

The case of the missing tax data



Arnout van Delden, Sander Scholtus, Peter-Paul de Wolf

The views expressed in this paper are those of the author(s)
and do not necessarily reflect the policies of Statistics Netherlands

Discussion paper (201115)



Explanation of symbols

.	= data not available
*	= provisional figure
**	= revised provisional figure
x	= publication prohibited (confidential figure)
—	= nil or less than half of unit concerned
—	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2010–2011	= 2010 to 2011 inclusive
2010/2011	= average of 2010 up to and including 2011
2010/'11	= crop year, financial year, school year etc. beginning in 2010 and ending in 2011
2008/'09–2010/'11	= crop year, financial year, etc. 2008/'09 to 2010/'11 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher

Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress

Statistics Netherlands - Grafimedia

Cover

TelDesign, Rotterdam

Information

Telephone +31 88 570 70 70
Telefax +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order

E-mail: verkoop@cbs.nl
Telefax +31 45 570 62 68

Internet

www.cbs.nl

ISSN: 1572-0314

© Statistics Netherlands, The Hague/Heerlen, 2011.
Reproduction is permitted. 'Statistics Netherlands' must be quoted as source.

The case of the missing tax data

Arnout van Delden, Sander Scholtus, Peter-Paul de Wolf

Summary: In this paper, we describe the controversy that arose between Statistics Netherlands and the Ministry of Economic Affairs in 2009 after the Dutch government announced a tax relief measure for businesses, which deteriorated the quality of tax data used by Statistics Netherlands for producing short-term statistics.

Keywords: short-term statistics, tax declaration data, controversy

1. Introduction

Like most national statistical institutes in recent years, Statistics Netherlands (SN) has been replacing its primary data collection by means of questionnaires more and more with the use of administrative data from external sources. In fact, since 2003 SN is by law only allowed to collect data for statistical production by means of questionnaires if no suitable administrative sources are available. This limitation is part of a larger operation carried out by the Dutch government, aimed at reducing the administrative burden on Dutch businesses.

In this paper, we describe a series of events that occurred in 2009 after the Dutch government announced a tax relief measure for businesses, which had an impact on the statistical quality of tax data used by SN for producing short-term statistics. SN felt that primary data collection had to be re-introduced in some economic sectors, in order to produce statistical output of sufficient quality. This led to a controversy between the Ministry of Economic Affairs and SN, which was finally resolved when SN decided to stop primary data collection again.

Section 2 of this paper outlines the events that preceded the involvement of the methodology department in the controversy. Section 3 describes some of the methodological research into the matter. Finally, section 4 briefly describes the outcome of the controversy.

2. Background

In April 2009, Mr. De Jager, the Dutch State Secretary of Finance, announced that, starting July 1st of that year, most Dutch businesses would be allowed to choose whether they wanted to declare tax on a monthly or quarterly basis. Until then, Dutch businesses had been assigned to a reporting cycle by the tax authorities, based on an estimate of their tax due to be paid: the largest businesses had to declare tax

every month, smaller businesses only every quarter.¹ With this measure, the Dutch government hoped to boost the economy, by giving businesses an opportunity to select the reporting frequency that is most attractive to them from a fiscal point of view.

The announcement by the State Secretary of Finance came as a surprise to SN, and sparked off some commotion. In the spring of 2009, the economic statistics division was about to complete a long-running redesign of the production process for short-term statistics. For small and medium-sized businesses, after the redesign, administrative data from the tax authorities were supposed to replace the traditional survey data collected by questionnaires. In fact, the short-term statistics on the building industry and international trade had already been partly based on tax data for some time. For the building industry, the use of tax data had replaced primary data collection among small businesses (with less than ten employed persons).

Since short-term statistics are published every month for some economic sectors, it was immediately clear that the governmental tax relief measure might adversely influence the statistical properties of the tax declaration data. In particular, businesses that keep reporting on a monthly basis might differ in some important respects from businesses that switch to reporting on a quarterly basis. For instance, businesses that are expecting a tax refund are more likely to report monthly than quarterly, and vice versa. This means that when businesses have been given the option to switch reporting cycles, the monthly tax return data might pertain to a selective subpopulation of businesses, so that monthly growth figures based on this data might no longer be representative for the development of the Dutch economy.

It soon became clear that the measure would be executed by the government as announced, leaving no room for modifications. SN therefore decided that the new production process for short-term statistics would not be implemented until the effects of the measure had been examined. Moreover, for the monthly statistics on the building industry and international trade, it was decided that primary data collection would be re-introduced starting with the reference month June. This was done because the effects of the tax relief measure were not yet known and the economic statistics division wanted to safeguard the quality of the statistical output. A random sample of, in total, 10,000 businesses would be drawn each month, to assess the quality of the tax data after some businesses had switched from monthly to quarterly reporting. This number breaks down into 7,700 small businesses from the building industry and 2,300 businesses from the international trade sector.

On April 28, the Dutch newspaper *De Telegraaf* ran a story on the planned re-introduction of primary data collection in the building industry and the international trade sector, using the headline “New pile of paperwork for entrepreneurs”. As a result, on May 1st, two members of parliament submitted formal questions to the

¹ There is also a limited group of very small businesses that have to report to the tax authorities once a year. Since these businesses have a negligible effect on the economy, we shall ignore them in this paper.

State Secretaries of Finance and Economic Affairs, querying the need for SN to re-introduce data collection by means of questionnaires. These questions were answered in writing by Mr. De Jager on May 27.

The first questionnaires were sent to businesses in the building industry and the international trade sector in mid-July, asking respondents to report their turnover for June. This event led to additional questions in parliament, which were submitted to the State Secretaries of Finance and Economic Affairs on July 22. In particular, members of parliament complained that the number of questionnaires (10,000) was large compared to the number of businesses that had switched from monthly to quarterly reporting at that moment (about 55,000 across all sectors). During consultations between SN and the two governmental departments about these questions, some pressure was put on SN to stop the primary data collection.

In the meantime, SN was looking into the effects of the tax relief measure on the statistical quality of the tax data. This research was done for both the building industry and the international trade sector, but in the remainder of this paper we focus on the building industry. Early efforts at assessing the effects of the measure suffered from a lack of available data. On June 19, after much insistence, the tax authorities provided SN with a list of all businesses in the building industry that had applied for a switch from monthly to quarterly reporting so far. The economic statistics division used this list to conduct a simulation study, by recomputing growth statistics for previous months under the assumption that the businesses on the list had stopped reporting on a monthly basis at an earlier time, and comparing the resulting figures to those that were obtained using all available data. No particularly large deviations were found, but SN hesitated to draw conclusions from this because: (a) the development of the Dutch economy in the period considered in the simulation study was atypical, since it coincided with a phase of stagnation and recession following the global financial crisis; (b) only a small number of businesses had switched so far, but this number might become much larger in the following months; and (c) there was no objective criterion to decide how large a deviation would be acceptable.

At the methodology department, meanwhile, some research was done into the sample design used for primary data collection in the building industry. This research is described in Section 3.1.

By early September, questionnaire data on the months June and July had been collected and edited at SN, and the methodology department was asked to analyse the data, to help decide whether primary data collection remained necessary. This analysis is described in Section 3.2.

3. Research by the methodology department

3.1 Sample design

In July, three methodologists wrote a report that explored possibilities of drawing a smaller sample of businesses from the building industry. The original design (a simple random sample of 7,700 businesses) had been adopted by the economic statistics division without consulting the methodology department. In the report, the methodologists showed that a sample of 5,000 businesses would actually lead to similar confidence intervals if a form of stratification were used. In the absence of actual data, model assumptions were used to obtain this result.

By July, only about 2,000 businesses in the building industry had used the option of switching from monthly to quarterly reporting to the tax authorities. The reason why the sample sizes just mentioned are much larger than the number of businesses that actually changed their reporting cycle – which, as mentioned in Section 2, puzzled some members of parliament – is that SN chose to draw a sample from the population of all businesses, rather than a sample of businesses that had decided to switch. In principle, it would have been sufficient to conduct primary data collection only among businesses that had switched to reporting on a quarterly basis, since tax data were still available for businesses that had not switched. This option was also considered in the methodological report written in July. However, it turned out that the tax authorities could not provide SN with an updated list of businesses that had switched in time for SN to use it as a sampling frame.

The initiative for writing the report on the sample design was taken by the methodology department itself. Its suggestions were not implemented in practice.

3.2 Data analysis

In September, the economic statistics division asked the methodology department to analyse whether primary data collection in the building industry was necessary, or whether the quality of the tax data was still sufficient to produce short-term statistics. For this analysis, edited questionnaire and tax data on the reference months June and July were available. By this time, the Ministry of Economic Affairs was putting a lot of pressure on SN to either stop primary data collection in the building industry, or else come up with concrete evidence that primary data collection was necessary. For this reason, the methodology department was only given just over two weeks to complete the analysis.

In their research, the methodologists used the tax data and the questionnaire data to compute two separate estimates of the turnover development in the building industry between June and July 2009. Since the estimate based on the questionnaire data was (asymptotically) unbiased, it could be used to assess the magnitude of the bias in the tax data estimate, due to a selective group of businesses having switched from monthly to quarterly reporting. An approximate 95% confidence interval was also computed for the growth estimate based on questionnaire data. It was decided that

the bias in the tax data estimate would be considered significant if this estimate fell outside the 95% confidence interval of the questionnaire data estimate.

The results of the analysis were not clear-cut, as there were some indications that the questionnaire data suffered from bias due to non-response. It was concluded, however, that the analysis had not revealed a substantial bias in the estimate based on tax data, and hence that it could not be used to show that primary data collection was necessary.

In a report that was completed on September 25, the methodology department advised SN to continue primary data collection for two more months, so that a full quarter of questionnaire data would be available for subsequent analyses. Assuming that these analyses would not lead to a different conclusion, primary data collection could be stopped after that period, although it might become necessary to start sending out questionnaires again in the future, if the number of businesses that used the option to switch would increase considerably.

4. The outcome

Following the methodological analysis described in Section 3.2, SN decided to stop primary data collection in the building industry starting with the reference month October. Primary data collection in the international trade sector had already been ended, following analyses by the methodology department and the economic statistics division.

Before sending the reports of the methodological analyses to the Ministry of Economic Affairs, SN asked two consultants from the RAND Institute to validate the findings in these reports. This was an initiative of the director of the economic statistics division, to forestall the Ministry's call for an independent enquiry into the matter. The validated reports were subsequently sent to the Ministry, along with the announcement that SN intended to stop collecting questionnaire data in the building industry and international trade sector.

After this announcement, there followed some controversy between the Ministry of Economic Affairs and SN about the precise formulation of a response to the parliamentary questions of July 22, which still remained to be answered. Formal answers to these questions were finally delivered in writing to parliament by Mr. Heemskerk, the State Secretary of Economic Affairs, on December 4. By that time, the interest of members of parliament in the matter had subsided, as SN had already ended the data collecting activities that concerned them.