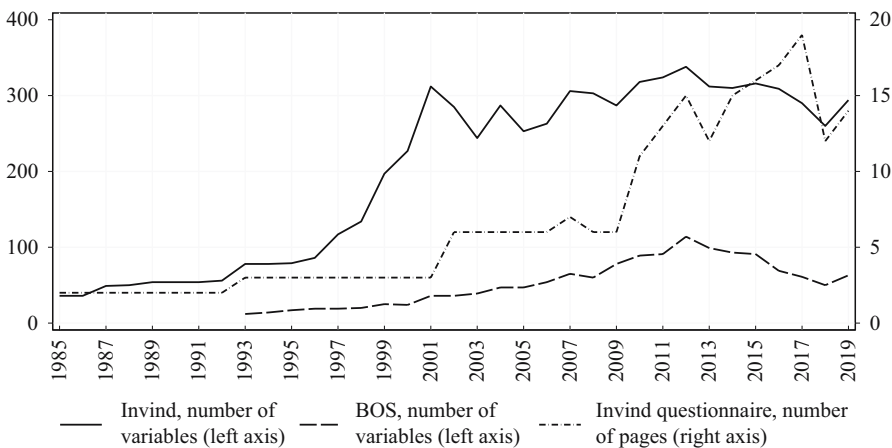


Fig. 1. The complexity of INVIND and BOS questionnaires over time.

factors to response burden, including (1) too many questions, (2) several people involved in answering the questions, (3) the use of unclear terms, (4) not exhaustive response options, and (5) difficulties in choosing the correct answer. Table 2 shows the 2017 to 2019 responses to these questions. Across the years, firms that perceived an excessive level of difficulty indicated higher ratings for the contributing factors of questionnaire length and the need to involve more people to obtain the required data.

Figure 2 shows the dynamics of the perceived response burden in the two surveys. Over the period, the percentage of firms reporting a “high” burden in the INVIND survey is always greater than 35%, but in a cyclical pattern. In the BOS survey, such a percentage decreases from about 15% in 2012 to around 5% in 2019. This descriptive evidence suggests two main results. First, the perceived burden moves quite a lot with the actual burden represented here by the number of variables in the questionnaire. The association between the measures of perceived and actual burden is further confirmed when considering the individual-specific measures of actual burden (the number of people involved, the percentage of external consultants and the time spent). The higher the perceived burden, the higher the actual burden (Table 3).

Second, we don’t find evidence to support the phenomenon of “hardening survey climate” according to which the complexity in conducting surveys increases over time, regardless of the survey’s characteristics. Figure 2 suggests that this phenomenon is not particularly relevant in the case of Italian firms. In the BOS survey, the perception of effort does not increase over time, even if the businesses participating in the survey are almost the same as those participating in the INVIND survey. This suggests that perceived burden is mainly affected by specific idiosyncratic survey factors. In any case, the econometric analyses performed below will take this phenomenon into account, whenever possible, by using a full set of yearly dummies that allows us to control for any possible time fixed effect.

It is worth stressing that since our measures of actual and perceived burden have been collected starting in different years, we will be using different datasets depending on the

Table 2. Average score for each factor disaggregated by the perceived response burden (INVIND 17–19).

Response burden	Too many questions	More people involved	Use of unclear terms	Not exhaustive response options	Difficulties in choosing the answer
2017					
Low	3.2	2.3	2.1	2.1	2.1
Average	4.8	4.2	2.9	3.0	3.2
High	7.0	6.1	4.0	3.4	4.0
Excessive	8.3	7.2	4.9	4.0	4.7
2018					
Low	2.6	2.1	1.8	1.9	1.9
Average	4.4	4.0	3.2	3.0	3.2
High	6.6	6.0	4.3	3.4	4.3
Excessive	8.5	7.1	5.3	3.9	5.3
2019					
Low	3.0	2.4	2.1	2.2	2.1
Average	4.7	4.2	3.2	3.3	3.5
High	6.9	6.1	4.3	3.8	4.5
Excessive	8.6	6.8	5.2	4.3	4.9

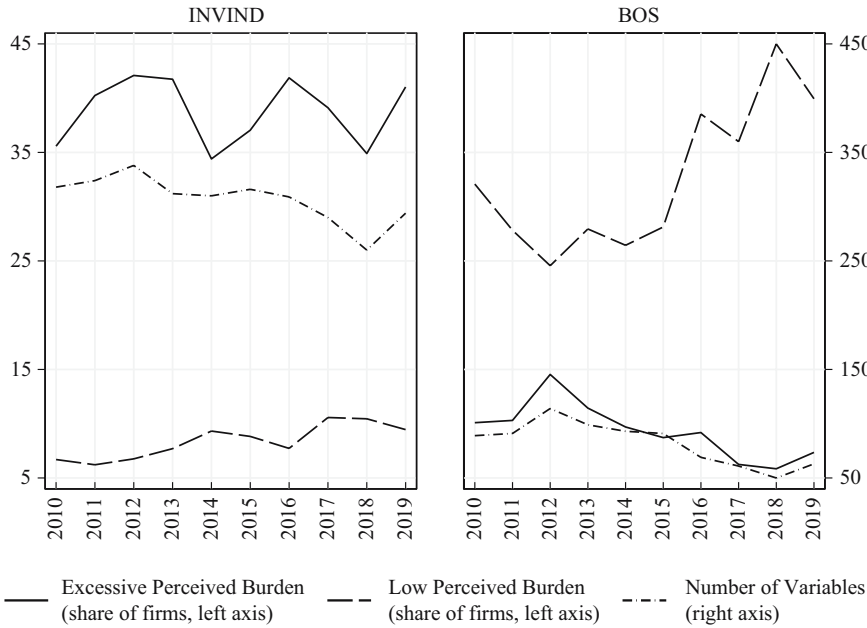


Fig. 2. Firms' perceived burden and number of variables over time.

Table 3. Average people, external consultant and time per level of perceived response burden.

Perceived response burden	Average number of people involved			
	2016	2017	2018	Total
Low	1.5	1.4	1.4	1.5
Average	2.4	2.3	2.4	2.3
High	3.4	3.0	3.1	3.1
Excessive	3.6	3.2	3.1	3.3
Perceived response burden	Firms using external consultant (%)			
	2017	2018	2019	Total
Low	16	14	14	15
Average	23	22	22	22
High	29	34	31	32
Excessive	45	41	41	42
Perceived response burden	Average completion time (hours)			
	2017	2018	2019	Total
Low	1.8	2.1	2.2	2.0
Average	3.7	3.6	3.6	3.6
High	6.4	5.3	5.3	5.7
Excessive	8.2	7.5	5.9	7.1

purpose of the analysis. Yet the two main datasets refer to the period 2004–2019, when only actual burden is studied, and to the period 2017–2019 when also the perceived burden is considered.

To better evaluate the relationship between the perceived response burden and the various factors that can affect it, we create a dummy variable equal to 1 if the perceived response burden is “high” or “excessive” and 0 otherwise. We then estimate two different logistic regressions (as a robustness check). The first one is run on waves from 2004 to 2019 for which only the measures of actual burden we have reconstructed are available. We use the following covariates (Table 4):

Table 4. Probability of reporting a high ('elevated' or 'excessive') perceived response burden (logit model).

	(1)	(2)	(3)
N. of variables	1.004** (0.001)	1.004*** (0.001)	1.004*** (0.001)
N. of pages	1.039*** (0.007)	1.040*** (0.008)	1.047*** (0.008)
Internal instruct.	0.786* (0.084)	0.796* (0.087)	0.776* (0.086)
N. of quantitatives	0.999 (0.001)	0.999 (0.001)	0.999 (0.002)
N. of waves	0.936*** (0.007)	0.929*** (0.007)	0.933*** (0.007)
N. of waves ²	1.002*** (0.000)	1.001*** (0.000)	1.001*** (0.000)
log(empl)		1.198*** (0.025)	1.059 (0.061)
Δ turnover _t		0.947 (0.046)	0.939 (0.048)
Δ employment _t		0.954 (0.112)	0.999 (0.118)
$\frac{\text{investment}_{t-1}}{\text{turnover}_{t-1}}$		0.862 (0.079)	0.890 (0.073)
$\frac{\text{investment}_t}{\text{turnover}_t}$		1.063 (0.099)	1.077 (0.097)
50–199 empl. x N. of waves		1.029*** (0.004)	1.000 (0.006)
200–499 empl. x N. of waves		1.022*** (0.006)	1.014* (0.006)
≥ 500 empl. x N. of waves		1.014* (0.007)	1.018** (0.007)
Δ per capita Val. Add. Sett.		1.830* (0.500)	1.433 (0.401)
Constant	0.172*** (0.065)	0.0905*** (0.035)	0.120*** (0.052)
Observations	5905	59028	59028
Pseudo R ²	0.005	0.013	0.022

Odds ratios. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Based on INVIND 2004–2019 waves. Column 3 includes dummies for industry, area and size class. Missing cases to the question on perceived response burden are excluded.

- indicators of the questionnaire's complexity: the total number of variables and pages, the number of questions requiring quantitative information and whether or not the instructions were placed inside the questionnaire;
- the total number of waves (for both the BOS and INVIND) in which firms have participated;
- firm characteristics such as firm size, the sector of activity, indicators of performance such as the variation in turnover and employment, the ratio between investments and turnover and the growth rate of the (per capita) annual sectorial value added.

The first set of variables refers to decisions about the survey design that are under the Bank of Italy's full control, while firm characteristics are mainly used as control variables that should account for the overall context. It is worth stressing that the indicators of the questionnaire's complexity do not vary among respondents in a given wave, they only change over time. To account for the possible confounding effects of other time-varying factors, we include the growth rate of the (per capita) annual sectorial value added as a proxy of the general economic outlook. This variable should capture possible effects of the economic cycle on the perceived response burden.

The second logistic regression model is based on waves from 2017 to 2019, for which we have information on actual burden directly collected from respondents. We model perceived burden as a function of [Table 5](#):

- three individual-specific measures of actual burden: the number of people involved in the survey, completion time, and whether or not the firm has used external consultancy. In particular, for the first two variables we have created a dummy variable equal to 1 if the value is greater than the 75th percentile of the corresponding variable, 0 otherwise;
- the total number of waves (for both the BOS and INVIND) in which firms have participated;
- the same firm's characteristics used in the previous model;
- time dummy variables and their interaction with all the above mentioned variables.

It is worth noting that even if questions on perceived and actual burden are placed in the same section at the end of the questionnaire, the item nonresponse for the former is around 10%, while it is about 18% for the latter. Since the decision to skip the questions on actual burden is probably an indicator of excessive burden endured by respondents, we include in all the regressions three dummy variables taking the value 1 if the business has not answered to such questions. Our results may be summarized as follows.

An increase in the total number of pages in the questionnaire is significantly linked to an increase in the probability of a large or excessive perceived response burden, with the former playing a stronger role than the latter. This result can be explained by the fact that firms may download the template of the questionnaire, scroll through it and look at the number of total pages before starting to fill it in. The number of pages is probably the most immediate measure that firms use to anticipate their burden. As expected, the number of variables collected in the survey is also positively associated with a higher perceived burden.

Table 5. Probability of reporting a high ('elevated' or 'excessive') perceived response burden (logit model).

	(1)	(2)	(3)
People inv.(> 75th)	2.056*** (0.170)	2.023*** (0.176)	1.945*** (0.170)
People inv. miss	1.129 (0.265)	1.125 (0.263)	1.037 (0.250)
External Cons: Y	1.642*** (0.132)	1.657*** (0.137)	1.711*** (0.146)
External Cons miss	1.698* (0.411)	1.688* (0.410)	1.668* (0.415)
Completion time (> 75th)	2.367*** (0.209)	2.345*** (0.208)	2.360*** (0.212)
Completion time miss	1.547** (0.229)	1.552** (0.231)	1.666*** (0.249)
N. of waves	0.945*** (0.013)	0.941*** (0.013)	0.944*** (0.014)
N. of waves ²	1.001* (0.000)	1.001* (0.001)	1.001* (0.001)
log(empl)		1.015 (0.055)	0.789 (0.109)
Δ turnover _t		0.891 (0.091)	0.891 (0.088)
Δ employment _t		1.108 (0.280)	1.161 (0.294)
$\frac{\text{investment}_{t-1}}{\text{turnover}_{t-1}}$		0.954 (0.190)	0.962 (0.147)
$\frac{\text{investment}_t}{\text{turnover}_t}$		0.897 (0.250)	0.920 (0.255)
50–199 empl. x N. of waves		1.011 (0.008)	0.991 (0.011)
200–499 empl. x N. of waves		1.004 (0.011)	0.994 (0.012)
≥ 500 empl. x N. of waves		0.997 (0.014)	0.973* (0.013)
Constant	0.423*** (0.038)	0.407*** (0.096)	1.131 (0.575)
Observations	11846	11842	11842
Pseudo R ²	0.081	0.082	0.094

Odds ratios. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Column 2 includes dummies for industry, area and size class; Column 3 adds time dummies, also interacted with all the dummies in column 2. Based on INVIND 2017–2019 waves. Missing cases to the question on perceived response burden are excluded.

Furthermore, we find that an excessive number of people involved and an excessive amount of time (greater than the respective 75th percentiles) as well as the use of external consultants increase the probability of reporting a high response burden. Moreover, all the three dummy variables indicating a nonresponse to the questions about actual burden are positively linked with the probability of declaring a higher perceived burden. This result may be explained by the fact that the respondent feels so stressed by the survey, that

she(he) prefers not to waste additional time responding the questions on actual burden (that are placed at the end of the questionnaire).

Interestingly, we do not find evidence that the number of quantitative variables is associated with perceived burden. This is probably because quantitative questions in the INVIND questionnaire mainly concern balance sheet data that are probably not so difficult for the respondent to obtain. On the other hand, the inclusion of instructions within the questionnaire is negatively correlated with the score assigned to the perceived burden: respondents may not feel like consulting a separate sheet of paper and therefore a question without explanation can be considered more burdensome, and, even if they do, this may take longer and more effort than having questions accompanied by corresponding instructions.

Perceived response burden is also associated with other factors. For instance, larger firms tend to report a high burden, especially when the number of surveys they have participated in increases. This result may seem to be counterintuitive since large-sized businesses are supposed to have good documentation systems and people who respond to surveys as part of their job. One possible explanation is that in the INVIND survey there is a big effort to enroll larger businesses; firms with more than 5,000 employees are always eligible to be included in the target population, even if they have refused to do so in the past. Moreover, large businesses are always selected in the business surveys conducted by national statistical offices or by other statistical agencies. Since they receive frequent enquiries and since they cannot easily refuse participation, they are likely to manifest their dissatisfaction by declaring a high burden.

The predicted probability of observing a high response burden, obtained by the first estimated logistic regression (Table 4), has a “U-shaped” relationship with the number of waves in which firms have participated (Figure 3): it initially decreases as the total number of

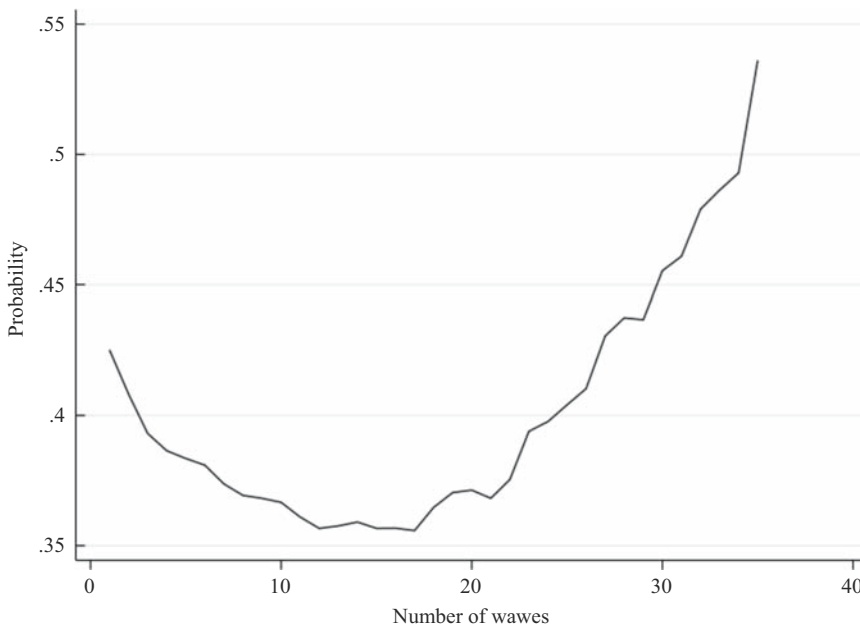


Fig. 3. Predicted probability of reporting a higher response burden obtained by the estimated logistic regression in Table 4.

waves increases but subsequently after around ten waves, it starts to rise. The initial decrease may be the result of two things. First, it could be due to a self-selection process: firms who find the questionnaire burdensome may decide to drop out the following year. As a result, the sample consists of collaborative firms that don't think that participation requires an excessive effort. Second, the decrease could reflect the presence of a learning process over multiple waves that makes it easier for firms to complete the questionnaire. However, the reduction in the response burden is less intense as the frequency of participation increases, meaning that, at some point, a certain level of stress could take over, thereby raising the burden.

A final point is worth mentioning. Even if we find a positive association between actual and perceived burden, the measures of actual burden explain only a small fraction of the total variability of the perceived burden as it is indicated by the low values of the Pseudo R-square indicator. The main reason is that, though we are using a very simple and coarse measure of perceived difficulty, this can capture many unobserved positive and negative effects relating survey participation, such as the respondent's interest in the topic, their ability to answer the questions or their opinion in the utility of the survey. Also, some unobserved factors relating to business activity, such as the internal organization or the documentation system may affect respondents' perceptions.

5. Perceived Response Burden and Data Quality

Data quality can be defined as "fitness for use" of statistical information. The European statistical system which provides guidelines for all European national statistical institutes, defines it as the result of eight dimensions: relevance, accuracy, timeliness, punctuality, accessibility, clarity, comparability and coherence ([European Commission. Statistical Office of the European Union 2014](#)).

In this article, we focus on data accuracy, that is, the degree to which the information correctly describes the phenomena it was designed to measure.

One of the main risks of a high response burden concerns the firms' decision not to respond. They may decide not to participate in the survey at all (unit nonresponse), or they may employ some response strategies that allow them to reduce the effort that they have to expand without leaving the survey altogether. Some of these strategies include skipping some questions (item nonresponse), using the "don't know" or "no opinion" response options, or choosing the first reasonable response. Other behaviors include speeding through the survey by giving low-effort responses or not fully answering open-text questions. In the following subsections, we provide some empirical evidence of how these aspects affect data accuracy using the INVIND survey.

5.1. Unit Nonresponse

Since BOS and INVIND are two longitudinal surveys, we can study the association between perceived response burden and attrition.

Descriptive analysis shows that the propensity to leave the panel significantly increases for firms that perceived a high response burden in previous surveys. In particular, from 2017 to 2019, 9% of businesses stating a low level of burden for INVIND do not participate in the BOS conducted in the same year; compared to 14% that stated an excessive difficulty rating. Similarly, around 16% of the firms declaring a low burden in

BOS refuse to participate in the following the INVIND survey (conducted a year later). This percentage rises to 21% for businesses that complained about excessive effort.

To further investigate the role of response burden on attrition, we run a logistic regression using as a dependent variable a dummy indicating whether firms that participated in waves 2016–2018, then decided to drop out in the subsequent surveys (2017–2019 respectively). We regress this variable on the perceived response burden declared in the last survey before dropping out, indicators of actual burden (the number of people involved, the percentage of external consultants and the time spent), a dummy variable for nonresponse at the question on perceived burden, the total number of waves in which firms have participated, as well as their characteristics. (Table 6).

The results show a significant increase in the probability of attrition for firms that report an excessively high response burden and for those that do not answer the question on response burden. The number of waves they have participated in plays a significant role in reducing the probability of attrition, confirming the existence of possible learning effects (as already shown in Figure 3).

Moreover, the variation in the number of employees that we use as a proxy of economic performance has a significant effect in reducing the probability of attrition. This may in part be due to the fact that firms with better performance have a lower probability of dropping out of the survey (a similar result is shown in D'Aurizio and Papadia 2019).

Finally, we find that including the three individual-specific indicators of actual burden in the regression doesn't affect the estimate of parameters: columns 3–5 of Table 6 show similar estimates as the remaining ones (columns 1–2). Moreover, the actual burden faced by each respondent is not significantly correlated with the probability of leaving the panel once controlled for the perceived burden and the other observables. One possible explanation is that when the respondent has been contacted again, he or she remembers the stress induced by taking part in the previous year's survey rather than its actual difficulty, and this affects the decision whether or not to participate in the new round of the survey. Hence, the actual burden would affect the probability of attrition only through the perceived burden.

5.2. Item Nonresponse

The response behavior of survey participants depends on many factors. Some respondents may decide to answer hastily and carelessly, since they perceive the survey as too time-consuming. In other cases, they may limit their attention to questions that are mandatory for completing the questionnaire (for which an alert signals the invalidity of the entire questionnaire if they are left blank). These behaviors may become more likely as the complexity of the questionnaire grows.

Figure 4 shows the cumulative distributions of firms according to their share of missing items in the 1999 and 2017 INVIND waves. In these two years, the questionnaires present very different levels of complexity: the number of variables increased from about 200 to about 300, the number of pages from about 5 to about 20. To make the two distributions comparable, we selected only industrial processing firms with more than 50 employees and excluded the compulsory variables. In the 1999 survey, 90% of businesses had a share of item nonresponse lower than 35%. In the 2017 survey this share rises to 60%. Moreover,

Table 6. Probability of attrition (logit model).

	Attrition in 2016–2018		Attrition in 2017–2018		
	(1)	(2)	(3)	(4)	(5)
High Perc. Burd.	1.425*** (0.133)	1.286** (0.123)	1.432** (0.165)	1.488*** (0.173)	1.389** (0.164)
Perc. Burd. miss.	2.308*** (0.319)	1.922*** (0.274)	2.083*** (0.381)	1.679* (0.374)	1.421 (0.329)
People inv.(> 75th)				0.842 (0.123)	0.837 (0.127)
People inv. miss				1.335 (0.408)	1.370 (0.434)
External Cons: Y				1.136 (0.143)	0.970 (0.127)
External Cons miss				1.436 (0.530)	1.344 (0.507)
Completion time (> 75th)				0.809 (0.122)	0.842 (0.128)
Completion time miss				0.641 (0.207)	0.667 (0.213)
N. of waves		0.884*** (0.016)			0.895*** (0.020)
N. of waves ²		1.003*** (0.001)			1.002** (0.001)
log(empl)		0.961 (0.179)			1.193 (0.282)
Δ turnover _t		0.813 (0.194)			0.584 ⁺ (0.172)
Δ employment _t		0.434 ⁺ (0.192)			0.504 (0.284)
$\frac{\text{investment}_{t-1}}{\text{turnover}_{t-1}}$		1.132 (0.211)			1.545 (0.588)
$\frac{\text{investment}_t}{\text{turnover}_t}$		0.942 (0.230)			0.681 (0.258)
Constant	0.152*** (0.009)	0.281 ⁺ (0.187)	0.151*** (0.011)	0.155*** (0.015)	0.112** (0.092)
Observations	12758	12756	8445	8445	8443
Pseudo R ²	0.011	0.044	0.009	0.013	0.045

Odds ratios; Standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Columns 2 and 5 include dummies for time, industry, area and size class.

firms that answer at least half of the non-compulsory questions decreases by about 20 percentage points over the same period.

Item nonresponse rates vary across different sections of the questionnaire. By informally inspecting the questionnaire content, we find that sections requiring qualitative information on easy-to-understand topics (such as the section about funding) get relatively low nonresponse rates. These rates increase when the complexity of the formulation of the question and of the terminology used seems to grow.

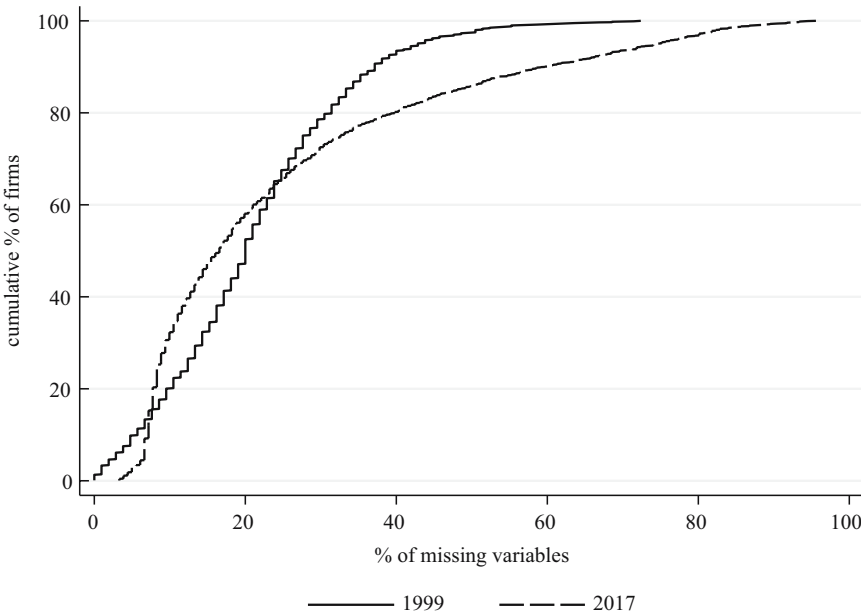


Fig. 4. Cumulative distributions of firms for shares of missing variables in INVIND 1999 and 2017.

The increase in the share of missing variables also involves questions that are of great importance for economic analysis. Figure 5 shows the time series of the percentage of missing data on questions about expectations on investment and turnover for the next year. Although their complexity has remained constant over the years in terms of formulation, the share of missing answers about investment plans reaches 15% in 2019, compared with less than one-third of that figure in the 1980s; the share related to expectations on turnover (collected since 1997) rises from about 3% to almost 10%. This implies that the response burden can involve the potential loss of information on historical and important variables, as well as on new ones.

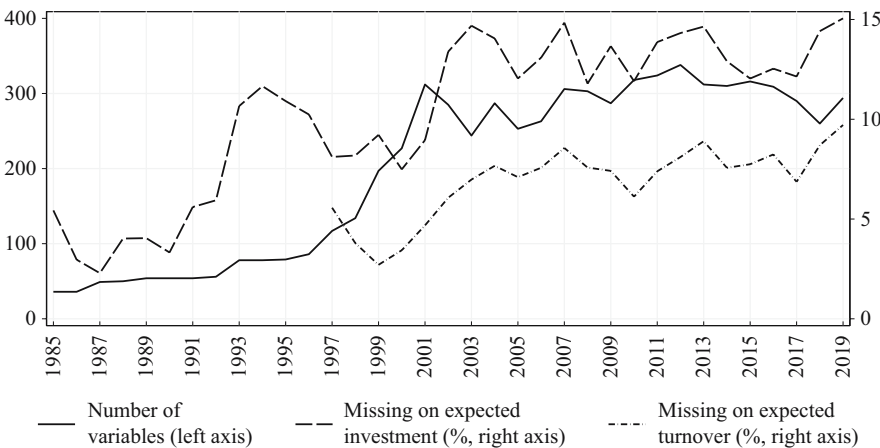


Fig. 5. Share of missing data on questions about expected investment and expected turnover.

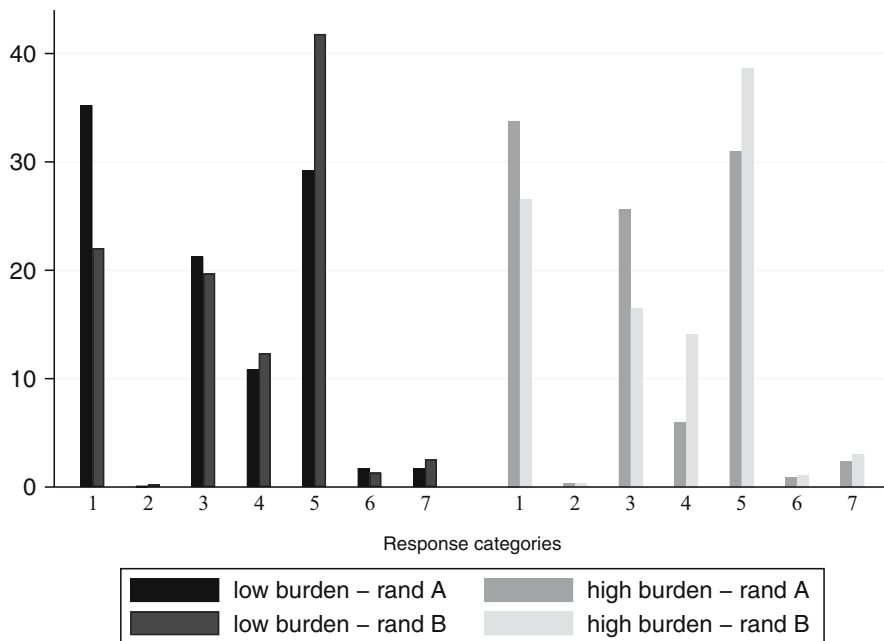


Fig. 6. Distribution of factors affecting the expected price dynamics of firms. Randomized experiment with reversed response categories (percentages).

Note: Response categories are labeled as: 1 = total demand; 2 = change in the financial burdens borne by the firm; 3 = competitors' prices; 4 = labor costs; 5 = raw materials prices; 6 = expectations for exchange rates; 7 = inflation expectations. For group A the response options are ordered as above. For group B the order is reversed. INVIND 2017 waves.

As further analysis, we estimate two regression models for the probability of not answering the questions about expected turnover and investment as a function of the perceived response burden, our measures of actual burden, some firm characteristics, indicators of firm's performance and time dummies (Table 7). The probability of nonresponse is significantly greater for firms that perceive an excessive response burden and is even greater when the respondent does not even provide answers to questions about the evaluation of the questionnaire. A missing value to such questions (which are located at the end of the questionnaire) may be an indication that the respondent thinks that the survey is too demanding. We also find that the longer the completion time, the higher the probability of item nonresponse.

We also find that perceived burden plays a much more important role in explaining nonresponse than actual burden. The more significant variables are those indicating that the respondent has not answered the questions about the time needed to complete the survey and the number of people involved.

5.3. Measurement Error

Our data also enable us to provide some evidence of the association between perceived burden and accuracy.

We use the CEBIL/CERVED company database that provides information on the balance sheet of all the joint stock companies operating in Italy. We link this database to

Table 7. Probability of item nonresponse (logit model).

	Nonresponse on expected investment		Nonresponse on expected turnover	
	(1)	(2)	(3)	(4)
High Perc. Burd.	1.230*** (0.033)	1.328*** (0.082)	1.330*** (0.047)	1.503*** (0.121)
Perc. Burd. miss.	3.077*** (0.097)	1.677*** (0.166)	3.949*** (0.151)	2.908*** (0.343)
People inv.(> 75th)		0.976 (0.077)		0.950 (0.100)
People inv. miss		1.779*** (0.284)		1.554* (0.321)
External Cons: Y		0.838* (0.070)		0.802 (0.095)
External Cons miss		0.902 (0.147)		0.823 (0.170)
Completion time (> 75th)		0.792** (0.067)		0.703** (0.080)
Completion time miss		2.119*** (0.226)		2.346*** (0.314)
log(empl)	1.213*** (0.042)	1.478*** (0.114)	1.251*** (0.050)	1.453*** (0.131)
Δ turnover _t	0.995 (0.010)	0.927 (0.123)	0.804** (0.062)	1.008 (0.153)
Δ employment _t	1.022 (0.077)	0.640 (0.167)	0.623** (0.092)	0.459* (0.168)
$\frac{\text{investment}_{t-1}}{\text{turnover}_{t-1}}$	1.014 (0.020)	1.090 (0.096)	1.034 (0.026)	1.170 (0.285)
$\frac{\text{investment}_t}{\text{turnover}_t}$	1.143** (0.057)	0.933 (0.197)	1.019 (0.043)	0.638 (0.250)
Constant	0.0687*** (0.009)	0.0197*** (0.006)	0.0315*** (0.005)	0.0108*** (0.004)
Observations	66861	12806	66861	12806
Pseudo R ²	0.041	0.095	0.075	0.154

Odds ratios; Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Column 2 includes dummies for time, industry, area and size class. Cols (1) and (3) are based on INVIND 2004–2019 waves; Cols (2) and (4) are based on INVIND 2017–2019 waves.

the INVIND survey and compute two proxies of response error for each business. The proxies are the difference (in absolute terms) between the value of turnover and investments reported in the survey and the corresponding values resulting from the administrative records (with the same reference year).

We then run a regression using these proxies as dependent variables and include perceived burden, actual burden and other controls (firm size, measures of performance, sector of activity and location) as covariates (Tables 8 and 9).

We find that the response error does not seem to be affected by perceived response burden. A high perceived burden is positively associated with the response error on

Table 8. Response error in turnover (linear model).

	(1)	(2)	(3)	(4)	(5)
High Perc. Burd.	272.9*** (72.583)	187.4 (167.842)	110.6 (149.348)	122.1 (144.143)	128.9 (155.375)
Perc. Burd. miss.	223.9* (113.277)	-62.47 (114.804)	-290.3 (266.805)	-394.5 (308.197)	-399.2 (309.690)
People inv.(> 75th)			352.8* (161.699)	399.5* (159.246)	423.7* (185.002)
People inv. miss			347.0 (286.148)	-100.9 (209.664)	-218.7 (214.320)
External Cons: Y				-170.4 (148.217)	-154.9 (163.175)
External Cons miss				519.5 (310.343)	372.9 (337.895)
Completion time (> 75th)					-88.87 (217.269)
Completion time miss					275.0 (231.697)
Part of a group	783.1*** (87.598)	508.3** (165.805)	475.7** (155.791)	462.6** (160.547)	467.9** (164.001)
Δ turnover _t	12.33 (13.671)	-4.616 (22.346)	-5.278 (22.317)	-5.492 (22.412)	-5.587 (22.265)
Δ employment _t	-237.3 (353.268)	111.1 (828.462)	138.3 (827.458)	151.8 (833.752)	155.9 (828.957)
log(empl)	1128.9*** (121.970)	900.1*** (265.727)	870.2** (274.881)	866.0** (274.458)	876.2** (268.983)
$\frac{\text{investment}_{t-1}}{\text{turnover}_{t-1}}$	-53.81 (34.949)	36.99 (79.125)	41.39 (80.716)	28.69 (78.115)	31.83 (76.291)
$\frac{\text{investment}_t}{\text{turnover}_t}$	307.5 (196.377)	-40.29 (161.974)	-56.10 (159.216)	-50.34 (166.366)	-57.17 (162.606)
Constant	-2893.7*** (460.174)	-2551.7** (928.094)	-2560.6** (934.283)	-2509.2** (923.869)	-2552.2** (912.223)
Observations	59598	11462	11462	11462	11462
R ²	0.091	0.121	0.121	0.122	0.122

The dependent variable is computed as the absolute value of the difference between the value declared in the survey and the one from administrative records. Amounts in thousands of euros. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include dummies for time, industry, area and size class. Col (1) based on INVIND 2004–2017 waves, cols (2–5) INVIND 2017–2019 waves.

turnover and investments only if we consider past waves for which measures of actual burden are not available (column 1). Its effect loses significance once the analysis is limited to the last waves (column 2). Moreover, controlling also for the three proxies of actual response burden shows that the average response error is associated with some indicators of actual burden (such as the number of people involved or the time required for completion). Therefore, it seems that it is the actual arduousness of the survey that affects the measurement errors rather than the one subjectively perceived. For example, involving more people means fragmentizing the questionnaire in multiple parts and probably sending only that part to the person. This can have two implications: first, different

Table 9. Response error in investment (linear model).

	(1)	(2)	(3)	(4)	(5)
High Perc. Burd.	33.86 (26.335)	49.63 (53.001)	31.20 (55.327)	24.25 (55.946)	41.02 (56.636)
Perc. Burd. miss.	111.3** (42.006)	29.92 (94.303)	-53.58 (118.245)	-90.45 (121.099)	-79.65 (121.294)
People inv.(> 75th)			78.27 (59.647)	69.53 (64.942)	102.9 (67.525)
People inv. miss			115.5 (87.413)	-37.75 (173.223)	-17.36 (178.286)
External Cons: Y				37.90 (67.706)	53.96 (68.004)
External Cons miss				197.6 (177.897)	207.8 (182.913)
Completion time (> 75th)					-133.6* (60.703)
Completion time miss					-51.32 (101.090)
Part of a group	113.2*** (28.403)	94.22 (58.214)	87.38 (58.396)	89.47 (58.963)	93.66 (58.947)
Δ turnover _t	18.54*** (3.020)	21.41** (7.087)	21.24** (7.079)	21.29** (7.118)	21.09** (7.201)
Δ employment _t	-738.9*** (107.365)	-842.6** (266.941)	-835.6** (266.586)	-838.8** (268.185)	-832.3** (271.272)
log(empl)	50.95 (48.359)	-144.7 (107.358)	-151.0 (107.385)	-150.5 (106.954)	-143.1 (106.991)
$\frac{\text{investment}_{t-1}}{\text{turnover}_{t-1}}$	-53.28 (37.132)	-49.65 (128.049)	-47.88 (130.444)	-52.99 (126.079)	-58.60 (124.371)
$\frac{\text{investment}_t}{\text{turnover}_t}$	820.6*** (113.049)	1465.0*** (313.774)	1460.7*** (315.787)	1470.8*** (308.416)	1478.5*** (307.242)
Constant	808.3*** (181.330)	961.2* (391.829)	951.5* (392.301)	936.9* (386.261)	924.7* (386.490)
Observations	59598	11462	11462	11462	11462
R ²	0.069	0.066	0.067	0.067	0.068

The dependent variable is computed as the absolute value of the difference between the value declared in the survey and the one from administrative records. Amounts in thousands of euros. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All regressions include dummies for time, industry, area and size class. Col (1) based on INVIND 2004–2017 waves, cols (2–5) INVIND 2017–2019 waves.

respondents will answer questions probably without having the corresponding instructions and the awareness of the complete task; second, the main respondent, that is, the one answering the evaluation final part, probably feels less stressed since they have accomplished only a residual part of the questionnaire.

We also find that response error is positively correlated with some characteristics of the business. For instance, as far as turnover is concerned, the error is greater for firms that are part of a group and for large businesses. In these cases, the presence of multiple establishments or multiple companies that are strictly related may lead to ambiguity concerning which unit to respond for.

6. Concluding Remarks

In this article, we have provided some empirical evidence of the relationship between perceived (the difficulty rating) and actual burden and of the association between perceived burden and data quality. Drawing from a unique and rich dataset, we have been able to conduct our analysis also controlling for other contextual factors relating to firms' characteristics.

Our key findings may be summarized as follows.

- The perceived difficulty in completing the survey is associated with measures of actual burden such as the complexity of the questionnaire. We find that it is not simply the number of questions that increases perceived burden, but also the number of pages of the questionnaire, which is probably used by businesses to anticipate their effort (regardless of the effective difficulty of the questions). We also find that the higher the number of people involved in the survey, the higher the perceived burden;
- perceptions are also driven by other firm-specific characteristics such as the number of employees. Large businesses tend to report a high burden. This is also probably because many efforts are devoted to preventing them from dropping out of the survey. Given their difficulties in terms of refusing to participate, they are likely to complain and complain that the questionnaire is burdensome;
- we find empirical evidence that supports our first assumption that the measure of difficulty rating captures different information from the one contained in the measures of actual burden. The latter and other contextual factors capture only a small fraction of the overall variability of perceived burden. The unexplained variability is probably associated with unobservable characteristics such as the respondent's interest in the topic, their ability to answer the questions or their opinion on the utility of the survey. Besides, unobserved factors relating to business activity, such as the organization of the business, may play a role;
- we also find that data quality is directly associated with perceived burden, even after controlling for actual burden and other characteristics (our second assumption). The probability of attrition increases with a higher perceived burden. Moreover, even if the firm participates in the survey, an excessive perceived burden is associated with a high probability that the respondent will not complete the whole questionnaire. The questions that are more likely to be skipped are those that are not compulsory, in the sense that without a response to those questions, the whole questionnaire is considered incomplete (and therefore corresponding data are treated as a unit nonresponse), as well as questions that require more effort to answer (such as the questions about firms' expectations on future investment and turnover). On the other hand we do not find evidence that an excessive difficulty rating is associated with more inaccurate answers, at least as far as turnover or investments are concerned. One possible explanation is that since firms know that their responses can be linked to register data, the best response behavior is to retrieve such information directly from their balance sheets in order to provide consistent information to the outside world.

In summary, our analysis shows that even a simple and coarse measure of difficulty rating, like the one we use in this article, is a good instrument for monitoring data quality.

It is easy to collect and captures many unobserved factors, which play a role in determining the final quality of survey data.

7. Appendix

7.1. Questions on Response Burden

English translation of the questions on response burden:

- How would you rate the level of complexity of the survey?
1 = modest; 2 = average; 3 = large; 4 = excessive.
 - To what extent do you think the following factors made it difficult to fill in the questionnaire?
(For each factor please assign a score ranging from 1 to 10, where 1 indicates that the factor played a very limited part in making the questionnaire difficult to fill in while 10 indicates that it played a very large part)
- | | Rate |
|---|--------------------------|
| A Too many questions | <input type="checkbox"/> |
| B It was necessary to seek the help of several people to answer the questions | <input type="checkbox"/> |
| C It was not always easy to understand the questions because some of the terms were not clear | <input type="checkbox"/> |
| D The possible answers did not include my situation | <input type="checkbox"/> |
| E For some questions, it was difficult to choose the correct answer | <input type="checkbox"/> |
- How many people from your firm, including yourself, were involved in completing the survey?
 - Was it necessary to involve external consultants (*e.g., accountant, labor consultant, and so on.*)? (Yes/No)
 - Could you please indicate how much time approximately it took your firm to collect the necessary information and fill in the questionnaire? (*please indicate the number of hours*)

8. References

- Bavdaž, M. 2010. “Sources of Measurement Errors in Business Surveys.” *Journal of Official Statistics* 26 (1): 24–42. Available at: <https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/sources-of-measurement-errors-in-business-surveys.pdf> (accessed September 2021).
- Bavdaž, M., D. Giesen, S.K. Černe, T. Löfgren, and V. Raymond-Blaess. 2015. “Response Burden in Official Business Surveys: Measurement and Reduction Practices of National Statistical Institutes.” *Journal of Official Statistics* 31 (4): 559–588. DOI: <https://doi.org/10.1515/jos-2015-0035>.
- Bergman, L.R., and R. Brage. 2008. “Survey Experience and Later Survey Attitudes, Intention and Behavior.” *Journal of Official Statistics* 24 (1): 99–113. Available at: <https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/survey-experiences-and-later-survey-attitudesintentions-and-behaviour.pdf> (accessed September 2021).

- Bradburn, N.M. 1978. "Respondent burden." In *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 35–40. Available at: http://www.asasrms.org/Proceedings/papers/1978_007.pdf (accessed March 2021)
- Crawford, S.D., M.P. Couper, and M.J. Lamias. 2001. "Web surveys: Perceptions of burden." *Social Science Computer Review* 19 (2): 146–162. DOI: <https://doi.org/10.1177/089443930101900202>.
- Dale, T., J. Erikson, J. Fosen, G. Haraldsen, J. Jones, and Ø. Kleven. 2007. *Handbook for Monitoring and Evaluating Business Survey Response Burdens*. European Commission. Available at: <https://ec.europa.eu/eurostat/documents/64157/4374310/12-HANDBOOK-FORMONITORING-AND-EVALUATING-BUSINESS-SURVEY-RESONSE-BURDEN.pdf/600e3c6d-8e8d-44f7-a8f5-0931c71d9920> (accessed March 2021)
- D'Aurizio, L., and G. Papadia. 2019. "Using Administrative Data to Evaluate Sampling Bias in a Business Panel Survey." *Journal of Official Statistics* 35 (1): 67–92. DOI: <https://doi.org/10.2478/jos-2019-0004>.
- Edwards, P., I. Roberts, M. Clarke, C. DiGuseppi, S. Pratap, R. Wentz, and I. Kwan. 2002. "Increasing response rates to postal questionnaires: systematic review." *BMJ* 324 (7347): 1183. DOI: <https://doi.org/10.1136/bmj.324.7347.1183>.
- European Commission. Statistical Office of the European Union. 2014. *Manuals and guidelines. ESS Handbook for Quality Reports*. DOI: <https://doi.org/10.2785/983454>.
- Fisher, S., and L. Kydoniefs. 2001. "Using a theoretical model of response burden to identify sources of burden in surveys." Paper presented at the 12th International Workshop on Household Survey Nonresponse, September 12–14, Oslo, Norway.
- Galesic, M. 2006. "Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey." *Journal of Official Statistics* 22: 313–328. Available at: <https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/dropouts-on-the-web-effects-of-interest-and-burden-experienced-during-an-online-survey.pdf> (accessed September 2021).
- Galesic, M., and M. Bosnjak. 2009. "Effects of questionnaire length on participation and indicators of response quality in a web survey." *Public Opinion Quarterly* 73 (2): 349–360. DOI: <https://doi.org/10.1093/poq/nfp031>.
- Giesen, D. 2012. "Exploring causes and effects of perceived response burden." In *Proceedings of the Fourth International Conference on Establishment Surveys: American Statistical Association*, June 11–14, Montréal, Canada. Available at: <https://www2.amstat.org/meetings/ices/2012/papers/302171.pdf> (accessed March 2021)
- Groves, R.M., M.P. Couper, S. Presser, E. Singer, R. Tourangeau, G.P. Acosta, and L. Nelson. 2006. "Experiments in producing nonresponse bias." *Public Opinion Quarterly* 70 (5): 720–736. DOI: <https://doi.org/10.1093/poq/nfl036>.
- Haraldsen, G. 2004. "Identifying and Reducing Response Burden in Internet Business Surveys." *Journal of Official Statistics* 20 (2): 393–410. Available at: <https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/identifying-and-reducing-response-burdens-in-internet-business-surveys.pdf> (accessed September 2021).
- Haraldsen, G. 2013. "Quality issues in business surveys." *Designing and Conducting Business Surveys*, 83–125. John Wiley & Sons. DOI: <https://doi.org/10.1002/9781118447895.ch03>.

- Haraldsen, G., and J. Jones. 2007. "Web and paper questionnaires seen from the business respondent's perspective." In Proceedings of the Third International Conference for Establishments Surveys, JUNE 18–21, Montréal, Canada. Available at: <https://ww2.amstat.org/meetings/ices/2007/proceedings/ICES2007-000259.PDF> (accessed March 2021)
- Haraldsen, G., J. Jones, D. Giesen, and L.-C. Zhang. 2013. "Understanding and coping with response burden." In *Designing and Conducting Business Surveys*, 219–252. John Wiley & Sons. DOI: <https://doi.org/10.1002/9781118447895.ch06>.
- Heerwegh, D., and G. Loosveldt. 2002. "An evaluation of the effect of response formats on data quality in web surveys." *Social Science Computer Review* 20 (4): 471–484. DOI: <https://doi.org/10.1177/089443902237323>.
- Holbrook, A.L., J.A. Krosnick, D. Moore, and R. Tourangeau. 2007. "Response order effects in dichotomous categorical questions presented orally: The impact of question and respondent attributes." *Public Opinion Quarterly* 71 (3): 325–348. DOI: <https://doi.org/10.1093/poq/nfm024>.
- Jones, J., J. Rushbrooke, G. Haraldsen, T. Dale, and D. Hedlin. 2005. "Conceptualising total business survey burden." In *Survey Methodology Bulletin* 55. UK Office for National Statistics, 1–10. Available at: <http://www.ons.gov.uk/ons/guide-method/method-quality/survey-methodology-bulletin/smb-55/index.html> (accessed March 2021)
- Peytchev, A. 2011. "Breakoff and Unit Nonresponse Across Web Surveys." *Journal of Official Statistics* 27 (1): 33. Available at: <https://www.scb.se/contentassets/ca21efb41-fee47d293bbee5bf7be7fb3/breakoff-and-unit-nonresponse-across-web-surveys.pdf> (accessed September 2021).
- Presser, S., J. Blair, and T. Triplett. 1992. "Survey sponsorship, response rates, and response effects." *Social Science Quarterly* 73 (5): 699–702. Available at: <http://www.jstor.org/stable/42863089>.
- Snijkers, G., G. Haraldsen, J. Jones, and D.K. Willimack. 2013. *Designing and conducting business surveys*. New York: Wiley.
- Tomaskovic-Devey, D., J. Leiter, and S. Thompson. 1994. "Organizational survey nonresponse." *Administrative Science Quarterly* 39 (3): 439–457. DOI: <https://doi.org/10.2307/2393298>.
- Tourangeau, R., T. Yan, and H. Sun. 2019. "Who Can You Count On? Understanding The Determinants of Reliability." *Journal of Survey Statistics and Methodology* 8 (5): 903–931. DOI: <https://doi.org/10.1093/jssam/smz034>.
- Willeboordse, A. 1997. "Minimizing response burden." In *Handbook on Design and Implementation of Business Surveys*, edited by Willeboordse. 111–118. Available at: <https://ec.europa.eu/eurostat/documents/3859598/5825949/CA-09-97-818-EN.PDF/f5ee3198-6fc0-4672-96a1-8fdb4a81ca93?version=1.0> (accessed March 2021).
- Yan, T., F.G. Conrad, R. Tourangeau, and M.P. Couper. 2010a. "Should I stay or should I go: The effects of progress feedback, promised task duration, and length of questionnaire on completing web surveys." *International Journal of Public Opinion Research* 23 (2): 131–147. DOI: <https://doi.org/10.1093/ijpor/edq046>.
- Yan, T., R. Curtin, and M. Jans. 2010b. "Trends in Income Nonresponse Over Two Decades." *Journal of Official Statistics* 26 (1): 145–164. Available at:

<https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/trends-in-income-nonresponse-over-two-decades.pdf> (accessed September 2021).

Yan, T., S. Fricker, and S. Tsai. 2014. "The impact of response burden on data quality in a longitudinal survey." In Proceedings of the International Total Survey Error Workshop, October 1–3, Washington, D.C. Available at: <https://www.niss.org/events/2014-international-total-survey-error-workshop-itsew-2014> (accessed March 2021)

Yan, T., S. Fricker, and S. Tsai. 2019. "Response burden: What is it and what predicts it?" In *Advances in Questionnaire Design, Development, Evaluation and Testing*, 193–212. John Wiley & Sons. DOI: <https://doi.org/10.1002/9781119263685.ch8>.

Received August 2019

Revised August 2020

Accepted March 2021