Journal of Official Statistics, Vol. 26, No. 3, 2010, pp. 559–575

Using XBRL in a Statistical Context. The Case of the Dutch Taxonomy Project

Marko Roos[1]

In this article the background of XBRL and the involvement of Statistics Netherlands in the Dutch Taxonomy Project are discussed. The discussion predominantly focuses on the statistical context of using XBRL and the Dutch Taxonomy for communicating data terms to companies. The aim of the article is to make a case for XBRL as a standard that statistical offices, in cooperation with other regulators, should use for communicating their data terms to companies. Statistics Netherlands is a pioneer in this field and has been faced with many practical and methodological difficulties. These are also discussed. In addition to this the methodological implications of XBRL (and the Dutch Taxonomy) based data collection are discussed. The latter will become important when XBRL acceptance becomes widespread.

Key words: Electronic data collection; XBRL; taxonomy; administrative software; accounts.

1. Introduction

Statistical offices are responsible for the publication of important national economic indicators such as gross national product, the national accounts, structural business statistics and short-term indicators. For producing those indicators statistical offices rely on data provided, either directly or indirectly, by companies and institutions (UN Economic and Social Council 2004). The received data reflect the economic activities of those companies.

Relevant data about the economic activities of companies are traditionally collected by statistical offices through paper or electronic forms in which companies provide the requested data. Statistics Netherlands is also increasingly relying on tax office data, such as value added tax and profit tax data (Göttgens et al. 2005). Still, a significant amount of data is provided directly by companies.

To a large extent, the economic activities of companies leave an administrative trail. The activity of purchasing stock, for example, is recorded in the ledger administration where individual transactions are balanced, for example the increase in stock equals the decrease in the "accounts payable." Besides the ledger administration, the activity can also be recorded in a stock administration and in an administration for the accounts payable. Physical evidence of the transaction is the actual stock itself and the receipt. Hiring staff gets administered in the wages administration and, again, in the account book.

[1] Statistics Netherlands, Department of Methodology, PO Box 4481, 6401 CZ Heerlen, The Netherlands. Email: mros@cbs.nl

Acknowledgments: The author thanks Jelke Bethlehem, Piet Daas, Deirdre Giesen, Marc van Hilvoorde, Frank van der Pol and Jean Ritzen for their invaluable feedback on earlier versions of this article.

Most of the data that reflect the economic activities of companies and that are of interest for statistical offices can be found in the account book (or general ledger) and the trial balance (Lammers 2004). Those are kept in the (nowadays largely digital) book-keeping systems of companies. Other data of interest for statistical offices can also often be found directly, or in a derived form, in administrative systems (see for instance Peltola 2007). In this article, a book-keeping system is regarded as a special case of an administrative system. The book-keeping system keeps track only of the financial transactions of economic actors, whereas administrative systems in general can also record nonfinancial transactions.

For statistical purposes, companies now often print out the data in their administration systems and then key those figures into the (electronic) form provided by the statistical office (Giesen 2007). Linking the data found in the (digital) administration systems to the data terms of the statistical office would provide a more efficient and less error-prone data collection process. It would also reduce the administrative burden imposed by a statistical office. A digital data exchange format would provide the computer language for that link. Such an exchange format typically defines the lay-out and contents of an electronic message and often also defines the transport protocol layer (e.g., the Open Systems Interconnection (OSI) model gives a description of data exchange format layers).

In the past various attempts have been made by Statistics Netherlands and other statistical offices to supply companies and their software providers with a digital data exchange format. Examples of such attempts are the International Trade Statistic data format (Centraal Bureau voor de Statistiek 2007), The Teler initiative (Knüppel and Kunzler 2001), the EDIsent tool (Piebinga 1999), the Dutch "Elektronische Heerendiensten" initiative (EIM 2001) and the EGS-POS pilot (Ministry of Economic Affairs 2005). The electronic data exchange format can be used by software providers to export the data required by the statistical office out of various administrative systems. Experience has shown that the implementation rates of these initiatives are low. This is probably caused by the fact that software providers do not expect a profitable return on investment. To date, of the mentioned initiatives, only the format for the International Trade Statistics is still supported. Exact numbers of software providers implement the data exchange formats are lacking, but from the low success rate of these Dutch initiatives it is apparent that numbers were not encouraging in the Netherlands.

The business case for implementing those standards is made unattractive by two factors. First of all, the content and layout of the digital formats is often explicitly described in documents called message implementation guides. The software provider has to interpret those guides and implement the requirements into the software. Changes in either content or layout mean that those guides and software have to be updated. It takes precious IT-development time to correctly interpret, implement and test the digital exchange format and updates.

Another critical factor in the lack of acceptance of digital exchange formats imposed by statistical offices, is that statistical offices are not the only organizations (in the XBRL community called "regulators") issuing such formats. Customs, tax offices, banks and insurance companies are just a few examples of regulators issuing their own, often unique, data exchange formats. Since a statistical office by itself does not always provide enough

potential customers (and customer pressure), the urge to implement the statistical digital exchange format is low for software providers.

Crucially, regulators often require similar or partly overlapping sets of data. So, despite a multitude of different layouts for the existing data exchange formats, often the same administrative facts are required. Standardization of the digital exchange format of the regulators could provide the leverage needed to make the business case for the software providers profitable. It would provide software providers with an opportunity to invest once in a single and flexible data format that could be used to send data to all regulators electronically. Implementation of such a standard would enable the software to interpret the data terms provided by the regulator (the data a regulator wishes to receive) in a standardized yet flexible way. This will make it unnecessary for software providers to "hard code" the data terms in their software.

XBRL, short for eXtensible Business Reporting Language, may very well be this standard format (Collet et al. 2003; PWC 2003). Developed and maintained by a consortium of regulators, accountants and software builders, it can offer a link between the data kept in book-keeping systems and the data terms of regulators, such as national statistical and tax offices. In the Netherlands, XBRL is introduced and used for this purpose. The initiative behind the use of XBRL is called the Dutch Taxonomy Project. It is a joint effort of the Dutch Ministry of Finance and the Ministry of Justice (2003). It aims at reducing the administrative burden on companies in the Netherlands to a large extent. Statistics Netherlands, the Dutch Tax Office and the Dutch Chambers of Commerce are the regulators involved in the Dutch Taxonomy Project. The prospect is that the project achieves its goals by harmonizing legislation surrounding mandatory regulatory reporting and by introducing a joint supported single electronic data format for expressing the data requirements of regulators.

XBRL is used because it is a standard for expressing administrative data. It is not a standard developed for electronic data transport. Although it plays a vital part in enabling administrative burden reduction, the infrastructure necessary for transporting XBRL messages is not within the scope of the article.

In this article I introduce XBRL and the concept of reusing variables in XBRL taxonomies. A distinction is made between a company driven XBRL approach and a regulatory XBRL approach. Next, the key features of the Dutch Taxonomy are presented. Relating the taxonomy to the administrative facts is at the heart of the matter and is discussed next. After that, the methodological implications of XBRL and the approach of the Dutch Taxonomy are presented and discussed. The article concludes with a look at the prospects of XBRL and the impact it can have on regulators and companies in general and statistical offices in particular.

2. An Introduction into XBRL and XBRL Taxonomies

2.1. Introducing XBRL

The acronym XBRL stands for eXtensible Business Reporting Language (XBRL.org 2003). It is an XML-based computer language specifically developed for the exchange of business facts between computer systems. In this article, business facts are defined as

administered events that are of economic interest to the company or related organizations. Examples are borrowing money from a bank, issuing stock, selling goods and depreciating the inventory of a company.

The XBRL standard provides a precise, predictable structure for describing and expressing those business facts in a way that can be used and processed by computer systems. The XBRL standard defines how software should generate and process XBRL documents. It is developed and maintained by a global nonprofit consortium of accountants, software vendors and regulators, among others (see www.xbrl.org).

XBRL and XBRL-enabled software make it possible for reporters to define and express the data they (have to) provide in a flexible format that can readily be used and interpreted by receiving parties without the need to agree on a fixed data structure (Collet et al. 2003). The actual business facts are contained in an XBRL instance. An instance contains data. The business facts reported in an XBRL instance are defined in an XBRL taxonomy. An XBRL taxonomy defines variables and the relations that may exist between those variables. A taxonomy may also refer to variables defined in other taxonomies.

Business facts in an XBRL instance are multi-dimensional in the sense that they relate a single variable (e.g., "profit and loss before tax") to different periods in time (e.g., January 1st 2006 to December 31st 2006 and January 1st 2007 to January 31st 2007). The variable can also relate to different entities such as subsidiaries (e.g., Acme Transportation Services, Acme Holding Inc.) and can be expressed in different measurements (e.g., dollars, euros, square meters).

In short, business facts expressed in XBRL instances must always relate to

(a) variables defined in taxonomies,
(b) specific time periods or points in time,
(c) specific (monetary or physical) units, and
(d) specific entities.

XBRL-enabled software uses the information contained in an instance and in a taxonomy to display the business facts in a user interface.

2.2. *XBRL Taxonomies and Reusing Variables*

Each variable used in an XBRL instance must be defined in an XBRL taxonomy. The XBRL taxonomy defines a variable by giving it a unique name, a specific data type (e.g., monetary, string, numeric) and a period type. The period type of a variable reflects whether it describes a period (e.g., "net profit" in the annual account which describes the profit made during a year) or a point in time (e.g., "products in stock," which describes the stock at a given point in time). In addition, a taxonomy can also provide a variable with labels in any number of languages, and with references to, for example, legislation textbooks or websites containing additional information.

The taxonomy also defines variables by relating them to each other. Housing costs for example can be related to energy costs, rent and cleaning costs. Assets can be related to property, plant and equipment, receivables, cash and inventories. The profit a company makes can be defined as total sales minus total costs. Business reports are in general organized into identifiable hierarchical data structures such as described above.

Those structures can also be defined in the taxonomy. XBRL distinguishes between the presentation of variables and the calculation of variables. In the presentation part of the taxonomy (called the presentation linkbase) structures (or "views") such as the "Balance sheet" and the "Profit and Loss accounts" can be defined. In these views the order and hierarchy of variables (both of a monetary nature and explanatory disclosures) are defined. In the calculation linkbase the way variables (of a monetary nature) add up is defined.

XBRL-enabled software uses the information provided by the taxonomy to present the data in XBRL instances in different languages and in different "views" (e.g., Balance Sheet, Cash Flow Statement, Profit and Loss Account). A taxonomy will generally be published on a website where it can be accessed directly by XBRL-enabled software. Based on the information contained in the taxonomy, the software checks whether a variable is expressed in the right format (e.g., 21 instead of 21.3) or whether the values of certain variables add up correctly. In short, XBRL taxonomies allow creators of taxonomies to take control over how data will be presented in XBRL instances.

As mentioned before, every variable used in an XBRL instance must be defined in an XBRL taxonomy. An XBRL instance may contain variables defined in different taxonomies. This important feature of XBRL enables the reuse of variables defined in existing taxonomies. The International Accounting Standards Board (IASB), for example, publishes an XBRL-taxonomy version of the International Financial Reporting Standards (IFRS, see www.iasb.org/xbrl). This XBRL taxonomy contains numerous variables relevant for the filing of the annual accounts of companies (e.g., balance, profit and loss account). In an XBRL instance several variables defined in the IFRS taxonomy can be combined with variables defined in another taxonomy. Usually the latter "extends on" the already existing taxonomy. For example the IFRS taxonomy can be referenced in the taxonomy created by "Acme Inc." for the filing of their annual account. Here, the "Acme Inc." taxonomy forms a so-called extension-taxonomy of the IFRS taxonomy. Variables not defined in the referenced taxonomy (for example a more detailed breakdown of the total cost structure) should be defined in the extension taxonomy and, if possible, related to variables of the taxonomy it extends on (e.g., Total Costs in the IFRS taxonomy). The instance containing both the IFRS variables and the extension-taxonomy variables is based on the extension-taxonomy. As such, the extension-taxonomy forms the "entrance point" of what is commonly called a "Discoverable Taxonomy Set" (DTS).

2.3. Publishing Taxonomies: Regulator Versus Company-driven Approach

XBRL was designed with the publication of business facts in mind. Typically a company filing its annual or quarterly reports would extend on a local GAAP (Generally Accepted Accounting Principles) taxonomy and add its own specific variables into an extension taxonomy. For instance "Acme Inc." would extend US-GAAP with variables concerning the appreciation principles of its unique employees. It would then file its annual or quarterly account on its website or send it to its shareholders (Teixera 2005; Coffin 2001).

However, a regulator is not interested in companies sending in business facts based on "company-specific" extended taxonomies. Data based on individualized taxonomies are harder to compare and it takes time and effort to interpret them. Often the knowledge needed to interpret the extension taxonomy of the company is possessed by the company

or its accountant. It is better for regulators to publish their specific data terms in a separate regulator taxonomy, extending an already published taxonomy if necessary. This enables the providers of the business facts to apply the taxonomy to the information contained in their own administrative systems, using their own (XBRL-enabled) software tools. As an example, a statistical office could create and publish a taxonomy that extends on the IFRS taxonomy with a more detailed and specific definition of the variables contributing to the costs and revenues. The tax office could create a(n extension) taxonomy for the wages and salaries declarations. Banks could create a taxonomy to assess the solvency of companies. The only thing companies would have to do is map their administrative data to the variables in the regulatory reporting taxonomies. The companies can subsequently produce an XBRL instance concerning a specific period, containing the mapped variables according to the data terms of the regulator as published in the regulator taxonomies.

Often the taxonomies of different regulators will contain similar (overlapping) information. It would be very convenient for companies if they could map this combined information only once instead of repeating this step for the different (but in principle the same) variables in each regulator taxonomy. A substantive reduction of the administrative burden, depending on the extent of overlap, of companies could be achieved if different regulators shared variables.

In the Netherlands, the central government acknowledged this possibility and started the Dutch Taxonomy Project. In this project the possibilities of XBRL are used to combine the data terms of the tax office, Statistics Netherlands and the Foundation for Annual Reporting (responsible for the content of the annual accounts that are filed with the Chambers of Commerce). The latter is responsible for maintaining the Dutch GAAP. Combining the data terms of these organizations in a single taxonomy, using one technical standard, could provide the leverage needed for software providers to build in the necessary functionality.

3. Description of the Dutch Taxonomy and Its Core Components

3.1. A Modular Approach

Each regulator participating in the Dutch taxonomy has a number of reports that companies have to complete periodically. Those reports serve different purposes, such as establishing the turnover tax to be paid, determining the profit of a company or the wage tax to be paid, providing a basis for estimating the production of goods or publishing an overview of business annual accounts. There are significant similarities in the data gathered in those reports. Part of the information is present in the administration of a company because of the mere fact that the information is requested by the regulator, but in a number of cases the information needs to be derived from other data .

The main goal of the Dutch Taxonomy Project was to achieve a large reduction in the administrative burden of companies caused by the national government (Ministry of Justice 2004). This goal should be achieved by using a single technical standard for expressing business facts (XBRL) and establishing a common vocabulary among the participating governmental parties (expressed in a combined taxonomy). The hope is that this goal can be achieved by eliminating or diminishing the need for different “niche”

software and specialized reporting accountants and by minimizing the time needed to complete reporting obligations. Through harmonizing legislation, normalizing data terms of regulators and establishing one technical standard, the project is expected to reach its targets.

The Dutch taxonomy is built up using a modular approach (Figure 1). This means that not all the information is put into one single (large) taxonomy, but in a set of small, easier to maintain, taxonomies. The Dutch Taxonomy has an architecture in which common taxonomies (with variables that are shared between parties) and domain-specific taxonomies are combined (Daas and Stroom 2006). In addition, a distinction is made between taxonomies in which variables are defined, data types are introduced, and taxonomies in which variables are related to each other in the presentation and calculation linkbase. These last taxonomies are the so-called formsets. Formsets closely resemble (parts of) the regular paper forms. In the formsets the relevant variables from the common and domain taxonomies are selected and grouped together.

The starting point for a company filing a report is the reportset. This is a taxonomy that bundles one or more formsets together. In a reportset all variables required to produce a filing for a specific regulator are combined. Companies obliged to send in a specific regulator report (e.g., structural business statistic) select the appropriate reportset in the Dutch Taxonomy and create an XBRL instance on the basis of the information available in that report.

Figure 1 displays an overview of this approach. The bottom layer is the IFRS layer, in which variables are defined that are used by all regulators. The layer above extends on the IFRS taxonomy and adds variables and data types that are used by more than one regulator. The layer above the common layer extends on the common taxonomy and defines variables unique to Statistics Netherlands. In the formset layer the variables that are relevant for a specific reporting obligation (short-term statistic: STS, structural business statistic: SBS and investment statistic) are selected and grouped together.

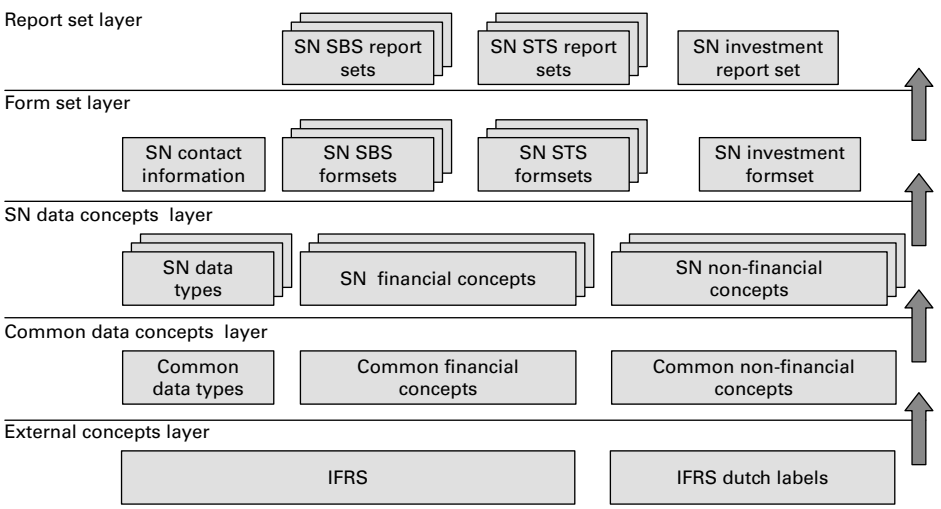

Fig. 1. An overview of the modular approach of the Dutch Taxonomy

In the reportset layer the formset "contact information" and a formset for a specific reporting obligation are grouped together.

3.2. Harmonization and Normalization Processes

As discussed in the previous section, regulators use variables out of their respective "domain bases" (taxonomies containing only regulatory-specific variables) and variables that are shared with other parties in their formsets. The most important shared taxonomy is the IFRS taxonomy. Other shared taxonomies are the NL-gen ("Dutch generic") base and data type taxonomy and the NL-cd ("Dutch common data") base and data type taxonomy. The NL-gen base taxonomy contains the shared financial variables. The NL-cd taxonomy contains the shared nonfinancial data such as the address and contact person information.

In the Dutch Taxonomy Project there are two separate processes for achieving as many shared variables as possible. They are called the normalization and the harmonization process. In the normalization process content matter experts of various domains compare definitions of variables and decide which variables to move to the shared base. In the harmonization process, legislation is adapted to overcome legal obstacles to cross-domain reporting. The normalization and harmonization processes play an essential part in achieving the goals set for the reduction of administrative burden.

3.2.1. Normalization

In the case of the Dutch Taxonomy Version 1.1, the normalization process was mainly done in cooperation between the Dutch tax office and the statistical office and also within these offices themselves. In the majority of cases, Statistics Netherlands unilaterally decided which tax office variable suited the statistical definition of a variable. In a few cases sharing variables was the result of coordination between the statistical office and the tax office. The normalization process resulted in an NL-gen base taxonomy (Version 1.1) containing 149 shared financial variables. As an example, the reportset taxonomy of the structural business statistic contains four IFRS variables and twenty-seven NL-gen base variables out of a total of eighty-seven variables. Overall, for Tax Office Netherlands and Statistics Netherlands combined, the percentage of shared variables compared to the total of variables used is thirteen.

In the 1.1 version of the Dutch Taxonomy, no normalization process was undertaken for the annual accounting part. The domain taxonomy for the annual accounting defines roughly 1,400 unique domain-specific, variables. Of course, for annual accounting purposes only subsets of the total number of variables have to be used.

3.2.2. Harmonization

The harmonization process was not applied to the 1.1 version of the Dutch Taxonomy. When it came to the 2.0 version of the taxonomy, however, the effects of the change in legislation were noticeable.

The change in legislation concerns small registered (with an annual accounting obligation) companies. Under the old legislation those companies were obliged to make a distinction between appreciation of their assets on a fiscal basis (historical costs) and

appreciation on a commercial basis (current value). Under the new law, small companies can limit themselves to the appreciation of their assets on a fiscal basis. In practice this means that the information used for the fiscal annual account under the new law can also be used legally for the filing of the company's commercial annual account.

4. Relating the Book-keeping Systems to the Dutch Taxonomy

A reportset in the Dutch Taxonomy defines which variables can be present in a specific report. The report itself, containing the actual business facts, is called an XBRL instance document. As was explained earlier, an instance document contains the values of the variables directly related to the variable name, the context of a value (the period and company or entity it describes) and the unit on which it is based (e.g., dollars, euros, square metres) (Daas and Stroom 2006).

The instance document will usually be created by XBRL-enabled software (although this could theoretically also be done in Notepad or any other text editor). Ideally administration software (book-keeping and other) with built-in XBRL functionality at the core of the system would be used. On a general ledger level, all accounts should be mapped to taxonomy variables; and in combination with XBRL Global Ledger (XBRL-GL) technology, an aggregation of the accounts will lead to the desired business facts (see, for example Prather and Ramin 2003). Once that is the case, reporters are able to create an XBRL instance document directly from their journal entries (where individual transactions are recorded) covering a certain period. To achieve this a mapping of the taxonomy variables to the accounts they use, a unit definition and a selection of the period the reporters have to report about are needed. For more complex companies additional consolidation information could also be required.

Unfortunately, as to date there are no systems available that have XBRL functionality at the core of their systems (Bottemanne 2007). In addition, the majority of the reportsets in the Dutch Taxonomy are a more or less direct translation of the traditional paper filings and variables defined in those filings. As such they are not always directly relatable to the accounts used by companies (Lammers 2004; Hilvoorde 2005).

Book-keeping software builders are now beginning to implement XBRL functionality. Traditionally in the software market external reporting is often a separated module from the book-keeping software. Usually most software builders who build-in XBRL functionality by adding XBRL-enabled reporting modules do not include XBRL in the core of their book-keeping system. The functionality is limited to a translation of the financial report into an XBRL instance. In addition, many software builders have traditionally specialized in either fiscal reporting or annual accounting (Bottemanne 2007) and very often "hard-coded" a fiscal filing report into their reporting module. This is a labour-intensive process from which especially fiscal reporting software builders derive a substantial percentage of their added value. So, although a number of software builders will have XBRL reporting functionality in their systems and will be able to generate XBRL instances, the existing external reporting modules will be mainly aimed at fiscal or annual reporting and not statistical reporting. The pick-up rate on the latter thus lags behind.

As an alternative for generating XBRL instance documents with administrative software, reporters could use XBRL-enabled (web)forms to generate valid XBRL instances. This approach is taken by the central bank of Belgium. In Belgium companies are obliged to file their annual accounts in XBRL format and are offered XBRL-enabled web forms. In these forms they can both import XBRL instances *and* they can also key in the requested data (Vanderhaegen 2006). The underlying web service then creates a complete XBRL instance document, which can be used by both the reporter and the regulator. Because the accounting system for companies is standardized in Belgium, a company can directly relate its administration to the government-prescribed accounts.

At Statistics Netherlands a proof of concept was carried out in cooperation with Adobe systems to demonstrate functionality to import and export XBRL data in and out of Adobe forms (Roos and Daas 2008). Once implemented, this would facilitate a flexible environment in which companies can import XBRL data generated by their external reporting modules and complete the missing information with the aid of the Adobe form.

5. Using XBRL Data in the Statistical Context: The Practical Approach

5.1. The Ideal Situation

In the ideal situation, companies map the complete content (all the variables used in the formsets) of the Dutch taxonomy to their general level accounts used in their book-keeping systems (using both XBRL-GL and the Dutch Taxonomy). Such book-keeping systems could be used to file all consolidated data (available for reporting) on a monthly basis (or longer) to a governmental gateway. The gateway would make the data available to all government agencies that are allowed to make use of them. Those agencies would translate or map the information contained in the reports into their own variables. All governmental agencies would use a single company identification methodology including consolidation and deconsolidation information. This information would be used to aggregate company data to the desired company structure (e.g., kind of activity unit). In this way, instead of information based on samples, information for the complete population would be available.

Besides the ethics of the "Big Brother" aspect, the current practice is still a number of obstacles away from the ideal situation described above. In this section the obstacles are described. The practical approach developed by Statistics Netherlands carefully avoided the obstacles and allowed the use of the Dutch Taxonomy without major problems.

The choices made by Statistics Netherlands may at first hand seem to be limiting the full use of all XBRL possibilities. When XBRL and the Dutch Taxonomy become more established, however, those choices are not irreversible. With sufficient market acceptance there will be enough leverage to implement the more challenging aspects of XBRL.

5.2. Reports

First of all, a reporter does not send in a single XBRL instance document to cover the needs of all regulators and all regulators' reporting requirements. Instead, a reporter is required to send in an XBRL instance document for each reporting obligation. The instance document is sent in via a governmental gateway and the gateway relays

the instance document to the designated regulator only. Other regulators are not able to use the information contained in that regulator-specific document.

At Statistics Netherlands six different reports for Version 1.1 of the Dutch Taxonomy were defined. Four are activity-specific short-term statistic reports and the others are the investment and lease statistic report and a structural business statistic (SBS) report. The different activity-specific structural business statistics forms (totalling 189 forms) that are sent in the traditional modes are not represented in the SBS reportset. The reason behind this is that Statistics Netherlands does not want to burden a company with a choice among a large number of reportsets. As a result, a significant number of detailed, form-specific variables were dropped from the Statistics Netherlands part of the Dutch Taxonomy. It was also decided for 2008 not to compensate the loss in detail by an additional demand of data from companies who had already sent in XBRL data. As the XBRL mode becomes more established it may very well become possible to make more detailed and specific taxonomies for companies.

5.3. Definition of a Company

Statistic Netherlands requests data from companies that reflect the economic activities of the company as defined by Statistic Netherlands. The definition of Statistics Netherlands of a company is based upon homogeneity, market orientation and independence in the decision-making process. The company as defined by Statistics Netherlands is called the enterprise. Often the enterprise comprises of one or more legal units, the number largely depending on the size of the company (Struijs and Ruyl 1995).

The company's book-keeping system will not necessarily reflect the transactions of the enterprise. In some cases a company might be keeping the books of more legal units than those included in the enterprise, and in some cases the book-keeping system will only contain a part of the legal units.

As XBRL is expected to enable extraction of data out of book-keeping systems, those data will not necessarily reflect the economic activity of the enterprise, but rather a superset or subset of the legal units making up the enterprise. For this reason, the decision was made that *for now*, only companies with 50 or fewer employees can send in XBRL data based upon the Statistics Netherlands' part of the Dutch Taxonomy. Beyond this size, companies tend to have complex enterprise structures. In the documentation describing the XBRL document instance structure, on the website of Statistics Netherlands and in the letters that companies receive from Statistics Netherlands, it is also stated that the (consolidated) data have to be based on the enterprise as defined by Statistics Netherlands.

The Dutch tax office already uses a more integrated approach: it requests (consolidated) data reflecting the fiscal unit, but it also requests data of all legal (fiscal) entities that comprise the fiscal unit in the same XBRL instance document.

When XBRL becomes more established, Statistics Netherlands may also allow more complex companies to send in XBRL data.

5.4. The Structure of the Variables: The Nonatomic Approach

As mentioned earlier, the Dutch tax office and Statistics Netherlands' parts of the taxonomy are more or less direct translations of the regular paper or electronic filings.

Because of the heterogeneous nature of accounting systems the business facts of companies may not often be directly related to the variables used in the taxonomy. This is especially difficult because at this stage it was not yet possible to achieve univocal definitions of what accounts contribute to a specific taxonomy variable and what accounts do not. For example, Statistics Netherlands defines "housing costs" by totalling the accounts "rent and lease costs," "energy costs" and "cleaning costs." Government grants for housing have to be added to the account "revenues." If the Dutch tax office were to have housing costs defined as rent and lease costs and energy costs, but not cleaning costs, those "smaller" more detailed variables could be used to establish a more harmonized financial data model. The small (atomic) building blocks of diverging variables could be used as a common vocabulary. As to date, the taxonomy lacks such an atomic structure.

Because of this, mapping of the variables to the (smaller, atomic) accounting systems of companies by companies or their accountants is difficult. Often interpretations need to be made by second-guessing or by looking up additional information. Once a choice has been made for a specific definition of a variable (where for example the account "housing costs" includes the gas bill, but not the cleaning costs) it will not be likely that accounts will be remapped for another reporting obligation.

At Statistics Netherlands we choose however not to define the variables down to the atomic level, because to the public this would look like additional administrative burden (the number of variables that Statistics Netherlands use would go up). The leverage of other (larger) parties such as the tax office is needed to introduce an atomic structure into the taxonomy. The political mandate of the Taxonomy Project lies in the reduction of the administrative burden, so the odds of introducing a mandatory taxonomy with an atomic structure (meaning more variables to report) are low. However, since take-up rates are still not up to expectations, it could be that what's missing from the business case for companies is the ability to directly map to the accounting systems. So there might be paradox in the case for an atomic structure: reporting more detailed information might mean lower administrative burden. A possible "hybrid" suggestion towards a solution is introduced in Section 6.2.

6. (Methodological) Issues in Using XBRL Data in a Statistical Context

The Dutch Taxonomy and XBRL are still in their early adoption phase. In 2007 only a small number of XBRL test instances were received. For 2008, however, we are aware of five software companies that have announced that they will "XBRL enable" their book-keeping and administration software. In contrast, for tax filing and annual accounting over 2 million XBRL messages are factored in. The difference is due to the fact that every enterprise has to report to the tax office, whereas statistical offices use samples and hence generate a smaller market for XBRL-enabled software. Therefore Dutch Taxonomy-based statistical reports are expected to pick up slower than fiscal reports. When the XBRL technology is correctly implemented, however, sending in XBRL-based statistical reports should not be much of an extra effort.

When the XBRL pick-up rate accelerates, Statistics Netherlands will be faced with some challenging methodological and process-based issues. Those issues are discussed in the following subsections.

6.1. Missing Data in the SBS Taxonomy

As was described above, a large number of detailed variables were dropped from the structural business statistic part of the taxonomy. Those variables relate largely to product-specific revenues and costs. These detailed variables are necessary for producing the national accounts and other statistics that are obligatory and prescribed by Eurostat. If this level of detail is missing from the taxonomy, it is obvious that companies cannot provide Statistics Netherlands with those data in XBRL instance documents. Somehow the detailed information has to be derived.

Stroeks (2007) describes two possible approaches for deriving the detailed information. He suggests using data of larger companies (that still have to provide the detailed information in questionnaires) for estimating this level of detail for smaller companies. Alternatively or combined with this approach, requesting detailed information could also be done not on a yearly basis, but rather on a 2–5-year basis. In both cases, a measure would be necessary to verify the stability of the estimator. Stroeks states that it is possible to still provide the detailed information within reliability constraints, but also notes that it would be time-consuming and costly to implement.

Of course, the taxonomy could also be extended so that the detailed variables are present. Technology for this is available yet costly in terms of development and implementation, especially for the software developers. The current approach of Statistics Netherlands of not requesting the detailed information from companies sending in XBRL data is very likely to be continued, largely because of reductions in administrative burden.

6.2. XBRL Data Not Based on the Enterprise

If adoption of the Dutch Taxonomy reaches majority status, a large number of companies and administrative companies will be able to easily send in XBRL data to the governmental gateway. As discussed above, messages would still be related to a single report for a single regulator, but that situation does not necessarily need to hold ground. In the future it is conceivable that companies periodically send in a "copy" of their book-keeping system in XBRL format to the governmental gateway. These data will reflect the activity of a company structure that does not necessarily follow the definition of Statistics Netherlands and would not always contain all the data required by Statistics Netherlands. Timeliness could for some companies be better, and for some worse than Statistics Netherlands desires.

It would confront Statistics Netherlands with a situation where an unprecedented large, but unpredictable and dynamic dataset becomes available. It would be highly desirable (if not obligatory) to use the available data, but it would mean more complex sampling methods, very flexible data collection processes (e.g., dynamic questionnaire generation) and a flexible data warehouse environment. Some of these approaches are foreseen at Statistics Netherlands, but some also remain beyond the horizon. The methodological issues surrounding data collection based on the Dutch Taxonomy resemble those surrounding data collection based on register and administrative sources (Statistics Finland 2004).

6.3. The Atomic Structure

As was discussed earlier, the Dutch taxonomy to a large extent lacks an atomic structure. This means that reported XBRL variables do not necessarily have to reflect the economic indicator Statistics Netherlands expects them to. A company will not always read the definition, and if it does not, will not always be able to map the correct accounts to a variable.

If an extension to the taxonomy were to be introduced in which atomic variables were related to the variables that Dutch regulators require, Statistics Netherlands would have to make an effort to deepen knowledge of the accounting systems used by companies in the Netherlands. It is likely that companies are very heterogeneous in the way the accounts are arranged, but it is also likely that they are more or less homogeneous within specific branches. These accounts would have to be mapped and used to relate the different accounts and arrangements of accounts to variables of Statistics Netherlands.

As mentioned earlier, the Netherlands lack a central system of accounts. Although this contributes to flexibility of the management of company finances, it makes external regulatory reporting for companies complex and expensive. The introduction of an atomic structure into the taxonomy (with nonmandatory variables) could cause a paradigm shift in the way book-keeping is done in that still different accounting systems could be used but companies would be more inclined to coordinate their accounting systems with the (atomic) variables in the Dutch taxonomy. A disadvantage of this approach is that implementation costs could get higher.

7. Conclusions

Data derived directly from company administrations reflect the administrative reality of companies more precisely than paper forms would do. Defining and publishing data terms of statistical offices and receiving matching data directly from company administrations has been a goal of statistical offices for a long time. Those data could be more accurate and timely and could quite possibly reduce the administrative burden imposed by statistical offices significantly.

In the introduction of this article it was stated that relating statistical data terms to the electronic administration of companies using a specific unilaterally imposed electronic format did not work, the reason for this being that there is no profitable business case for builders of administrative software. The business case is not profitable because of the costs for implementing such a format and the relatively small market of companies that have to provide statistical data.

A standard for defining data terms that would both be generic and would provide an extensive market should make the business case for software builders more attractive. With XBRL such a standard seems to be available.

Besides offering a standard electronic format for defining and publishing regulatory (financial) data terms it also offers the possibility of reusing data and data terms. This makes it possible that the regulatory reporting requirements of several regulators are combined in one taxonomy (or set of taxonomies).

The Dutch Government recognized the possibilities and started the Dutch Taxonomy Project. This initiative aimed at reducing the administrative burden of companies imposed

by the Dutch government. It combined the data terms of the Dutch statistical office, the Dutch tax office and the Foundation for Annual Reporting in a set of XBRL taxonomies (called "The Dutch Taxonomy"). The data terms went through a normalization and harmonization process which resulted in a substantial reduction of the number of variables in the taxonomies. Together with the creation of the taxonomy, an infrastructure for sending in XBRL data was established.

The 1.1 version of the taxonomy, published in March 2007, did not result in many XBRL instances. Despite the signing of a covenant, software providers seemed to be hesitant and perhaps somewhat intimidated by the requirements of the XBRL technology. In 2008 the tax office and Chambers of Commerce received some reports, all send in by a single intermediary. In 2009 the Dutch government has changed its policy. The focus is now on increasing the number of users instead of expanding the scope of SBR.

The change of policy seems to be paying off: accountancy firms are requesting XBRL-enabled software. The larger software builders now all claim to be implementing the Dutch Taxonomy and XBRL technology. Larger banking organizations such as Rabobank, ABN-Amro and ING participated in the banking taxonomy and now are preparing for processing XBRL reports (Accountancynieuws 2010). Tax Office Netherlands and the Chambers of Commerce have reported an increase in a number of XBRL instances received (Accountancynieuws 2009).

For Statistical Offices, XBRL could mean quite a change in the data collection process. If XBRL is widely implemented, larger numbers of administrative data will become available, and much faster, than with the traditional methods (including register-based data). Measuring the "real-time economy" comes one step closer.

However, there are a number of pitfalls and obstacles. For instance, the way variables are defined (using smaller building blocks on the one hand or more uncalibrated ones on the other), the way companies are defined and the way missing data are handled.

XBRL and the Dutch Taxonomy approach could eventually lead to a paradigm shift in regulatory reporting: the Dutch government defines collectively and univocally all required regulatory variables, software builders "XBRL-enable" their general ledger software and integrate the (different) financial reporting modules, and companies and accountancy firms adjust their accounting systems so that they can be mapped well-defined to the Dutch Taxonomy variables. Such a paradigm shift would in the long run benefit all parties involved. The more parties are involved, the more likely the paradigm shift becomes.

8. References

Accountancynieuws (2009). Kamer van Koophandel verwacht in 2010 forse toename aantal XBRL deponeringen. Available at http://www.accountancynieuws.nl/actueel/branche/kamer-van-koophandel-verwacht-in-2010-forse.86163.lynkx?PostedField%5Bkeyword%5D=xbrl+toename&Confirmed=Zoeken [In Dutch]

Accountancynieuws (2010). De bankentaxonomie: banken automatiseren en standaardiseren kredietrapportages. Available at: http://www.accountancynieuws.nl/actueel/branche/de-bankentaxonomie-banken-automatiseren-en.90803.lynkx [In Dutch]

Bottemanne, G. (2007). Marktanalyse: software voor accountantskantoren [market analysis of software for accountancy offices]. (on online request) Available at www.GBNed.nl [In Dutch]

Centraal Bureau voor de Statistiek (2007). International Trade, Manual for International Trade Statistical Declarations: Record Layout for Electronic Reporting. Available at E-mail: infosupport@cbs.nl

Coffin, Z. (2001). Top Ten Effects of XBRL. Management Accounting, 82(12), 64–67.

Collet, L., Kirsch, P., and Cuvelier, S. (2003). Trusted and Efficient Financial Reporting. Working paper PWC. Available at: http://www.pwc.com/en_GX/gx/xbrl/assets/pwc_trusted_efficient_financial_reporting.pdf

Daas, P. and Stroom, A. (2006). Taxonomy and Instance Creation and Control. Discussion Paper Statistics Netherlands. Available at http://www.cbs.nl/nl-NL/menu/methoden/research/discussionpapers/archief/2006/2006-02-x10-pub.htm

EIM (2001). Onderzoek naar administratieve lasten en uitvoeringskosten [Research into administrative burden and costs of execution]. Available at http://www.eim.nl/index.cfm/11,615,c,html/Administratieve-lasten-uitvkosten.pdf [In Dutch]

Giesen, D. (2007). The Response Process Model as a Tool for Evaluating Business Surveys. Third International Conference on Establishment Surveys (ICES-III), Montreal, Canada.

Göttgens, R., Snijkers, G., Beukenhorst, D.J., van Dam, G., and Verbruggen, M. (2005). Waarneemstrategie Centraal Bureau voor de Statistiek. Unpublished Report. Statistics Netherlands. [In Dutch]

Hilvoorde, M. (2005). XBRL en het getrouwe beeld [XBRL and the faithfull representation]. De Accountant, February, 32–37. [In Dutch]

Knüppel, C. and Kunzler, W. (2001). Influence of the Internet on Data Collection and Dissemination in the European Statistical System. Paper presented at the IAOS meeting. Available at http://www.stat.go.jp/english/info/meetings/iaos/pdf/knuppel.pdf

Lammers, J. (2004). The Transformation Procedures from the Enterprises' Accounting File to Financial Statistics. 55th ISI Conference, Australia.

Ministry of Economic Affairs Department (2005). Voortgangsrapportage administratieve lastendrukverlichting [Progress report on administrative burden reduction]. Letter to Parliament. [In Dutch]

Ministry of Justice (2003). Kabinetsplan aanpak administratieve lasten. Available at http://ikregeer.nl/document/kst-29515-8?format=pdf [In Dutch]

Peltola, I. (2007). The Experiences of Web Based Data Collection from Enterprises in Finland. Available at http://www.oecd.org/dataoecd/10/1/39608007.pdf

Piebinga, J. (1999). EDISENT – Use of Electronic Questionnaires. http://europa.eu.int/en/comm/eurostat/research/conferences/etk-99/papers/edisent.pdf

Prather, D. and Ramin, K. (2003). Building an IAS/IFRS Taxonomy Using XBRL. Available at http://www.ccbn.com/_pdfs/XBRL_09_02.pdf

PriceWaterhouseCoopers (2003). Standards Based Regulatory Reporting (white paper). Available at http://www.pwc.com/Extweb/service.nsf/docid/D8BEAA7056E6D7CC85256BF0005786A9/$file/GlobalRegulatoryPaper.pdf

Roos, M. and Daas, P. (2009). Results of Proof of Concept XBRL-Adobe. Unpublished report, Statistics Netherlands.

Statistics Finland (2004). Use of Registers and Administrative Data Sources Finland; Best Practices of Statistics Finland. Helsinki.

Stroeks, L. (2007). XBRL in de productiestatistiek [XBRL in the production statistic]. Unpublished report, Statistic Netherlands. [In Dutch]

Struijs, P. and Ruyl, H. (1995). Implementing the Concept of Autonomy. Paper presented at the Eurostat Working Group on Statistical Units.

Teixera, A. (2005). Implications of XBRL; Financial Reporting Research Opportunities. Available at http://www.iasb.org/XBRL/images/xbrllab/past_projects/xbrl_paper_for_iasb_2005.pdf

UN Economic and Social Council (2004). Implementation of the Fundamental Principles of Official Statistics. Official Records of the Economic and Social Council, (2003), Supplement No. 4 (E/2003/24), Chapter I.A. Available at http://unstats.un.org/unsd/statcom/doc04/2004-21e.pdf

Vanderhaegen, E. (2006). Internetverslaggeving met XBRL [Internetreporting with XBRL]. Maandschrift Accountancy & Bedrijfskunde, 10, 3–19. [In Dutch]

XBRL.org (2003). Extensible Business Reporting Language (XBRL) 2.1. RECOMMENDATION – 2003-12-31 + Corrected Errata – 2005-11-07. Available at http://www.xbrl.org/Specification/XBRL-RECOMMENDATION-2003-12-31+Corrected-Errata-2005-11-07.htm

Received April 2008
Revised February 2010