

Introduction to CRSP (WRDS)

Authors:

Kenneth In Kyun Jørgensen

Last updated:

August 2008

Table of contents

1. CONTENTS OF THE DATABASE	1
1.1 Monthly Stocks (Annually/Quarterly).....	1
1.1.1 Stocks	1
1.1.2 Indices.....	1
1.2 Daily Stocks (Annually/Quarterly)	1
1.2.1 Stocks	1
1.2.2 Indices.....	1
1.2.3 Events.....	1
1.2.4 Daily Stock Header	2
1.3 Indices and Deciles (Annually/Quarterly)	2
1.4 Monthly Treasuries (Annually/Quarterly)	2
1.5 CRSP/Compustat Merged (Annually/Quarterly)	2
2. HOW TO GET ACCESS?.....	2
3. HOW DOES IT WORK?	2
3.1 Overview of company codes and tickers.....	2
3.1.1 PERMNO	2
3.1.2 PERMCO	3
3.1.3 CUSIP/NCUSIP	3
3.1.4 TICKER	3
3.2 Example of searching in CRSP	4
3.3 Quick-Search: Monthly Stocks - securities	5
3.4 Conditional search	7
4.1 FURTHER RESOURCES.....	8

1. Contents of the database

Center for Research in Securities and Prices (CRSP) contains a comprehensive database on security prices, returns and volume data for the NYSE, AMEX and NASDAQ exchanges. CRSP contains information on stocks, bonds and indices. The information can be categorized within the following groups.

Financial markets and prices

- Stock prices
 - o CRSP Stock Monthly/Daily
- Market indices and factors
 - o CRSP Indices and Deciles
- Bonds
 - o CRSP Monthly/Daily treasuries

The information is suitable for time-series analysis and event studies. The identification tickers offer reliable tracking features for uninterrupted time-series analysis.

1.1 Monthly Stocks (Annually/Quarterly)

1.1.1 Stocks

This section offers data on monthly securities. Relevant data items are Opening/Closing prices, Volume, Holding period returns with and without dividends, Shares outstanding etc.

1.1.2 Indices

This section offers data on monthly equity indices. The data available covers returns for Value-Weighted and Equally-Weighted indices with and without dividends. Furthermore the return on the S&P Composite index is available. The data is convenient for calculating Beta values by regressions.

1.2 Daily Stocks (Annually/Quarterly)

1.2.1 Stocks

This section offers data on daily securities. Relevant data items are the same as mentioned in section 0, and furthermore Beta Excess Return, Standard Deviation Excess Return. These added categories are necessary items in an Event study.

1.2.2 Indices

The information is the same as in section 0 only with daily frequency.

1.2.3 Events

In order to construct an event study information on the specific events must be obtained. This section provides the opportunity to search for specific company related events by Names, which tracks changes to Company Names, CUSIP numbers, Share Classes etc.

1.2.4 Daily Stock Header

Stock header provides the opportunity for an advanced screening of company events by e.g. Delisting structures, Names changes structures, No. of issued shares structures etc. A complete listing of the data contents for Daily Stock Header is available through.

1.3 Indices and Deciles (Annually/Quarterly)

The section offers various combinations on different portfolios based on CRSP securities. A feature of this section is the categorization of index deciles based on e.g. a market cap of size, a categorization according to Beta values and a categorization according to standard deviations. A nice feature of the categorizations is that they allow a user to specify the specific market indices in order for a more targeted approach in e.g. Beta calculations.

1.4 Monthly Treasuries (Annually/Quarterly)

This section provides Bond data for the U.S. market. Data includes returns, prices, Yield to maturities etc. It is possible to specify, which bonds to analyze e.g. non-callable or callable bonds. The section FAMA risk free rates contains short term risk free rates based on different maturities. It is also possible to find information on Bond portfolio returns.

1.5 CRSP/Compustat Merged (Annually/Quarterly)

The CRSP/Compustat Merged database is a combination of CRSP and Compustat North America in order to establish a unique tracking of listed securities from both databases. In combination with an event study it might be relevant to investigate the event's impact on different parts of the balance sheet or income statements. The linkage between the two databases enables the user to be sure that the companies at stake are uniquely identifiable. For an introduction to the Compustat North America database see the Introduction Manual available.

2. How to get access?

Access to CRSP is part of the *Wharton Research Databases Services* (WRDS) and is available for students and employees at AU. It can be accessed through <https://wrds-web.wharton.upenn.edu/wrds/index.cfm> after registration has been completed.

3. How does it work?

The following example will illustrate a search on security prices for Novo Nordisk, which is listed on New York Stock Exchange, 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. A translation of different codes and tickers can be done with the company search tool in WRDS. A specific translator for PERMNO and PERMCO is available under Tools in the left hand side menu.

3.1.1 PERMNO

Every stock issue covered within CRSP is assigned a unique number, its PERMNO. While a company may change its name, ticker, exchange, or CUSIP from time to time, the PERMNO of a stock within CRSP remains unchanged. Consequently, the PERMNO is the principal identifier of a stock in the CRSP database, and provides a reliable way of tracking a stock over

time. As discussed immediately below, a PERMNO identifies an individual stock. For the most part, companies will have only one issue of common stock on the market at any time, so there is usually a unique link between PERMNO and company. However, a company with multiple common stock issues on the market will be associated with multiple PERMNOs.

3.1.2 PERMCO

If a company has issued multiple tranches of publicly traded common stock, CRSP will track all of them individually with different PERMNOs. The CRSP PERMCO is a number that uniquely identifies each company within the CRSP database. This makes it possible to associate different stock issues from the same company. Most companies have only a single common stock issue outstanding at any time, so there is usually a one-to-one correspondence between PERMNO and PERMCO. As of December 2003, for example, CRSP was covering three different series of common stock (A, E, and H) related to General Motors. These three stocks have different PERMNOs within CRSP, but the common PERMCO indicates that the stocks belong to the same company.

3.1.3 CUSIP/NCUSIP

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 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 CRSP, 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). CUSIP codes can change, which can pose a problem, when tracking data series over a long date range. In order to account for any changes it is recommended to use NCUSIP, since this will keep track of the same company throughout the date range specified.

3.1.4 TICKER

Ticker symbols can change over time. Since the Symbols can be reused for different companies any user of ticker symbols must be aware of this and consequently make sure to validate any data extracted on the basis of Ticker symbols.

3.2 Example of searching in CRSP

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.

3.3 Quick-Search: Monthly Stocks - securities

In our search we want information on the development in security prices for clicking on Monthly Stocks in the left-hand side the following menu appears.

The screenshot shows the WRDS datasets interface for the CRSP dataset. The search process is outlined with three callout boxes:

- Specify the data range:** Points to the "Step One: Date Range" section where "Beginning" is set to Jan 1993 and "Ending" is set to Dec 2007.
- Specify Ticker: PERMNO:** Points to the "Step Two: Search" section where "Search By" is set to PERMNO and the value "63263" is entered in the "Code Lookup" field.
- Specify company name:** Points to the search bar at the bottom left containing the text "Novo Nordisk".

At the bottom, the results table shows one match found:

PERMNO	CUSIP	NCUSIP	HSICCD	TICKER	COMNAM	DATE RANGE
63263	67010020	67010020	2833	NVO	NOVO NORDISK A S	19810430-20071231

The first part of our search specification is now completed. In the next step we need to specify, which parts of the financials we want information on.

In the next we need to specify the export process.

Step Three: Variables

Identifying Information	Select All Clear			
<input type="checkbox"/> Cusip	<input type="checkbox"/> Ncusip	<input checked="" type="checkbox"/> Company Name	<input type="checkbox"/> Ticker	<input checked="" type="checkbox"/> CRSP Permanent Company Number
<input type="checkbox"/> Share Code	<input type="checkbox"/> Share Class	<input type="checkbox"/> Nasdaq Issue Number	<input type="checkbox"/> Exchange Code	
<input type="checkbox"/> Header Exchange Code	<input type="checkbox"/> SIC Code	<input type="checkbox"/> Header SIC Code	<input type="checkbox"/> Header SIC Major Group	
<input type="checkbox"/> Header SIC Industry Group	<input type="checkbox"/> Names Ending Date	<input type="checkbox"/> Trading Symbol		
<input type="checkbox"/> North American Industry Class System	<input type="checkbox"/> Primary Exchange	<input type="checkbox"/> Trading Status		
<input type="checkbox"/> Security Status				
Time Series Information	Select All Clear			
<input checked="" type="checkbox"/> Price	<input type="checkbox"/> Price Alternate	<input type="checkbox"/> Price Alternate Date	<input type="checkbox"/> Ask or High	<input type="checkbox"/> Bid or Low
<input checked="" type="checkbox"/> Closing Bid	<input checked="" type="checkbox"/> Closing Ask	<input type="checkbox"/> Share Volume		
<input type="checkbox"/> Holding Period Return	<input type="checkbox"/> Holding Period Return without Dividends			
Share Information	Select All Clear			
<input type="checkbox"/> Number of Shares Outstanding	<input type="checkbox"/> Shares Observation End Date	<input type="checkbox"/> Share Flag		
Delisting Information	Select All Clear			
<input type="checkbox"/> Delisting Code	<input type="checkbox"/> New CRSP Permo	<input type="checkbox"/> Date of Next Available Information		
<input type="checkbox"/> Amount After Delisting	<input type="checkbox"/> Delisting Price	<input type="checkbox"/> Date of Delisting Payment	<input type="checkbox"/> Delisting Return	
<input type="checkbox"/> Delisting Return without Dividends				

Select company name and CRSP permanent company number for identification

Select price, closing bid and closing ask under Time series information

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 CPRT)	
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	201126508
Libraries/Data Sets	crsp/msf /
Frequency/Date Range	mon / Jan 2000 - Dec 2007
Search Variable	PERMO
Input Codes 1 item(s)	63263
Conditional Statements	n/a
Output format/Compression	sas32 /
Variables Selected	COMMAM PERMO PRC BID ASK
Extra Variables and Parameters Selected	

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

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

[q201126508.sas7bdat \(24 KB, 96 observations 7 variables\)](#)

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

3.4 Conditional search

Another option is to specify conditional search through the entire database. If we were interested in locating companies with a holding period return above 10 % for the specified period, we could enter the following search conditions.

CRSP Monthly Stock

[Data Query](#) | [Data Contents](#) | [Documentation](#) | [Retrieve from myWRDS Queries](#)

Step One: Date Range

Beginning <input type="button" value="Jan"/> <input type="button" value="2006"/>	Select frequency and range of data.
Ending <input type="button" value="Dec"/> <input type="button" value="2007"/>	

Step Two: Search

Search By <input type="button" value="PERMNO"/> Using Ticker Symbols and CUSIPs	Select the variable used to search the database. Enter the companies using 1 of 3 methods.
Choose 1 Of 3 Methods	1. Codes separated by a single space. Example: ibm msft dell
<input checked="" type="radio"/> 1. Company Codes <input type="text"/> Code Lookup	2. A text file (File Format) on your local computer with codes entered one per line. Example: ibm msft dell
<input type="radio"/> 2. File Containing Company Codes <input type="text"/> Browse...	3. All companies
<input checked="" type="radio"/> 3. Entire Database	
Conditional Statements (Optional)	Optional Build conditional statements using the list of variables, comparison operators, desired values, and logical operators. (see data contents for more information). Example: RET > 0.02 AND VOL > 1000
Holding Period Return(RET) <input type="button" value=""/> > <input type="text" value="0.10"/>	Note If the closing price is not available for any given period, the number in the price field is replaced with a bid/ask average. Bid/ask averages have dashes placed in front of them. These do not wrongly reflect negative prices, but using condition like PRC > 5 will filter out the observations with bid/ask average > 5. To solve this issue, you can use PRC > 5 OR PRC < -5.
<input checked="" type="radio"/> AND <input type="radio"/> OR	
--- Select a Variable --- <input type="button" value=""/> > <input type="text"/>	

4.1 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.

- [Annual Update](#) [Quarterly Update](#)
- CRSP Monthly Stock**
- Data Query | Data Contents | Documentation | Retrieve from myWRDS Queries
- ▶ [Monthly Stocks \(A\)](#)

The following manuals are available online for CRSP

The screenshot shows the WRDS datasets page for the CRSP database. The main navigation bar includes links for home, my.wrds, tools, and support. The left sidebar contains a navigation tree for CRSP data types, including Annual, Quarterly, and Monthly updates, as well as sections for Daily Stocks, Indices and Deciles, Monthly Treasuries, Daily Treasuries, CRSP/Compustat Merged, Tools, Other Products, Mutual Funds, and CRSP/Ziman Real Estate. The right panel displays the 'CRSP Monthly Stock' manual, which includes links for Data Overview, Data Manuals, Tools, User Guides, and Programmers Guide. The 'Data Overview' section lists links for the CRSP/COMPSTAT Merged Database Overview and CRSP Indices Overview. The 'Data Manuals' section lists links for the CRSP/COMPSTAT Merged Database (CCM), CRSP Daily US Treasury Guide, CRSP Monthly US Treasury Guide, CRSP Survivor-Bias-Free US Mutual Funds Guide, and CRSP/Ziman Real Estate Data Series Guide. The 'Tools' section lists links for Data Descriptions Guide, Utilities Guide, and FORTRAN/77 to FORTRAN/95 Migration Guide. The 'User Guides' section lists links for User Guide for FORTRAN-95 and The SASECRSP Engine. The 'Programmers Guide' section lists links for C and FORTRAN programming information for the stock and indices database.

THIS GUIDE HAS BEEN PRODUCED BY**ANALYTICS 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

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