Introduction to Compustat Global and EMDB (WRDS)

Authors:

Kenneth In Kyun Ernst Jørgensen

Last updated: **August 2008**

AARHUS
UNIVERSITY

Table of contents

AARHUS
UNIVERSITY

1. Contents of the database

Compustat Global and Emerging Markets are two databases within WRDS. Global contains information on companies from more than 80 countries, which covers about 90 % of the world's capitalization. Historical data are available up to 12 years back. Aarhus School of Business does not have access to the Emerging Markets database. The information can be categorized within the following groups

Company Financials

- Balance sheets and income statements

- Compustat Global Industrial/Commercial (World)

- Compustat Global Financial Services (World)

Financial markets and prices

- Stock prices

- Compustat Global Issue (World)

- FX Currencies

- Compustat Global Currency

The information in the database is primarily accounting related, which e.g. can be used for valuation purposes.

1.1 Compustat Global Currency

Global Currency covers exchange rates for different cross currencies. The amount of currencies is limited compared to e.g. Datastream, which make Datastream more suited in this area.

1.2 Compustat Global Financial Services

In accounting there is a difference between balance sheets and income statements for industrial companies and financial companies. Compustat Global Financial Services covers accounting related information from financial companies within the sectors of Banking, Insurance, and Investments etc. Compared to e.g. Compustat North America Banks the section of Global Financial Services is more general and experience a lesser degree of detail. Furthermore the frequency is only annual for Global Financial Services.

1.3 Compustat Global Industrial/Commercial

Global Industrial/Commercial offers a broad access to accounting related financials. The Database is more general than Compustat North America, where North America has a higher degree of detail for the specific data items. Furhtermore is the frequency only annual for Global Financial Services, compared to Annually and Quarterly for North America. Compared to e.g. Datastream the Global Industrial/Commercial still experience a higher degree of detail. An example is items like Exchange Rate Effect and more detailed items for Inventories/Stocks.

AARHUS
UNIVERSITY

1.4 Compustat Global Issue

Global Issue offers security prices, dividends, Earnings per share, adjustment factors etc. It is less comprehensive compared to CRSP and Compustat North America Securities.

2. How to get access?

Access to Compustat North America is part of the *Wharton Research Databases Services* and is available through the computers at ASB. WRDS is a database portal, which is accessible through http://lib.asb.dk/linkdb/. Besides Compustat North America the portal offers access to Compustat Global and EMDB, CRSP and NYSE TAQ.

3. How does it work?

The following example will illustrate a search on company financials for Novo Nordisk, which is listed in Europe, and thereby available in the database. The primary searching is quite alike in WRDS in general, and the search will thereby act as an example for further searches.

3.1 Overview of company codes and tickers

Searching in WRDS is available based on different ticker symbols and codes. The following sections will outline the difference and advantages of the different codes and tickers.

3.1.1 GVKEY

Compustat has defined a proprietary identifier, the GVKEY, for each company in the database. The GVKEY can be used to track a company over time, while the company name, CUSIP, or ticker may change over time.

3.1.2 SMBL

SMBL can usually, but not always, be identified with the ticker symbol of the company's publicly traded common stock. However, the same ticker can be assigned to different companies on different exchanges and company tickers can change over time. For example, the ticker "C" has been used by both Chrysler and Citigroup. Consequently, the ticker symbol is often an unreliable way to identify companies over time and in different databases.

3.1.3 CUSIP/CNUM/CIC

CUSIP is a nine-character alphanumeric identifier assigned to individual financial assets by an independent agency. As a rule, the first six characters of the CUSIP can be identified with a company; the next two characters (the seventh and

AARHUS
UNIVERSITY

eighth) identify a particular asset (e.g., a class of stock or a bond issue) issued by the company; and the ninth digit is a "check digit" to improve the accuracy of electronic transmission of CUSIPs.

Within Compustat, the first six characters of the CUSIP are referred to as the CNUM (CUSIP issuer code), and the last three characters of the CUSIP are called the CIC (CUSIP issue code). Throughout much of the history of Compustat data files, the CNUM was used as a company identifier, but it has recently been supplanted in this function by the GVKEY. In any event, the Compustat files do not track a company's CUSIP history. Rather, only the most recent CUSIP or CNUM is included.

3.1.4 DNUM/ NAICS

Until recently, Compustat used its DNUM variable to indicate a company's main line of business. This code was generally based on the SIC (Standard Industrial Classification) code created by the U.S. Census Bureau. The Census Bureau is replacing SIC codes with NAICS (North American Industry Classification System) codes. Further, Standard & Poor's and Morgan Stanley have developed GICS (Global Industry Classification System) codes.

A company's industrial classification can change from time to time, but the DNUM in Compustat data files generally reflects only the current classification, not the history of a company's industry affiliations.

3.1.5 Country

This descriptor indentifies the country in which the company is incorporated or legally registered. The 3-character alpha codes are established by the International Standards Organization (ISO). Each company has a country of incorporation code.

AARHUS
UNIVERSITY

3.2 Example of searching in Compustat Global.

The startup in WRDS gives you the following options

By clicking on Compustat Global and EMDB the following screen appears, where the data options are located in the left-hand side.

3.2.1 Quick-Search: Global Industrial/commercial for specific company

In our search we want information on the development in Long-Term debt for Novo Nordisk. By clicking on Global Industrial/Commercial in the left-hand side the following menu appears.

The first part of our search specification is now completed. In the next step we need to specify, which parts of the indices we want to extract. In this example we want the specific development in long-term debt for Novo Nordisk.

Step Three: Variables

Company Information Select All Clear
☑ CUSIP Issuer Code/GINS Code ☐ Ticker Symbol ☑ Company Name
☐ ISO Country Code of Incorporation ☐ Industry Code
☐ N. American Industrial Classification System ☐ ISO Currency Code - Financial Data
☐ Scaling Factor - Financial Data ☐ Fiscal Yearend Month of Data ☐ Calendar Year
☐ Inactive Company Status Marker ☐ Inactive Company Status Date
☐ Company Status Alert

Accounting Related Information Select All Clear
☐ Level of Consolidation ☐ Accounting Standard ☐ Accounting Method
☐ Auditor Opinion #1 ☐ Auditor Opinion #2 ☐ Auditor Opinion #3
☐ Income Statement Model Number ☐ Balance Sheet Presentation Code
☐ Inventories/Stocks Valuation Method ☐ Format Code - Flow of Funds Statement
☐ Comparability Status ☐ Acquisition Method ☐ Record Number ☐ Number of Sets
☐ Total Number of Sets ☐ Period Duration - Weeks ☐ Period Duration - Months
☐ Update Code ☐ Source Document Code

Data Items
DATA104 -- Current Liabilities - Total
DATA105 -- Deferred Taxes (Balance Sheet)
DATA106 -- Long-Term Debt
DATA107 -- Minority Interest (Balance Sheet)
DATA108 -- Reserves - Untaxed
DATA109 -- Liabilities - Other

Data Footnotes
FTNT19 -- Depreciation, Depletion, & Amortization (Accumulated)(#78)
FTNT20 -- Debt in Current Liabilities (#94)
FTNT21 -- Accounts Payable/Creditors - Trade (#97)
FTNT22 -- Income Taxes Payable (#100)
FTNT23 -- Long-Term Debt (#106)
FTNT24 -- Taxes - Noncurrent (#112)

Specify the no. of company related information and accounting related information to supplement the data extraction. In this example we just select the CUSIP identifier and the company name.

Select Long-Term debt under data items and Long-term debt under data footnotes.By CTRL+Click you can select more items.Footnotes can contain vital information on specific cases, and is therefore included.

AARHUS UNIVERSITY

In the next we need to specify the export process.

Step Four: Output

Output Format

SAS Windows_32 dataset (*.sas7bdat)	⌄
fixed-width text (*.txt)	
comma-delimited text (*.csv)	
tab-delimited text (*.txt)	
dBase file (*.dbf)	
SAS Windows_32 dataset (*.sas7bdat)	
SAS Solaris_64 dataset (*.sas7bdat)	
SAS transport file (*.trp) (PROC CPORT)	
SAS transport file (*.xpt) (PROC COPY)	
HTML table (*.htm)	
STATA file (*.dta)	
SPSS file (*.sav)	set

How 'Save Query' works -- Using myWRDS Queries

Under output format it is possible to specify, which data format to export to. In our example we will export to a SAS dataset. WRDS offers numerous output formats into the most popular statistical packages and text files.

By clicking submit request the following window appears.

Data Request Summary

Data Request ID	200906157
Libraries/Data Sets	comp/gvic /
Frequency/Date Range	ann / 1989 - 2007
Search Variable	GVKEY
Input Codes 1 item(s)	8020
Conditional Statements	n/a
Output format/Compression	sas32 /
Variables Selected	CNUM CONAME DATA106 FINT23
Extra Variables and Parameters Selected	

Your output is complete. Click on the link below to open the output file.

g200906157.sas7bdat (16 KB, 19 observations 6 variables)

Download instructions
Netscape users... Shift-click
Internet Explorer users... Right-click and select "Save Target As..."

Right-click on the link, and choose save as. The SAS dataset can now be imported into the SAS System.

AARHUS
UNIVERSITY

3.3 Conditional search

Another option is to specify conditional search through the entire database. If we were interested in locating companies with long-term debt above 10,000 (mill.), we could enter the following search conditions.
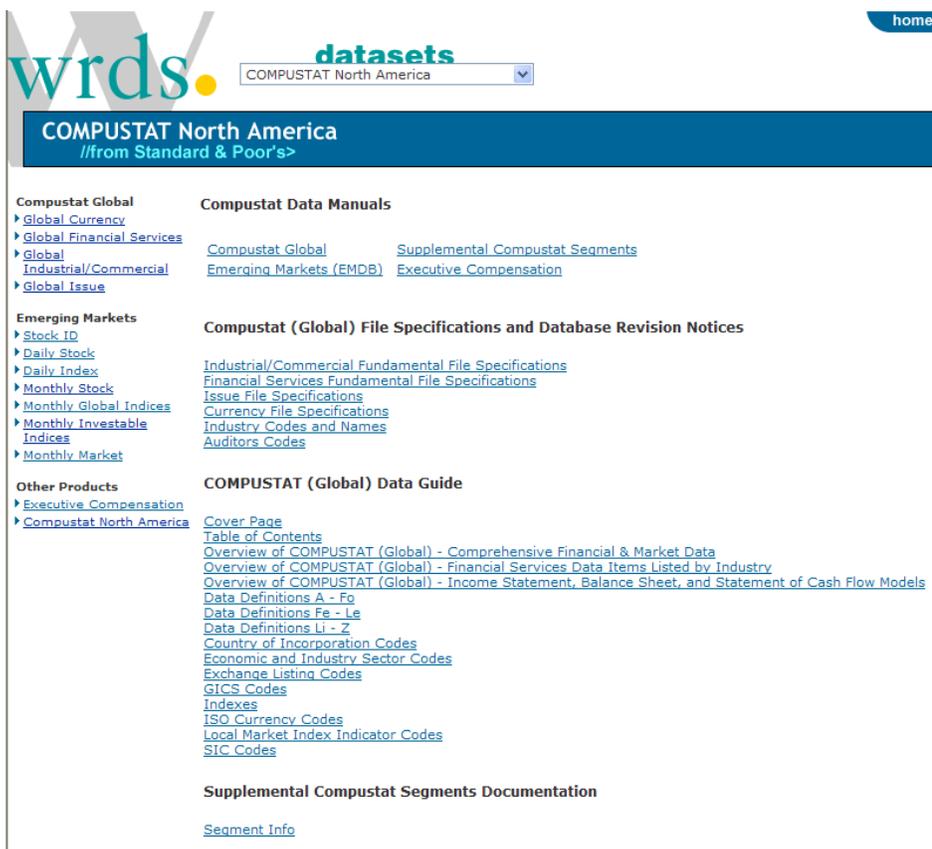
4. Further resources

WRDS offers various manuals online that readily explains the different data types and how to export the data. Usually the manuals are available at the top of each category.

The following manuals are available online for Compustat Global by clicking on Data Manuals.

THIS GUIDE HAS BEEN PRODUCED BY

ANALYTICS GROUP

ADVANCED MULTIMEDIA GROUP

Analytics Group, a division comprised of student instructors under AU IT, primarily offers support to researchers and employees.

Our field of competence is varied and covers questionnaire surveys, analyses and processing of collected data etc. AG also offers teaching assistance in a number of analytical resources such as SAS, SPSS and Excel by hosting courses organised by our student assistants. These courses are often an integrated part of the students' learning process regarding their specific academic area which ensures the coherence between these courses and the students' actual educational requirements.

In this respect, AG represents the main support division in matters of analytical software.

Advanced Multimedia Group is a division under AU IT supported by student instructors. Our primary objective is to convey knowledge to relevant user groups through manuals, courses and workshops.

Our course activities are mainly focused on MS Office, Adobe CS and CMS. Furthermore we engage in e-learning activities and auditive and visual communication of lectures and classes. AMG handles video assignments based on the recording, editing and distribution of lectures and we carry out a varied range of ad hoc assignments requested by employees.

In addition, AMG offers solutions regarding web development and we support students' and employees' daily use of typo3.

PLEASE ADDRESS QUESTIONS OR COMMENTS REGARDING THE CONTENTS OF THIS GUIDE TO

ANALYTICS@ASB.DK

AARHUS
UNIVERSITY