Kalemli-Ozcan, Sebnem; Sørensen, Bent; Villegas-Sanchez, Carolina; Volosovych, Vadym; Yesiltas, Sevcan

Working Paper

How to construct nationally Representative Firm Level Data from the ORBIS Global Database

Tinbergen Institute Discussion Paper, No. 15-110/IV

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Kalemli-Ozcan, Sebnem; Sørensen, Bent; Villegas-Sanchez, Carolina; Volosovych, Vadym; Yesiltas, Sevcan (2015) : How to construct nationally Representative Firm Level Data from the ORBIS Global Database, Tinbergen Institute Discussion Paper, No. 15-110/IV, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/125116

TI 2015-110/IV
Tinbergen Institute Discussion Paper

How to construct nationally Representative Firm Level Data from the ORBIS Global Database

Şebnem Kalemli-Özcan[1]
Bent Sørensen[2]
Carolina Villegas-Sanchez[3]
Vadym Volosovych[4]
Sevcan Yeşiltaş[5]

[1] *University of Maryland, United States;*
[2] *University of Houston, United States;*
[3] *Universitat Ramon Llull, Spain;*
[4] *Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, the Netherlands.*
[4] *Johns Hopkins University, United States.*

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

How to Construct Nationally Representative Firm Level Data from the ORBIS Global Database

Şebnem Kalemli-Özcan
University of Maryland, NBER, and CEPR

Bent Sørensen
University of Houston and CEPR

Carolina Villegas-Sanchez
ESADE – Universitat Ramon Llull

Vadym Volosovych
Erasmus University Rotterdam,
Tinbergen Institute, and ERIM

Sevcan Yeşiltaş
Johns Hopkins University

September 2015

Abstract

Firm-level data on productivity, financial activity and firms' international linkages have become essential for research in the fields of macro, international finance and growth. This paper describes the development of a firm-level global panel dataset for public and private companies based on the administrative micro-dataset ORBIS, provided commercially by Bureau van Dijk Electronic Publishing (BvD). The ORBIS database provides data on firms' financial and productive activities from balance sheets and income statements together with detailed information on firms' domestic and international ownership structure for over 130 million companies across the world. Researchers need to overcome several challenges before making the database usable for research. First, the database is not designed for large downloads that is essential for an econometric analysis. Second, there are several inherent biases in the database that affect the download process and lead to missing information. Third, the raw data may contain a number of irregularities which, if not dealt with, will result in data loss during a standard cleaning procedure. In combination, these issues cause minimal coverage of small firms, extensive missing data, and poor national representation. We give detailed instructions on the data gathering process from ORBIS in terms of downloading methodology and cleaning procedure so that a researcher can construct a database that is nationally representative with minimal missing information. We provide examples from several European countries to present the process and discuss the resulting dataset in detail.

Contents

Part I

Introduction

Firm-level data is becoming increasingly important in international macro/finance, international development and productivity research. Policy prescriptions in these fields are based on the predictions of macro models and many such macro models rely on firm-level heterogeneity for their predictions. Understanding the effects of globalisation on domestic production, investigating the role of access to finance for productivity, comparing multinationals' productivity to that of domestic firms and measuring knowledge spillovers to the domestic economy, analyzing the determinants of firm-level productivity and connecting these determinants to aggregate productivity, investigating innovation and entrepreneurship, determining which businesses succeed and become competitive and what role financial frictions play in such competitiveness, linking firms' financing decisions to their own investment and employment decisions and then in turn to aggregate economy outcomes—these are issues that can only be satisfactorily studied with firm-level data. Investigating the implications of real-financial linkages at the micro level is more imperative than ever in order to be able to formulate the right policy responses in the aftermath of important events such as financial crises.

There are several administrative firm-level databases of *listed* firms from many countries that can be used to answer some of these questions. Because listed firms have to file with official agencies the information on them is more readily available. The most commonly used data source is Worldscope database by Thomson Financial that contains balance sheet information on large listed companies. Compustat Global by S&P is similar to Worldscope. Compustat North America contains more detailed and specified information than Compustat Global but again it only covers listed firms in the U.S. and Canada. All these databases are widely used in the macro, finance and development literatures.

The drawback is that all these databases cover large firms and, unlike the United States, the bulk of the real economic activity—measured as output and employment—in most countries may not be fully accounted for by large listed firms.[1] In order to link firm-level outcomes to aggregate outcomes, we need information on both small and large companies that ultimately make up the aggregate economic activity. In terms of firm-level heterogeneity, it is also important to use both small and large firms because many macroeconomic models rely

[1]To make this point clear, we provide examples from official data based on national censuses from several countries in part 3, chapter 6.

on rich firm heterogeneity in terms of firm size and/or sector.

The practice in the literature so far is to use data from national censuses in order to cover a large number of small companies in the economy. Although national censuses cover all firms and, therefore, are fully representative, they have two main disadvantages. First, they only focus on the real side of firm activity and completely lack the financial side reflected in variables such as debt, equity, assets, or bank loans. These variables are reported in company balance sheets and income statements and will generally not be available in census data. This will prevent researchers from linking firms' financing decisions to their output and productivity. Second, because censuses are surveys which are not conducted every year, researcher who need annual data will need to impute the data for the missing years. Accounting data has the advantage of making it possible to connect the real and financial sides of firm decision making over time, for both large and small firms and for a large set of countries.

The ORBIS database (compiled by the Bureau van Dijk Electronic Publishing, BvD) is a commercial dataset, which contains administrative data on 130 million firms worldwide. ORBIS is an umbrella product that provides firm-level data covering around 100+ countries, developed and emerging, since 2005. Certain subsets of the database, marketed separately, cover individual countries or regions (such as AMADEUS for Europe or ORIANA for Asia-Pacific region, or the national products, such as FAME for UK and Ireland or REACH for the Netherlands) or company types (for example, the global database ISIS on insurance companies or the global database on banks BANKSCOPE). The financial and balance-sheet information in ORBIS comes from business registers collected by the local Chambers of Commerce to fulfill legal and administrative requirements and are relayed to BvD via over 40 different information providers (see Table A.1 for a list of the information providers to BvD). BvD organizes the public data from administrative sources and arranges them in a standard "global" format to facilitate company comparisons. ORBIS formats have been derived from the most common formats used for the presentation of business accounts in Europe. These harmonized balance sheets and income statements cover firms in all sectors of the economy (manufacturing, retail, financial sector, etc.).

Coverage of small firms and balance sheet variables change from country to country

given the filing requirements by business registers in each country (see Table A.1 for these requirements). Although most countries require limited liability companies to register once they are formed, requirements in terms of who reports (above certain firm size) and what to report from the balance sheet items vary across countries (only sales and employment are required, for example, for most companies in the United States). In most European countries, it is a regulatory requirement to file most of the balance sheet variables for firms of all sizes, therefore, firm coverage is superior.

Even for European countries where filing is mandatory for small firms and firm coverage is superior in business registers, the ORBIS/AMADEUS database will fail to deliver the full information to the researcher given the built-in features and restrictions of the BvD softwares/disks and its not user-friendly interface. Many researchers have experienced that, once they have access to ORBIS (or AMADEUS) via the web browser interface or via purchases of disks and download the data, they see a large number of unique firm identifiers but many financial or real variables are missing, especially going back in time. There are several reasons for this. First, there is a reporting lag in BvD products of roughly two years, meaning that a firm's filing in 2010 will appear fully in the media issued/accessed in 2012. Second, depending on the BvD product, certain companies are erased from the database if there is no reporting done for some time, even if the firm continues operating (but not reporting). Third, there is a download cap imposed by BvD on web interfaces, and most of the time this cap translates into missing data rather than termination of the download job. Fourth, BvD collection efforts improve over time, such that the firm information from early years that is not available in the older BvD disks appears in the disks issued subsequently. In addition to these considerations there are certain issues with the quality and the harmonization of the data by BvD and hence, a certain cleaning and checking procedure has to be implemented.

In this paper, we describe a detailed approach to overcome these problems and to make the ORBIS data as representative as possible for a given country. We show examples for select European countries, where the data we constructed from ORBIS-AMADEUS covers 75-80 percent of the economic activity reported in Eurostat. Our data also matches the official size distribution of firms provided by Eurostat.[2]

[2]Eurostat provides these statistics based on national censuses but does not provide the underlying firm-level data itself.

We proceed as follows. Section 2 describes preliminaries. Section 3 explains the methodology used to put together the financial data of firms in panel form and presents comparisons to official Eurostat data in terms of firm coverage and size distribution. Section 4 explains the methodology used to put together the ownership data of firms in panel form and presents comparisons to official data on international firm linkages from OECD. Section 5 concludes with detailed appendices.

Part II

Preliminaries

Chapter 1

Accessing BvD Products

BvD's best known two products for firm-level data are the global database ORBIS and the European database AMADEUS. There are other products that are either country specific or region specific as mentioned in the introduction (for example for the UK and for the Asia-Pacific region). All these products cover listed and unlisted firms. Most of the information in this paper relates to ORBIS and AMADEUS because we focus our examples on Europe. In fact, AMADEUS was the original flagship product of BvD with many features incorporated later on into ORBIS database. In what follows we will write about ORBIS and stress the unique features of AMADEUS separately.

A researcher can access ORBIS and European AMADEUS databases in three ways.

1. BvD proprietary browser online (orbis.bvdinfo.com and amadeus.bvdinfo.com).
2. BvD historic (CD/DVD-ROM, Blu-Ray) disks.
3. Through the Wharton Research Data Services (WRDS) from the Wharton School at the University of Pennsylvania (AMADEUS only).

Each of the access methods has its benefits and costs for a researcher; the decision about which one is the most optimal depends on the research budget and the type of the data one expects to work with.

We refer to the information available from BvD disks as "vintages" of the BvD data. We

use the same term when we talk about the data retrieved from an online access to BvD or WRDS at a given point of time. In each case, "vintage" will mean the release date of the disk or the time of online access, respectively.

Each product is split by the type of information provided. For example, ORBIS contains the "sections" ORBIS Financials with firm financial information and ORBIS Ownership with ownership information. There are other sections in ORBIS with valuable information such as Auditors and Advisors, Board Members, Patents, etc.

The users who obtain AMADEUS through WRDS (Method 3) should be aware that the internal organization of the whole database (AMADEUS Financials and AMADEUS Ownership) consists of three non-overlapping files corresponding to three company size "tiers." The tiers are defined by BvD in terms of the size of company sales, operating revenue, and employment. The thresholds of these variables for the companies to be classified by BvD as Very Large & Large, Medium, or Small vary by country (e.g., companies in Eastern Europe may have lower sales but still are qualified as Very Large & Large). This matters when querying and downloading the data with the SAS software available under the WRDS subscription.

In the following pages we describe what we regard as the "best practice" for obtaining the most comprehensive financial data and ownership data over time.

1.1 Financial Module

ORBIS Financials sub-database includes detailed information about numerous balance sheet items, profit and loss account items and financial ratios over time and also static descriptive variables. The descriptive information includes, among other items, official national identification number, address (country, region, city, street), legal form, year of incorporation (entry), status of the company (active/liquidation/merger-acquisition), number of employees, quoted/unquoted indicator, industry and activity codes (4 digit level) and, when available, the description of the nature of business in the local language and English.[1]

[1]The default option is to download textual information, such as company names, in the original format. We encountered problems with the encoding of non-Latin alphabets and non-standard national letters. The user should choose the available option to show the textual information in international alphabet instead of

The access to historic (time-series) financial information is available by downloading the data from a one-time web access or from a single disk by selecting several historical years though there are several issues with such a methodology:

- The *download speed and cap issue*. BvD platform or disks are not designed for the purposes of academic research involving working with large volumes of data. Extracting large amounts of data from any BvD platform is in general slow. In addition, BvD puts a cap on the amount of information researchers can download both from disks and also from its own web-site. Most of the time this cap does not turn into a termination of the download job but rather the downloaded files will have missing information. WRDS is the most user-friendly platform because the imbedded WRDS browser allows to run optimized queries and compresses the data at the time of download; WRDS also allows the researcher to retrieve the data by running a SAS for UNIX code directly at their servers. However, WRDS covers only AMADEUS, and ORBIS and AMADEUS do not overlap 100 percent in terms of companies and variables even for a given European country (any country outside Europe will not be in AMADEUS).[2]

- The *survivorship bias*. Both ORBIS and AMADEUS contain a number of years of historic financial data. However, the two databases follow somewhat different rules regarding the inclusion of companies and years. AMADEUS provides at most 10 recent reporting years of data for the same company while ORBIS de facto reports data for only up to 5 recent reporting years. AMADEUS will also delete the company from the database if the company did not report anything in the last 5 years, ORBIS will keep this company as long as the company is active in the business register.

- There is a *reporting lag* of about 2 years, on average, and there are differences in the coverage of particular variables depending on when the BvD product has been released. Hence, in the 2010 vintage, a company may not show up for its' 2010 filings but the 2010 filings of this same company will appear in the 2012 vintage. BvD data collection improves over time and hence this lag might vary by country and by time.

original alphabet before download.

[2]For the same Italian or German company for a given year, it is possible to have more non-missing variables in ORBIS vs AMADEUS. It is also possible to obtain more firms for the same country and year in ORBIS than AMADEUS. We have confirmed that both cases are widespread.

- Issues with the *presentation format*. Certain variables, such as employment, will not be on the balance sheet but rather in memorandum items.

- *Merging issues*. BvD identifies each company by a unique company ID which is designed to trace the same company across all their products. However, a researcher merging the time-series financial information coming from several BvD historic disks or the online downloads done at various points of time may encounter occasional *BvD ID changes* over time. The BvD ID number incorporates either the national ID number or the ID provided by their information providers (IP). According to BvD, the ID numbers may change when the national ID numbers change in the official data sources or the BvD IPs decide to switch their ID numbers. The ID changes are related to changes of address, legal form, or M&A activity. In acquisitions, acquiring company will keep its ID and the target's ID is blocked. BvD mentions that Spanish companies encounter a BvD ID change if they change legal form while companies incorporated in Germany, Austria or Italy may in some cases see their BvD ID change if the company changes address.[3] Finally, BvD itself can initiate the ID change when an entity is available on more than one product, or is provided by more than one IP and BvD harmonizes the IDs across databases using a set of priority rules. As long as BvD does not know that a certain company is the same entity, it will have several different BvD ID numbers on ORBIS . Because it is hard to keep track of all these idiosyncracies, the researcher should request the "correspondence table" of BvD IDs from their BvD representative. BvD ID changes can also be obtained by subscribing institutions via the dedicated BvD ID Change Lookup tool at idchanges.bvdinfo.com.

1.2 Ownership Module

ORBIS Ownership sub-database contains information on each company's equity ownership structure: the names of owners, their respective ownership shares, the level of ownership (direct or ultimate cross-ownership), their countries of origin. For each owner of every target

[3]If a company moves from area 1 to area 2 and area 2 has a different office collecting the information, the company will get a new national ID, thus the BvD ID number changes. If in area 2 the same office is responsible as in area 1, the company keeps its national ID.

firm there is one observation (we refer to such record as "ownership link"). There are two major issues of concern for the construction of time-series ownership information.

- The *vintage issue* for ownership. BvD proprietary browser online and WRDS contain only the *latest* available ownership information. If one were to access ownership information through the browser, through a specific vintage disk of AMADEUS or ORBIS or through WRDS, ownership information will be static ("as of date"). The only option to reconstruct the historic (time-series) ownership information is by purchasing the historic ORBIS disks from BvD.

- *Merging issues.* The same issue related to BvD ID changes emerges if one downloads ownership data from several vintages of ORBIS. The issue is more acute because one needs to rely on annual vintages to track annual ownership changes.

  There is a separate BvD product that tracks all mergers and acquisitions, that is *changes* in ownership, at the transaction level over time (ZEPHYR). In principle, one can supplement the ownership stakes from ORBIS Ownership with the transaction data from ZEPHYR by adjusting the equity stakes reported in ORBIS Ownership prior to transaction. The data we present in this paper incorporates such an adjustment.

Next, we propose a detailed approach in order to construct the most comprehensive financial and ownership information for a representative set of firms over time.

Chapter 2

Downloading from ORBIS and AMADEUS Databases

2.1 Download Methodology

While AMADEUS and ORBIS has an impressive number of unique firm IDs, as many researchers have discovered, a large number of those IDs contain only information on company name and a few other variables. Once one requests some key variables, such as total assets, sales or employment, these turn out to be missing. As we detailed out above there are several reasons for this. It is necessary to download data in a non-standard way to overcome these problems. As we have explained in the previous chapter, there are three different ways to access BvD data:

1. Through BvD's proprietary web platform available by the direct subscription.
2. Through BvD's historical vintages, available on historic CD-ROM disks (or harddrives/blue-ray disks).
3. Through WRDS archives.

The standard and the most commonly used methodology is Method 3. To avoid the pitfalls mentioned above one must follow Method 2, however. Let us explain the advantages

of Method 2 over other methods and how one can maximize the coverage and representation of small firms and also recover many financial variables by using Method 2.

To maximize coverage (for European countries), a researcher must use both ORBIS and AMADEUS and several vintages from both databases.[1] The reason is that these databases follow different rules regarding the inclusion of companies and years.[2] AMADEUS provides at most 10 recent years of data for the same company while ORBIS, de facto, only reports data for up to 5 recent years, despite the possibility of choosing 10 years of data going back in the ORBIS software. The justification given to us by BvD was that the global ORBIS database contained much more information and the information included in a given vintage had to be limited because of the media capacity. Moreover, AMADEUS drops firms from the database if they did not report anything during the last 5 years while ORBIS keeps the information for these companies as long as companies are still in the business register. This problem makes it clear why Method 2 is superior to other methods. A company might file information with BvD for the last time in year 2007. However, in the business registry, this company is still active. Due to non-reporting in the last 5 years, in AMADEUS-2013 vintage this company will not be included, but the same company's information for the period 2002–2007 will still be reported in ORBIS-2013 disk. In addition, because there is a *reporting lag* of financial data of usually 2 years (it varies by country), the coverage of, for example, years 2007 and 2008 from the 2009 ORBIS disk (or an online download done in the year 2009) will be very poor. For this purpose, again Method 2 will be better since recent database vintages will complement earlier ones and hence, one can get more firms for the years 2007 and 2008 from the 2010 vintage (or 2010 WRDS download). It is also the case that information is updated over time and some variables that were not available in early disks is made available in later vintages.

There are differences in coverage in certain variables across ORBIS and AMADEUS going back in time. The reason for this might be a combination of the issues mentioned above or the fact that all the access methods will *cap the amount of information* one can download in

[1] For countries outside Europe, the only option is ORBIS. Our instructions below related to ORBIS will apply to other countries, such as the United States, in terms of maximizing coverage and representation.

[2] One also needs to keep in mind that AMADEUS, being a regional database, includes some Europe-specific variables that the ORBIS Global Standard Format may not contain. Plus some variables may be coded slightly differently in the two databases; for example, the type of owner is textual in AMADEUS while in ORBIS this variable contains standardized single-letter codes.

one run (the number of firms and the number of variables). This cap unfortunately translates into missing observations in the resulting download instead of termination of the download job. This issue can also be dealt with if we use Method 2. To illustrate the problem consider a researcher who in April 2015 wants to obtain data on Spanish firms for the year 2006. The researcher goes to WRDS (Method 3) and downloads the data by choosing the year 2006. Table 2.1 presents the distribution by size category in 2006 of firms in Spain in our data constructed based on the methodology we prescribe in this paper (RAW) and the same size distribution based on the data downloaded from WRDS on April 2015 for the year 2006 for Spanish companies. Panel A uses employment to measure firm size distribution, whereas Panel B uses wage bill. In Panel A, in column (1) we see that firms with less than 20 employees account for 25 percent of total employment, whereas firms who have between 20–249 employees account for almost 50 percent and firms who have more than 250 employees account for 26 percent of employment. Panel B delivers a similar size distribution. The point we want to make is that if we use our data as shown in column (1), put together from several vintages using Method 2 for download or we use direct download from WRDS for the year 2006 where we access WRDS on April 2015 as shown in column (2), we obtain a similar size distribution.

What is the problem then if our data and a single shot download from WRDS delivers the same data? The problem starts when the researcher wants to have full information on a certain set of variables. For example, if we want to calculate total factor productivity, then we need to have the variables output, employment, capital stock and materials reported and hence, we need to go down to a subset of firms that report all these variables. When we do that column (3), which is our data, performs very well and delivers a similar size distribution. However, column (4) which is the direct download from WRDS as of April 2015 performs clearly worse since practically, there are no small firms reporting information on materials. This is the artifact of the one-time download from WRDS. As we show in chapter 6 of part 3, the size distribution based on our data for several countries (including the one shown in Table 2.1) matches the official size distribution provided by Eurostat based on national censuses.

Table 2.1: Company coverage comparison in Spanish manufacturing sector in 2006 obtained from our data (RAW) vs AMADEUS online from WRDS

Sample	RAW	WRDS	RAW-TFP	WRDS-TFP
Panel A: Employment				
1-19	24.7%	24.0%	24.2%	0.6%
20-249	49.2%	49.1%	49.7%	50.1%
250+	26.1%	26.9%	26.1%	49.3%
Panel B: Wage Bill				
1-19	19.6%	18.0%	19.2%	0.6%
20-249	46.6%	44.7%	47.0%	44.4%
250+	33.8%	33.5%	33.8%	53.4%

NOTES: RAW refers to the sample of firms after basic cleaning in our data. WRDS refers to the sample of firms as obtained from the WRDS after basic cleaning. TFP refers to the sample of firms after basic cleaning and with the required information to compute total factor productivity (TFP), i.e., non-missing values for employment, output, capital stock, and materials. WRDS-TFP refers to the sample of firms in WRDS after basic cleaning and with available information to compute TFP.)

In order to maximize the coverage of firms and variables by country over time our download strategy (Method 2) **for financials** makes use of several vintages of BvD products: ORBIS disk 2005, ORBIS disk 2009, ORBIS disk 2013, AMADEUS online 2010 (from WRDS; accessed in May) and AMADEUS disk 2014. We chose the vintages to ensure a time overlap to get around the reporting rules in AMADEUS and ORBIS.

For ownership our download strategy is more involved because, as we explained, any of the three access methods—browser access (Method 1 or 3) or the current vintage in hand (Method 2)—would give access to the ownership information "as of date". We use the Method 2 download strategy but construct the historic (time-series) ownership information by purchasing the historic ORBIS disks from BvD *for every year*. Because we prefer to record the ownership information as of the end of each calendar year, we choose the ORBIS disk issued as close as possible to the end of the desired year. For example, to obtain the ownership as of end of 2010 we use the ORBIS disk issued in January 2011, and so on (disks are issued monthly but firms report yearly, though ownership can change within a year). To take full advantage of slight differences in ownership data in AMADEUS and ORBIS we combine bi-annual vintages of AMADEUS Ownership with annual vintages of ORBIS Ownership.

These differences will be discussed in part IV.

2.2 Financial Reporting in BvD Databases

2.2.1 Time Stamp

Before downloading, one has to choose how time stamp the year of financial data. One option is the conventional "absolute years," where the year of the financial accounts explicitly refers to calendar years, 2006, 2007, and so on. An alternative is the so called "relative years," where the most recent year of non-missing financials (as available to the BvD) is referenced as the "Latest Year" and the earlier observations are referenced as the Latest Year –1, Latest Year –2, and so on.

We advise to downland financials via the "relative years" option. While a priori it seems counter intuitive, the relative year option is superior to the absolute year option because of the reporting lag and the survivorship bias issues detailed above. The distinction in the dating option is extremely important for the companies which report irregularly, with gaps in their time series. For such companies, 5 relative years might cover a longer calendar period than 5 years requested explicitly (recall that ORBIS de facto reports only the most recent 5 years of a given company, with a reporting lag of 1–2 years). For example, consider a company reporting data for year 2007 and then for each year between 2009–2012. Then, asking for the 5 latest absolute years 2008–2012 would leave us with just 4 observations (2009–2012) while asking for 5 relative years would retain the 2007 value too (the 2007 is referenced as "Latest Year –4" in this case). To obtain the correct assignment of the relative years in terms of calendar years we use the special BvD index variable (Closing date, Latest Year; Closing date, Year –1 and so on).

2.2.2 Consolidation of Subsidiaries

For a given company, the BvD databases report one or more financial statements, whose type is specified by the "Consolidation Code." Most of the large companies with good coverage report either consolidated accounts (the statement of a parent company integrating

the statements of its controlled subsidiaries) or unconsolidated accounts (the statement not integrating the statements of the controlled entities). Some companies report both kinds of accounts. Finally, there are entities with limited financial data, no recent financials (where the last available accounts are more than 48 months old), and the combination of the two. In most of these cases, only the number of employees and the operating revenue are available. The type of account reported is related to country filing requirements for particular size or the legal type of companies, as detailed in Table A.1 (e.g., the non-independent branches (establishments) are often included in ORBIS for the U.S.). To speed-up the process we downloaded the data separately for non-limited financial accounts (which includes most companies) and limited financials accounts. We download both consolidated and unconsolidated accounts and so far, use unconsolidated accounts in all of our applications. Any user can choose which account to use by looking at "U" or "C" letters at the end of the firm ID (or by using the BvD variable consolidation code), for unconsolidated and consolidated statements respectively. Consolidated accounts will involve double counting when both consolidated account of the parent (with all its subsidiaries) and the unconsolidated account of the parent (without subsidiaries) are reported.

2.2.3 Units and Currency of Financials

Balance sheet financial variables are at book value. Care should be exercised when choosing how the financial data will be downloaded both in terms of the monetary units and the currency of financials. By default, the formatted export from BvD disks will result in the units and currency in which a particular company originally filed its financials. This means that a given company may report in thousands in some years, and then millions of the same currency in other years. To avoid the spurious jumps in the data the best practice is to choose the units (units, thousands, millions) in Formatted Export Wizard explicitly, rather that using the default. In case the default was chosen, the variable `UNITS` contains the reference to what units a given observation corresponds to (the values are textual in ORBIS and numeric powers of 10 in AMADEUS). In the latter case, the harmonization is achieved by the transformation $x = x * 10^{\text{UNITS}}$. But note that the `UNITS` variable might have errors in certain disks, as we have discovered, therefore the best practice is to force the

download to be in specific units (such as thousands) and not use the default.[3]

We choose the data in the original currency filed by the company, which might be the national currency of the country or sometimes a foreign currency. The currency of the given account is available in the variable "Account published in currency." It is tempting to use the available BvD variable "Exchange Rate" in order to convert the data from all companies into a single currency. We advise against using this variable because different products use it for different purposes. In the ORBIS or AMADEUS disks, once we choose to download the data in "local currency" the values of the variable "Exchange Rate" are always set to 1. However, in the AMADEUS downloads from WRDS this variable contains the actual exchange rate of the currency of account against the US dollar. Hence, in general, the variable "Exchange Rate" does not contain the rate to, say, US dollar (as was the case with WRDS AMADEUS) but has the exchange rate of the currency of an account to the currency chosen by the person downloading the data.

2.3 Data Selection and Exporting

Before discussing the merging and cleaning steps, we touch upon some other subtle issues in the setup of the download process. The process starts from selecting the sample of companies in a given BvD disk, typically by country.[4] To overcome the download limits, large countries may be downloaded by regions. An alternative is to use sectoral splits in a given country. The software then displays the set of the selection criteria, combined by the "AND" expression, and the resulting number of unique companies satisfying all the criteria. One can replace the "AND" expression to adjust the selection criteria and/or display the list of companies. Regardless of the variable used for splitting the large country data for separate downloads, it is necessary to watch out for the cases when the variable used for splitting is not available for a subset of companies. For that, one needs to select the criterion of the entire country

[3]By errors we mean the cases when the value of the UNITS switches from, say, thousands to millions but the corresponding financial variables do not show the 1000x decrease in the order of magnitude. We describe a filter we developed to check for these issues in Section 4.

[4]One must chose industrial companies from the start since there are also banks and insurance companies in ORBIS. "Type" variable helps this selection. In addition global format is what is available for private firms whereas detailed format is available for listed firms.

(such as "all companies in the U.S.") and add the criterion where *all* regions or sectors are explicitly listed but precede the latter condition with the "NOT" expression (AND NOT "companies in sectors A, B, C...").[5]

To download from the *older disks* (issued prior to January 2012 for ORBIS or December 2010 for AMADEUS) one has to navigate to the Formatted Export menu (`File`>`Export`>`Formatted Export To` on the upper-left) in order to adjust the default list of variables; the user does not need to display the list of companies after this. In the Formatted Export menu one can also choose the time period of the data (absolute/relative years), currency, units of financials, and then the export layout. See Section C.1 in Appendix C for illustration of this interface. An alternative, "List Export" method is accessible from the list of companies. To adjust the default list of variables and make other choices, one has to navigate to `Options`>`List Format`>`New Format` menu. We do not discuss this method because we did not see a clear advantage over the Formatted Export method in older BvD disks. If anything, a potential disadvantage is the increase in the waiting time for the list of companies to be displayed before the download can start.

In *newer disks* (issued after January 2012 for ORBIS or December 2010 for AMADEUS), which emulate the BvD online interface (Method 1), more download options are available after the resulting number of unique companies satisfying the criteria is revealed. After selecting the Formatted Export menu on the right, one now needs to further choose from the Database export, Statistical export (not useful), or Custom export. The closest analog to the Formatted Export interface in the older disks is the Custom Export where one can adjust the variables, data time stamp (absolute/relative years), units, currency, and layout. Section C.2 in Appendix C shows the details of its interface. However, for unknown reasons the list of variables one can choose from does not include all the variables available in ORBIS.[6] In particular, none of the ownership data is available. A potential remedy is the Database Export (another sub-heading of Formatted Export), specifically designed for downloading large amounts of data according to BvD. For the interface of this method see Section C.3

[5]This additional download can be ignored if the number of companies lost due to the missing splitting variable is negligible.

[6]Since all the variables were available to choose in older disks under Formatted Export, we believe the reason they are not under Custom Export (which is a sub-heading of Formatted Export) is due to space limitations in the newer disks with the increase in the amount of information given more years.

in Appendix C. Database Export lets one chose all the variables in ORBIS, currency, and units. The download works much faster than other methods, which is an advantage for large downloads such as ownership data. In our experiments, we saw roughly a twofold time gain compared to other methods. However, a serious drawback of the Database Export is the inability to download financials with the option of "relative" years (it only downloads in terms of "absolute years"). For this reason, we have to choose the Custom Export option as the only way to force the relative years download. For ownership downloads the Database Export is acceptable because, as we explained in Section 2.1, the time dimension is irrelevant given the availability of ownership data as a snapshot in a given release. A final option is the List Export, illustrated in Section C.4 in Appendix C. One can chose absolute and/or relative years and all of the variables under the List Export option. Hence, this method works for both financial and ownership data downloads. The only drawback of List Export is extra time needed to first reveal the list of the companies before the download can commence.

Part III

Financial Panel Data

Chapter 3

Vintage Raw Data

The following are the key steps we take in constructing the firm-level financial database. We execute these steps for each vintage of the BvD database we use.

1. Extract the data from ORBIS with 5 latest "relative years" in ASCII comma-separated value format,[1] transform it to Stata using Statransfer, name and label the variables using the database codes and names. The data comes in Stata wide format with the rows consisting of *unique* records (lines) per company consisting of BvD ID and similar non-time varying data, the last available year per firm `LASTYEAR`; and the blocks of time-varying variables in the form X in the Latest Year, X in Latest Year –1, and so on, until X in the Latest Year –4. We rename the variables using internal BvD codes to become, correspondingly, $X1, ..., X5$.

2. Delete the observations with just a name of company and no other information and the observations with missing BvD ID or BvD Account number (the main account identifier).

3. Notice that at this stage the data does not have a harmonized time variable showing the precise calendar year of each data point. In order to reshape the data from Stata wide format to the convenient long format we use the row number as the cross-section identifier and specify in Stata that the number following the stub X of the time-varying

[1] For AMADEUS data we select 10 relative years for the reasons explained in Section 2.1.

variables $X1, ..., X5$ has to be treated as the "fake" time identifier `YEAR`. The resulting `YEAR` variable is a mere marker of the observation per firm because it does not account for the possible gaps in reporting that we have discussed in Section 2.2.1. There are also a lot of redundant empty observations because the `-reshape-` command creates a full panel and the observations in the periods when a given firm does not report the data will be empty.

4. Use the variable "Account Closing Date" `CLOSEDATE` to correctly assign the calendar year. We re-construct the `YEAR` variable based on the following convention. If the closing date is after or on June 1st, the current year is assigned (if `CLOSEDATE` is 4th of August, 2003, the year is 2003). Otherwise, the previous year is assigned (if `CLOSEDATE` is 25th of May, 2003, the year is 2002).[2,3]

5. Create our main company identifier `ID_NUMBER`, which is a copy of the BvD ID number. The difference between the BvD account number and the BvD ID number is the single letter, U or C, in the end of the BvD account number, reflecting what account type the record represents (consolidated and unconsolidated). If the BvD account number is missing, we concatenate the BvD ID number with the first letter of non-missing BvD Consolidation Code, following the BvD convention for the BvD account numbers (U, C and nothing for limited financials accounts). We create the country code based on the first two letters of the `ID_NUMBER` which by BvD convention starts by a two-letter country code (BE for Belgium, US for the U.S., GB for the UK and so on).

6. All the financial variables are transformed from strings into numeric data type. In

[2]The raw data may contain duplicates in terms of ID-YEAR. These duplicates arise for two reasons: i) The presence of both quarterly and annual reports. ii) Firms switching from presenting their end of accounting year balance sheet information in one month to some other month (from December to May, for example). We would like to keep the most recent reports for these companies. One way to handle duplicates is to retain the data for the closing date closest conceptually to the end of year. An additional step to handle duplicates is to use a flow variable with good coverage such as Operating Revenue to identify potential quarterly reports (quarterly sales should be much lower than yearly sales). Then we drop duplicates whose revenue are less than the maximum per firm-year. For example, in 2005 vintage, there are around 34 thousand duplicates like that out of over 18 million observations. We eliminate a small number of remaining duplicates for which we cannot know whether these remaining reports refer to annual or monthly data.

[3]One might think that BvD correctly marks the lags in relative years taking into account the possible time gaps in financial data. As we discovered, this is not the case. Hence it is essential that one downloads the CLOSEDATE variable, together with BvD ID number, BvD account number, and the last available year variable.

the vintages where we downloaded the data in the "original units" we transform the variable UNITS from textual form to the integer power of ten to express all financials in the same units. We developed a filter to check for mistakes in the UNITS variables as explained below.

7. Clean the raw vintage data.

   - Drop duplicates in terms of `ID_NUMBER` and `YEAR`.[4]
   - Drop observations with no financial information. We verified that these company-year points do not have an account closing date and are mostly missing observations generated during the reshaping stage.
   - Drop observations for which country code created based on the BvD ID numbers does not correspond to BvD's country ISO code.[5]
   - Drop observations with missing currency.
   - Drop observations with missing Account Closing Date.

[4]These will have the same financial data but different industry codes. In ORBIS the first observation per duplicate represents the main industry, hence, we keep the main industry per company for companies that had multiple industries reported.

[5]As we mentioned, BvD IDs start from the 2-letter code corresponding to the company's country. BvD country ISO code is the same for all the companies in a given country. Hence, if a company has a different country code, we interpret this as a mistake. This can also be a tax front. Note that these are not the multinationals since the country code attached to BvD ID numbers reflect where the company operates. If the company is part of a multinational, this will be reflected in the ownership variables not in the ID numbers.

Chapter 4

Merging Vintages

At this stage each observation in each vintage is uniquely identified by the variables `ID_NUMBER` and `YEAR`. The following steps are performed at each consecutive merge, with vintage specifics highlighted.

1. As explained in Section 1.1, in some instances the unique company identifiers change over time which would affect the success of the merge. Before the merge, we replace the old IDs for the countries which had their ID changed over time with the most recent ones. We retain the old "legacy IDs" on the data.

   A slight complication emerges for IDs in former Yugoslavia, coming from earlier ORBIS disks. The same companies "belong" to Yugoslavia in the early 2000s (country code and the first two letters of IDs are YU), then to Serbia-Montenegro (CS) in disks from the mid-2000s, then to, separately, Serbia (RS) and Montenegro (ME), or even Kosovo (KV). In the IDs of these companies only the letter part changes over years – the numeric part does not. Where it is possible, we assign YU and CS to RS and ME using the numeric part of IDs; where we do not know, we assume they are in RS. We do this after the replacement of IDs using the correspondence table because some firms with ID from CS are present in that table and are already assigned to RS or ME.

2. Before the merge, we harmonize the names of the variables across vintages and products.

3. Before the merge, we check the consistency of the Units of Financials variable (`UNITS`). Recall that at this stage the financial data is reported in various units of local currency (units, thousands, etc.) for different companies and even for the same company over time. This is due to BvD non-harmonization across products. Downloading in "default" option for the units requires the researchers to use the UNITS variable to harmonize the units. Another way is to force the download to be in a particular unit such as units or thousands. We employ the following strategy to deal with this issue if the researcher downloads the data with the "default" option. For each company we check if the moment of switch in units coincides with a "reasonable" move of total assets (can also do other financial variables); if not—we drop the entire firm. For the lower threshold of assets growth we choose -99% because the 1000x decrease (due to a change in the UNITS variable) of otherwise unchanged assets is -99.9% growth. For the upper bound of assets growth we choose 19,800% because the 1000x increase of otherwise unchanged assets is 99900% growth. If we allow the company to have a large (70 percent) drop in assets in the year when the units switch 1000X (X*0.3*1000), this is close to 19900% growth.[1]

   Then we calculate the actual growth of assets and verify if this growth lies outside of these thresholds *in the years when UNITS change*. We drop the entire company where the moment of switch in UNITS does not coincide with the reasonable move in assets.[2] One can keep the companies which are marked by BvD as "inactive" because the assets of those firms can genuinely go down to (almost) zero.[3]

4. Before the merge, we express all financial variables in harmonized units by the transformation $X = X \cdot 10^{\text{UNITS}}$, except in the 2013 ORBIS vintage that is downloaded with the explicit "in units of local currency" option. The UNITS variable is then discarded.

5. Before the merge, we create the textual identifier for the vintage to keep track of the vintage which contributes to a given observation after all merging steps are done.

[1] Recall that all the balance sheet values are book values for the non-listed firms.

[2] As a result we eliminate about 3% of observations in 2005 and 2009 ORBIS vintage; less than 1% in 2010 WRDS AMADEUS vintage; and less than 0.5% in 2014 AMADEUS vintage.

[3] The variable Status takes the values Dissolved, Dissolved (merger or take-over), In liquidation, Inactive (no precision), Dissolved (liquidation), Dissolved (merger or take-over), Inactive (no precision), Bankruptcy, Dissolved (bankruptcy), Dissolved (demerger) or some peculiar active types (Active (default of payments), Active (dormant), Unknown.

6. An additional methodological complication arises when we combine the 2013 ORBIS vintage with earlier vintages. This is related to the change in sectoral classification in 2008 from NACE Rev 1.1 to NACE Rev 2 by the Eurostat.[4] In 2005 and 2009 ORBIS vintages, the sector is identified by the "NACE Rev. 1.1, Core code (4 digits)" (`NACECD`); in 2010 WRDS AMADEUS, 2013 ORBIS, and 2014 AMADEUS vintages it is already "Rev. 2 Core code (4 digits)" (`NACEREV2CCODE`). Both codes are unique per company.[5] We prefer to use the more recent NACE Rev. 2 classification in our data.

   We went through a detailed process of matching the industry classifications pre- and post-2008. We start from the official Eurostat correspondence table between NACE Rev 1.1 and NACE Rev 2. To supplement the official correspondence table in a way that we have a one-to-one match for every sector, we proceed as follows. Most often, multiple NACE Rev 2 codes correspond to a given NACE Rev 1.1 code. In the official correspondence tables, when multiple NACE Rev 2 codes are matched to a unique NACE Rev 1.1, they are sorted in the ascending order of the numeric NACE Rev 2 code. The first code is the *most closely related sector* to the one in NACE Rev 1.1 classification. We retain the first NACE Rev 2 code provided in the official table and discard the rest.[6] This approach attains a good match for manufacturing sectors (codes 10 and higher in NACE Rev 1.1) but not as successful for agricultural sectors (codes below 10 in NACE Rev 1.1). We manually match codes by reading the long descriptions of the codes. We do the same if there are sectors that are completely missing in the official correspondence tables. Our own correspondence table is available upon request.

   Before merging with the 2013 ORBIS vintage we merge the earlier vintage data with our sector correspondence table. We keep the original sector classification from each vintage just in case.

7. Care should be exercised when combining the disk data from ORBIS with AMADEUS down-

[4] Table A.2 lists sectors classified by NACE Rev 2, Level 2. For space considerations we do not report the 4-digit industry classification.

[5] There are numerous "secondary" codes in BvD. All of the above is for the "main" code. We also keep the alternative unique industry codes such as SIC and NAICS.

[6] For example, the NACE Rev 1.1. code 10.20: Mining and agglomeration of lignite is matched to three NACE Rev 2 codes: 05.20: Mining of lignite; 09.90: Support activities for other mining and quarrying; 19.20: Manufacture of refined petroleum products. We retain the first line from the correspondence table and matches "10.20: Mining and agglomeration of lignite" to "05.20: Mining of lignite."

loads from the WRDS (in our case, we used the 2010 download of AMADEUS from the WRDS), mostly with respect to the harmonization of variable names. Here are some peculiarities which the users are advised to verify. i) As we explain above, the "Exchange Rate" variable of WRDS should not be used. ii) In May 2010 WRDS AMADEUS the label of the core sector variable NACECD states NACE Revision 1.1. but we discovered from observing the values that it is in fact Revision 2. We verified this with WRDS who in turn confirmed this by contacting BvD. For this reason, no sector correspondence table needs to be applied to this vintage. iii) In WRDS, the listed company identifier is a binary variable ("Quoted company" `LISTED`) while in ORBIS it is textual having the values like "Delisted," "Listed," "Unlisted." We add the textual variable to WRDS AMADEUS vintage before merging with the other data.

8. We use the values from the later vintages to supplement missing values. A non-missing value, however, will never be replaced with a missing. (In Stata language, we merge with update and replace options). Depending on the order of merge and the computing power (RAM) availability the users may merge all the vintages at once or, if the data gets big to fit the RAM, merge countries one-by-one and then append (stack) the country data.

9. Check for duplicates by ID and `YEAR` and, in case of duplicates, retain only observations coming from the most recent vintage.

Chapter 5

Cleaning Merged Data

After we merge the financial data from individual vintages, we perform some further data harmonization and very mild cleaning for obvious data mistakes.

1. Companies in several countries report financials in multiple currencies. We always retain the accounts in major currencies, such as, US dollar, Euro, UK Pound, but delete the observations with missing or unreasonable currencies which probably are mistakes (for example South African Rand or Canadian dollar for European companies).

2. We express the financial variables in real dollars 2005 base. To convert from the units of the nominal currency of accounts we i) convert the currency of accounts to the official currency of the country; ii) deflate the series by the national GDP deflator with the 2005 base from the World Bank; and iii) divide by the exchange rate of the official currency to the U.S. dollar in the year 2005. A number of complications arise at this stage.

   Because companies in several countries report in multiple currencies in order to add the official exchange rate, we do not use the country code but the currency code.[1] The problem with the World Bank data (or the IMF's IFS data) is that the source does not

[1] As we explained in Section 2.2, the BvD variable "Exchange Rate" is useless for conversion of the data downloaded in the "original currency." Users may choose to download the raw data in some other currency, such as U.S. dollars, in which case the Exchange Rate might report the exchange rate to the currency of accounts. We prefer to convert the data ourselves and not rely on the internal BvD data convertor.

report the ISO currency code even if the country changes the currency; we just observe a jump in the exchange rate. We obtain the Compustat Global exchange rates, which uses currency ISO symbols (USD, GBP, etc.) as the main identifier of the existing and legacy currencies. We also supplement and harmonize the Compustat currency series rates with currencies actually observed in our data in particular country-years. In case the legacy currencies are missing in Compustat in some years (e.g., in the former Soviet Union, in Yugoslavia, etc.) we refer to the national central banks data.

For the conversion step i), we need to make sure that the currency of financials for all firm-years is the "official" local currency as of today and not other currencies. This is important because we will use the deflator in the official local currency. In particular, we need to decide what official currency to use with the recent Eurozone members in conjunction with the GDP deflator data. In our sample, Estonia, Slovenia, Slovakia, and Malta are such countries; Latvia who adopted the euro on 1 January 2014 is still marked as having currency as Latvian Lats in our data. Consider Slovenia as an example. The country adopted the euro on 1 January 2007. In Slovenian data we observe the companies which report the data in single currency (Euro or Slovenial Tolar) in all years, pre- and post-2007. The notes to the WB GDP deflator data for Slovenia says "A simple multiplier is used to convert the national currencies of EMU members to euros. The following irrevocable euro conversion rate entered into force on January 1, 2007: 1 euro = 239.64 Slovenian tolar." This implies that the deflator is effectively in tolar (SIT) until 2007 because it is a ratio of year 2006 to year 2005 (the base year in WB data) and both of these years are multiplied by the same number 239.64 (the fixed parity rate). Hence, the official currency of Slovenia is considered to be SIT before 2007, and all financial data is recalculated to SIT before 2007. From 2007 the deflator reflects the dynamics of local prices in Euro and we express all the financials of Slovenia in euro.

3. Drop company-years with missing information on total assets and operating revenue and sales and employment (simultaneously).

4. Drop the entire company (all years) if total assets is negative in any year.

5. Drop the entire company if employment (in persons) is negative in any year and com-

panies with employment larger than that of Walmart (2 million) in any year.

6. Drop the entire company if sales are negative in any year. Of note, we do not perform this filter in terms of Operating Revenue because this P&L account item is equal to sales + Other operating revenues + Stock variations. While sales indeed cannot be negative, revenue in principle can be negative if a company has a sizable financial loss (say, loss due to hedging, etc.). For countries, like Denmark, whose firms do not report sales but only operating revenue, we will effectively not implement this filter.

7. Drop the entire company if Tangible Fixed Assets (such as buildings, machinery, etc.) is negative in any year.

8. For a given company ID year, we replace missing strings which are unlikely to change over time with values for this company for other years. We complement information on country, company name, city, region, postal code, legal form, and date of incorporation with lagged/lead values in the years where such info is present. This is reasonable because if a company changes the legal form it obtains a new BvD ID and will be treated as a new entity. If information is missing in all years, they remain missing.

These steps complete the cleaning of financial data. Next, we explore how our data compares to official sources in terms of aggregates.

Chapter 6

Coverage of Financial Panel Data

We examine how well our data cover the universe of firms compared to official statistics from Eurostat along several dimensions (see Appendix B for details of official statistics). We first show how much of aggregate economic activity (as reported by Eurostat) is accounted for by the firms in our data. We also compare number of firms and we examine the coverage for the manufacturing sector in terms of fraction of activity accounted for and in terms of the firm size distribution.

6.1 Total Economy

Tables 6.1 and 6.2 show how much of the official gross output data from Eurostat is covered by the firms in our data for the total economy both for non-Eurozone and Eurozone countries. Each cell is the ratio of value of total output produced by our firms relative to value of total output produced as in the official data. BvD provides firm-level information on gross output (`OPRE`) for all sectors of a given European country between 1999-2012, however Eurostat SBS data provides information on gross output (Turnover) for a subset of sectors (see Table B.3). So, for a given country-year, ratios are computed by taking the ratio of aggregated gross output values where aggregated gross output is computed by totalling gross output over these sectors for which the gross-output related variable is available in both data sets (see step 2 in section B.2 for further details). Missing ratios still appear in some country-year due to

missing Eurostat data. As shown in Tables 6.1 and 6.2, with the exception of Germany, UK, Ireland, and the Netherlands, our data can account more than 50 percent of the aggregate output in all countries and around 80-90 percent in most countries.

Another way to check for the representativeness of the data is to look at the number of firms instead of how much each firm produces in terms of output. Tables 6.3–6.4 show these statistics. In particular, Table 6.3 shows the firm coverage of CompNet Database[1] and our database constructed from BvD data relative to Eurostat with respect to the number of firms in each country-year. We define the total economy as the overlapping sectors of our data with Eurostat data. Each cell corresponds to the number of firms in total economy from the relevant data source relative to number of firms given by Eurostat. Although this type of comparison is less informative because many firms in Eurostat have zero employment (self-employed), we still show it in order to be able to compare our data to the alternative sector-level database CompNet. Neither our database nor CompNet includes self-employed and hence a comparison based solely on number of firms might be misleading. Even with this caveat though, our data captures most of the firms in the total economy in terms of number of firms.

Table 6.4 is a better comparison. It shows the size and sector distribution of firms in CompNet and BvD, compared to the Eurostat data.[2] Each cell corresponds to the share of indicated category's number of firms in total economy from the relevant data source for the given country-year (%). For example, in Belgium, 88 percent of the firms have less than 10 employees in our data and CompNet, whereas the official number from Eurostat is 96 percent. According to this table, our number of firms in terms of employment are very close to what is reported by the Eurostat. In terms of sectors, based on our data in Belgium, 13.3 percent of firms in its total economy operates in manufacturing sector as in Eurostat data.[3]

[1] CompNet is a sector-level database that is constructed by ECB from similar sources as AMADEUS.

[2] CompNet and Eurostat numbers come from ECB WP 1634, February 2014, Tables 6 and 7.

[3] We pick these countries for comparison purposes because only these countries' statistics are reported in ECB WP 1634, February 2014, Tables 6 and 7. Notice that ECB WP 1634, Tables 6 and 7 also report coverage numbers on AMADEUS which are clearly different than our numbers reported here, indicating worse coverage. We suspect that this is an artifact of the way the AMADEUS data was put together by the authors of the ECB WP. As we detail out in the current paper for the best coverage a certain procedure has to be followed.

6.2 Manufacturing

Tables 6.5–6.8 show how much of the official gross output and employment data from Eurostat we cover in our data for the manufacturing sector for Eurozone and non-Eurozone countries. We show these tables in two different samples, Total Sample and TFP Sample: The Total Sample consists of firms that report data with positive values of the corresponding measure (i.e. employment (`EMPL`) and gross-output (`OPRE`)), whereas the TFP Sample consists of firms that report positive values on employment (`EMPL`) or wage bill (`STAF`), and tangible fixed assets (`TFAS`), gross-output (`OPRE`), and materials (`MATE`) so that researchers can calculate total factor productivity.

The table clearly shows that the coverage improves over time for all countries until 2005 and is stable thereafter. In the case of Germany, on average we observe worse coverage than in other countries which is explained by the under-representation of small firms in Germany prior to 2006. Only recently, EU harmonization laws made reporting by small firms compulsory also in Germany. There are other countries (Ireland, the Netherlands, Slovakia if we use employment, and Luxembourg for the data needed for TFP calculation), with problems similar to Germany but overall we cover 70-80 percent of the real economy in manufacturing in the bulk of European countries. Denmark and the UK perform worse in the TFP sample because firms do not report materials use in these countries. Missing percentages appear in some country-years because there is no data available for the corresponding measure in Eurostat for manufacturing sector.[4]

Tables 6.9–6.10 report the size distribution in the manufacturing sector. Tables present the share of economic activity (gross-output and employment) accounted for by firms belonging in three size categories in a given year. We randomly pick 2006. In each panel, the first three rows report the measures from ORBIS-AMADEUS and the next three are same numbers from Eurostat's SBS data. Each column is a different country. Row entries denote the fraction of total economic activity accounted for by firms belonging to each size class. As before in total economy in terms of number of firms, here we also match well the official statistics in terms of size distribution of economic activity undertaken in manufacturing sector.

[4]There are few country-years with missing information due to non-available Eurostat data.

Table 6.1: Coverage in Total Economy Based on Gross Output: Eurozone

	AT	BE	EE	FI	FR	DE	GR	IE	IT	LV	LT	LU	NL	PT	SK	SI	ES
1999	0.21	0.60	0.76	0.35	0.54	0.29	0.95	0.13	0.45	0.54	0.16		0.17	0.45	0.26	0.29	0.63
2000	0.31	0.65	0.85	0.40	0.57	0.46	0.93	0.32	0.50	0.51	0.18		0.21	0.49	0.33	0.66	0.64
2001	0.47	0.63	0.87	0.44	0.65	0.49		0.36	0.50	0.58	0.19		0.20	0.48	0.40	0.66	0.67
2002	0.54		0.88	0.47	0.62	0.58	0.24	0.32	0.54	0.62	0.39		0.22	0.59	0.51	0.77	0.67
2003	0.49	0.63	0.93	0.50	0.67	0.56	0.47	0.35	0.53	0.63	0.53		0.21	0.60	0.75	0.75	0.68
2004	0.47	0.63	0.92	0.51	0.70	0.64	0.48	0.36	0.57	0.84	0.52	0.63	0.22	0.63	0.77	0.74	0.68
2005	0.47	0.63	0.94	0.53	0.69	0.81	0.45	0.34	0.57	0.87	0.54	0.59	0.23	0.80	0.82	0.74	0.69
2006	0.61	0.62	0.99	0.51	0.68	0.68	0.47	0.44	0.58	0.87	0.59	0.48	0.25	0.82	0.79	0.74	0.71
2007	0.68	0.62	0.95	0.55	0.70	0.76	0.47	0.44	0.60	0.86	0.78	0.43	0.25	0.83	0.78	0.73	0.68
2008	0.72	0.73	0.78	0.57	0.79	0.68	0.54	0.49	0.72	0.88	0.62	0.44	0.28	0.82	0.97	0.80	0.80
2009	0.69	0.69	0.89	0.53	0.76	0.66	0.46	0.46	0.68	0.80	0.60	0.28	0.31	0.83	0.93	0.80	0.80
2010	0.71	0.65	0.91	0.59	0.80	0.65	0.45	0.59	0.64	0.79	0.69	0.34	0.35	0.86	0.92	0.95	0.80
2011	0.73	0.74	0.93	0.58	0.81	0.61	0.45	0.54	0.71	0.82	0.57	0.80	0.30	0.81	0.95	0.84	0.82
2012	0.75	0.78	0.96	0.51	0.81	0.47	0.44	0.40	0.61	0.78	0.48	0.72	0.28	0.93	0.91	0.96	0.81

Notes: Tables 6.1 and 6.2 present the ratios that are calculated based on gross output for the countries of Eurozone and Non-Eurozone. The total sample consists of firms that report data with positive values of the corresponding measure (i.e. gross-output). The country codes within these classifications are as follows: AT (Austria), BE (Belgium), EE (Estonia), FI (Finland), FR (France), DE (Germany), GR (Greece), IE (Ireland), LV (Latvia), LT (Lithuania), LU (Luxembourg), NL (Netherlands), PT (Portugal), SK (Slovakia), SI (Slovenia), ES (Spain), IT (Italy); BG (Bulgaria), CH (Switzerland), NO (Norway), CZ (Czech Republic), DK (Denmark), HU (Hungary), RO (Romania), SE (Sweden), GB (Great Britian), PL (Poland), HR (Croatia). BvD provides firm-level information on gross-output (`OPRE`) for all sectors of a given European country between 1999–2012, however Eurostat SBS data provides information on gross output (Turnover) with the exceptions of some sectors (See Table B.3). So, for a given country-year, total economy percentages are computed by taking the ratio of aggregated gross output values where aggregated gross output is computed by totalling gross output over these sectors for which gross-output related variable is available in both data sets.

Table 6.2: Coverage in Total Economy Based on Gross Output: Non-Eurozone

YEAR	BG	HR	CZ	DK	GB	HU	NO	PL	RO	SE
1999			0.39	0.04	0.46	0.79	0.63	0.36	0.58	0.52
2000			0.46	0.23	0.45	0.84	0.63	0.40	0.53	0.56
2001			0.56	0.29	0.46	0.71	0.77	0.47	0.68	0.60
2002	0.57		0.65	0.29	0.48	0.73	0.79	0.44	0.70	0.63
2003	0.60		0.67	0.29	0.48	0.66	0.65	0.47	0.75	0.65
2004	0.79		0.77	0.35	0.48	0.76	0.67	0.57	0.83	0.67
2005	0.83		0.79	0.34	0.51	0.80	0.59	0.54	0.82	0.68
2006	0.84		0.75	0.35	0.53	0.81	0.67	0.61	0.78	0.71
2007	0.91		0.79	0.35	0.48	0.79	0.71	0.65	0.82	0.70
2008	0.94	0.88	0.75	0.40	0.64	0.75	0.59	0.55	0.86	0.73
2009	0.92	0.84	0.79	0.38	0.78	0.87	0.78	0.66	0.92	0.84
2010	0.96	0.00	0.93	0.37	0.68	0.76	0.79	0.61	0.92	0.88
2011	0.88	0.82	0.85	0.42	0.83	0.72	0.72	0.56	0.92	0.82
2012	0.88	0.81	0.79	0.49	0.83	0.82	0.67	0.59	0.91	0.76

Notes: See the notes in Table 6.1

Table 6.3: Coverage relative to Eurostat for Number of Firms in Total Economy

Country	Year	CompNet	BvD
Belgium	2008	26.5	65.2
Estonia	2007	65.9	97.1
France	2009	30.6	85.2
Germany	2008	3.1	63.6
Hungary	2007	3.6	37.9
Italy	2008	2.2	58.8
Poland	2007	1.2	12.3
Slovakia	2008	12.8	40.3
Slovenia	2007	28.4	19.5
Spain	2008	23.6	41.7

Notes: Each cell corresponds to the number of firms in total economy from the relevant data source relative to the number of firms in Eurostat for the given country-year (%).

Table 6.4: Firm Size Distribution in Total Economy by Sector
Based on Number of Firms

		Size Distribution									Sector Distribution					
		0 to 19 employees			20 to 249 employees			250 + employees			Manufacturing			Non-manufacturing		
Country	Year	CompNet	BvD	Eurostat	CompNet	BvD	Eurostat	CompNet	BvD	Eurostat	CompNet	BvD	Eurostat	CompNet	BvD	Eurostat
Belgium	2008	88.9	87.6	96.3	10.3	11.5	3.5	0.8	0.9	0.2	20.7	13.3	14.8	79.3	86.7	85.2
Estonia	2007	86.1	86.4	90	13.2	12.9	9.5	0.6	0.7	0.5	24.5	17.8	20.6	75.5	82.2	79.4
France	2009	86.3	87.8	96.4	12.8	11.1	3.4	0.9	1.1	0.2	19.2	16.6	14.9	80.8	83.4	85.1
Germany	2008	24.8	81	91.8	61.6	17.4	7.6	13.6	1.6	0.6	52.3	18.5	17	47.7	81.5	83
Hungary	2007	57.6	84.2	96.9	37.9	14.5	2.9	4.5	1.6	0.2	72.8	23.7	16.4	27.2	76.3	83.6
Italy	2008	56	85.1	97.4	43	14.1	2.5	1	0.8	0.1	58.2	28.3	19.5	41.8	71.7	80.5
Poland	2007	0.0	64.2	93.7	95.2	32.4	5.7	4.8	5.4	0.6	65.4	37.8	45	34.6	62.2	55
Slovakia	2008	11.1	85.6	87.5	78.3	12.8	11.2	10.6	1.6	1.3	53.4	24.7	27.3	46.6	75.3	72.7
Slovenia	2007	87.9	76.3	95.9	11.1	25.2	3.7	1	3.5	0.4	26	21.6	24.8	74	78.4	75.2
Spain	2008	91.5	86.8	96.1	8.1	12.6	3.7	0.4	0.6	0.2	20.3	21.3	19.4	79.7	78.7	80.6

Notes: Each cell corresponds to the share of indicated category's number of firms in total economy from the relevant data source for the given country-year (%). Number of firms is summed over overlapping sectors with Eurostat SBS data.

Table 6.5: Coverage in terms of Manufacturing Sector: Eurozone, Total Sample

	AT	BE	EE	FI	FR	DE	GR	IE	IT	LV	LT	LU	NL	PT	SK	SI	ES
Panel A: Gross Output																	
1999	0.2	0.75	0.83	0.3	0.64	0.21	0.95	0.12	0.61	0.56	0.13		0.19	0.59	0.32	0.3	0.75
2000	0.29	0.8	0.9	0.34	0.76	0.36	0.93	0.27	0.66	0.56	0.15		0.19	0.65	0.42	0.77	0.77
2001	0.47	0.78	0.92	0.36	0.79	0.5		0.38	0.65	0.63	0.17		0.2	0.63	0.54	0.83	0.78
2002	0.61		0.93	0.37	0.82	0.51		0.36	0.71	0.62	0.35		0.23	0.71	0.62	0.9	0.8
2003	0.59	0.81	0.93	0.39	0.79	0.57	0.92	0.33	0.7	0.61	0.43	0.46	0.26	0.73	0.83	0.92	0.79
2004	0.47	0.8	0.97	0.41	0.83	0.64	0.73	0.36	0.73	0.73	0.4	0.47	0.27	0.75	0.88	0.89	0.79
2005	0.45	0.8	0.95	0.41	0.82	0.9	0.69	0.29	0.77	0.79	0.41	0.43	0.24	0.91	0.78	0.91	0.78
2006	0.67	0.78	0.94	0.4	0.84	0.73	0.66	0.43	0.79	0.8	0.44	0.3	0.31	0.93	0.78	0.91	0.83
2007	0.72	0.79	0.96	0.45	0.87	0.77	0.68	0.42	0.79	0.81	0.57	0.33	0.32	0.92	0.76	0.91	0.81
2008	0.77	0.78	0.95	0.49	0.9	0.64	0.64	0.5	0.9	0.85	0.51	0.21	0.29	0.94	0.99	0.92	0.85
2009	0.72	0.74	0.92	0.46	0.89	0.6	0.51	0.47	0.86	0.81	0.57	0.24	0.3	0.92	0.95	0.91	0.87
2010	0.78	0.78	0.92	0.47	0.92	0.6	0.47	0.58	0.87	0.83	0.52	0.92	0.31	0.93	0.94	0.99	0.9
2011	0.8	0.77	0.95	0.5	0.96	0.57	0.45	0.57	0.89	0.87	0.5	0.9	0.3	0.93	0.98	0.83	0.85
2012	0.76	0.77	0.96	0.51	0.95	0.48	0.4	0.4	0.86	0.83	0.5		0.31	0.92	0.91	0.97	0.83
Panel B: Employment																	
1999				0.33	0.51				0.44			0.31	0.17	0.07			0.54
2000			0.81	0.41	0.57				0.52	0.43	0.11	0.36	0.25	0.08	0.23	0.57	0.57
2001			0.85	0.43	0.59				0.53	0.48	0.14	0.33	0.19	0.07	0.27	0.53	0.61
2002			0.88	0.46					0.6	0.51	0.29		0.32	0.08	0.33	0.74	0.67
2003			0.83	0.46	0.68				0.59	0.52		0.01	0.52	0.08	0.4	0.8	0.68
2004	0.7		0.84	0.48	0.67	0.53	0.49		0.53	0.76	0.41	0.01	0.44	0.08	0.48	0.8	0.67
2005	0.71		0.86	0.49	0.63	0.58	0.49		0.43	0.79	0.42	0.2	0.4	0.07	0.69	0.8	0.69
2006	0.73	0.79		0.5	0.62	0.65	0.43		0.52	0.82	0.48	0.02	0.45	0.81	0.76	0.84	0.7
2007	0.76	0.8		0.53	0.67	0.74	0.43		0.53	0.84	0.71	0.34	0.56	0.83	0.74	0.83	0.69
2008	0.77	0.82		0.48	0.73	0.74		0.53	0.63	0.87	0.69	0.5	0.52	0.89	0.81	0.84	0.78
2009	0.65	0.78	0.74	0.44	0.45	0.73		0.43	0.58	0.87	0.75	0.6	0.51	0.88	0.84	0.8	0.79
2010	0.72	0.79	0.86	0.45	0.49	0.66		0.41	0.58	0.89	0.79	0.68	0.54	0.89	0.85	0.91	0.84
2011	0.72	0.81	0.85	0.49	0.44	0.66		0.39	0.76	0.9	0.62	0.67	0.77	0.89	0.8	0.93	0.87
2012	0.69	0.82	0.82		0.35	0.61		0.23	0.75	0.9		0.59	0.69	0.9	0.14	0.92	0.82

Notes: Tables 6.5–6.8 present the percentages that are calculated based on different measures of economic activity (employment and gross output) for the countries of Eurozone and Non-Eurozone, respectively. The country codes within these classifications are as follows: AT (Austria), BE (Belgium), EE (Estonia), FI (Finland), FR (France), DE (Germany), GR (Greece), IE (Ireland), LV (Latvia), LT (Lithuania), LU (Luxembourg), NL (Netherlands), PT (Portugal), SK (Slovakia), SI (Slovenia), ES (Spain), IT (Italy); BG (Bulgaria), NO (Norway), CZ (Czech Republic), DK (Denmark), HU (Hungary), RO (Romania), SE (Sweden), GB (Great Britian), PL (Poland), HR (Croatia). Percentages correspond to the ratio of aggregation of the corresponding measure over all firms in ORBIS-AMADEUS sample to that in Eurostat sample. The percentages based on employment are obtained by the comparison of our data and Eurostat's Eurostat's Business Demography (BD) data, which excludes self-employed workers. The percentages based on gross output are obtained by comparison of our data and Eurostat's Structural Business Statistics (SBS) data which includes self-employed workers. Those percentages are produced for two different samples, Total Sample and TFP Sample: The Total Sample consists of firms that report data with positive values of the corresponding measure (i.e., employment and gross-output), whereas the TFP Sample consists of firms that report positive values on employment or wage bill, and tangible fixed assets, gross-output, materials. Further details on comparison exercises are available in Section B.2.

Table 6.6: Coverage in terms of Manufacturing Sector: Eurozone, TFP Sample

	AT	BE	EE	FI	FR	DE	IE	IT	LV	LU	NL	PT	SK	SI	ES
Panel A: Gross Output															
1999		0.71	0.81	0.25	0.6	0.01		0.6			0.07	0.51	0.02		0.75
2000	0.01	0.76	0.88	0.27	0.72	0.04	0.03	0.66	0.03		0.	0.46	0.06		0.76
2001	0.11	0.74	0.89	0.29	0.71	0.17	0.03	0.65	0.04		0.08	0.46	0.23	0.08	0.77
2002	0.27		0.9	0.3	0.78	0.22	0.05	0.7	0.05		0.1	0.56	0.55	0.79	0.79
2003	0.3	0.78	0.6	0.31	0.75	0.27	0.01	0.69	0.06	0.17	0.11	0.69	0.83	0.78	0.78
2004	0.33	0.77	0.56	0.32	0.78	0.28	0.01	0.73	0.05	0.1	0.13	0.71	0.86	0.85	0.78
2005	0.33	0.77	0.56	0.32	0.76	0.42		0.76	0.04	0.25	0.09	0.87	0.76	0.89	0.77
2006	0.55	0.75	0.56	0.31	0.8	0.39		0.78	0.04	0.24	0.1	0.91	0.73	0.9	0.77
2007	0.5	0.76	0.56	0.32	0.81	0.42		0.78	0.05	0.3	0.11	0.9	0.71	0.9	0.77
2008	0.55	0.76	0.55	0.34	0.85	0.34		0.88	0.05	0.19	0.09	0.93	0.78	0.92	0.72
2009	0.47	0.73	0.58	0.31	0.76	0.33		0.84	0.04	0.22	0.02	0.91	0.75	0.9	0.75
2010	0.54	0.76	0.59	0.31	0.83	0.37	0.01	0.84	0.06	0.87	0.05	0.92	0.75	0.97	0.74
2011	0.55	0.76	0.64	0.34	0.84	0.35	0.06	0.87	0.06	0.8	0.01	0.92	0.93	0.80	0.75
2012	0.38	0.76	0.65	0.35	0.82	0.31	0.06	0.84	0.05			0.91	0.88	0.94	0.72
Panel B: Employment															
1999				0.29	0.5			0.44		0.14	0.03	0.07			0.54
2000			0.79	0.36	0.55			0.51	0.01	0.18	0.04	0.08	0.04		0.56
2001			0.83	0.38	0.57			0.53	0.04	0.17	0.04	0.07	0.09		0.6
2002			0.85	0.4				0.6	0.04		0.04	0.07	0.26	0.68	0.65
2003			0.58	0.4	0.65			0.54	0.05		0.05	0.07	0.39	0.72	0.66
2004	0.21		0.59	0.41	0.64	0.13		0.52	0.04		0.07	0.08	0.44	0.74	0.65
2005	0.25		0.61	0.42	0.61	0.19		0.43	0.04		0.06	0.07	0.66	0.77	0.67
2006	0.34	0.62		0.43	0.6	0.24		0.52	0.04		0.05	0.79	0.7	0.81	0.68
2007	0.36	0.63		0.44	0.64	0.25		0.53	0.03	0.29	0.06	0.8	0.68	0.81	0.67
2008	0.37	0.64		0.39	0.69	0.24		0.6	0.03	0.41	0.05	0.86	0.72	0.8	0.67
2009	0.34	0.61	0.46	0.36	0.43	0.24		0.56	0.03	0.51	0.01	0.85	0.78	0.78	0.71
2010	0.41	0.62	0.55	0.36	0.47	0.24	0.01	0.57	0.04	0.59	0.01	0.86	0.73	0.88	0.7
2011	0.38	0.63	0.55	0.4	0.42	0.25	0.14	0.73	0.04	0.56		0.86	0.67	0.9	0.72
2012	0.27	0.65	0.53		0.34	0.21	0.14	0.73	0.04	0.48		0.86		0.88	0.67

Notes: See notes to Table 6.5.

Table 6.7: Coverage in terms of Manufacturing Sector: Non-Eurozone, Total Sample

	BG	HR	CZ	DK	GB	HU	NO	PL	RO	SE
Panel A: Gross Output										
1999			0.54	0.04	0.6	0.84	0.6	0.51	0.66	0.6
2000			0.63	0.29	0.6	0.88	0.6	0.56	0.61	0.67
2001			0.73	0.31	0.68	0.91		0.62	0.76	0.73
2002	0.57		0.7	0.34	0.66	0.97	0.75	0.57	0.74	0.76
2003	0.58		0.77	0.34	0.66	0.84	0.68	0.59	0.83	0.74
2004	0.72		0.84	0.41	0.62	0.91	0.72	0.71	0.96	0.77
2005	0.83		0.86	0.43	0.7	0.88	0.69	0.67	0.95	0.75
2006	0.89		0.81	0.44	0.72	0.91	0.75	0.7	0.84	0.79
2007	0.92		0.88	0.43	0.69	0.88	0.76	0.74	0.91	0.78
2008	0.99	0.92	0.8	0.43	0.67	0.81	0.69	0.6	0.9	0.74
2009	0.96	0.99	0.85	0.36	0.81	0.9	0.88	0.74	0.97	0.87
2010	0.97	0.95	0.94	0.38	0.84	0.84	0.83	0.69	0.93	0.88
2011	0.89	0.99	0.9	0.43	0.89	0.77	0.83	0.6	0.91	0.85
2012	0.85	0.98	0.87	0.43	0.83	0.86	0.88	0.62	0.93	0.78
Panel B: Employment										
1999					0.56					0.5
2000			0.62		0.58	0.03			0.76	0.57
2001			0.72		0.58	0.05			0.71	0.58
2002			0.77		0.59	0.22			0.74	0.59
2003			0.71		0.56	0.09	0.69		0.8	0.62
2004	0.9		0.77	0.63	0.54	0.17	0.68	0.5	0.85	0.64
2005	0.87		0.73	0.66	0.55	0.32	0.77	0.51	0.92	0.65
2006	0.92		0.75	0.68	0.56	0.42	0.77	0.51	0.76	0.66
2007	0.95		0.77	0.68	0.58	0.82	0.72	0.49	0.97	0.69
2008	0.84		0.65		0.63	0.66	0.75	0.52	0.91	0.75
2009	0.83		0.82	0.58	0.58	0.89	0.73		0.78	0.7
2010	0.83		0.94	0.56	0.61	0.89	0.72	0.82	0.93	0.74
2011	0.84		0.95	0.56	0.63	0.89	0.78	0.88	0.94	0.78
2012	0.94	0.83	0.86	0.54	0.63	0.92	0.82	0.72	0.96	0.77

Notes: See notes to Table 6.5.

Table 6.8: Coverage in terms of Manufacturing Sector: Non-Eurozone, TFP Sample

	BG	HR	CZ	DK	GB	HU	NO	PL	RO	SE
Panel A: Gross Output										
1999			0.07	0.01		0.72	0.47	0.39	0.66	0.21
2000			0.19	0.01		0.72	0.47	0.43	0.60	0.22
2001			0.45	0.01		0.78		0.46	0.75	0.25
2002	0.51		0.54	0.01		0.83	0.6	0.35	0.74	0.27
2003	0.52		0.73			0.7	0.56	0.36	0.82	0.26
2004	0.45		0.82	0.04		0.82	0.68	0.43	0.95	0.29
2005	0.58		0.81	0.05		0.82	0.66	0.42	0.94	0.28
2006	0.75		0.79	0.05		0.85	0.71	0.44	0.83	0.31
2007	0.85		0.87	0.05		0.84	0.72	0.47	0.9	0.31
2008	0.9	0.9	0.79	0.05	0.15	0.74	0.65	0.39	0.89	0.28
2009	0.87	0.87	0.84	0.05	0.22	0.8	0.84	0.5	0.96	0.34
2010	0.88	0.83	0.87	0.05	0.22	0.76	0.79	0.46	0.92	0.34
2011	0.87	0.86	0.84	0.05	0.21	0.69	0.79	0.4	0.9	0.32
2012	0.83	0.86	0.83	0.04	0.2	0.76	0.83	0.42	0.92	0.27
Panel B: Employment										
1999										0.23
2000			0.19			0.02			0.76	0.25
2001			0.42			0.03			0.7	0.26
2002			0.58			0.09			0.73	0.27
2003			0.65			0.08	0.52		0.79	0.28
2004	0.4		0.73	0.05		0.15	0.63	0.27	0.84	0.31
2005	0.39		0.69	0.06		0.28	0.72	0.28	0.9	0.32
2006	0.56		0.72	0.06		0.38	0.73	0.3	0.72	0.33
2007	0.68		0.75	0.06		0.77	0.68	0.29	0.94	0.34
2008	0.66		0.64		0.16	0.6	0.71	0.32	0.89	0.38
2009	0.61		0.8	0.07	0.16	0.65	0.69	0.3	0.76	0.35
2010	0.61		0.83	0.07	0.16	0.66	0.68	0.19	0.91	0.37
2011	0.75		0.85	0.08	0.16	0.66	0.74	0.2	0.92	0.39
2012	0.86	0.81	0.75	0.08	0.18	0.65	0.78	0.11	0.93	0.31

Notes: See notes to Table 6.5.

Table 6.9: Size Distribution in terms of Manufacturing Sector: Eurozone, Total Sample, 2006

		AT	BE	EE	FI	FR	DE	GR	IE	LV	NL	PT	SI	ES	IT
Panel A: Gross Output															
ORBIS-AMADEUS															
	1 to 19 employees	0.03	0.05	0.15	0.08	0.05	0.06	0.03	0.14	0.04	0.13	0.11	0.07	0.13	0.12
	20 to 249 employees	0.23	0.30	0.67	0.38	0.23	0.23	0.97	0.62	0.45	0.44	0.37	0.32	0.40	0.49
	250 + employees	0.74	0.66	0.18	0.54	0.72	0.70	0.00	0.24	0.51	0.43	0.52	0.61	0.47	0.40
Eurostat (SBS)															
	0 to 19 employees	0.07	0.08	0.12	0.06	0.09	0.06	0.02	0.12	0.09	0.14	0.05	0.12	0.14	0.20
	20 to 249 employees	0.32	0.27	0.60	0.21	0.27	0.22	0.10	0.54	0.33	0.42	0.23	0.32	0.38	0.41
	250 + employees	0.61	0.65	0.28	0.74	0.64	0.72	0.88	0.34	0.59	0.43	0.72	0.56	0.49	0.39
Panel B: Employment															
ORBIS-AMADEUS															
	1 to 19 employees	0.08	0.13	0.15	0.16	0.1	0.05	0.13	0.06	0.19	0.21	0.26	0.08	0.25	0.13
	20 to 249 employees	0.36	0.41	0.58	0.43	0.34	0.32	0.57	0.49	0.56	0.54	0.54	0.38	0.49	0.55
	250 + employees	0.56	0.45	0.27	0.41	0.56	0.63	0.3	0.46	0.25	0.24	0.2	0.55	0.26	0.32
Eurostat (SBS)															
	0 to 19 employees	0.17	0.18	0.17	0.14	0.19	0.15	0.5	0.1	0.18	0.25	0.32	0.18	0.31	0.4
	20 to 249 employees	0.38	0.37	0.54	0.34	0.34	0.32	0.28	0.43	0.52	0.42	0.49	0.36	0.43	0.38
	250 + employees	0.46	0.45	0.3	0.51	0.47	0.53	0.22	0.47	0.3	0.33	0.19	0.45	0.26	0.22

Notes: Tables 6.9–6.10 present the share of economic activity (employment and gross output) accounted for by firms belonging in three size categories in a given year. We pick 2006. In each panel, the first three rows report the measures from ORBIS-AMADEUS for Total Sample which consists of firms that report data with positive values of the corresponding measure, the next three rows are same numbers from Eurostat's SBS data. Row entries denote the fraction of total economic activity accounted for by firms belonging to each size class. Each column is a different country. The country codes within these classifications are as follows: AT (Austria), BE (Belgium), EE (Estonia), FI (Finland), FR (France), DE (Germany), GR (Greece), IE (Ireland), LV (Latvia), LT (Lithuania), LU (Luxembourg), NL (Netherlands), PT (Portugal), SK (Slovakia), SI (Slovenia), ES (Spain), IT (Italy); BG (Bulgaria), NO (Norway), CZ (Czech Republic), DK (Denmark), HU (Hungary), RO (Romania), SE (Sweden), GB (Great Britian), PL (Poland), HR (Croatia). Further details on comparison exercises are available in Section B.2.

Table 6.10: Size Distribution in terms of Manufacturing Sector: Non-Eurozone, Total Sample, 2006

		BG	HR	NO	CZ	HU	RO	SE	PL	DK	GB
Panel A: Gross Output											
ORBIS-AMADEUS											
	1 to 19 employees	0.08	0.12	0.11	0.04	0.01	0.09	0.09	0.02	0.03	0.03
	20 to 249 employees	0.31	0.36	0.40	0.32	0.14	0.32	0.26	0.32	0.29	0.23
	250 + employees	0.61	0.52	0.49	0.64	0.85	0.59	0.65	0.67	0.68	0.74
Eurostat (SBS)											
	0 to 19 employees	0.08		0.13	0.10	0.07	0.08	0.09	0.09	0.10	0.08
	20 to 249 employees	0.31		0.36	0.31	0.21	0.32	0.28	0.28	0.34	0.29
	250 + employees	0.61		0.51	0.59	0.72	0.60	0.63	0.62	0.55	0.63
Panel B: Employment											
ORBIS-AMADEUS											
	1 to 19 employees	0.16	0.18	0.05	0.14	0.02	0.12	0.17	0.01	0.03	0.12
	20 to 249 employees	0.48	0.47	0.41	0.48	0.26	0.38	0.34	0.35	0.4	0.39
	250 + employees	0.36	0.35	0.53	0.38	0.72	0.5	0.48	0.64	0.57	0.5
Eurostat (SBS)											
	0 to 19 employees	0.16	0.2	0.19	0.15	0.2	0.12	0.18	0.19	0.21	
	20 to 249 employees	0.49	0.42	0.37	0.4	0.36	0.4	0.33	0.38	0.38	
	250 + employees	0.35	0.38	0.44	0.46	0.44	0.48	0.49	0.43	0.41	

Notes: See notes to Table 6.9.

Part IV

Ownership Panel Data

Chapter 7

Vintage Raw Data

The following are the key steps we take in constructing the firm-level ownership database. In some respect, the process of preparing the raw vintage data is more straightforward than that for the financials because each vintage corresponds to a single time observation. However, compared to the financial data, the ownership data has more than two identifiers (the company ID and year). In the case of ownership, the additional dimension comes from the fact that each company could have multiple owners or subsidiaries.

The raw ASCII data has the rows consisting of non-time varying data and the blocks of variables corresponding to the groups in ORBIS Ownership database. They are company's shareholders, domestic ultimate owners (UOs), global UOs, and subsidiaries. In order to have the data for a given country in one file we download data using the Formatted Export method in older disks and the List Export method in newer disks as discussed in Section 2. Of note, we download some variables describing the company itself such as name, BvD ID, size category, type of company, location information, core industry, and three key financials (employment, total assets, and operating revenue in the original currency of accounts and in the last available year, together with Units and Currency Code) in case we would need to identify the company better. The key identifier is still company BvD ID.

Assume a Company A has 2 shareholders, 1 global UO, and 3 subsidiaries. The data for this company will have 3 observations (rows) in a given vintage, corresponding to the largest number of observations across the groups in ORBIS Ownership database:

NAME*	BvD ID*	Core Industry*	Shareholder Name	Shareholder % Stake Direct	GUO Name	GUO % Stake	Subsidiary Name	Subsidiary % Stake
Company A	ZZ1234U	6123	Company B	70	Company B	100	Company C	100
Company A	ZZ1234U	6123	Mr. Smith	30			Company D	90
Company A	ZZ1234U	6123					Company E	WO

The variables marked with (*) are unique per company-vintage. Because we export the data asking the disk to repeat each single item all these variables will be repeated for each company record in the raw downloads. The order of the non-unique variables corresponds to the internal order in the ORBIS Ownership or AMADEUS Ownership databases.

We execute the following steps for each vintage of the ORBIS Ownership or AMADEUS Ownership database we use.

1. We extract the data in ASCII comma-separated value format, transform it to Stata using Statransfer, name and label the variables following the database names.

2. Delete the observations with just a name of company and no other information.

3. Generate the YEAR variable marking the time as of which the ownership information is recorded in a given vintage of BvD product. Following our assumption the YEAR takes the value of one less the year when the ORBIS disk was issued. Recall that we choose the vintages that are closest but subsequent to that year. For example, the ownership data coming from the ORBIS disk no.27 issued in January 2011 is assigned to the YEAR 2010, and so on.

4. As in case of financial data, we create our main identifier `ID_NUMBER`, which is a copy of the BvD account number. It is a copy of the BvD ID number if the BvD account number is missing.

5. To keep the file size manageable we split the data generating a separate file for a given country in a given year and the group of variables in ORBIS Ownership database. As a result we have the following files per country-year.

    - In the file with the information about the company itself we keep all the identifiers mentioned above, removing the duplicates in terms of all remaining variables and then removing duplicates by ID and Employment.

- In the file with information on company shareholders we keep ID, YEAR and all variables from this group, dropping observations with missing information on owners and then removing the duplicates in terms of all remaining variables.
- We do the same in the file with separate information for company's immediate shareholders, for its domestic ultimate owners, for its global ultimate owners, and for its subsidiaries.

Chapter 8

Merging and Cleaning Ownership Data

In this section we illustrate how we create the country panels of the company shareholder data; that is, the information on the direct ownership of the equity rights for the universe of companies covered in ORBIS Ownership dataset, "Shareholders" group of variables. The preparation of the other ownership data involves similar steps. We first describe how we prepare the company panels of the direct ownership "links", that is the data where the unit of observation is the company-its shareholder pairs in a given year. After that we discuss how we build the company-year level data where the "links" information is aggregated to the company level, again, by year.

8.1 Data with Company Shareholder Links

Each of the following steps is performed for individual country ownership files coming either from ORBIS Ownership or AMADEUS Ownership databases.

1. Combine (stack) all annual ownership data files for a given country. Recall that we have bi-annual vintages of AMADEUS Ownership since 2000 and annual vintages of ORBIS Ownership since 2005.

2. Create the country code based on the first two letters of the `ID_NUMBER` which by BvD convention should start from two-letter country code (BE for Belgium, US for the U.S., GB for the UK and so on).

3. Convert the character variables Percentage Owned Direct (`ODIRECT`) and Percentage Owned Total (`OTOTAL`) into numeric format, replacing some special character values they may take in the raw data. In particular, we replace percentage with a leading <,>, ± with the percentage after the symbol; eliminate possible % sign; replace special codes "WO" (wholly owned) with 100%, "MO" (majority owned) with 50.01% (because by the GAAP practice the majority ownership involves 50% plus one share but the smallest stakes reported by BvD are 0.01%), "CQP1" (50% plus 1 share) with 50.01%, "NG" (negligible) with 0.01% (again, the smallest observed stake according to BvD), "-" (not significant) or "n.a." (not available) with missing; "BR" (branch, ORBIS Ownership only) with 100%; "JO" (jointly owned, AMADEUS Ownership only) with 50% (our exploration of such cases shows that there is always exactly two owners in case of the JO code). We keep the character versions of these variables.

4. Implement company ID changes following the procedure described in Section 4 for financial data, saving the legacy IDs on the dataset.

5. To determine the total foreign ownership at the company level, we proceed as follows. Whenever the variable Shareholder Country ISO code (`OCOUNTRY` in AMADEUS or `SHARCOUN` in ORBIS) is different from the company own Country code we consider the link foreign. By default, we assume that the shareholder with missing country code or with `SHARCOUN` taking the values of "-" and "n.a." is located in the same country as the given company as it is done in the literature.

6. For AMADEUS Ownership we further improve the above simple rule by the manual assignments by country, based on the variable Shareholder Name (`ONAME`). That variable contains some indication that the owner is foreign even when its is missing.[1] We determine the unique values of `ONAME` and then manually replace the ownership links which have the missing `OCOUNTRY`. One can further try to compare the company

[1] We make this decision mostly based on the parts of the names reflecting the legal type of the shareholder. We compare that to what is typical for the company's own country.

country code to the first two letters of the available Shareholder BvDEP ID number (`OID` in AMADEUS or `SHARSIR` in ORBIS), but not every shareholder has the BvD ID available. In addition one can investigate the values of the variable Shareholder Type (the textual variable `OTYPE` in AMADEUS or the standardized letter code `SHARTYPE` in ORBIS), assuming that certain owner types are domestic (such as company employees or management, or the entries like "private individual(s)" or "unnamed private shareholders", and so on) unless it is clear from the name they are foreign (for example, "foreign investors").

7. Combine shareholders information from our AMADEUS Ownership and ORBIS Ownership, making sure we do not have duplicates in overlapping years. In overlapping years we establish which database has more recorded shareholders with non-missing direct ownership stakes for a given company-year (non-missing variable Shareholder - Percentage of ownership, direct, `SHARDPER` and use the record with more data.[2] In case of exact match we use ORBIS but retain the variable Shareholder NACE Rev. 2 Core Code (`ONACE`) and other useful information available only in AMADEUS. Actual elimination of the duplicates is done later because we use some information across two products in a given year to supplement the data.

8. Supplement the missing textual information (Shareholder Type in textual form from AMADEUS and 1-letter code from ORBIS; Shareholder NACE Rev. 2 Core Code) using the unique shareholder IDs and shareholder names. Supplement missing Shareholder Type in textual form from AMADEUS using non-missing 1-letter code from ORBIS and vice versa.

9. Because we first establish the foreign ownership link status at the individual vintage level, we can take advantage of the full panel and revise the foreign link status based on *other* years when this info is available for a given company ID. As above, using the

[2]We encountered the case when the same observation (ownership links) got the value "-" in AMADEUS Ownership and the value of exactly 0% in ORBIS Ownership. We think that in AMADEUS Ownership the code "-" actually means "negligible" while it is "missing" in ORBIS Ownership . Since we could not confirm or refute this distinction we recoded "-" as missing. Because the 0% stakes do not bear any useful information, we treat the observations with Shareholder Percentage=0% as missing when counting the shareholders with available ownership stake.

unique shareholder IDs and shareholder names, we replace links ever found to be as foreign in at least one year as such in all the years.

10. After all the information across products has been used, we delete the duplicated observations (all the links for a given company-year) from ORBIS and AMADEUS keeping the ones with the best coverage. We save the data files of direct ownership links for individual countries.

8.2 Direct Ownership Types

Once we have the shareholder links data for country-years, we can aggregate it in the variety of ways because each link record has a number or variables describing not only the equity stake held by a given (direct) investor but also investor's location, type, or industry.

For illustration, we will discuss our work to aggregate ownership stakes by foreign/domestic status, further split by the type of the shareholder involved. In particular, this data can be merged with the financials data panel, described in part III, also uniquely identified by the company ID (IDNUMBER") and YEAR". Each of the following steps is performed over individual country data files.

1. We start from the ownership links data obtained as described in section 8.1 and delete the link records which have no usable information.

2. Identify foreign and domestic links, specific to the owners of a particular type. The exercise is similar to what we did to find foreign/domestic links but here we use the information on whether a link is foreign or domestic and, in addition, take advantage of the information in the variable Shareholder Type (the textual variable OTYPE" in AMADEUS or the standardized letter code SHARTYPE" in ORBIS). Recall that in the codes preparing raw links data we supplemented the values of both variables with one another and across all the years. In particular we define the following indicators:

    - *Foreign (Domestic) Owner-Industrial Type*, =1 if Foreign (Domestic) owner has the (textual) type Industrial company, Corporate, Self-owned, Branch, or types

reflecting the individuals working for the company (such as, employees, personnel, managers, directors, self-ownership) because these owners are likely to bring similar types of "expertise" as industrial owners for the majority of the companies in our financials database.[3]

- *Foreign (Domestic) Owner-Financial Type*, =1 if Foreign (Domestic) owner has the type Bank, Financial company, Insurance company, Other financial institution, Mutual & Pension Fund/Nominee/Trust/Trustee, Foundation/Research Institute, Private Equity firms, Venture capital, Hedge funds.
- *Foreign (Domestic) Owner-Government Type*, =1 if Foreign (Domestic) owner has type Governments, State, Public authority.
- *Foreign (Domestic) Owner-Individuals Type*, =1 Foreign (Domestic) owner has the 'individuals' type with known names. Besides single private individuals or families, this category includes shareholders designated by more than one *named* individual or families (the entries like "Mr Gregory Edward Bailey & Mrs Margaret Ethel Bailey" or "Mme Bringaud et son fils". The idea behind this is that they would probably exert their voting power alone or together.
- *Foreign (Domestic) Owner-Other Types*, =1 if Foreign (Domestic) owner has the *aggregated* types including unnamed individuals, the entries indicating that there are more than one private shareholders, collectively designated (for example, "Individual(s) or family(ies)"), unknown types of owners ("Unnamed private shareholders, aggregated", Miscellaneous; Undefined company, Unknown, "n.a.", NA), or simply missing owner type.
- *Owner is Public*, =1 if the company is owned by numerous shareholders, collectively designated as "public." The owner type "Public" is possible only for publicly quoted companies.[4]

3. In addition we specify two sub-types of the owners of financial type:

[3]ORBIS has some companies in the financial intermediation and insurance sector. Our assumption is less valid for such entities.

[4]Notice, that for public companies BvD may report some owners of more specific types (banks, individuals, industrial companies, etc.) with their corresponding stakes, as long as those stakes are known. We do not assign the owners designated collectively as "Public" to either foreign or domestic type because we do not know how the shareholder base is split.

- *Foreign (Domestic) Owner-Active Financial Type*, =1 if Foreign (Domestic) owner has the type Financial company, Insurance company, Other financial institution, Mutual & Pension Fund/Nominee/Trust/Trustee, Foundation/Stichting, Private Equity firms, Venture capital, Hedge funds.[5]
- *Foreign (Domestic) Owner-Passive Financial Type*, =1 if Foreign (Domestic) owner has the type Bank.

4. Once we identify the stakes held by the direct investors of certain type with the help of those dummies, we can generate the company-level foreign and domestic ownership variables out of the links data. We "collapse" data by summing up the stakes of the same owner type, separately domestic and foreign, by year and company ID. We also generate the simple count of number of owners, foreign owners as well as dummies identifying all owner types at the company-year level. For convenience we also create the variable "Check 100 Ownership" which is the sum of all known ownership stakes. If this variable is less than 100 we have unassigned ownership percentages in certain company-years which we assign as domestic. After summation by collapse, the ownership stake percentages larger than 100% are possibly due to rounding and replacing of some special codes such as "NG" with 0.01%, or "WO" with 100%, or simply due to ownership data mistakes. Hence, we perform some cleaning steps.

5. We round all the direct ownership percentages to the second digit after the decimal (to repeat, the smallest stake observed by ORBIS is 0.01%).

6. Remove duplicates in terms of `ID_NUMBER` and `YEAR` by retaining that of the two duplicates which has larger number of ownership observations, hoping that there is a bigger chance it provide useful data for percentages, etc.

7. After this, we delete the observations with the ownership percentages larger than 103% and then replace the values 'slightly more' than 100%, that is in $(100, 103]$ range, with exactly 100%. At this stage, we have constructed an unbalanced country panel of company-level direct ownership data.

[5]Unfortunately we cannot separate such arguably passive types of institutional investors as pension funds from the combined type "Mutual & Pension Fund/Nominee/Trust/Trustee" or corresponding code SHARTYPE="E".

8. The ownership panel is merged with financials panel by `ID_NUMBER` and `YEAR`.

9. In the combined dataset, we fill-in the missing ownership time series, regardless of the time coverage of financials. In particular, we supplement ownership data in missing years using previous or consecutive non-missing values (using Stata `-carryforward-` command) assuming the following: 1. We use the existing earliest value of ownership to carry forward until i) the new non-missing value is reached or ii) the end of the time-series is reached for that company. 2. For the missing initial years of ownership, we assume that it is the same as in the first observation of non-missing data.[6]

10. After we filled-in ownership data using lags and leads we assign the companies with financial data but no ownership data to the category of companies with Domestic Industrial ownership.

[6] We can keep track of the filled-in ownership data using the "Check 100 Ownership" variable which will be missing for filled-in observations.

Chapter 9

Comparison to OECD Foreign Ownership (FDI) Data

We compare our ownership panel to the alternative source on the inward and outward activities of multinationals. The OECD provides data on the activities of foreign affiliates of multinationals in OECD countries in the AMNE database (Activity of Multinational Enterprises). They key variables presented are production, employment, value added, research and development, labor compensation and exports. The data is broken down by country of origin of the ultimate owner (inward investment), location (outward investment), and main sector of economic activity of the multinational company following ISIC Rev.4 sector classification. AMNE covers 28 OECD host countries from 2008, onwards, although the coverage varies by country and over time.

For historical data, users need to refer to two prior databases that used the ISIC Rev. 3 classification: AFA (Activities of Foreign Affiliates) and FATS (Foreign Affiliates Statistics). The AFA database presents detailed data on the performance of foreign affiliates in the *manufacturing* industry for 28 OECD countries . The FATS database gives detailed data on the activities of foreign affiliates in the *services* sector for 25 OECD countries. These databases can all be accessed using the OECD portal http://www.oecd.org/sti/ind/amne.htm.

The AMNE database, as well as the AFA and FATS databases, are based on data reported to the OECD the Eurostat in form of annual surveys on the activities of foreign-controlled

enterprises and foreign affiliates abroad controlled by residents of the compiling country. Surveys are conducted in most cases by the national statistical office of each country or the central bank. While the key variables in the survey are common across countries, the target sample varies across countries. In order to compare ORBIS-AMADEUS ownership data to the official OECD data, the following issues have to be considered:

- Both AMNE and ORBIS-AMADEUS report data at the firm level, while AFA-FATS, covering the pre-2008 period, report the key variable the "number of enterprises or establishments." Therefore, in certain countries, the number of foreign affiliates is not comparable with ORBIS-AMADEUS since establishments are not included in BvD database. All other key variables refer to a *firm* in all datasets.

- The notion of *Foreign Affiliate* is based on the concept of controlling interest. According to AFA-FATS, a single institutional unit (another corporation, a household, or a government body) secures control over a corporation by owning more than half the voting shares or otherwise controlling more than half the shareholder voting power. However, the definition of controlling interest varies across countries. In most countries, controlling interest is based on direct majority ownership (50%) while others (Hungary and US) also consider minority control (between 10% and 50%). Moreover, some countries include *indirectly* owned foreign affiliates in addition to the directly controlled ones. In Table 9.1, we provide a summary of the characteristics of the AMNE database by country, whether indirect foreign control is considered, multiple owners and the main data source. For comparison purposes, for ORBIS-AMADEUS, we provide statistics based on two definitions of foreign affiliates: firms with direct foreign ownership shares greater than 10 or 50 percent, respectively. Multiple owners are considered.

- In AMNE database, starting in 2008, the total economy is defined as sectors B–N at ISIC Rev.4. Prior to 2008, the total economy includes additional sectors. Therefore, to have a consistent series over time we sum only data corresponding to these sectors B–N. For the same reasons, in AMADEUS-ORBIS, we only sum sectors B–N to compute the total economy statistics.

- As shown in Table 9.1, in the OECD database, some countries do not sample all firms and surveyed firms depend on certain economic thresholds. In AMADEUS-ORBIS, we always use all information available, regardless of firm size.

Table 9.1: OECD AMNE database sources

Country Code	Indirect Ownership	Multiple Ownership	Main Sample Characteristics
AT	from 2007		The two major sources are the structural business statistics collected by Statistics Austria and the direct investment survey conducted by the OeNB (including a question on the ultimate parent company). The OeNB submits a list of all resident foreign-controlled enterprises to Statistics Austria. Next, Statistics Austria aligns this list of enterprises with the entries in the statistical business register, with a view to creating a data set that matches the scope of the relevant annual structural business statistics.
CZ	no	yes	Data are extracted from the annual structural survey (for non-financial enterprises) and from the annual survey in the financial sector. No special survey on enterprises with foreign participation is conducted, they are identified in the Business Register. The information on the share of foreign participation is updated by the Czech National Bank.
DK	yes	yes	From 2000 data, a fourth data source was added: the register of foreign-owned companies held by Kbmandstandens Oplysningsbureau (KOB). This new data source allowed the identification of small foreign-owned companies.
EE	yes		20+ employees
HU	yes	yes	Survey: sorted separately those companies (NACE Rev. 1. A-O) with 10-49% of foreign direct investment, companies with 50-50 % of foreign direct investment and companies with more than 50% of foreign direct investment from the FDI register.
FI	yes	no	The FDI survey data from the Bank of Finland provides information on direct foreign owners. This information is obtained by a universe inquiry at a five years interval. It is supplemented every year with a limited survey. Other sources are also used to update this information (annual reports of enterprises, information on corporate acquisitions). The Enterprise Group Register is used to identify indirectly foreign-owned enterprises. It provides information on ownership relations between enterprises belonging to a group. The size threshold is approximately 60 persons employed in a group.
FR	yes		Up to 2001, source data are derived from the annual business survey (Enqute Annuelle d'Entreprise- EAE) complemented with results from the Institut National de la Statistique et des tudes conomiques (INSEE) survey on financial ties (LiFi). From 1999, information from the Diane database (coedition Bureau Van Dijk Electronic Editions / Coface SCRL) has been added. The data for food industries come from the annual business survey from the SCEES, Ministry of Agriculture. They have been included from 1999 onwards.
DE	from 2002		From 2002 to 2006, the submission of reports is required of every German enterprise with a balance sheet total of more than EUR 3 million; from 1999 until 2001, enterprises in Germany with a balance sheet total of more than EUR 500 000 were covered in which a non-resident (or several economically linked non-residents) holds 50% or more of the shares or voting rights of the German enterprise; reports were also required of German enterprises with a balance sheet total of more than EUR 5 million in which a non-resident (or several economically linked non-residents) held at least 10% but less than 50% of the shares or voting rights in the German enterprise concerned. Prior to 1999, the enterprises covered were those with foreign participating interests of more than 20% and with a balance sheet total exceeding EUR 500 000.

Continued on next page

Table 9.1 – continued from previous page

Country Code	Indirect Ownership	Multiple Ownership	Main Sample Characteristics
GB			NOTE Information directly from UK statistical office, OECD does not provide info: The surveys are run from a register, which is compiled primarily from administrative information such as VAT details from HM Customs and Excise and PAYE from the Inland Revenue. The register holds a number of variables including information on the country of ownership for each group. It also holds information on which UK groups have foreign affiliates.The main source of information on these foreign links for the latest annual survey was a Dun and Bradstreet publication. This was supplemented with information from ONS surveys into acquisitions and mergers of companies. These surveys are conducted on a continuous basis,collecting information when a UK company acquires or disposes of a foreign company and similarly when a foreign company acquires or disposes of a UK company. Work is currently being undertaken to review the register sources used for the survey to ensure completeness. In particular, Dun and Bradstreet's 'WorldBase' information has been used to give a better estimate of the population of companies with foreign links. For both populations, in order to maximise the survey coverage of foreign direct investment assets, all groups in the top strata (containing the largest businesses) are sent questionnaires. However, in strata containing smaller businesses only a proportion are selected. Additionally, the sample of smaller businesses is rotated to minimise burden on the respondents.
ES	yes		
IT	yes		The average response rate to the survey exceeds 50% in terms of firms and 65% in terms of persons employed. After manual controls on non respondents, the response rate raises up to 80% in terms of persons employed.
PL	yes	considered	The data come from the Enterprise Department of the Central Statistical Office of Poland (CSO). They are collected via the annual statistical survey on entities with foreign capital. When there is a case of non-response of a huge company (in terms of size class or capital value), data are imputed by using data from other surveys and ownership structure is taken from known sources (previous years' questionnaires, media etc.).
PT	yes		Structural Business Statistics are the main data source (no separate data collection for statistics on foreign-controlled enterprises). In order to define which companies are controlled from abroad, administrative data are used, but it is not possible to know the last unit of institutional control. Thus, INE takes some steps to meet the last foreign unit that actually exercises control over the affiliates in Portugal, by phone contacts with the resident enterprise.
SE			The target population consists of all active foreign-controlled enterprises in Sweden and all foreign-controlled enterprises of major economic importance. The statistics cover all enterprises identified where more than 50% of the voting rights are controlled by foreign investors. The reporting unit is, in most cases, the enterprise. Statistics Sweden's Structural Business Statistics: all non-financial enterprises are surveyed annually. These statistics are based on data from annual reports, tax returns and questionnaires on revenues and costs, etc. Collection of these data is mandatory for all Swedish enterprises. Growth Analysis's survey on ownership: from reference year 1996, data on foreign-owned affiliates are based on annual questionnaires to all parent companies, subsidiaries as well as to all branches located in Sweden. The Growth Analysis register of international enterprises/groups and the business database and foreign trade statistics of Statistics Sweden are combined and merged annually.

Continued on next page

Table 9.1 – continued from previous page

Country Code	Indirect Ownership	Multiple Ownership	Main Sample Characteristics
SI	yes		Foreign direct investment data of the Bank of Slovenia are used to identify foreign affiliates. The Structural Business Statistics database is the source for foreign affiliates' variables.
SK	yes	considered	Data come from the annual structural business survey.
NL	no		The ultimate controlling unit is determined on an annual basis by combining enterprise information from various sources, notably two surveys conducted by Statistics Netherlands, the Financial Statistics of Large Enterprise Groups (SFGO) and the Community Innovation Survey (CIS). However, from 2006 onwards, the UCI list of enterprises in the General Business Register (GBR), which is based on SFGO and CIS information, has been completed by the addition of information from an external source (Dun & Bradstreet database), providing a better insight into the total share of foreign enterprises in the Netherlands. In addition, a redesign of the Social Statistical Database in that year has resulted in a better match of the key indicators on employment, via the unique enterprise identifier (BEID), with the concurrent economic indicators. These improvements mean that the locus of control can now be established for over 90 percent of the total population of enterprises (with registered jobs in the Social Statistical Database) in the Netherlands. A weighting procedure was developed for the remaining share of enterprises that could not be matched to the GBR.
NO	yes		Statistics Norway's Structural Business Statistics are the main data source (no separate data collection for statistics on foreign-controlled enterprises). In order to define which companies are controlled from abroad, the register of foreign assets and liabilities in Norway (the SIFON register) and the balance of payments reporting are used. The Directorate of Taxes' register of shareholders has been the main source for updating the SIFON register. The press and Internet etc. are also used to map new foreign-controlled enterprises in Norway, and to map the ultimate country of ownership.

Notes: Source: OECD AMNE-database documentation online and national statistical office when not available. AT – Austria, CZ – the Czech Republic, DE – Germany, EE – Estonia, ES – Spain, FI – Finland, FR – France, GB – the United Kingdom, GR – Greece, HU – Hungary, IT – Italy, NL – Netherlands, NO – Norway, PL – Poland, PT – Portugal, SI – Slovenia, SK – Slovakia, SE – Sweden.

Bearing all these caveats in mind, we proceed in the following tables to compare the data we compile based on ORBIS-AMADEUS to the data collected by the OECD in AMNE, FAS, and FATS. Missing ratios can either appear because there is no data available in Eurostat or there is no data on a particular sector (remember we only consider sectors B-N to define the total economy) and therefore, we would not have a comparable sample.

Tables 9.2 and 9.3 report the coverage of our dataset ORBIS-AMADEUS compared to the OECD data for the total economy while Tables 9.4 and 9.5 report the comparison figures only for the manufacturing sector. In all tables, panel A reports the ratio of the number of firms in our dataset relative to that in the OECD, Panel B reports the representativeness in terms of output, and Panel C in terms of number of employees.

From Tables 9.2 and 9.3, it is clear that ORBIS-AMADEUS have better coverage in countries such as the UK, Germany, Italy, Ireland and Poland, where we observe ratios bigger than 1. According to Table 9.1, these are countries in which the survey by the national statistical office is done on a sub-sample of firms based on some size threshold: Germany limits to firms with balance sheets of more than 3 million euro, the UK sends questionnaires to the largest businesses and subsamples smaller ones and the response rate in Italy is not 100%. Some methodological changes are apparent for example, in the case of Hungary after 2008. The share of number of foreign affiliates covered by BvD drastically drops. As we have outlined before, in the OECD data, starting in 2008 some countries report both the number of firms and establishments therefore, creating a discontinuity in the time series of this variable. From the online documentation provided by the OECD it is not possible to know which ones are these countries but Hungary seems a clear case. According to the OECD data the number of foreign firms in 2006 is 3,423 while in 2008 it is 18,698. Exploring the data by sector of economic activity the largest jumps are in wholesale (from 857 firms in 2006 to 6,436 in 2008) and real state (from 905 in 2006 to 4,148). Given these methodological changes we prefer to asses the coverage of the dataset using measures of output and employment.

Similar to the financial data module, in Tables 9.4 and 9.5, we provide separate coverage statistics for the manufacturing sector. The same patterns found in the total economy are evident in the manufacturing sector. In most countries our data covers more than 50 percent

of the multinational economic activity reported in the OECD and in many countries ORBIS-AMADEUS does better.

Table 9.2: Foreign Affiliate Coverage (EUROZONE, Total Economy)

	AT		DE		EE		ES		FI		FR		IE		IT		LU		NL		PT		SI		SK	
	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50
Panel A: Number of Firms																										
1999	.	.	.	.	.	.	.	.	.	.	1.04	0.95	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2000	.	.	.	.	.	.	.	.	.	.	1.14	1.04	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2001	.	.	.	.	.	.	.	.	.	.	.	.	0.62	0.55	0.77	0.59	.	.	.	.	.	.	.	.	.	.
2002	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.93	0.71	.	.	.	.	.	.	.	.	.	.
2003	1.31	1.14	.	.	.	.	.	.	0.87	0.77	.	.	0.87	0.77	0.80	0.62	.	.	1.42	1.36	0.89	0.73	.	.	1.36	.
2004	.	.	2.02	1.68	.	.	2.08	1.88	0.90	0.80	.	.	1.20	1.07	1.37	1.03	.	.	1.34	1.28	.	.	.	.	0.74	0.67
2005	.	.	3.65	3.04	.	.	1.48	1.34	0.85	0.75	0.60	0.54	3.13	2.80	1.44	1.08	.	.	0.95	0.91	0.92	0.79	.	.	0.99	0.89
2006	.	.	5.33	4.49	.	.	1.61	1.46	.	.	0.62	0.56	3.77	3.33	1.54	1.14	.	.	.	.	0.77	0.66	.	.	1.16	1.04
2007	1.14	0.99	9.34	7.95	.	.	0.02	.	.	.	0.70	0.63	4.07	3.59	2.26	1.65	.	.	.	.	.	.	.	.	1.37	.
2008	1.15	1.01	2.38	2.03	.	.	1.02	0.93	0.88	0.78	0.75	0.66	1.57	1.40	2.41	1.74	.	.	1.20	1.15	0.79	0.68	0.45	0.39	0.64	0.57
2009	1.05	0.92	2.19	1.86	.	.	0.83	0.75	0.76	0.68	0.74	0.65	1.42	1.26	2.15	1.55	0.58	0.52	1.12	1.08	0.69	0.59	0.43	0.38	0.63	0.57
2010	1.04	0.91	1.79	1.52	.	.	0.76	0.69	0.84	0.74	0.67	0.59	1.44	1.27	2.27	1.60	0.51	0.46	0.88	0.85	0.60	0.52	0.49	0.44	0.65	0.59
2011	1.02	0.89	1.64	1.39	.	.	0.71	0.65	0.81	0.71	0.61	0.53	1.23	1.09	2.17	1.54	0.39	0.35	0.72	0.70	0.46	0.41	0.25	0.22	0.59	0.54
Average	1.12	0.98	3.54	2.99	.	.	1.06	1.10	0.85	0.75	0.76	0.68	1.93	1.71	1.65	1.21	0.49	0.44	1.09	1.05	0.73	0.62	0.41	0.36	0.90	0.69
Panel B: Output																										
1999	.	.	.	.	.	.	.	.	.	.	0.54	0.50	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2000	.	.	.	.	.	.	.	.	.	.	0.62	0.56	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2001	.	.	.	.	.	.	.	.	.	.	.	.	7.84	7.66	0.65	0.56	.	.	.	.	.	.	.	.	.	.
2002	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.64	0.52	.	.	.	.	.	.	.	.	.	.
2003	0.69	0.55	.	.	.	.	0.79	0.71	0.44	0.38	.	.	6.00	5.78	0.67	0.56	.	.	2.65	2.19	0.64	0.49	.	.	.	.
2004	.	.	0.51	0.41	.	.	0.74	0.66	1.25	0.69	0.45	0.41	4.34	4.13	0.75	0.64	.	.	2.52	2.13	.	.	.	.	0.87	0.72
2005	.	.	0.55	0.44	.	.	0.77	0.69	1.25	0.69	0.44	0.40	2.29	2.09	0.81	0.67	.	.	1.52	1.25	1.49	1.07	.	.	0.93	0.76
2006	.	.	0.67	0.57	.	.	0.74	0.67	.	.	0.50	0.45	2.66	2.47	0.82	0.71	.	.	.	.	1.56	1.14	.	.	0.98	0.80
2007	0.82	0.57	1.31	1.17	.	.	0.40	0.36	.	.	0.48	0.44	2.58	2.32	0.74	0.64	.	.	.	.	.	.	.	.	.	.
2008	0.82	0.54	0.92	0.41	.	.	1.10	0.66	1.02	0.62	0.90	0.47	0.80	0.63	1.15	0.68	.	.	1.46	1.26	1.59	1.17	0.66	0.61	0.90	0.78
2009	0.84	0.51	1.34	0.51	.	.	1.10	0.64	1.68	0.47	0.97	0.48	0.65	0.47	1.08	0.60	1.29	1.13	1.25	0.99	1.57	1.08	0.79	0.74	0.70	0.55
2010	0.75	0.50	1.24	0.52	.	.	1.04	0.59	1.47	0.64	0.94	0.45	0.74	0.55	1.11	0.60	1.34	1.17	1.26	1.06	1.64	0.93	0.87	0.74	0.67	0.53
2011	0.84	0.54	0.97	0.50	.	.	0.95	0.52	1.18	0.66	0.97	0.49	0.82	0.55	1.22	0.65	1.05	0.88	1.13	0.92	1.63	0.68	0.81	0.70	0.67	0.53
Average	0.79	0.53	0.94	0.57	.	.	0.85	0.61	1.18	0.59	0.68	0.47	2.87	2.67	0.87	0.62	1.22	1.06	1.68	1.40	1.44	0.94	0.78	0.70	0.82	0.67
Panel C: Employment																										
1999	.	.	.	.	.	.	.	.	.	.	0.38	0.35	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2000	.	.	.	.	.	.	.	.	.	.	0.39	0.36	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2001	.	.	.	.	.	.	.	.	.	.	.	.	0.26	0.20	0.80	0.67	.	.	1.56	1.16	.	.	.	.	.	.
2002	.	.	.	.	.	.	.	.	.	.	.	.	.	.	0.75	0.61	.	.	1.59	1.30	.	.	.	.	.	.
2003	0.73	0.64	0.36	0.26	.	.	0.77	0.69	0.47	0.40	.	.	0.33	0.21	1.00	0.88	.	.	2.44	1.62	0.13	0.09	.	.	.	.
2004	.	.	0.47	0.36	.	.	0.69	0.61	1.56	0.76	0.39	0.36	0.14	0.03	0.70	0.60	.	.	2.03	1.37	.	.	.	.	0.63	0.56
2005	.	.	0.56	0.44	.	.	0.76	0.66	1.65	0.79	0.35	0.32	0.23	0.10	0.65	0.56	.	.	1.29	1.02	0.34	0.15	.	.	0.87	0.79
2006	.	.	0.66	0.53	.	.	0.70	0.61	.	.	0.37	0.34	0.52	0.38	0.91	0.82	.	.	.	.	0.90	0.66	.	.	0.94	0.87
2007	0.88	0.50	1.28	1.06	.	.	0.19	0.17	.	.	0.39	0.35	0.85	0.67	0.65	0.56	.	.	.	.	.	.	.	.	.	.
2008	0.85	0.48	1.14	0.47	.	.	0.86	0.54	1.07	0.50	1.12	0.38	0.91	0.71	1.05	0.62	.	.	0.90	0.59	0.91	0.76	0.75	0.65	0.73	0.67
2009	0.82	0.38	1.48	0.56	.	.	0.95	0.53	1.69	0.39	1.09	0.27	1.09	0.72	1.04	0.59	1.47	0.98	0.80	0.54	1.10	0.71	0.81	0.74	0.74	0.67
2010	0.66	0.40	1.31	0.49	.	.	0.85	0.47	1.62	0.51	1.12	0.28	0.97	0.67	1.07	0.57	1.20	0.73	0.69	0.52	1.08	0.64	0.93	0.72	0.78	0.71
2011	0.72	0.40	0.97	0.44	.	.	0.82	0.45	1.22	0.45	1.07	0.24	1.26	0.69	1.15	0.67	1.43	0.94	0.76	0.55	1.11	0.53	0.86	0.68	0.63	0.59
Average	0.78	0.47	0.91	0.51	.	.	0.73	0.53	1.32	0.54	0.67	0.32	0.66	0.44	0.89	0.65	1.37	0.88	1.34	0.96	0.80	0.51	0.84	0.70	0.76	0.69

Notes: The table presents the ratio of outcome reported in BvD to that in OECD data. *FO*10 refers to companies with more than 10% foreign ownership. *FO*50 refers to companies with more than 50% foreign ownership. Panel A reports the number of firms, Panel B shows percentages in terms of output and finally, Panel C shows percentages in terms of employment. Prior to 2008, OECD reports the number of employees for all countries. After 2008 in the following countries only data on number of persons employed is available: Germany, Spain, France, Ireland, Luxembourg, Norway and Poland. The country codes correspond to the following countries: AT (Austria), BE (Belgium), EE (Estonia), FI (Finland), FR (France), DE (Germany), GR (Greece), IE (Ireland), LU (Luxembourg), NL (Netherland), PT (Portugal), SK (Slovakia), SI (Slovenia), ES (Spain), IT (Italy).

Table 9.3: Foreign Affiliate Coverage (NON-EUROZONE, Total Economy)

	GB		CZ		DK		HU		NO		PL		SE	
	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50
Panel A: Number of Firms														
1999	.	.	.	.	.	.	0.76	0.59	.	.	.	.	.	.
2000	.	.	.	.	.	.	0.75	0.58	0.80	0.66	.	.	.	.
2001	.	.	.	.	.	.	0.47	0.36	0.82	0.67	.	.	.	.
2002	.	.	.	.	.	.	0.82	0.66	0.85	0.68	.	.	0.28	0.26
2003	.	.	.	.	.	.	1.37	1.13	0.82	0.65	.	.	0.25	0.24
2004	1.68	1.55	0.15	0.14	.	.	1.68	1.34	0.88	0.70	1.08	0.98	0.26	0.24
2005	2.86	2.47	0.15	0.14	.	.	1.70	1.37	0.87	0.68	1.24	1.13	0.26	0.24
2006	3.14	2.63	0.29	0.27	.	.	0.31	0.26	0.93	0.71	1.52	1.37	0.25	0.23
2007	.	.	0.45	0.42	.	.	0.40	0.35	1.07	0.80	.	.	0.24	0.22
2008	2.58	2.19	0.53	0.50	0.87	0.77	0.05	0.04	0.96	0.71	1.86	1.71	0.22	0.20
2009	2.39	2.01	0.40	0.37	0.85	0.75	0.06	0.05	1.05	0.76	2.33	2.14	0.21	0.19
2010	2.34	1.96	0.39	0.36	0.83	0.74	0.06	0.05	0.90	0.66	1.95	1.79	0.21	0.19
2011	2.33	1.95	0.53	0.49	0.74	0.66	0.06	0.05	0.87	0.64	1.77	1.63	0.19	0.18
Average	2.48	2.11	0.36	0.34	0.82	0.73	0.65	0.53	0.90	0.69	1.68	1.54	0.24	0.22
Panel B: Output														
1999	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2000	.	.	.	.	.	.	0.97	0.70	0.58	0.52	.	.	.	.
2001	.	.	.	.	.	.	0.78	0.56	0.56	0.51	.	.	.	.
2002	.	.	.	.	.	.	0.82	0.70	0.46	0.41	.	.	0.29	0.27
2003	.	.	.	.	.	.	4.79	4.73	0.50	0.43	.	.	0.34	0.32
2004	0.90	0.50	0.87	0.81	.	.	0.49	0.32	1.23	0.61	0.94	0.82	0.81	0.36
2005	0.85	0.50	0.92	0.85	.	.	0.65	0.46	1.38	0.71	0.95	0.83	0.83	0.38
2006	0.89	0.53	0.87	0.82	.	.	0.49	0.34	1.40	0.65	1.00	0.91	0.78	0.35
2007	.	.	0.81	0.77	.	.	0.92	0.69	1.40	0.66	2.09	1.91	0.81	0.36
2008	1.39	0.57	0.86	0.82	0.68	0.46	0.53	0.40	1.45	0.71	1.01	0.94	0.77	0.32
2009	1.23	0.57	0.80	0.77	0.50	0.39	0.40	0.38	1.43	0.60	0.96	0.89	0.82	0.46
2010	1.31	0.58	0.81	0.77	0.47	0.33	0.49	0.46	1.51	0.60	0.90	0.85	1.04	0.45
2011	1.31	0.58	0.76	0.73	0.55	0.36	0.48	0.32	1.64	0.62	0.90	0.86	0.95	0.33
Average	1.13	0.55	0.84	0.79	0.55	0.39	0.98	0.84	1.13	0.59	1.09	1.00	0.74	0.36
Panel C: Employment														
1999	.	.	.	.	.	.	.	.	.	.	.	.	.	.
2000	.	.	.	.	.	.	0.05	0.03	0.55	0.51	.	.	.	.
2001	.	.	.	.	.	.	0.06	0.03	0.59	0.55	.	.	.	.
2002	.	.	.	.	.	.	0.25	0.18	0.78	0.74	.	.	0.32	0.30
2003	.	.	.	.	.	.	0.10	0.07	0.75	0.66	.	.	0.33	0.30
2004	0.90	0.57	0.95	0.88	.	.	0.20	0.14	1.32	0.92	0.84	0.72	1.38	0.39
2005	0.90	0.56	0.94	0.88	.	.	0.28	0.21	1.17	0.61	0.82	0.72	1.48	0.43
2006	0.96	0.59	0.80	0.75	.	.	0.25	0.19	1.00	0.50	0.81	0.71	1.42	0.39
2007	.	.	0.72	0.69	.	.	0.44	0.36	1.11	0.56	1.35	1.22	1.45	0.41
2008	.	.	0.68	0.64	0.75	0.46	0.32	0.26	1.07	0.60	0.78	0.72	1.35	0.38
2009	1.34	0.65	0.82	0.77	0.53	0.42	0.32	0.30	1.13	0.52	0.92	0.84	1.08	0.52
2010	1.18	0.56	0.83	0.78	0.58	0.36	0.34	0.32	1.15	0.45	0.66	0.61	1.36	0.49
2011	.	.	0.84	0.80	0.57	0.32	0.32	0.25	1.15	0.44	0.49	0.46	1.28	0.32
Average	1.06	0.59	0.82	0.78	0.61	0.39	0.24	0.19	0.98	0.59	0.83	0.75	1.14	0.39

Notes: The table presents the ratio of outcome reported in BvD to that in OECD data. $FO10$ refers to companies with more than 10% foreign ownership. $FO50$ refers to companies with more than 50% foreign ownership. Panel A reports the number of firms, Panel B shows percentages in terms of output and finally, Panel C shows percentages in terms of employment. Prior to 2008, OECD reports the number of employees for all countries. After 2008 in the following countries only data on number of persons employed is available: Germany, Spain, France, Ireland, Luxembourg, Norway and Poland. The country codes correspond to the following countries: CH (Switzerland), NO (Norway), CZ (Czech Republic), DK (Denmark), HU (Hungary), RO (Romania), SE (Sweden), GB (Great Britain), PL (Poland).

Table 9.4: Foreign Affiliate Coverage (EUROZONE, Manufacturing)

	AT		DE		EE		ES		FI		FR		IE		IT		LU		NL		PT		SI		SK	
	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50
Panel A: Number of Firms																										
1999	.	.	1.62	1.33	.	.	1.45	1.16	0.61	0.48	0.60	0.55	0.13	0.11	.	.	0.37	0.31	0.64	0.61	0.81	0.65	.	.	.	.
2000	.	.	1.51	1.22	.	.	1.44	1.16	0.65	0.51	0.64	0.59	0.21	0.19	.	.	0.35	0.30	0.72	0.68	0.90	0.72	.	.	.	.
2001	.	.	3.25	2.61	.	.	1.31	1.07	0.63	0.50	0.66	0.60	0.19	0.18	0.92	0.70	0.84	0.67	0.75	0.71	1.39	1.14	.	.	.	.
2002	.	.	3.89	3.22	.	.	1.03	0.87	0.83	0.68	0.53	0.49	0.23	0.21	1.08	0.81	1.05	0.89	0.72	0.68	1.49	1.19	.	.	.	.
2003	0.89	0.76	.	.	2.33	1.94	1.06	0.90	0.78	0.64	0.51	0.47	0.25	0.22	0.98	0.74	1.54	1.31	0.75	0.71	0.63	0.50	0.68	0.49	0.72	0.65
2004	.	.	5.47	4.43	2.35	1.97	0.95	0.83	0.97	0.79	0.48	0.44	0.31	0.28	1.23	0.91	1.71	1.42	0.68	0.65	0.72	0.58	0.65	0.48	0.62	0.55
2005	.	.	8.26	6.75	2.44	2.05	0.87	0.76	0.95	0.77	0.46	0.42	0.68	0.63	1.23	0.91	2.03	1.71	0.65	0.62	0.75	0.61	0.70	0.52	0.74	0.66
2006	.	.	10.48	8.61	2.40	1.98	0.93	0.81	0.87	0.70	0.49	0.44	0.87	0.79	1.39	1.02	.	.	.	.	0.70	0.57	0.71	0.54	0.84	0.75
2007	0.88	0.75	10.12	8.39	2.26	1.87	0.89	0.77	0.82	0.67	0.53	0.47	1.01	0.91	1.63	1.18	.	.	0.71	0.69	0.68	0.56	0.71	0.55	0.74	0.66
2008	0.99	0.85	1.31	1.08	2.18	1.84	0.84	0.73	0.90	0.74	0.57	0.50	1.15	1.03	1.89	1.35	.	.	0.66	0.64	0.77	0.65	0.71	0.55	0.59	0.53
2009	0.90	0.77	1.34	1.11	2.28	1.96	0.87	0.75	0.79	0.64	0.58	0.51	1.16	1.04	1.70	1.21	0.73	0.65	0.61	0.58	0.74	0.62	0.66	0.54	0.63	0.57
2010	0.88	0.75	1.15	0.94	2.44	2.08	0.75	0.65	0.96	0.76	0.57	0.49	1.10	0.99	1.82	1.26	0.67	0.59	0.53	0.51	0.62	0.52	0.68	0.55	0.62	0.56
2011	0.90	0.77	1.08	0.89	2.23	1.90	0.68	0.59	0.93	0.75	0.56	0.49	1.01	0.91	1.75	1.22	0.59	0.52	0.49	0.46	0.56	0.48	0.43	0.35	0.60	0.54
Average	0.91	0.78	4.12	3.38	2.32	1.95	1.00	0.85	0.82	0.67	0.55	0.50	0.64	0.58	1.42	1.03	0.99	0.84	0.66	0.63	0.83	0.68	0.66	0.51	0.68	0.61
Panel B: Output																										
1999	.	.	0.39	0.32	.	.	0.76	0.60	0.19	0.18	0.37	0.34	.	.	.	.	0.42	0.38	1.27	1.07	0.95	0.82	.	.	.	.
2000	.	.	0.46	0.40	.	.	0.78	0.61	0.22	0.20	0.42	0.39	.	.	.	.	0.38	0.33	2.09	1.83	0.93	0.79	.	.	.	.
2001	.	.	1.42	1.33	.	.	0.79	0.64	0.26	0.24	0.42	0.38	.	.	0.62	0.51	0.72	0.65	1.67	1.48	1.01	0.87	.	.	.	.
2002	.	.	1.11	0.92	.	.	0.66	0.60	0.35	0.34	0.40	0.36	.	.	0.66	0.54	0.61	0.53	1.42	1.26	0.88	0.69	.	.	.	.
2003	0.57	0.55	.	.	1.11	0.95	0.64	0.58	0.46	0.37	0.38	0.35	.	.	0.67	0.54	1.06	1.00	1.28	1.13	0.59	0.49	0.94	0.71	0.86	0.74
2004	.	.	1.61	1.27	1.70	1.36	0.64	0.59	2.63	1.21	0.32	0.30	.	.	0.80	0.68	10.16	9.94	1.15	1.01	0.99	0.66	0.94	0.75	0.81	0.65
2005	.	.	1.54	1.17	1.85	1.47	0.65	0.59	2.37	1.06	0.31	0.29	.	.	0.93	0.81	12.14	11.89	1.03	0.91	1.00	0.63	0.94	0.78	0.92	0.72
2006	.	.	1.88	1.54	2.06	1.62	0.71	0.66	2.22	0.99	0.38	0.34	.	.	0.85	0.73	.	.	.	.	0.95	0.60	0.94	0.81	0.92	0.73
2007	0.55	0.48	1.95	1.79	2.05	1.61	0.71	0.65	1.92	0.85	0.39	0.35	.	.	0.83	0.73	.	.	0.97	0.89	0.84	0.57	0.85	0.75	0.85	0.68
2008	0.55	0.44	1.00	0.38	1.92	1.74	0.86	0.70	1.76	0.86	0.77	0.37	0.44	0.39	1.18	0.84	.	.	0.96	0.88	1.03	0.69	0.82	0.73	0.77	0.61
2009	0.53	0.36	1.39	0.38	1.05	0.98	0.94	0.75	3.74	0.49	0.83	0.38	0.53	0.40	1.13	0.77	.	.	0.84	0.77	0.94	0.66	0.81	0.76	0.74	0.58
2010	0.50	0.36	1.28	0.45	1.08	1.02	0.77	0.61	2.99	0.88	0.87	0.41	0.52	0.46	1.05	0.66	2.75	2.35	0.88	0.79	0.99	0.76	1.01	0.76	0.69	0.55
2011	0.47	0.36	1.11	0.42	0.25	0.23	0.61	0.45	2.08	0.89	0.86	0.38	0.75	0.47	1.18	0.76	2.79	2.46	0.81	0.72	0.97	0.75	0.95	0.70	0.76	0.61
Average	0.53	0.42	1.26	0.87	1.45	1.22	0.73	0.62	1.63	0.66	0.52	0.36	0.56	0.43	0.90	0.69	3.45	3.28	1.20	1.06	0.93	0.69	0.91	0.75	0.81	0.65
Panel C: Employment																										
1999	.	.	1.28	0.97	.	.	0.76	0.63	0.19	0.17	0.31	0.30	0.07	0.06	.	.	0.38	0.35	1.58	1.27	0.19	0.15	.	.	.	.
2000	.	.	1.18	0.93	.	.	0.78	0.64	0.20	0.18	0.32	0.30	0.24	0.17	.	.	0.42	0.38	1.79	1.43	0.24	0.20	.	.	.	.
2001	.	.	3.44	2.85	.	.	0.81	0.67	0.26	0.24	0.34	0.32	0.25	0.17	0.80	0.66	0.91	0.88	1.53	1.20	0.21	0.17	.	.	.	.
2002	.	.	3.38	2.87	.	.	0.68	0.61	0.35	0.32	0.39	0.36	0.28	0.19	0.78	0.63	0.60	0.55	1.79	1.54	0.25	0.20	.	.	.	.
2003	0.55	0.53	.	.	1.07	0.92	0.69	0.61	0.42	0.34	0.40	0.37	0.29	0.17	1.22	1.08	0.80	0.80	1.93	1.68	0.16	0.10	2.18	1.93	0.56	0.51
2004	.	.	5.23	3.97	1.10	0.93	0.66	0.60	2.84	1.19	0.37	0.34	0.14	0.03	0.71	0.60	8.58	8.27	1.71	1.45	0.22	0.12	0.91	0.69	0.52	0.47
2005	.	.	5.68	4.43	1.12	0.96	0.67	0.60	2.56	1.09	0.35	0.32	0.17	0.05	0.78	0.67	9.14	8.79	1.76	1.49	0.20	0.13	0.89	0.67	0.81	0.75
2006	.	.	6.91	5.55	1.15	0.94	0.69	0.63	2.49	1.06	0.38	0.34	0.38	0.26	0.79	0.68	.	.	.	.	0.68	0.54	0.92	0.73	0.89	0.84
2007	0.69	0.52	7.00	5.79	1.14	0.91	0.72	0.63	2.27	0.96	0.39	0.34	0.74	0.59	0.75	0.64	.	.	0.94	0.76	0.67	0.58	0.81	0.66	0.81	0.76
2008	0.65	0.47	.	.	1.10	0.90	1.02	0.67	1.97	0.84	.	.	.	.	1.18	0.72	.	.	0.70	0.60	0.83	0.68	0.80	0.66	0.76	0.71
2009	0.62	0.37	.	.	.	.	.	.	3.59	0.50	.	.	.	.	1.17	0.69	.	.	0.74	0.56	0.79	0.68	0.82	0.72	0.81	0.76
2010	0.62	0.37	.	.	1.18	1.09	.	.	3.28	0.71	.	.	.	.	1.15	0.64	.	.	0.54	0.37	0.80	0.69	1.08	0.71	0.80	0.75
2011	0.56	0.38	.	.	1.08	0.98	.	.	2.18	0.55	.	.	.	.	1.40	0.87	.	.	0.75	0.57	0.83	0.69	1.00	0.68	0.68	0.64
Average	0.62	0.44	4.26	3.42	1.12	0.95	0.75	0.63	1.74	0.63	0.36	0.33	0.28	0.19	0.97	0.72	2.98	2.86	1.31	1.07	0.47	0.38	1.05	0.83	0.74	0.69

Notes: The table presents the ratio of outcome reported in BvD to that in OECD data. *FO*10 refers to companies with more than 10% foreign ownership. *FO*50 refers to companies with more than 50% foreign ownership. Panel A reports the number of firms, Panel B shows percentages in terms of output and finally, Panel C shows percentages in terms of employment. Prior to 2008, OECD reports the number of employees for all countries. After 2008 in the following countries only data on number of persons employed is available: Germany, Spain, France, Ireland, Luxembourg, Norway and Poland. The country codes correspond to the following countries: AT (Austria), BE (Belgium), EE (Estonia), FI (Finland), FR (France), DE (Germany), GR (Greece), IE (Ireland), LU (Luxembourg), NL (Netherland), PT (Portugal), SK (Slovakia), SI (Slovenia), ES (Spain), IT (Italy).

Table 9.5: Foreign Affiliate Coverage (NON-EUROZONE, Manufacturing)

	GB		CZ		DK		HU		NO		PL		SE	
	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50	FO10	FO50
Panel A: Number of Firms														
1999	0.96	0.93	0.21	0.18	.	.	0.22	0.17	0.37	0.29	0.48	0.43	0.30	0.29
2000	1.03	1.00	0.13	0.11	.	.	0.22	0.17	0.55	0.44	0.57	0.50	0.30	0.29
2001	0.81	0.78	0.16	0.15	.	.	0.17	0.13	0.52	0.41	0.65	0.58	0.25	0.24
2002	0.84	0.81	0.21	0.19	.	.	0.26	0.21	0.52	0.41	0.72	0.64	0.26	0.25
2003	0.85	0.82	0.24	0.22	.	.	0.91	0.75	0.52	0.40	0.71	0.63	0.21	0.19
2004	0.97	0.91	0.24	0.22	.	.	1.13	0.90	0.56	0.42	0.83	0.75	0.22	0.19
2005	1.35	1.17	0.24	0.22	.	.	1.12	0.89	0.57	0.43	0.90	0.81	0.22	0.19
2006	1.48	1.24	0.26	0.24	.	.	0.28	0.24	0.56	0.42	0.98	0.87	0.22	0.19
2007	1.52	1.28	0.25	0.23	.	.	0.10	0.09	0.71	0.51	1.03	0.93	0.21	0.18
2008	1.55	1.31	0.49	0.46	0.68	0.59	0.09	0.08	0.83	0.57	1.22	1.11	0.23	0.20
2009	1.49	1.26	0.44	0.41	0.64	0.55	0.11	0.09	0.93	0.63	1.37	1.25	0.23	0.20
2010	1.53	1.30	0.43	0.40	0.65	0.55	0.11	0.10	0.87	0.60	1.22	1.11	0.23	0.20
2011	1.56	1.32	0.46	0.42	0.65	0.54	0.11	0.10	0.85	0.60	1.14	1.04	0.24	0.20
Average	1.23	1.09	0.29	0.27	0.66	0.56	0.37	0.30	0.64	0.47	0.91	0.82	0.24	0.22
Panel B: Output														
1999	0.74	0.47	0.64	0.49	.	.	.	.	0.34	0.28	0.61	0.57	0.19	0.19
2000	0.94	0.56	0.58	0.46	.	.	.	.	0.43	0.36	0.66	0.62	0.24	0.24
2001	0.83	0.49	0.60	0.50	.	.	.	.	0.39	0.34	0.73	0.69	0.24	0.23
2002	0.77	0.46	0.67	0.59	.	.	.	.	0.32	0.27	0.76	0.72	0.23	0.22
2003	0.66	0.37	0.67	0.60	.	.	0.41	0.38	0.47	0.28	0.73	0.70	0.25	0.24
2004	0.64	0.37	0.81	0.76	.	.	0.23	0.20	1.67	0.89	0.75	0.72	0.89	0.22
2005	0.61	0.32	0.90	0.85	.	.	0.48	0.43	1.77	0.89	0.81	0.77	0.96	0.24
2006	0.58	0.30	0.86	0.81	.	.	0.50	0.46	1.54	0.77	0.86	0.83	0.95	0.26
2007	0.59	0.31	0.80	0.77	.	.	0.41	0.39	1.69	0.84	0.87	0.84	1.08	0.30
2008	2.00	0.36	0.88	0.84	1.01	0.65	0.39	0.36	1.75	1.00	0.91	0.87	1.16	0.31
2009	1.68	0.34	0.74	0.70	0.55	0.41	0.37	0.34	1.93	0.48	0.88	0.84	1.43	0.76
2010	1.65	0.33	0.76	0.72	0.67	0.32	0.39	0.37	1.85	0.43	0.83	0.81	1.65	0.62
2011	1.70	0.33	0.72	0.70	0.56	0.22	0.34	0.32	1.96	0.43	0.84	0.81	1.54	0.31
Average	1.03	0.39	0.74	0.68	0.70	0.40	0.39	0.36	1.24	0.56	0.79	0.75	0.83	0.32
Panel C: Employment														
1999	1.09	0.64	0.89	0.76	.	.	.	.	0.27	0.25	0.65	0.59	0.24	0.24
2000	1.32	0.74	0.77	0.67	.	.	.	.	0.37	0.34	0.68	0.63	0.27	0.26
2001	1.00	0.57	0.81	0.73	.	.	.	.	0.31	0.28	0.70	0.65	0.26	0.25
2002	1.00	0.55	0.86	0.78	.	.	.	.	0.30	0.27	0.72	0.66	0.24	0.24
2003	0.83	0.44	0.69	0.63	.	.	0.07	0.06	0.47	0.27	0.74	0.69	0.26	0.25
2004	0.81	0.42	0.75	0.69	.	.	0.12	0.09	1.29	0.54	0.72	0.67	1.17	0.24
2005	0.90	0.44	0.77	0.73	.	.	0.22	0.18	1.93	0.58	0.72	0.67	1.17	0.25
2006	0.89	0.41	0.76	0.71	.	.	0.25	0.22	1.65	0.57	0.73	0.68	1.19	0.25
2007	1.01	0.46	0.67	0.62	.	.	0.24	0.22	2.04	0.72	0.70	0.66	1.29	0.28
2008	.	.	0.67	0.63	1.20	0.51	0.27	0.22	.	.	.	.	1.35	0.29
2009	1.48	0.46	0.79	0.74	0.71	0.51	0.27	0.23	.	.	.	.	1.82	0.77
2010	1.32	0.45	0.80	0.75	1.07	0.40	0.29	0.25	.	.	.	.	2.38	0.72
2011	.	.	0.79	0.76	0.81	0.28	0.28	0.25	.	.	.	.	2.35	0.27
Average	1.06	0.51	0.77	0.71	0.95	0.42	0.22	0.19	0.96	0.42	0.71	0.66	1.08	0.33

Notes: The table presents the ratio of outcome reported in BvD to that in OECD data. $FO10$ refers to companies with more than 10% foreign ownership. $FO50$ refers to companies with more than 50% foreign ownership. Panel A reports the number of firms, Panel B shows percentages in terms of output and finally, Panel C shows percentages in terms of employment. Prior to 2008, OECD reports the number of employees for all countries. After 2008 in the following countries only data on number of persons employed is available: Germany, Spain, France, Ireland, Luxembourg, Norway and Poland. The country codes correspond to the following countries: CH (Switzerland), NO (Norway), CZ (Czech Republic), DK (Denmark), HU (Hungary), RO (Romania), SE (Sweden), GB (Great Britain), PL (Poland).

Part V

Appendices

Appendix A

Filing Requirements and Sector Correspondence

Table A.1: BvD Company Filing Requirements and Data Providers for Selected Countries

Country Code	Which companies have to file accounts?	How many companies does that represent?	Data Provider
AT	AG, starting 1994 also GmbH and very large companies. Based on their size, companies may file shortened balance sheet and no PL account.	50,000	Creditreform
BE	Depends on the legal form: • Companies that must file their accounts are: SA; SPRL; SCRL (socit cooprative responsabilit limite); SE (Socit europenne); GEIE (Groupement europen dintrt conomique); GIE (Groupement dintrt conomique); Foreign companies located in Belgium. • Companies that have to file their accounts under certain conditions are: SCS (socit en commandite simple) if the company is large and one of the associates is an individual; SCRI (socit cooprative responsabilit illimite) if the company is large and one of the associates is an individual; SNC (socit en nom collectif) if the company is large and one of the associates is an individual; ASBL and Foundations if they are large or very large; Other (there are some other specific cases).	420,000	National Bank of Belgium, Coface Services Belgium
BG	All companies, which match 2 of the following 3 criteria: at least 50 persons staff, total assets at least eur 500.000, turnover at least eur 1.000.000	Less 10% of all active companies	Creditreform
HR	Private and public limited liability companies, general and limited partnerships, cooperatives have to file accounts to the State Authorities (State Register of accounts, established 2003).	Approximately 100,000 legal subjects filed their accounts for 2012, although there are more than 300,000 registered subjects	Creditreform
CY	All Cypriot Companies, whether local or international, must maintain accurate books of accounts, which should reflect the true and correct position of their conduct, as well as give adequate explanation of their operations. Audited financial statements and an Income Tax Return are required for all companies, even companies with no taxable income and/or dormant companies. Registered Branches (in Cyprus) of foreign companies are not legally bound to compile full separate branch accounts however when taxed on the island, are obliged to do so for income tax purposes. The following types of companies are obliged to file their financial statements: 1. *Limited Liability Companies.* They are obliged to submit a copy of their annual report, including their audited, financial statements for the year 2. *Public Companies.* They are obliged to submit a copy of their annual report, including their audited, financial statements for the year. In addition, public companies that are listed in the Cyprus Stock Exchange are obliged to publish their quarterly financial statements as well. Partnerships are exempt from any requirement to prepare audited accounts, but they are legally bound to keep proper books of account which must be available for scrutiny by individual partners.	This represents all the active companies operating in Cyprus (approximately 90.000 companies). However, this does not include the International Business Companies (IBCs) which are registered in Cyprus for taxation purposes and are actually operating abroad as: 1. There is no clear indication of which companies are actually IBC companies. 2. The vast majority of these companies avoid submitting their annual financial statements	Infocredit Group
CZ	All companies filled in Business Register. There is duty to fill full financial statements for companies with obligatory audit. Other companies in BR have duty to fill shortened FS. Obligatory audit refer joint stock companies (A.S.) and Limited liab. companies (S.R.O.) od cooperatives with (1) assets over 40 mil. CZK, (2) turnover 80 mil. CZK, (3), over 50 empl. For A.S. if is fulfilled any condition, for other companies fulfilled two of three conditions.	330,000 (many companies could be inactive, only registered with no financial statement). There is currently 330.000 companies in BR, 10% in liquidation or bankruptcy, and about 25% (80.000) companies registered but without turnover, not registered as VAT payers, etc. so suspicious.	Soliditet- main source is Business register where companies publish FS. They permanently go through the list of all companies in BR and seek newly published FS. This source is updated daily but some companies put documents into BR with big delays.
EE	Private limited companies, joint-stock companies, non-profitmaking associations, cooperative societies, general partnerships, limited partnerships, foundations.	C.a. 124,000 125,000	Krediidiinfo AS

Notes: Filing requirements were taken from Orbis Online Manual on February 3d, 2014.

Table A.1 (Cont'd.): Filing Requirements and Data Providers

Country Code	Which companies have to file accounts?	How many companies does that represent?	Data Provider
FI	All joint-stock companies and all co-operatives; - Limited partnerships, partnerships and private firms, which meet two of the following three conditions: • turnover over 7.30 million EUR; • balance sheet total over 3.65 million EUR; • number of personnel over 50.	The exact number is not known, but the estimate is approx. 120,000 companies	Suomen Asiakastieto Oy
FR	All of the following: • les socits responsabilit limite (SARL et EURL) ; • les socits de personnes (socits en nom collectif et socits en commandite simple), sous certaines conditions : les socits en nom collectif (SNC) dont au moins l'un des associs est une personne physique ne sont pas dans l'obligation de dposer leurs comptes annuels (pour plus de prcisions, se rfrer l'article L. 232-21 du Code de Commerce) ; • les socits par actions (socits anonymes, socits par actions simplifies et socits en commandite par actions) ; • les socits commerciales dont le sige est situ l'tranger qui ont ouvert un ou plusieurs tablissements en France ; • les socits d'exercice libral (SELARL, SELAFA, SELCA, SELAS) ; • les socits coopratives et unions sous certaines conditions (pour plus de prcisions, se rfrer l'article R. 524-22-1 du Code Rural).	1,400,000	Ellisphere
DE	Corporate enterprises (AG, GmbH) and cooperatives (e.G).	Breakdown: • *small cooperate enterprises:* approx. 980.000. Definition (267 HGB): staff: $\leq$ 50 individuals turnover: $\leq$ 9.680 TEUR total assets: $\leq$ 4.840 TEUR at least two criteria must apply They have to announce only the balance sheet information and the notes on the accounts. • *medium sized cooperate enterprises:* approx. 80.000. Definition (267 HGB): staff: between 50 and 250 individuals turnover: between 9.680 TEUR and 38.500 TEUR total assets: between 4.840 TEUR and 19.250 TEUR at least two criteria must apply They have to announce the balance sheet information as well as the statement of income and the notes on the accounts. • *big cooperate enterprises:* approx. 33.000. Definition (267 HGB): staff: more then 250 individuals turnover: more then 38.500 TEUR total assets: more then 19.250 TEUR at least two criteria must apply They have to announce the balance sheet information as well as the statement of income and the notes on the accounts. • *cooperatives:* approx. 7.500. They have to announce the balance sheet information as well as the statement of income and the notes on the accounts.	Creditreform and Creditreform Rating AG

Table A.1 (Cont'd.): Filing Requirements and Data Providers

Country Code	Which companies have to file accounts?	How many companies does that represent?	Data Provider
GB	Limited, PLC, LLP, LP.	1,000,000	Jordans Limited
GR	Societe Anonyme and Limited Liability Companies.	The Societe Anonyme and Limited Liability Companies that publish Balance Sheets represent approximately the 5% of the total active Business Universe in Greece.	ICAP
HU	All companies have to file accounts, except private enterprises. The companies have to send the accounts to the Ministry of Justice and to the Registry Court. The one-person firms and the limited deposit companies do not have to send it to the Ministry of Justice.	About 40%.	Creditreform
IS	Sameignarflag (fulfil two out of the three following prerequisites : total assets > 230 ML ISK, operating revenue > 460 ML ISK, average number of employees >50, or if the mother company has to file accounts), Samvinnufelag, Samlagsfelag (if mother company has to file accounts), Einkahlutaflag, Hlutaflag.	Approx. 20,000 companies were to file their accounts.	Icecredit Info.
IE	Limited.	100,000	Jordan Limited
IT	Includes: • S.p.A. (Societ per Azioni), • S.r.l. (Societ a responsabilit limitata), • Sapa (Societ in accomandita per azioni), • Societ Cooperative, • Societ Consortili, • G.e.i.e, Societ di persone (only consolidated accounts), • Consorzi con qualifica di Confidi. • Societ a responsabilit a socio unico e societ per azioni a socio unico.	Approximately 900,000	
LV	All companies, except sole proprietor enterprises, peasant farms and fishers farm whose annual turnover does not exceed LVL 200.000 (EUR 284.6 thousand).	Approximately 100,000.	Creditreform
LT	Includes: • Limited liability companies; • Joint stock companies; • State enterprises; • Municipal enterprises; • Agricultural companies; • Cooperative companies.	79,823 (excluding bankrupted, liquidating and inactive companies).	Creditreform
LU	Public (S.A.), limited (S.A.R.L.)	Not possible to find out : there are in total about 25,000 companies and 15,000 Holdings, total 40,000.	Creditreform

Table A.1 (Cont'd.): Filing Requirements and Data Providers

Country Code	Which companies have to file accounts?	How many companies does that represent?	Data Provider
NL	All limited companies (B.V.s and N.V.s) and some sole traders and cooperations.	680,000	LexisNexis Benelux, Graydon and Chambers of Commerce
NO	Limited.	Approximately 120,000.	Creditreform
PL	*Filing of the accounts:* All companies registered at the National Court Registry (KRS): joint-stock companies (S.A.), limited liability companies (Sp. z o.o.), cooperatives, state enterprises, etc., except for general partnership (sp.j.), professional partnership (sp.p.) that do not reach the annual turnover of 800,000 EUR. *Publishing of the accounts:* Joint-stock companies, banks, insurance companies, investment funds, plus others (limited liability companies, cooperatives, large private companies, etc.) complying with 2 of the following criteria: • average annual employment > 50 • total assets at the end of a financial year > 2.5 million euro • annual net profit > 5 million euro	About 200,000 companies. Among these companies, approx. 10,000 companies are also obliged to publish accounts.	InfoCredit collects information from National Court Registry, journals for entities and cooperatives, Judicial and Business Journal ("Monitor Sdowy i Gospodarczy", directly from the companies, as well as other alternative sources (if necessary).
PT	The Portuguese law compels all the companies to deposit the Balance Sheet. Therefore, about 57,500 balance sheets are public information.	These 57,500 balance sheets correspond to a small percentage of the totality of Portuguese companies, if we verify that the totality number of the universe has about 800.000 companies. That means that practically the companies do not deliver financial information.	Coface MOPE.
RO	Joint stock companies, partnerships limited by shares, limited liability companies, state owned concerns, co-operative companies.	500,000	Chamber of Commerce and Industry of Romania
RU	All juridical persons have to represent their accounts (individual entrepreneurs (manufacturers) and farms are not juridical persons)	Approximately 40% of all active companies file their accounts. So, if there are 1,500,000 registered active companies in Russia, the accounts are available for approximately 800,000 companies . Most of these are included in Ruslana.	
SK	All companies filed in Business Register.	About 70-80,000 (many companies could be inactive, only registered with no financial statement).Many companies do not file their accounts, there are no penalties for it.	Soliditet, s.r.o.
SI	All companies and sole proprietors	Around 160,000 companies.	Coface Slovenia
ES	S.A., S.L.	776,000	INFORMA
SE	Limited companies.	About 348,000.	UC AB
CH	There are no legal requirements to file their accounts in Switzerland (like in the UK). Public quoted companies make their financial statements available and all are collected, analysed and provided to the ORBIS database.	N/A	Worldbox AG
UA	All local companies and trade or representative offices of foreign companies have to file accounts (except companies mentioned in 3.) *Notes:* • In the database of the Central Statistical Department there are about 330,000 accounts of the biggest Ukrainian companies . • The remaining companies are split in 800 regional Departments of Statistic: there are currently no means by which we can get them for the time being.	There are more than 1,000,000 companies and trade representative offices registered.	Creditreform

Table A.2: NACE Revision 2, Level 2 Classification.

Code	Name of the Level 2 NACE sector
01	Crop and animal production, hunting and related service activities
02	Forestry and logging
03	Fishing and aquaculture
05	Mining of coal and lignite
06	Extraction of crude petroleum and natural gas
07	Mining of metal ores
08	Other mining and quarrying
09	Mining support service activities
10	Manufacture of food products
11	Manufacture of beverages
12	Manufacture of tobacco products
13	Manufacture of textiles
14	Manufacture of wearing apparel
15	Manufacture of leather and related products
16	Manufacture of wood and of products of wood and cork, except furniture, etc.
17	Manufacture of paper and paper products
18	Printing and reproduction of recorded media
19	Manufacture of coke and refined petroleum products
20	Manufacture of chemicals and chemical products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
22	Manufacture of rubber and plastic products
23	Manufacture of other non-metallic mineral products
24	Manufacture of basic metals
25	Manufacture of fabricated metal products, except machinery and equipment
26	Manufacture of computer, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment n.e.c.
29	Manufacture of motor vehicles, trailers and semi-trailers
30	Manufacture of other transport equipment
31	Manufacture of furniture
32	Other manufacturing
33	Repair and installation of machinery and equipment
35	Electricity, gas, steam and air conditioning supply
36	Water collection, treatment and supply
37	Sewerage
38	Waste collection, treatment and disposal activities; materials recovery
39	Remediation activities and other waste management services
41	Construction of buildings
42	Civil engineering
43	Specialised construction activities
45	Wholesale and retail trade and repair of motor vehicles and motorcycles
46	Wholesale trade, except of motor vehicles and motorcycles
47	Retail trade, except of motor vehicles and motorcycles
49	Land transport and transport via pipelines
50	Water transport
51	Air transport
52	Warehousing and support activities for transportation
53	Postal and courier activities
55	Accommodation
56	Food and beverage service activities
58	Publishing activities
59	Motion picture, video and television programme production, sound recording and music publishing
60	Programming and broadcasting activities
61	Telecommunications
62	Computer programming, consultancy and related activities
63	Information service activities
64	Financial service activities, except insurance and pension funding
65	Insurance, reinsurance and pension funding, except compulsory social security
66	Activities auxiliary to financial services and insurance activities
68	Real estate activities
69	Legal and accounting activities
70	Activities of head offices; management consultancy activities
71	Architectural and engineering activities; technical testing and analysis
72	Scientific research and development
73	Advertising and market research
74	Other professional, scientific and technical activities
75	Veterinary activities
77	Rental and leasing activities
78	Employment activities
79	Travel agency, tour operator and other reservation service and related activities
80	Security and investigation activities
81	Services to buildings and landscape activities
82	Office administrative, office support and other business support activities
84	Public administration and defence; compulsory social security
85	Education
86	Human health activities
87	Residential care activities
88	Social work activities without accommodation
90	Creative, arts and entertainment activities
91	Libraries, archives, museums and other cultural activities
92	Gambling and betting activities
93	Sports activities and amusement and recreation activities
94	Activities of membership organizations
95	Repair of computers and personal and household goods
96	Other personal service activities
97	Activities of households as employers of domestic personnel
98	Undifferentiated goods- and services-producing activities of private households for own use
99	Activities of extraterritorial organizations and bodies

Appendix B

Eurostat Data for Aggregate Economic Activity

Eurostat provides data on main economic indicators and some additional variables for all EU member states, Norway, Switzerland, and some candidate and potential candidate countries. The data is collected by national statistical agencies via surveys or by drawing on business registers or other administrative sources. It provides information within its two categories: Structural Business Statistics (SBS) and Business Demography (BD). Both parts cover variables related to business demography, while the scope of the data differs somewhat between SBS and BD. We first describe the differences and then explain how we utilize these data sets to compare the Eurostat data to our data from ORBIS-AMADEUS.

B.1 SBS and BD Data by Eurostat

Starting in 1995, the **SBS** data provides information that describes the structure, conduct, and performance of economic activities at the great level of detail (several hundred economic sectors).[1] The SBS coverage was limited to Sections C to K of NACE Rev. 1.1 until 2007. Starting from the reference year 2008, the data is available for sectors B to N and sub-sector S95 of NACE Rev. 2 classification. Some of these sectors, such as NACE Rev. 2 sector K

[1] See http://ec.europa.eu/eurostat/cache/metadata/en/sbs_esms.htm for a detailed description.

and NACE Rev. 1.1. sector J are only partially covered with data for insurance services, credit institutions, and pension funds but not all financial intermediation activities.

The main variables in the SBS data are business demographic variables (e.g., number of enterprises), "output related" variables (e.g., turnover, value added), "input related" variables such as labour input (e.g. employment, hours worked), goods and services input (e.g., total purchases), and capital input (e.g., material investments). This information is available for different company sizes. Among those available, we are interested in the categories TOTAL, 0-19, 20-249, and 250+ employees.

In 1997, the **BD** data started collecting variables mostly related to the demography of the business population.[2] The BD data is not as detailed in terms of variables is in the SBS data, which allows to cover a larger number of sectors in the BD data. Starting with the reference year 2008, the BD data covers NACE Rev. 2 sections B to N (excluding activities of holding companies, K64.2). Data for sections P, Q, R and S are provided on a voluntary basis. NACE Rev. 1.1 was used up to the reference year 2007 covering the data for sections C to K (excluding activities of holding companies-K74.15). Sectors M, N, and O were reported on a voluntary basis and, therefore, not available for all countries.

The main indicators of the BD data category include population of active enterprises, number of enterprise births, number of enterprise survivals up to five year, number of enterprise deaths, related variables on employment, derived indicators such as birth rates, death rates, survival rates and employment shares, and an additional set of indicators on high-growth enterprises and "gazelles" (high-growth enterprises that are up to five years old). This information is available for firms of different size and legal form. The three main categories by size are All, Zero, 10+.

As it is immediately seen, there are several differences between BD and SBS, important for our purposes. The BD data counts only active enterprises. Unlike the BD data, the SBS data counts both active and inactive companies. Moreover, the SBS data provides the information for more firm size groups based on employment than the BD data. Another advantage of the SBS data over the BD data is that the SBS data provides information on monetary values such as turnover, wage bill, investment, etc., and all these variables are

[2]See http://ec.europa.eu/eurostat/cache/metadata/EN/bd_esms.htm for a detailed description.

available for different firm sizes. The BD data covers more sectors but contains less variables. It also provides information separately for "Zero" firm size group (self-employed workers) that the SBS data lacks. For these reasons, in cases we do comparison based on number of enterprises and employment for either "Total" or "Zero" categories, we always use the BD data. In cases we do comparison based on gross output, we always use the SBS data. For all other cases where we do comparison for SMEs, we always use the SBS data. We couldn't use the BD data for comparisons based on employment for SMEs because, as above mentioned, the BD data provides information for employment but not for SMEs. In what follows, we present the details of the comparison of our data to statistics computed using Eurostat data.

B.2 ORBIS-AMADEUS versus Eurostat

We conduct different exercises to compare ORBIS-AMADEUS data with Eurostat data. Given the fact that the structure of ORBIS-AMADEUS and Eurostat data sets is different in many aspects, we tried to compare these data sets in a diligent manner.

We list the key steps we take to make Eurostat data sets and our BvD data comparable before conducting comparison exercises.

1. We identify Eurostat sectors based on NACE Rev. 2 classification because our ORBIS-AMADEUS data uses this classification (see step 6 in Chapter 4). In order to do this, we went through the process of matching the industry classifications pre- and post-2008 within each Eurostat data category. We first created a variable `NACE1` in both Rev. 1.1. and Rev. 2 files. We then filled in the values of this variable by values of NACE Rev. 2 codes based on our augmented BvD sector variable (`NACEREV2CCODE`) using the official NACE Rev. 2 Level 1 classification. For example, we replaced `NACE1` with C if it was coded as D in NACE Rev. 1.1. file. Then, we merged these two files by `NACE1` for each country-year-company category (company category is based either on size and/or legal form). Thus, this common `NACE1` variable in ORBIS-AMADEUS and Eurostat enabled us to compare them based on any sector defined at NACE Rev. 2 Level 1 classification.

2. We determine overlapping sectors across ORBIS-AMADEUS and Eurostat datasets

based on the `NACE1` variable. While ORBIS-AMADEUS data, by construction, covers all sectors for a given country-year, Eurostat data provides information for the business economy with the exception of some sectors. Table B.2 lists the sectors with available information in Eurostat data sets. However, this list is differential based on the variable of interest. For example, the variable Turnover provided by Eurostat SBS data is not available for all sectors given in Table B.2. The availability of this variable differs across country-sector-year triplets, and this fact should be taken into account especially in total economy comparison exercises.

To exemplify, Eurostat data sets provide information for the construction sector (NACE Rev. 2 sector F) between 1995–2012. However, according to Table B.3, Eurostat SBS data lacks information on gross-output for this sector in Belgium for the years 2008, 2009, 2011 and 2012. If we had disregarded such detail and had computed total gross output over BvD Belgium firms for all overlapping sectors by just considering the correspondence given in Table B.2, we would have inflated the BvD aggregate for 2008, 2009, 2011 and 2012. To put it differently, since BvD aggregates are used as numerators in the ratios we use to construct percentages, overlooking such details would produce artificial higher percentages. In order to avoid such mistakes, we instead followed the sector correspondence given in Table B.3, and construct the percentages accordingly. To be more exact, in case we computed the percentage based on gross output for Belgium in 2009, we aggregated gross-output over the sectors which have non-missing gross output in *both* Eurostat SBS and BvD data sets. The sectors we used in this case exactly referred to the ones coded with 1 in Table B.3.

3. Eurostat BD data provides information on the main indicators for the categories of different firm size and legal form. Among the categories of legal forms, we were forced to use "Total" which refers to all firms with different legal forms assuming that BvD collects data from firms of all legal firms. We did not want to deal with legal firm issue since it is hard to identify the legal form information for each country in ORBIS-AMADEUS data given differential filing policy across countries (See Table A.1). Among size categories, we used "Total" and "Zero" to construct a new size category i.e. "AllminusZero" which refers to all firms excluding the *self employed workers*. This enabled us to compare manufacturing employment of ORBIS-AMADEUS to

that of Eurostat BD in the most appropriate manner since BvD excludes self-employed workers by construction.

Eurostat SBS data provides information only for different size categories. However, it doesn't provide information under "Zero" size category, which forced us to use "Total" size category in all comparison exercises where Eurostat SBS data was used.

4. Our comparison exercises mainly depend on two economic activity measures; i.e., employment and gross-output. Eurostat data provides relevant indicators with different definitions. In order to precisely compare the coverage based on those measures, we picked the ones defined similar to our BvD variables. To illustrate, in our comparison exercises based on gross-output we used `OPRE` and V12110-Turnover from BvD and Eurostat SBS data, respectively. We express these financial variables in real dollars 2005 base using original values in Eurostat SBS data (see Step 2 in Chapter 5). Additional correspondences as well as the variable definitions are given in Table B.1.

5. In comparison exercises based on number of enterprises, we use either Eurostat SBS or Eurostat BD data. In case we use the BD data, we dropped *inactive* companies checking the variable `STATUS` in our data because the BD data counts only active companies.[3]

6. We constructed two different samples using our BvD data and repeated our comparison exercised for each sample: Total Sample and TFP Sample. The Total Sample consists of firms that report data with positive values of the corresponding measure (i.e. employment (`EMPL`) and gross output (`OPRE`)), whereas the TFP Sample consists of firms that report positive values on employment (`EMPL`) or wage bill (`STAF`), and tangible fixed assets (`TFAS`), gross output (`OPRE`), materials (`MATE`).

7. For a given company, ORBIS-AMADEUS provides financial statements regarding different consolidation codes i.e. C1, C2, U1 and U2.[4] Given this fact, we first dropped

[3] If a given BvD company is read as "Inactive," "Dissolved," "In liquidation," and "Bankruptcy" we count it as inactive.

[4] C1: account of a company- headquarter of a group, aggregating all companies belonging to the group (affiliates, subsidiaries, etc.), where the company headquarter has no unconsolidated account, C2: account of a company-headquarter of a group, aggregating all companies belonging to the group (affiliates, subsidiaries, etc.) where the company headquarter also presents an unconsolidated account, U1: account of a company with no consolidated account, and U2: account of a company with a consolidated account.

C2 accounts to avoid double accounting in our comparison exercises. For some specific cases, we further dropped C1 accounts. To illustrate, in cases we used the Total Sample, we dropped C1 accounts for all countries except Spain and Italy. In cases we used the TFP Sample, we dropped C1 accounts for all countries except Spain, Italy, Cyprus, Denmark, Great Britain, Greece, Ireland, and Lithuania.

8. As a last step before constructing percentages by which we compare the coverage between BvD data and Eurostat data, we winsorized the distribution of relevant economic activity measure within the corresponding country-sector-year triplet if needed.[5]

We picked the most important variables, reported in Table B.1 . This table can be used as a guide to follow our comparison tables within the text.

[5]For space considerations, the list of country-sector-year triplets to which we applied winsorization as well as the details of the winsorization (which varies between 0.01% and 0.5%) are not reported here. All these details are available upon request.

Table B.1: Details of Comparison Exercises

Tables 6.1–6.2		
Data	ORBIS-AMADEUS	Eurostat SBS
Year	1999-2012	1999-2012
Variable	OPRE	V12110-Turnover
Size	All	TOTAL - Total
Sector	See Table B.3	See Table B.3
Table 6.3		
Data	ORBIS-AMADEUS	Eurostat BD
Year	2007, 2008, 2009	2007, 2008, 2009
Variable	ID_NUMBER	V11910-Population of active enterprises in t
Size	All	Country specific
Sector	Total economy	Total economy
Table 6.4		
Data	ORBIS-AMADEUS	Eurostat SBS
Year	2007, 2008, 2009	2007, 2008, 2009
Variable	ID_NUMBER	V11110-Number of enterprises
Size	All, 1-19, 20-249, 250 + employees	TOTAL-Total, 0-19, 20-249, 250 + employees
Sector	Total economy	Total economy

Notes: Table B.1 presents the details of the comparison exercises. The sectors are compared based on NACE Rev. 2 Level 1 Classification (`NACE1`). In Tables 6.3–6.4, total economy in the reference country-year corresponds to all common available sectors with Eurostat BD and SBS data, respectively. The definitions of the variables are as follows: **OPRE**: Total operating revenues (Net sales + Other operating revenues+ Stock variations). The figures do not include VAT. Local differences may occur regarding excises taxes and similar obligatory payments for specific market of tobacco and alcoholic beverage industries; **EMPL**: Total number of employees included in the company's payroll, **V16910**: Number of persons employed in the population of active enterprises in t, **V13310**: Personnel costsare defined as the total remuneration, in cash or in kind, payable by an employer to anemployee(regular and temporary employees, as well as home-workers) in return for work done by the latter during the reference period, **V12110:** Turnover comprises the totals invoiced by theobservation unitduring thereference period, and this corresponds to the total value of market sales of goods and services to third parties, **V16110:** Number of persons employed. In BvD data, for the construction of TFP sample, we also used the variables **STAF**, **TFAS**, and **MATE**. **STAF**: All the employees costs of the company (including pension costs), **TFAS**: Book value of tangible fixed assets i.e. plant, equipment and machinery, and **MATE**: Material Costs.

Table B.1 (Cont'd.): Details of Comparison Exercises

Panel A in Tables 6.5–6.8		
Data	ORBIS-AMADEUS	Eurostat SBS
Year	1999-2012	1999-2012
Variable	OPRE	V12110-Turnover
Size	All	TOTAL - Total
Sector	C	C
Panel B in Tables 6.5–6.8		
Data	ORBIS-AMADEUS	Eurostat BD
Year	1999-2012	1999-2012
Variable	EMPL	V16910
Size	All	AllminusZero
Legal form	All	All
Sector	C	C
Tables 6.9–6.10		
Data	ORBIS-AMADEUS	Eurostat SBS
Year	2006	2006
Variable	OPRE	V12110-Turnover
Size	1-19, 20-249, 250 + employees	0-19, 20-249, 250 + employees
Sector	C	C
Tables 6.9–6.10		
Data	ORBIS-AMADEUS	Eurostat SBS
Year	2006	2006
Variable	EMPL	V16110
Size	1-19, 20-249, 250 + employees	0-19, 20-249, 250 + employees
Sector	C	C

Notes: See the notes above.

Table B.2: Sector Coverage in Eurostat Data

		SBS		BD	
Code	Definition	1995-2007	2008-2012	1997-2007	2008-2012
A	Agriculture, forestry and fishing	0	0	0	0
B	Mining and quarrying	1	1	1	1
C	Manufacturing	1	1	1	1
D	Electricity, gas, steam and air-conditioning supply	1	1	1	1
E	Water supply, sewerage, waste management and remediation	1	1	1	1
F	Construction	1	1	1	1
G	Wholesale and retail trade, repair of motor vehicles and motorcycles	1	1	1	1
H	Transportation and storage	1	1	1	1
I	Accommodation and food service activities	1	1	1	1
J	Publishing, audiovisual and broadcasting activities /telecommunications/IT and other information services	1	1	1	1
K	Financial and insurance activities	1	1	1	1
L	Real estate activities	1	1	1	1
M	A Legal, accounting, management, architecture, engineering, technical testing and analysis activities	1	1	1	1
N	Administrative and support service activities	1	1	1	1
O	Public administration and defence, compulsory social security	0	0	0	0
P	Education	0	0	1	1
Q	Human health services and residential care and social work activities	0	0	1	1
R	Arts, entertainment and recreation	0	0	1	1
S	Other services	1	1	1	1
T	Activities of households as employers; undifferentiated goods and services producing activities of households for own use	0	0	0	0
U	Activities of extra-territorial organisations and bodies	0	0	0	0

Notes: Table B.2 presents sector coverage for Eurostat SBS and BD data sets based on NACE Rev. 2 Level 1 classification. Sector-period pair is labeled as 1 if any information is available in the corresponding data category, 0 otherwise. In the SBS data, sector K is only covered partially by the data on insurance services, credit institutions and pension funds, and sector S is partially covered by the data on repair of computers and personal and household goods (S95) in Eurostat SBS data starting from 2008. In the BD data, sectors P, Q, R, and S are provided on voluntary basis and Sector K64.2 (activities of holding companies) is excluded.

Table B.3: Sector Coverage of Eurostat Countries across Years: 1999–2012
Based on Gross Output

AT	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	1	0

BE	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	0	0	0	0	0	0	0	1	0	0
C	1	1	1	0	1	1	1	1	1	1	1	1	1	1
D	1	1	1	0	1	1	1	1	1	0	0	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	0	1	1	1	1	1	0	0	1	0	0
G	1	1	1	0	1	1	1	1	1	0	0	1	0	0
H	1	1	1	0	1	1	1	1	1	1	1	1	1	0
I	1	1	1	0	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	0	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	0	0	0	0

BG	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	1	1	1	1	1	1	0	0	0	0	0
C	0	0	0	1	1	1	1	1	1	1	1	1	1	1
D	0	0	0	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	0	0	0	0	0	0	1
F	0	0	0	1	1	1	1	1	1	0	0	0	0	0
G	0	0	0	1	1	1	1	0	1	0	0	1	0	0
H	0	0	0	1	1	1	1	1	1	1	1	1	1	0
I	0	0	0	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	0	0	0	1	1	1	1	1	1	1	0	0	0	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	0	0	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: Table B.3 presents the coverage by economic activity based on NACE Rev. 2 Level 1 categorization for a given country-year. The variable of interest is gross-output so the information comes from Eurostat SBS data. These tables are used as reference for total economy comparisons based on gross-output. In each table, the corresponding country-year is labelled as 1 if information on gross-output is available in Eurostat SBS data, 0 otherwise. The country codes reported are as follows: AT (Austria), BE (Belgium), BG (Bulgaria), CH (Switzerland), CZ (Czech Republic), DE (Germany), DK (Denmark), EE (Estonia), ES (Spain), FI (Finland), FR (France), GB (Great Britian), GR (Greece), HR (Croatia), HU (Hungary), IE (Ireland), IT (Italy), LT (Lithuania), LU (Luxembourg), LV (Latvia), NL (Netherlands), NO (Norway), PL (Poland), PT (Portugal), RO (Romania), SE (Sweden), SI (Slovenia), and SK (Slovakia).

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

CH	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	0	0	0	0	0	0	0	0	0	1
C	0	0	0	0	0	0	0	0	0	0	1	1	1	1
D	0	0	0	0	0	0	0	0	0	0	1	0	0	1
E	0	0	0	0	0	0	0	0	0	0	0	0	0	1
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	0	0	1	0	0
H	0	0	0	0	0	0	0	0	0	0	1	1	1	0
I	0	0	0	0	0	0	0	0	0	0	1	0	0	0
J	0	0	0	0	0	0	0	0	0	0	1	0	1	0
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	1	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

CY	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	1	1	1	1	1	1	1	1	1	1	1	0
C	0	1	1	1	1	1	1	1	1	1	1	1	1	1
D	0	0	0	0	0	0	0	0	0	0	1	1	1	0
E	0	0	0	0	0	0	0	0	0	1	1	1	0	0
F	0	1	1	1	1	1	1	1	1	0	0	1	0	0
G	0	0	1	1	1	1	1	1	1	0	0	1	0	0
H	0	0	0	1	1	1	1	1	1	1	1	1	0	0
I	0	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	1	1	1	1	1	1	1	0
L	0	0	0	0	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	1	1	1	1	1	1	1	0
N	0	0	0	0	0	0	1	1	1	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	1	1	0	0

CZ	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	1	1	1	1	0	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	0	0	0
I	1	1	1	1	1	1	1	1	1	1	1	0	0	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	1	1	0	1	1	1	1	1	1	1	1	1	0	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	0	0	0	0

DE	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	0	0	0	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	0	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	0	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: See the notes above.

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

DK	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	1	1	1	0

EE	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	0	0	0	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	0	0	0	0	0	0	1	1	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	0	0	0	0	0	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	1	1	1	1	1	1	1	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	1	1	1	0
N	0	0	0	0	0	0	0	0	0	0	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ES	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	0	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

FI	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	1	1	1	1	1	1	1
F	1	1	1	1	1	1	0	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	1	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	0	0
M	0	0	0	0	0	0	1	0	0	1	1	1	1	0
N	0	0	0	0	0	0	1	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	1	1	0	0

Notes: See the notes above.

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

FR	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	0	0	0	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	1
G	1	1	1	0	1	1	1	1	1	0	0	1	0	1
H	1	1	1	1	1	1	1	1	1	1	1	1	1	1
I	1	1	1	1	1	1	1	1	1	1	1	1	1	1
J	0	0	0	0	0	0	0	0	0	1	1	1	1	1
L	1	1	1	1	1	1	1	1	1	1	1	1	1	1
M	0	0	0	0	0	0	0	0	0	1	1	1	1	1
N	0	0	0	0	0	0	0	0	0	1	1	1	1	1
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

GB	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
K	1	1	1	0	0	0	0	0	0	0	0	0	0	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

GR	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	1	1	1	1	1	1	1	1	1	1
C	1	1	0	0	1	1	1	1	1	1	1	1	1	1
D	0	0	0	0	1	1	1	1	1	0	1	1	1	1
E	0	0	0	0	0	0	0	0	0	0	1	1	1	1
F	0	0	0	1	1	1	1	1	1	0	0	0	0	0
G	0	0	0	0	1	1	1	1	1	0	0	0	0	0
H	0	0	0	0	1	1	1	1	1	1	1	1	0	0
I	0	0	0	0	1	1	1	1	1	1	1	1	0	0
J	0	0	0	0	0	0	0	0	0	1	1	1	0	0
L	0	0	0	0	1	1	1	1	1	1	1	1	0	0
M	0	0	0	0	0	0	0	0	0	1	1	1	0	0
N	0	0	0	0	0	0	0	0	0	1	1	1	0	0
S	0	0	0	0	0	0	0	0	0	1	0	1	0	0

HR	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	0	0	0	0	0	0	1	1	1	1
C	0	0	0	0	0	0	0	0	0	1	1	1	1	1
D	0	0	0	0	0	0	0	0	0	0	0	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	0	0	0	0	0	0	0	0	0	0	0	1	0	0
G	0	0	0	0	0	0	0	0	0	0	0	1	0	0
H	0	0	0	0	0	0	0	0	0	1	1	1	1	0
I	0	0	0	0	0	0	0	0	0	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	0	0	0	0	0	0	0	0	0	0	0	0	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: See the notes above.

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

HU	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	0	1	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	0	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	1	0	0	0

IE	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	0	1	1	0	0	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	0	0	0	0	0	0	1	1	0	0	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	0	0	0	0	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	0	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	0	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

IT	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	1
G	1	1	1	1	1	1	1	1	1	0	0	1	0	1
H	1	1	1	1	1	1	1	1	1	1	1	1	1	1
I	1	1	1	1	1	1	1	1	1	1	1	1	1	1
J	0	0	0	0	0	0	0	0	0	1	1	1	1	1
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	1	1	1	1	1	1	1	1	1	0	0	0	1	1
M	0	0	0	0	0	0	0	0	0	1	1	1	1	1
N	0	0	0	0	0	0	0	0	0	1	1	1	1	1
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

LT	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	0	0	0	0	0	0	0	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	0	0	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	0	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	0	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	0	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Notes: See the notes above.

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

LU	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	1	1	1	1	1	1	1	1	1	0
C	0	0	0	0	1	1	1	1	1	1	1	1	1	0
D	0	0	0	0	1	1	1	1	1	1	1	1	1	0
E	0	0	0	0	0	0	0	0	0	1	1	1	1	0
F	0	0	0	0	1	1	1	1	1	0	0	1	0	0
G	0	0	0	0	1	1	1	1	1	0	0	1	1	0
H	0	0	0	0	0	1	1	1	1	1	1	1	1	0
I	0	0	0	0	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	0	0	0	0	1	1	1	1	1	0	0	0	1	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

LV	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	0	0	0	0	0	0	0	0	0	1	1	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	0	1	1	1	1	1
E	0	0	0	0	0	0	0	1	1	0	0	1	0	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

NL	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	1	1	1	0	0	0	0	0	1	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	0	0	1	1	1	1	0	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	0	0	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	0	0	0	0	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

NO	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	0	0	0	0	1	1	1	1	1	1	1	1	1	1
C	1	1	0	1	1	1	1	1	1	1	1	1	1	1
D	0	0	0	0	0	0	0	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	0	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	0	1	0	0	0	1	1	1	1	1	1	1	1	0
I	0	1	0	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	0	1	0	0	1	1	1	1	0	1	0	0	0	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: See the notes above.

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

PL	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	1	1	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	1	0	0	0

PT	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	1	1	1	1	1	1	1
F	0	0	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
K	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

RO	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	1	1	1	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	0	1	1	1	1	1	1	1	1	0	0	1	0	0
H	0	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	0	0	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

SE	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1	1
E	0	0	0	0	0	0	0	0	0	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	1	1	1	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	1	1	1	1	1	1	1	0
N	0	0	0	0	0	0	1	1	1	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Notes: See the notes above.

Table B.3 (Cont'd.): Sector Coverage of Eurostat Countries across Years: 1999–2012

Based on Gross Output

SI	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	0	0	1	1	1	1	0	1	0	1	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	0	0	0	0	1	1	1	1	1
E	0	0	0	0	0	0	1	1	1	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	1	1	1	1	0
M	0	0	0	0	0	0	0	0	0	1	1	1	1	0
N	0	0	0	0	0	0	0	0	0	1	1	1	1	0
S	0	0	0	0	0	0	0	0	0	1	1	0	0	0

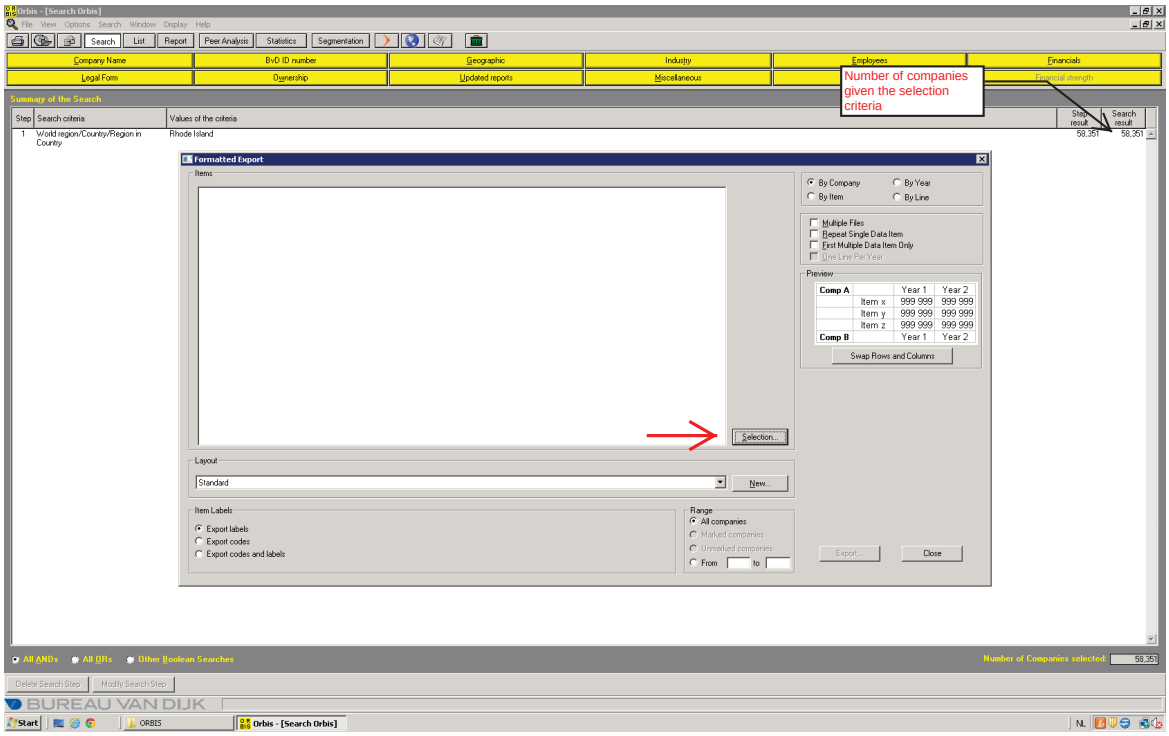
SK	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
B	1	1	1	1	1	1	0	1	0	0	0	0	0	0
C	1	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	0	0	0	0	0	1	0	1
E	0	0	0	0	0	0	0	0	0	0	1	1	1	1
F	1	1	1	1	1	0	0	0	0	0	0	1	0	0
G	1	1	1	1	1	1	1	1	1	0	0	1	0	0
H	1	1	1	1	1	1	1	1	1	1	1	1	1	0
I	1	1	1	1	1	1	1	1	1	1	1	1	1	0
J	0	0	0	0	0	0	0	0	0	1	1	1	1	0
L	1	1	1	1	1	1	1	1	1	0	0	1	0	0
M	0	0	0	0	0	0	0	0	0	1	1	1	0	0
N	0	0	0	0	0	0	0	0	0	1	1	1	0	0
S	0	0	0	0	0	0	0	0	0	1	0	0	0	0

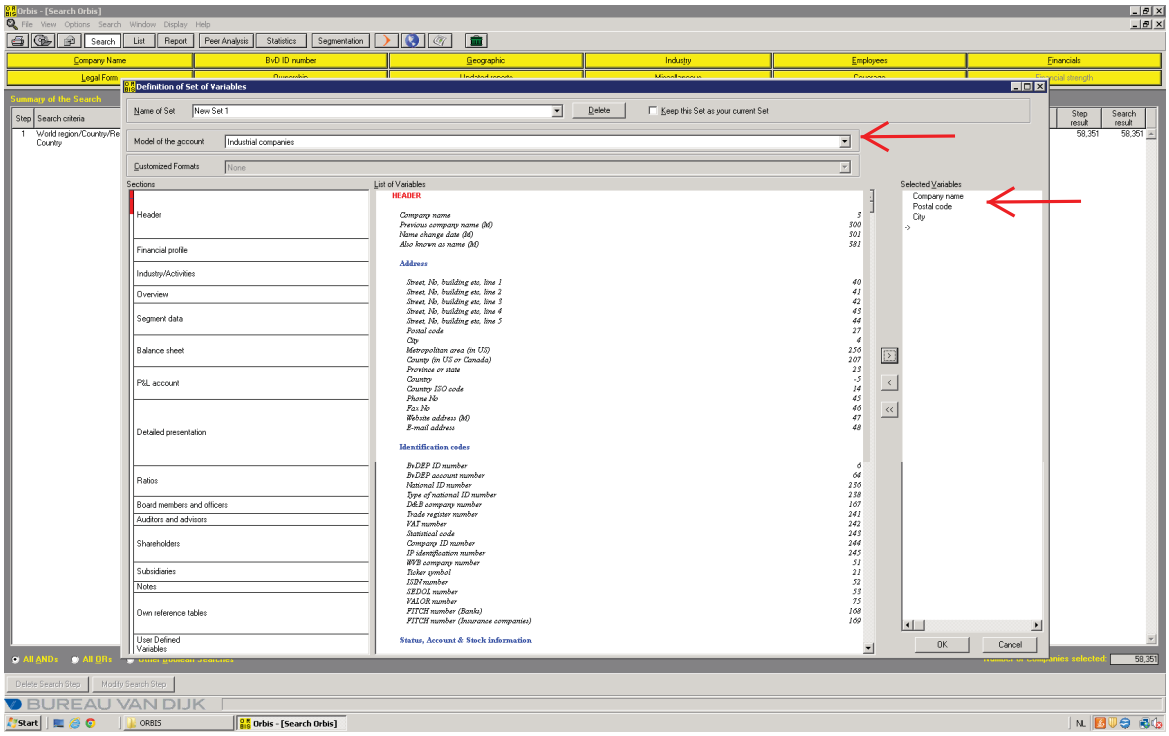
Notes: See the notes above.

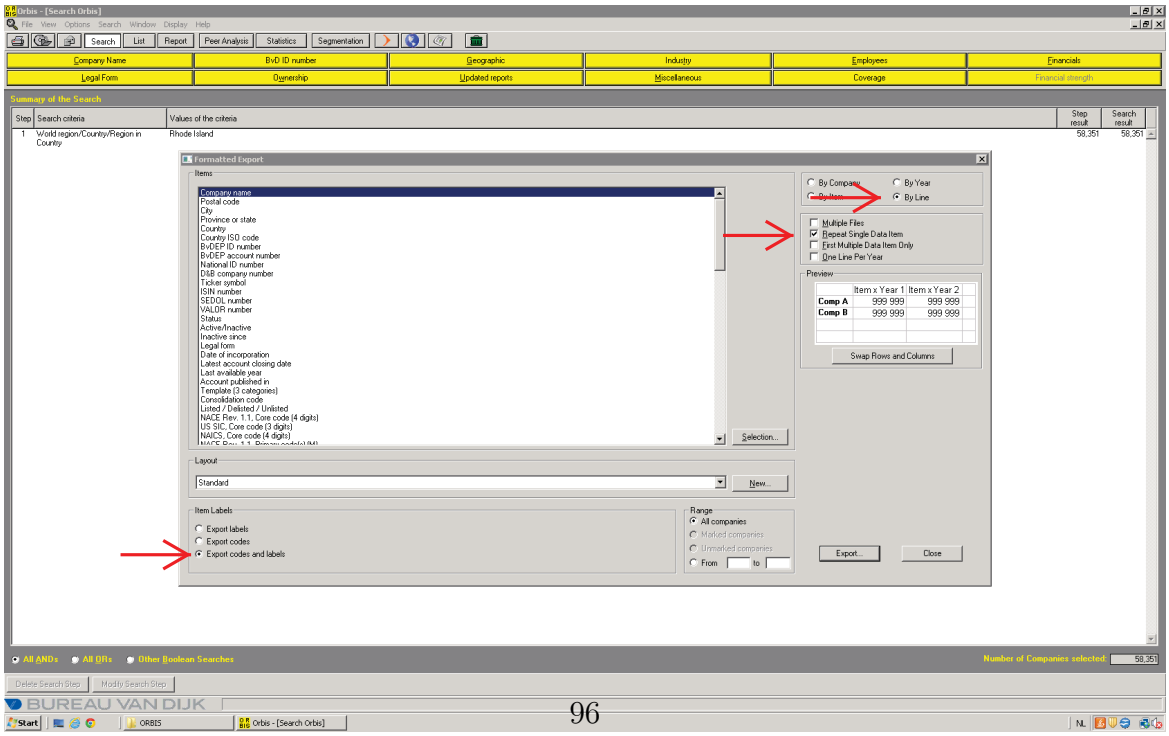
Appendix C

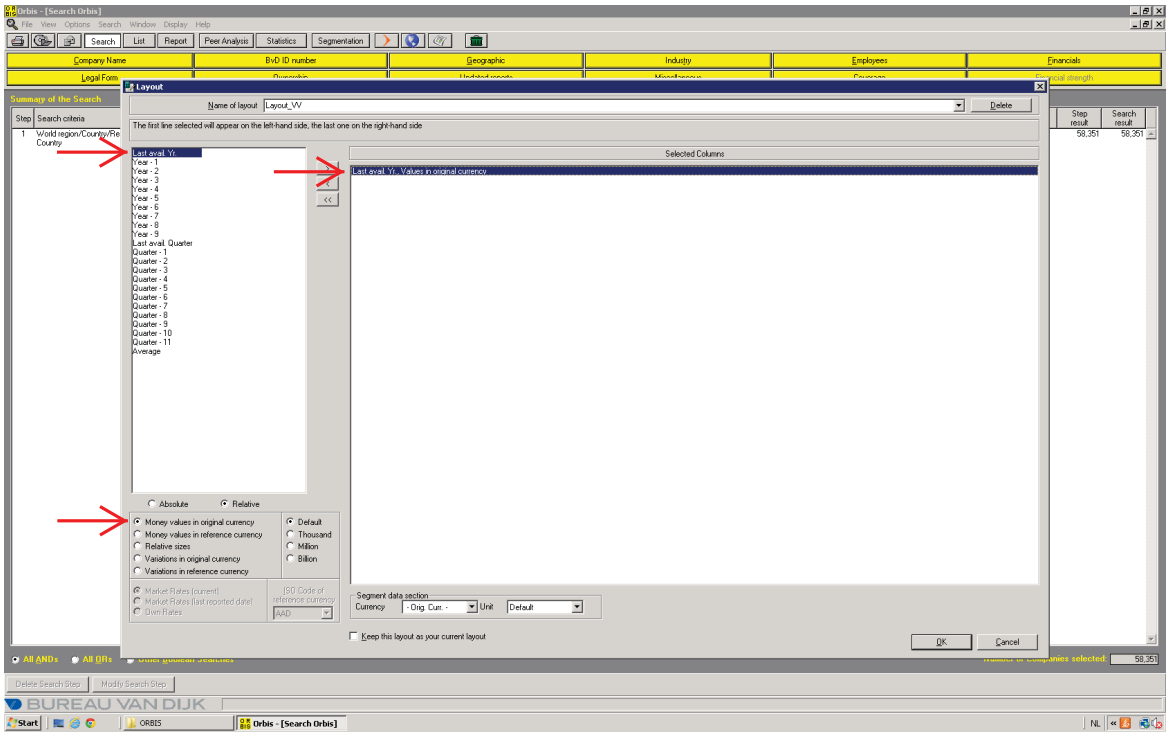
BvD Disks Interface

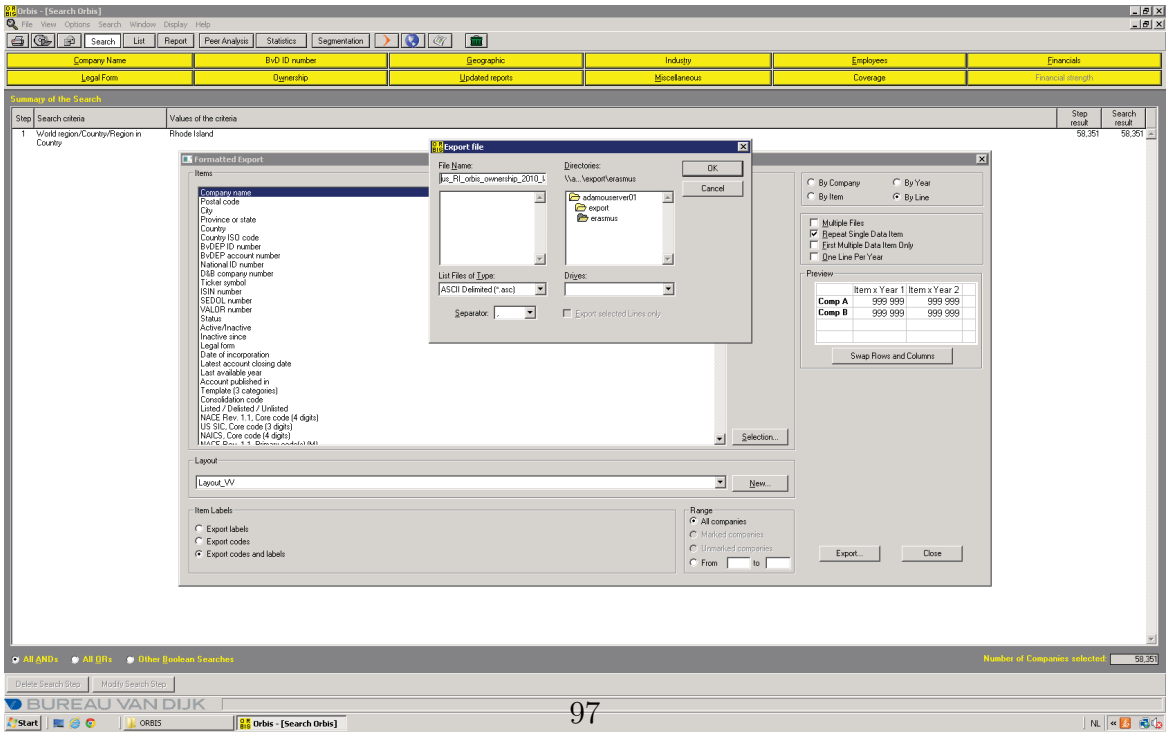
C.1 Formatted Export Interface of Older BvD Disks



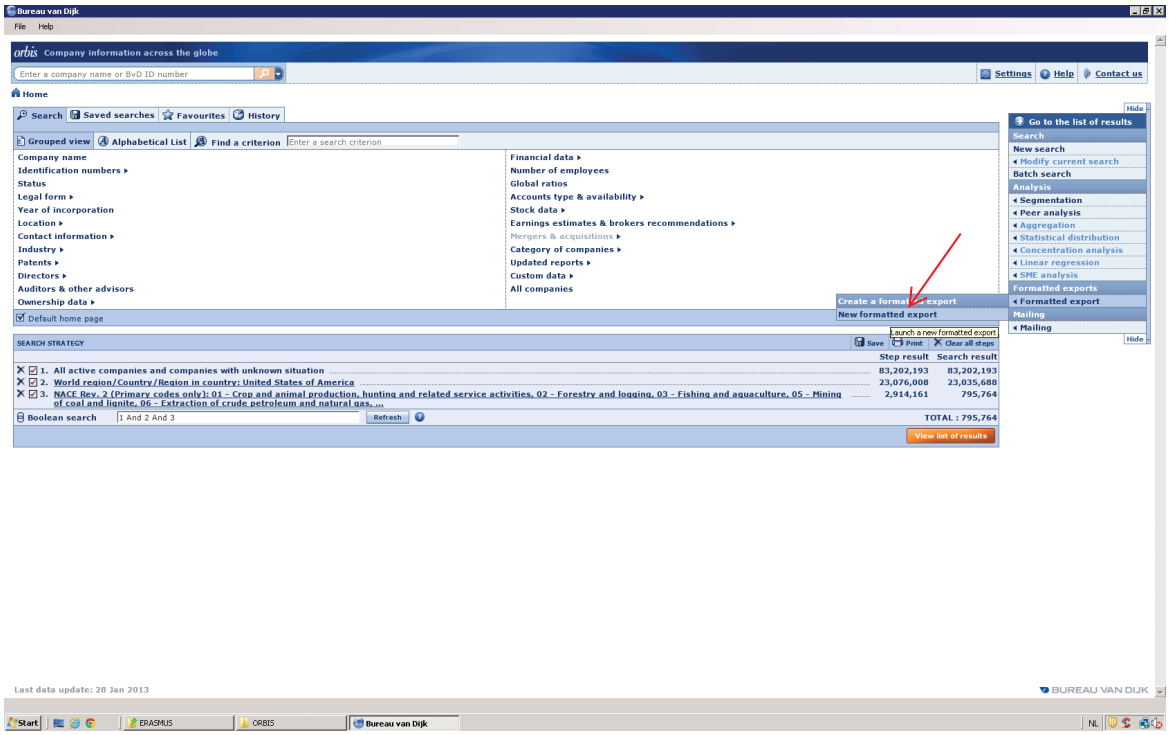

Definition of Set of Variables
Name of Set
New Set 1
Delete
Keep this Set as your current Set
Model of the account
Industrial companies
Customized Formats
None
Sections
List of Variables
Selected Variables
Company name
Postal code
City
OK
Cancel

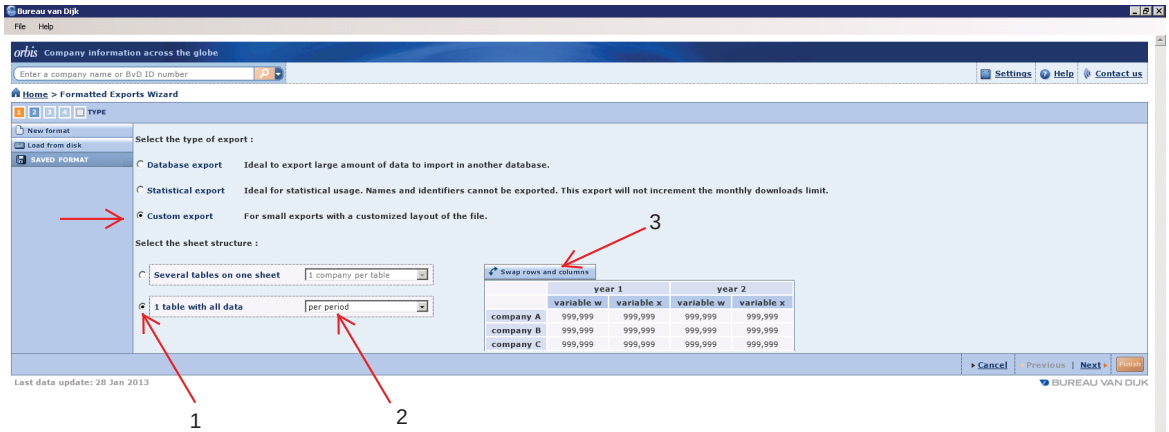

Formatted Export
Items
Company name
Postal code
City
Province or state
Country
Country ISO code
BvDEP ID number
BvDEP account number
National ID number
D&B company number
Ticker symbol
ISIN number
SEDOL number
VALOR number
Status
Active/Inactive
Inactive since
Legal form
Date of incorporation
Latest account closing date
Last available year
Account published in
Template (3 categories)
Consolidation code
Listed / Delisted / Unlisted
NACE Rev. 1.1, Core code (4 digits)
US SIC, Core code (3 digits)
NAICS, Core code (4 digits)
Selection...
Layout
Standard
New...
Item Labels
Export labels
Export codes
Export codes and labels
Range
All companies
Marked companies
Unmarked companies
From
to
By Company
By Year
By Item
By Line
Multiple Files
Repeat Single Data Item
First Multiple Data Item Only
One Line Per Year
Preview
Item x Year 1
Item x Year 2
Comp A
Comp B
999 999
Swap Rows and Columns
Export...
Close

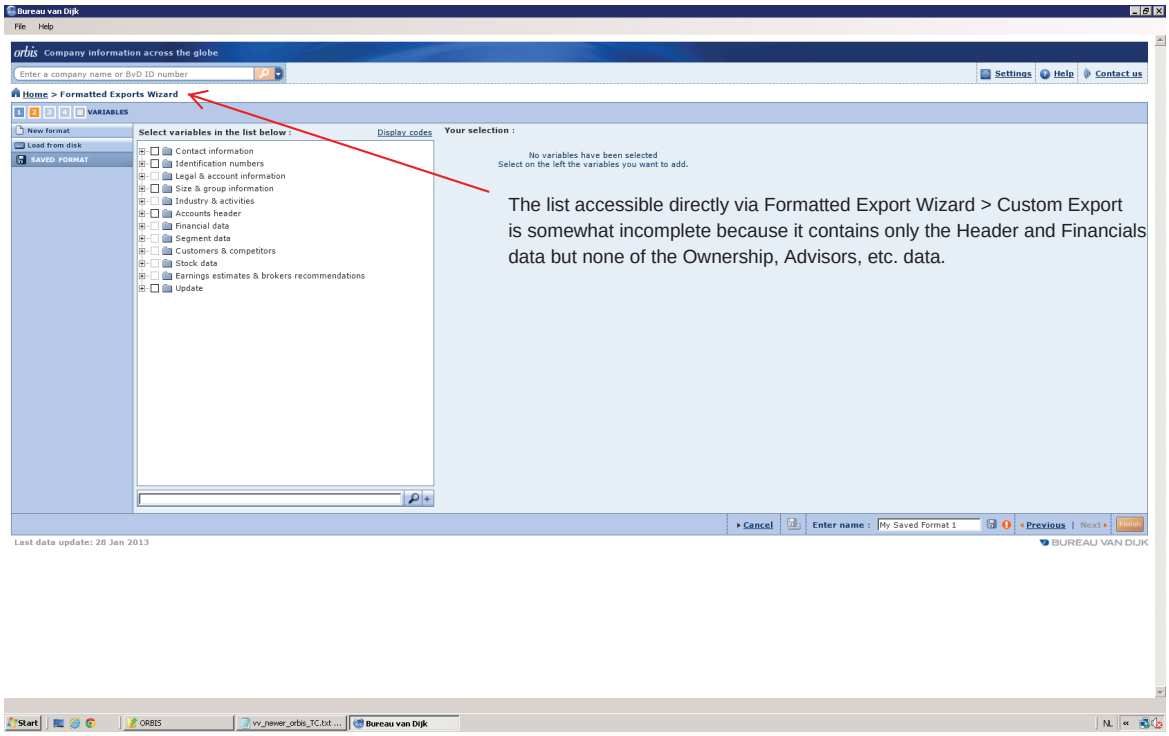

Orbis - [Search Orbis]
Search
List
Report
Peer Analysis
Statistics
Segmentation
Company Name
BvD ID number
Geographic
Industry
Employees
Financials
Legal Form
Ownership
Updated reports
Miscellaneous
Coverage
Financial strength
Summary of the Search
Layout
Name of layout
Layout_VV
Delete
The first line selected will appear on the left-hand side, the last one on the right-hand side
Selected Columns
Last avail. Yr.
Year - 1
Year - 2
Year - 3
Year - 4
Year - 5
Year - 6
Year - 7
Year - 8
Year - 9
Last avail. Quarter
Quarter - 1
Quarter - 2
Quarter - 3
Quarter - 4
Quarter - 5
Quarter - 6
Quarter - 7
Quarter - 8
Quarter - 9
Quarter - 10
Quarter - 11
Average
Last avail. Yr. Values in original currency
Absolute
Relative
Money values in original currency
Money values in reference currency
Relative sizes
Variations in original currency
Variations in reference currency
Default
Thousand
Million
Billion
Segment data section
Currency
- Orig. Curr. -
Unit
Default
Keep this layout as your current layout
OK
Cancel
All ANDs
All ORs
Delete Search Step
Modify Search Step
BUREAU VAN DIJK
Start
ORBIS
Orbis - [Search Orbis]
NL

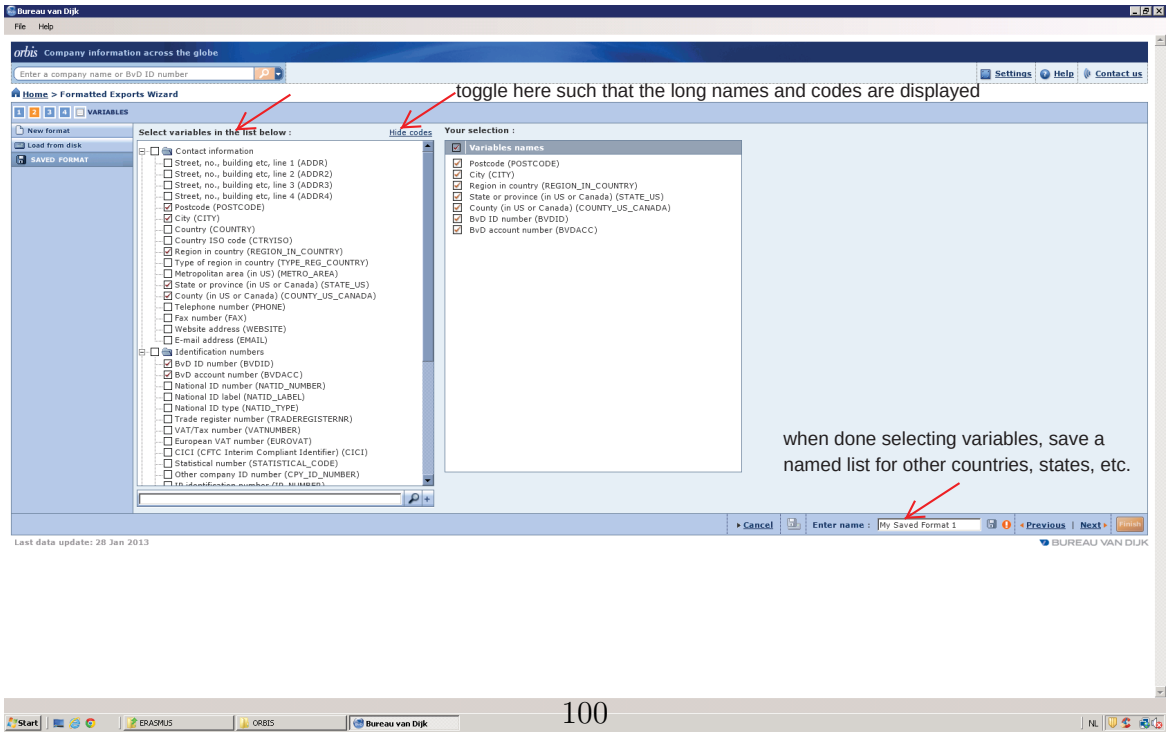

Orbis - [Search Orbis]
Search
List
Report
Peer Analysis
Statistics
Segmentation
Company Name
BvD ID number
Geographic
Industry
Employees
Financials
Legal Form
Ownership
Updated reports
Miscellaneous
Coverage
Financial strength
Summary of the Search
Step
Search criteria
Values of the criteria
1
World region/Country/Region in Country
Rhode Island
58,351
Formatted Export
Items
Company name
Postal code
City
Province or state
Country
Country ISO code
BvDEP ID number
BvDEP account number
National ID number
D&B company number
Ticker symbol
ISIN number
SEDOL number
VALOR number
Status
Active/Inactive
Inactive since
Legal form
Date of incorporation
Latest account closing date
Last available year
Account published in
Template (3 categories)
Consolidation code
Listed / Delisted / Unlisted
NACE Rev. 1.1, Core code (4 digits)
US SIC, Core code (3 digits)
NAICS, Core code (4 digits)
Selection...
Layout
Layout_VV
New...
Item Labels
Export labels
Export codes
Export codes and labels
Range
All companies
Marked companies
Unmarked companies
From
to
Export...
Close
Export file
File Name:
Directories:
OK
Cancel
List Files of Type:
ASCII Delimited (*.asc)
Drives:
Separator:
Export selected Lines only
By Company
By Year
By Item
By Line
Multiple Files
Repeat Single Data Item
First Multiple Data Item Only
One Line Per Year
Preview
Item x Year 1
Item x Year 2
Comp A
Comp B
999 999
Swap Rows and Columns
All ANDs
All ORs
Other Boolean Searches
Number of Companies selected
58,351
Delete Search Step
Modify Search Step
BUREAU VAN DIJK
Start
ORBIS
Orbis - [Search Orbis]
NL

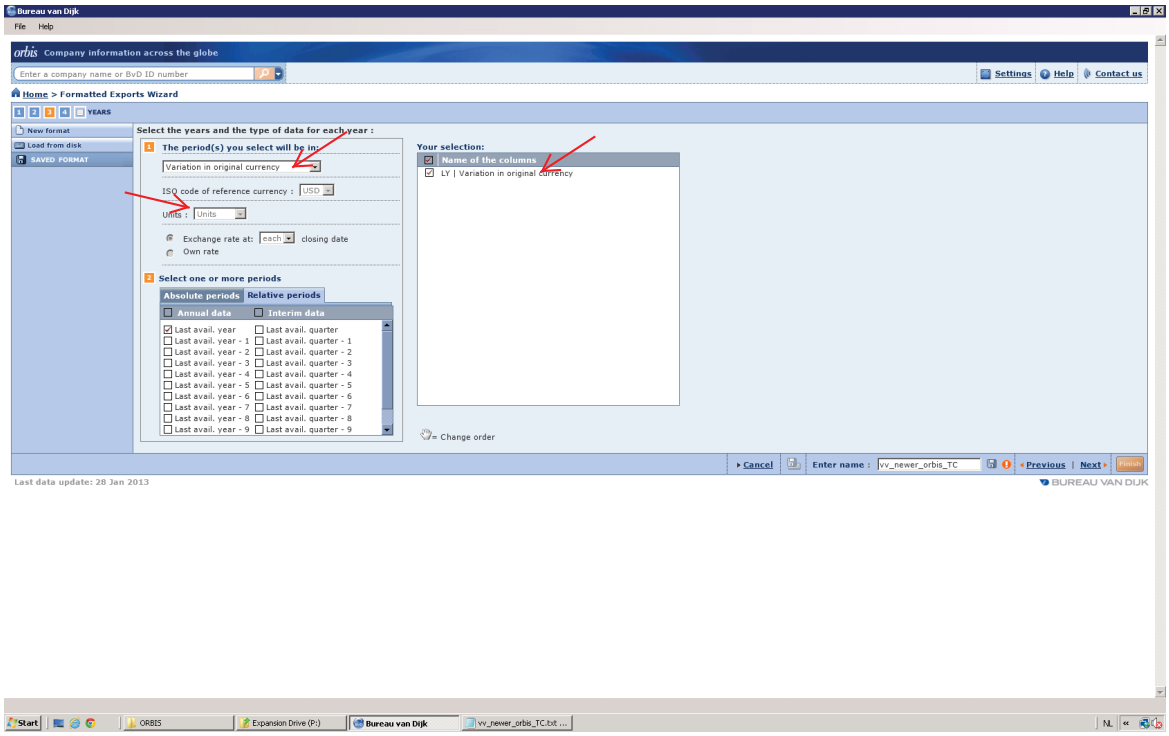
C.2 Custom Export Interface of Newer BvD Disks

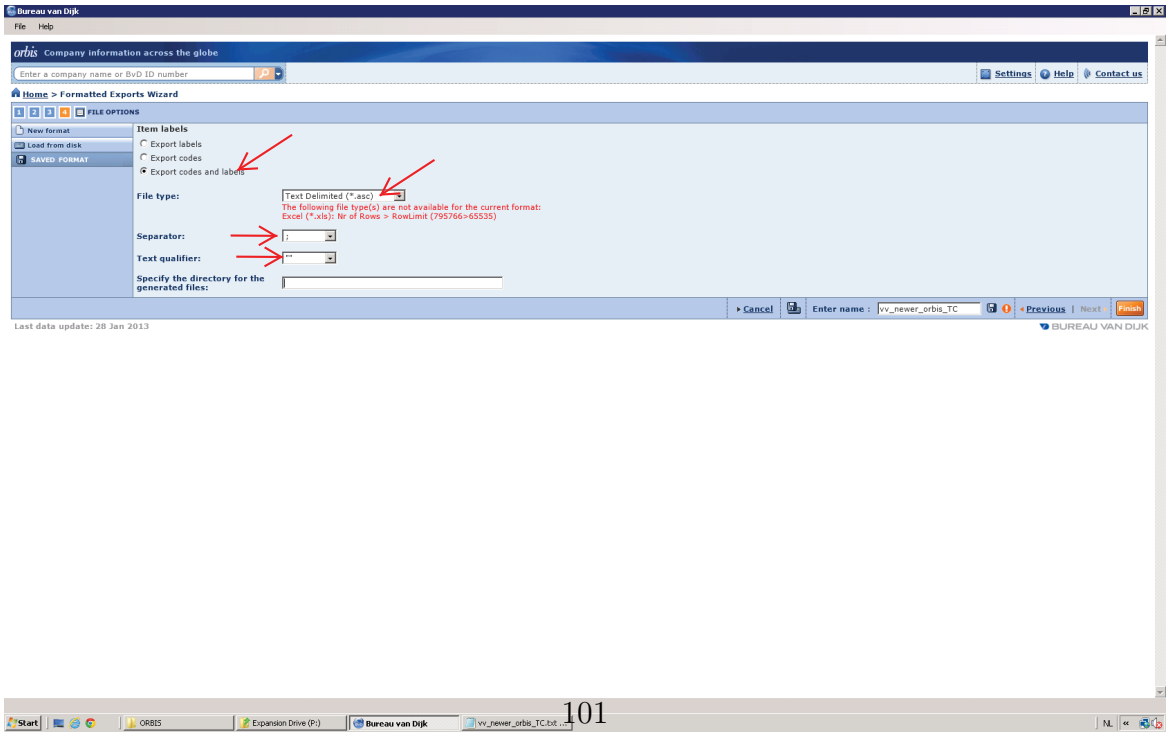




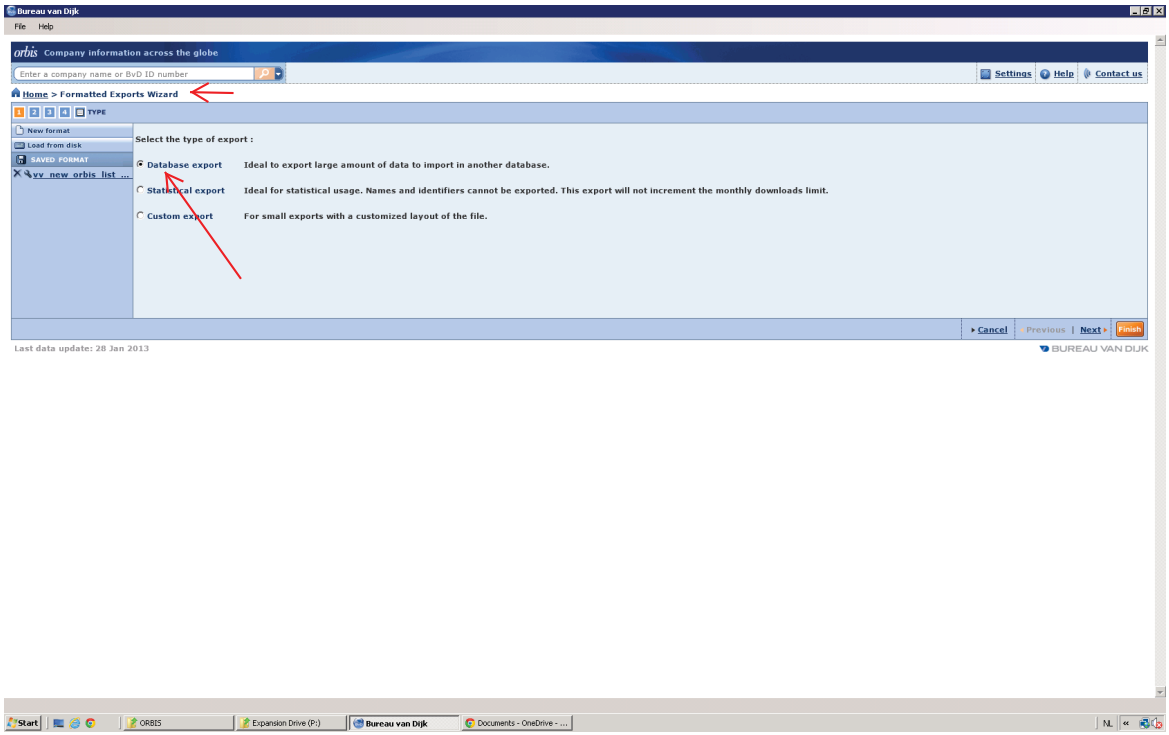

Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > Formatted Exports Wizard
VARIABLES
New format
Load from disk
SAVED FORMAT
Select variables in the list below :
Display codes
Your selection :
No variables have been selected
Select on the left the variables you want to add.
Contact information
Identification numbers
Legal & account information
Size & group information
Industry & activities
Accounts header
Financial data
Segment data
Customers & competitors
Stock data
Earnings estimates & brokers recommendations
Update
The list accessible directly via Formatted Export Wizard > Custom Export is somewhat incomplete because it contains only the Header and Financials data but none of the Ownership, Advisors, etc. data.
Cancel
Enter name :
My Saved Format 1
Previous
Next
Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK

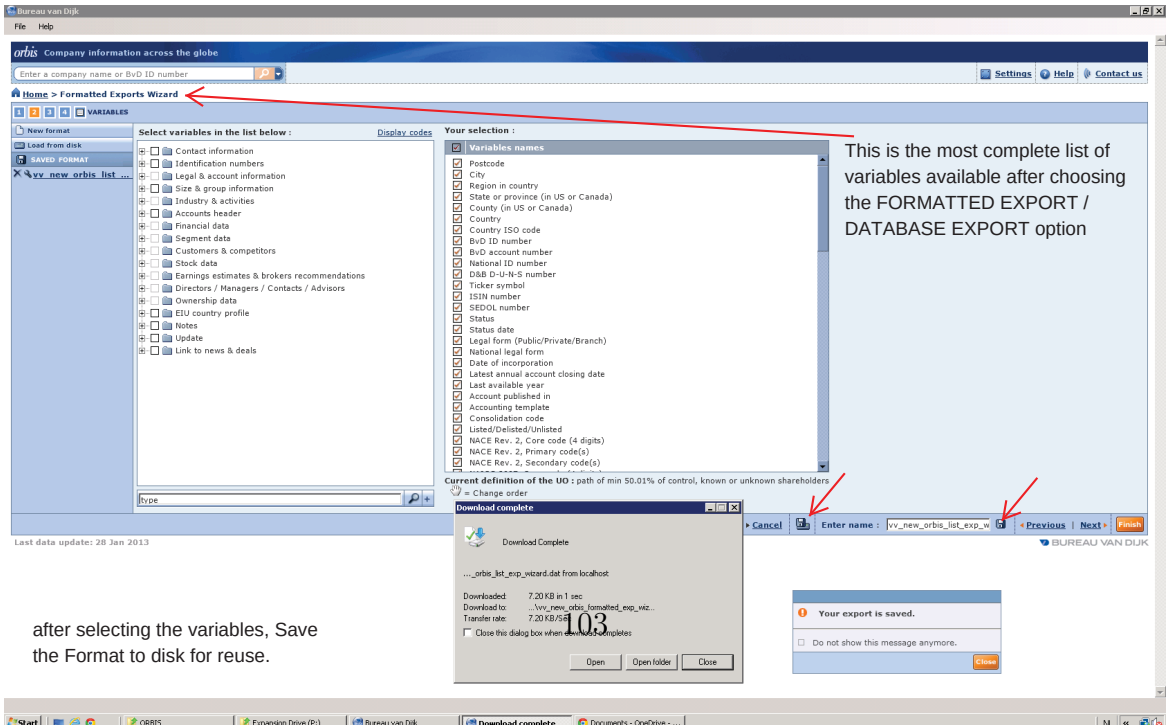

Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > Formatted Exports Wizard
toggle here such that the long names and codes are displayed
VARIABLES
New format
Load from disk
SAVED FORMAT
Select variables in the list below :
Hide codes
Your selection :
Variables names
Contact information
Street, no., building etc, line 1 (ADDR)
Street, no., building etc, line 2 (ADDR2)
Street, no., building etc, line 3 (ADDR3)
Street, no., building etc, line 4 (ADDR4)
Postcode (POSTCODE)
City (CITY)
Country (COUNTRY)
Country ISO code (CTRYISO)
Region in country (REGION_IN_COUNTRY)
Type of region in country (TYPE_REG_COUNTRY)
Metropolitan area (in US) (METRO_AREA)
State or province (in US or Canada) (STATE_US)
County (in US or Canada) (COUNTY_US_CANADA)
Telephone number (PHONE)
Fax number (FAX)
Website address (WEBSITE)
E-mail address (EMAIL)
Identification numbers
BvD ID number (BVDID)
BvD account number (BVDACC)
National ID number (NATID_NUMBER)
National ID label (NATID_LABEL)
National ID type (NATID_TYPE)
Trade register number (TRADEREGISTERNR)
VAT/Tax number (VATNUMBER)
European VAT number (EUROVAT)
CICI (CFTC Interim Compliant Identifier) (CICI)
Statistical number (STATISTICAL_CODE)
Other company ID number (CPY_ID_NUMBER)
when done selecting variables, save a named list for other countries, states, etc.
Cancel
Enter name :
My Saved Format 1
Previous
Next
Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK


Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > Formatted Exports Wizard
YEARS
New format
Load from disk
SAVED FORMAT
Select the years and the type of data for each year :
The period(s) you select will be in:
Variation in original currency
ISO code of reference currency : USD
Units : Units
Exchange rate at: each closing date
Own rate
Select one or more periods
Absolute periods Relative periods
Annual data Interim data
Last avail. year
Last avail. year - 1
Last avail. year - 2
Last avail. year - 3
Last avail. year - 4
Last avail. year - 5
Last avail. year - 6
Last avail. year - 7
Last avail. year - 8
Last avail. year - 9
Last avail. quarter
Last avail. quarter - 1
Last avail. quarter - 2
Last avail. quarter - 3
Last avail. quarter - 4
Last avail. quarter - 5
Last avail. quarter - 6
Last avail. quarter - 7
Last avail. quarter - 8
Last avail. quarter - 9
Your selection:
Name of the columns
LY | Variation in original currency
= Change order
Cancel
Enter name : vv_newer_orbis_TC
Previous Next Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK
Start
ORBIS
Expansion Drive (P:)
Bureau van Dijk
vv_newer_orbis_TC.txt ...
NL


Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > Formatted Exports Wizard
FILE OPTIONS
New format
Load from disk
SAVED FORMAT
Item labels
Export labels
Export codes
Export codes and labels
File type:
Text Delimited (*.asc)
The following file type(s) are not available for the current format:
Excel (*.xls): Nr of Rows > RowLimit (795766>65535)
Separator: ;
Text qualifier: "
Specify the directory for the generated files:
Cancel
Enter name : vv_newer_orbis_TC
Previous Next Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK
Start
ORBIS
Expansion Drive (P:)
Bureau van Dijk
vv_newer_orbis_TC.txt ...
NL

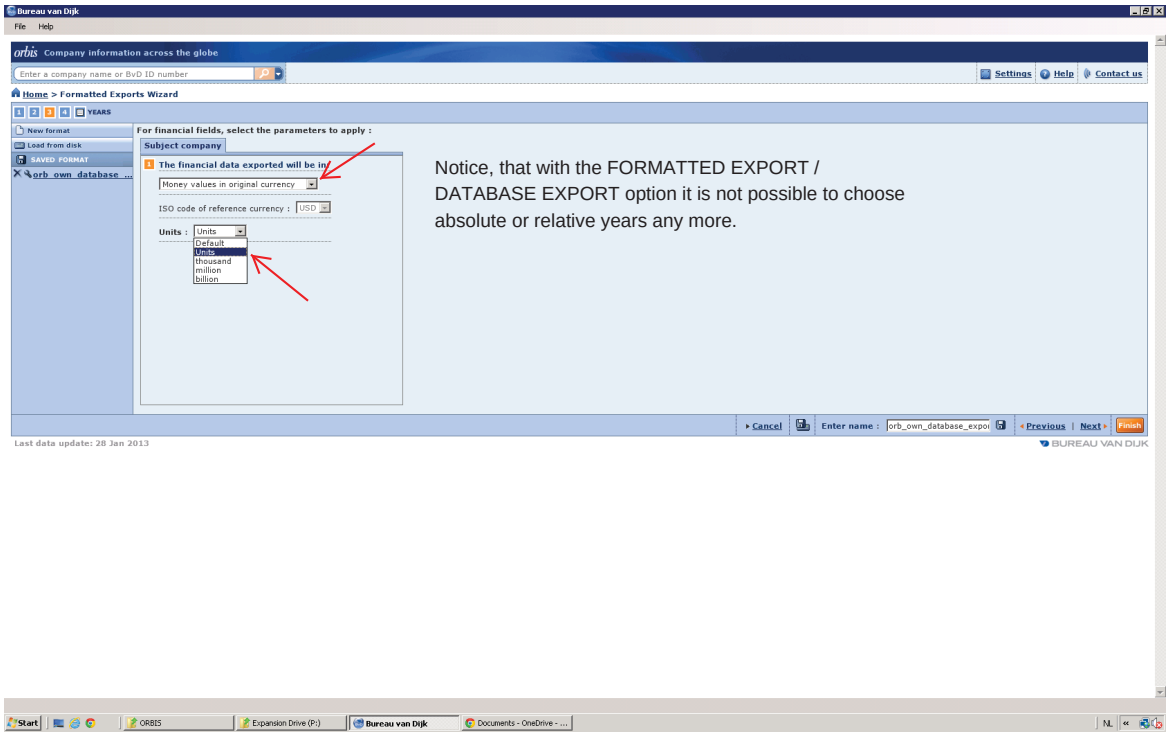
C.3 Database Export Interface of Newer BvD Disks

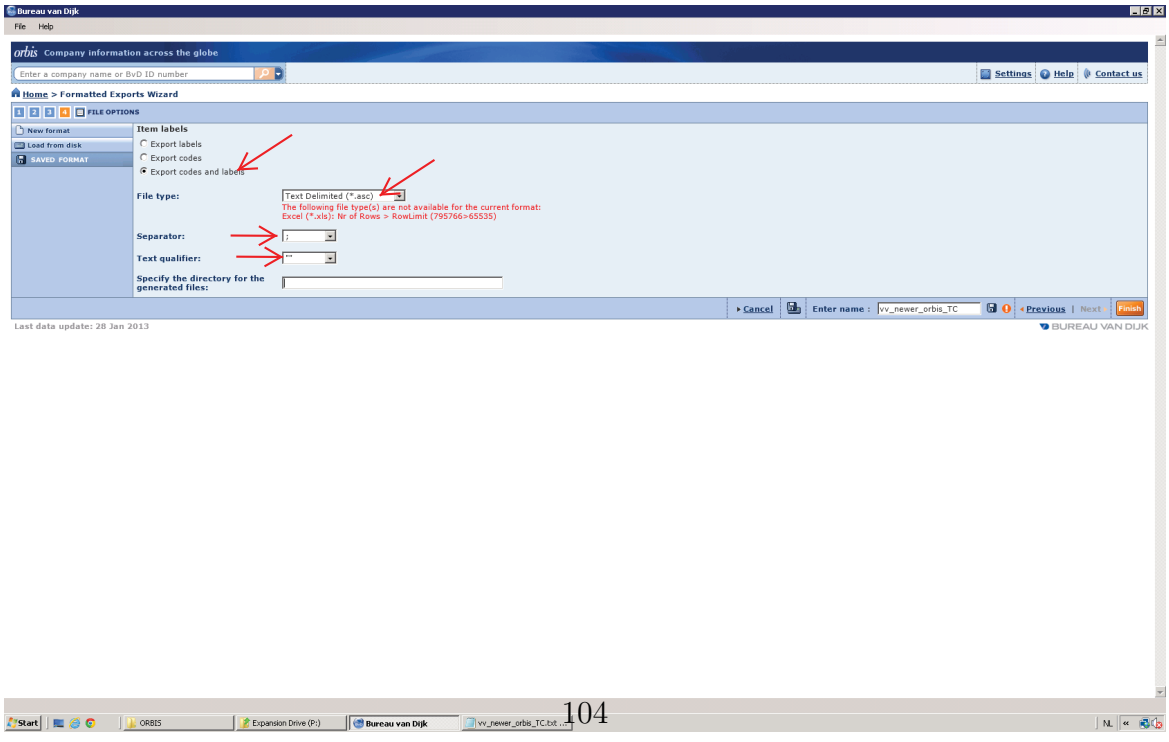




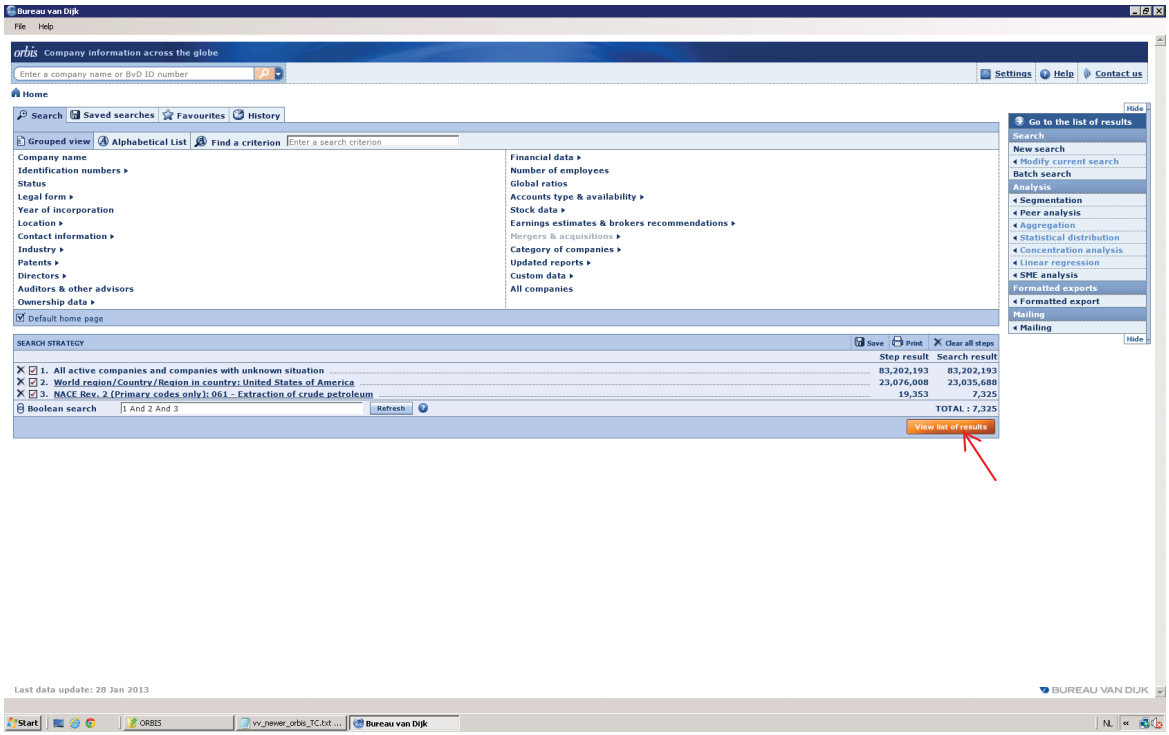
This is the most complete list of variables available after choosing the FORMATTED EXPORT / DATABASE EXPORT option

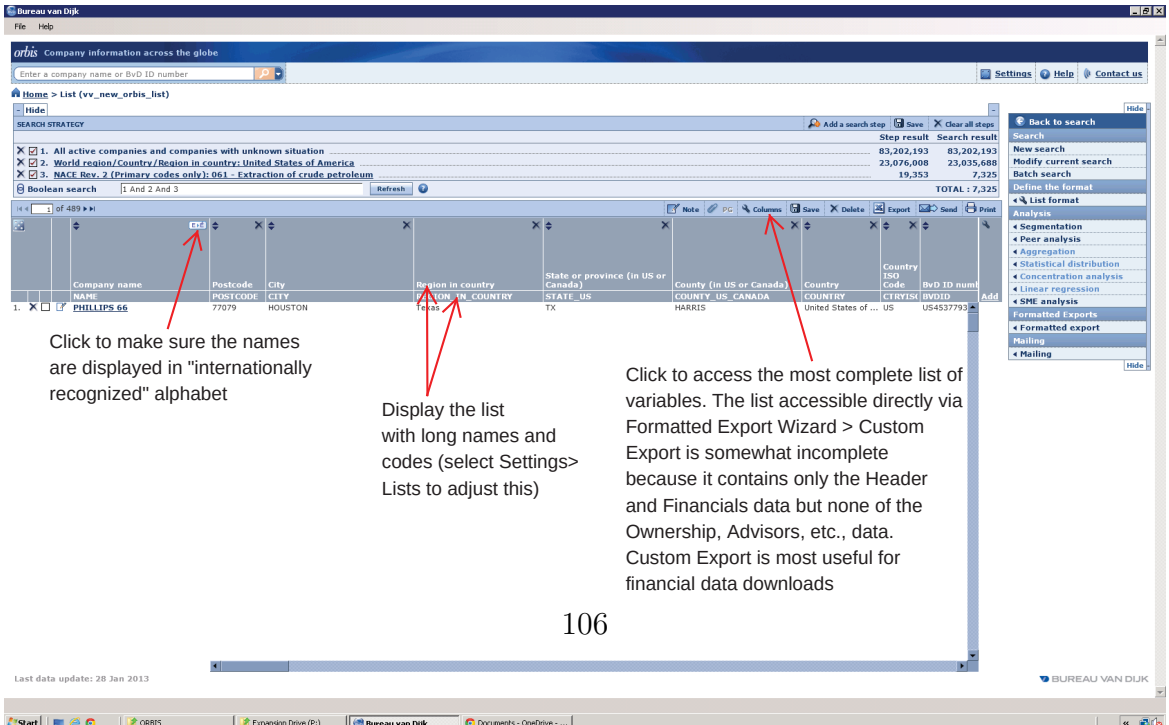
after selecting the variables, Save the Format to disk for reuse.

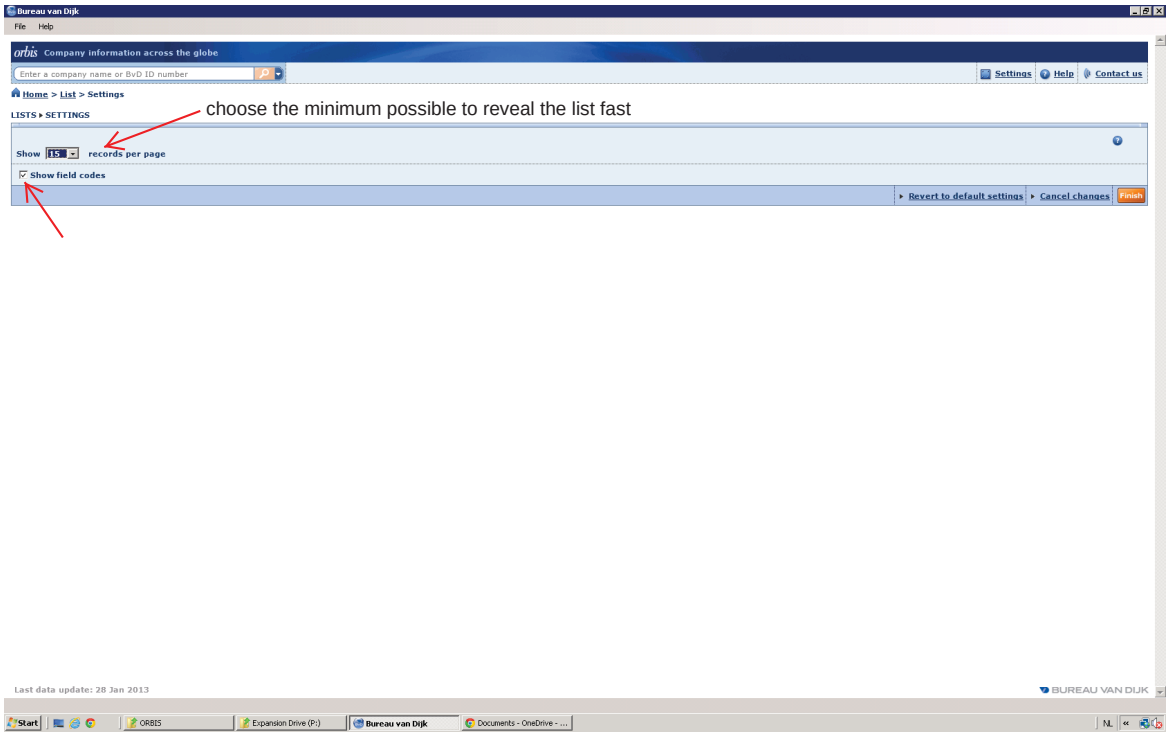

Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > Formatted Exports Wizard
YEARS
New format
Load from disk
SAVED FORMAT
orb_own_database ...
For financial fields, select the parameters to apply :
Subject company
The financial data exported will be in:
Money values in original currency
ISO code of reference currency : USD
Units : Units
Default
Units
thousand
million
billion
Notice, that with the FORMATTED EXPORT / DATABASE EXPORT option it is not possible to choose absolute or relative years any more.
Cancel
Enter name : orb_own_database_expor
Previous | Next
Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK
Start
ORBIS
Expansion Drive (P:)
Bureau van Dijk
Documents - OneDrive - ...
NL

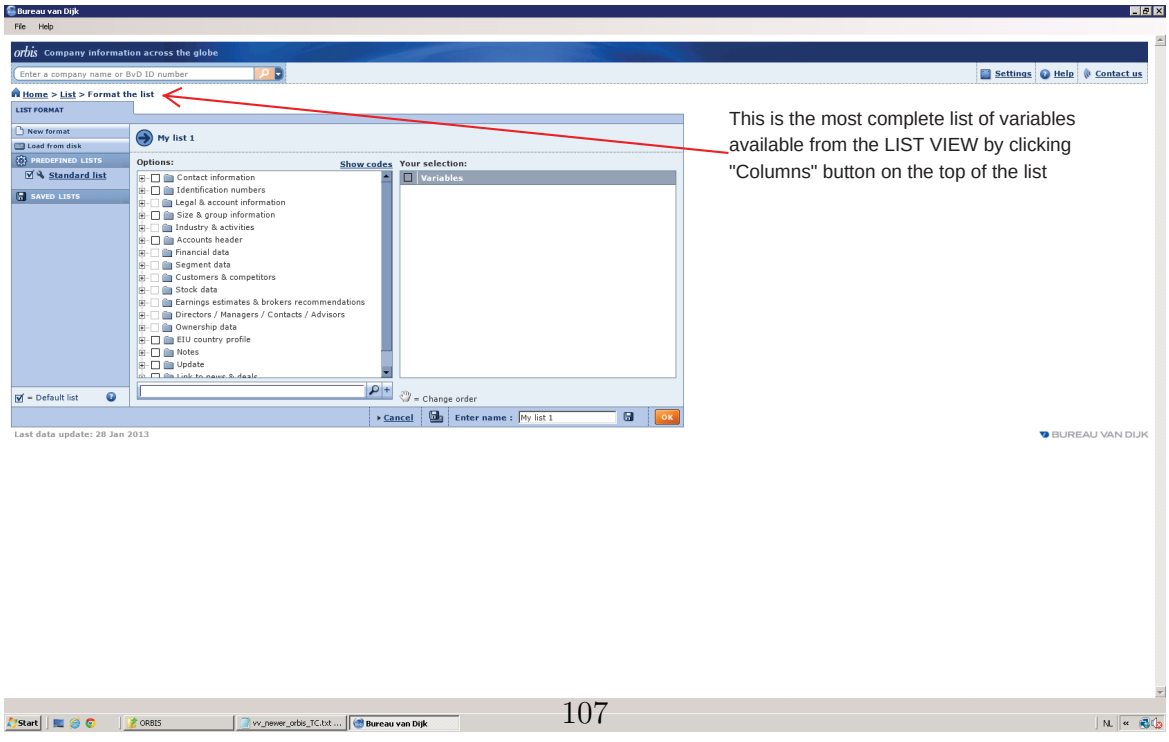

Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > Formatted Exports Wizard
FILE OPTIONS
New format
Load from disk
SAVED FORMAT
Item labels
Export labels
Export codes
Export codes and labels
File type:
Text Delimited (*.asc)
The following file type(s) are not available for the current format:
Excel (*.xls): Nr of Rows > RowLimit (795766>65535)
Separator:
;
Text qualifier:
"
Specify the directory for the generated files:
Cancel
Enter name : vv_newer_orbis_TC
Previous | Next
Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK
Start
ORBIS
Expansion Drive (P:)
Bureau van Dijk
vv_newer_orbis_TC.txt - ...
NL

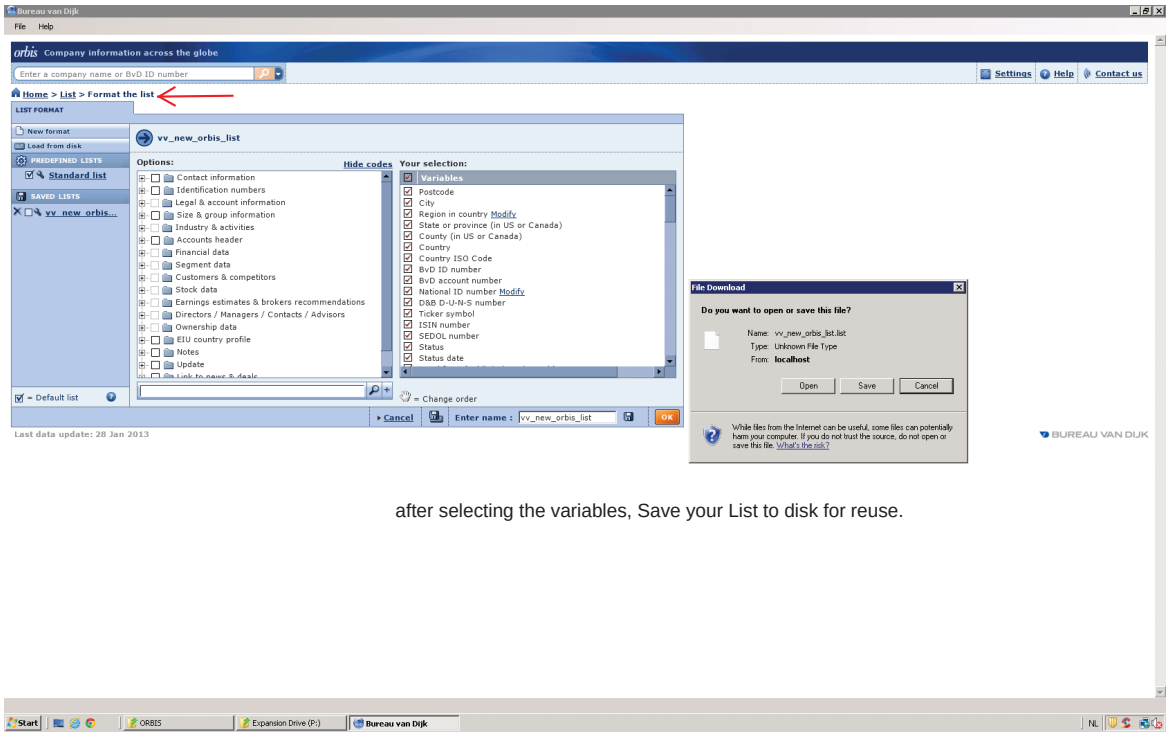
C.4 List Export Interface of Newer BvD Disks






Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > List > Settings
LISTS ▸ SETTINGS
choose the minimum possible to reveal the list fast
Show 15 records per page
Show field codes
Revert to default settings Cancel changes Finish
Last data update: 28 Jan 2013
BUREAU VAN DIJK
Start ORBIS Expansion Drive (P:) Bureau van Dijk Documents - OneDrive - ...
NL


Bureau van Dijk
File Help
orbis Company information across the globe
Enter a company name or BvD ID number
Settings Help Contact us
Home > List > Format the list
LIST FORMAT
New format
Load from disk
PREDEFINED LISTS
Standard list
SAVED LISTS
My list 1
Options:
Show codes
Your selection:
Variables
Contact information
Identification numbers
Legal & account information
Size & group information
Industry & activities
Accounts header
Financial data
Segment data
Customers & competitors
Stock data
Earnings estimates & brokers recommendations
Directors / Managers / Contacts / Advisors
Ownership data
EIU country profile
Notes
Update
= Default list
= Change order
Cancel
Enter name : My list 1
OK
This is the most complete list of variables available from the LIST VIEW by clicking "Columns" button on the top of the list
Last data update: 28 Jan 2013
BUREAU VAN DIJK
Start ORBIS vv_newer_orbis_TC.txt ... Bureau van Dijk
NL



after selecting the variables, Save your List to disk for reuse.