

CRSP_DP.exe Utility

September 2011



Université Lille 2
Droit et Santé



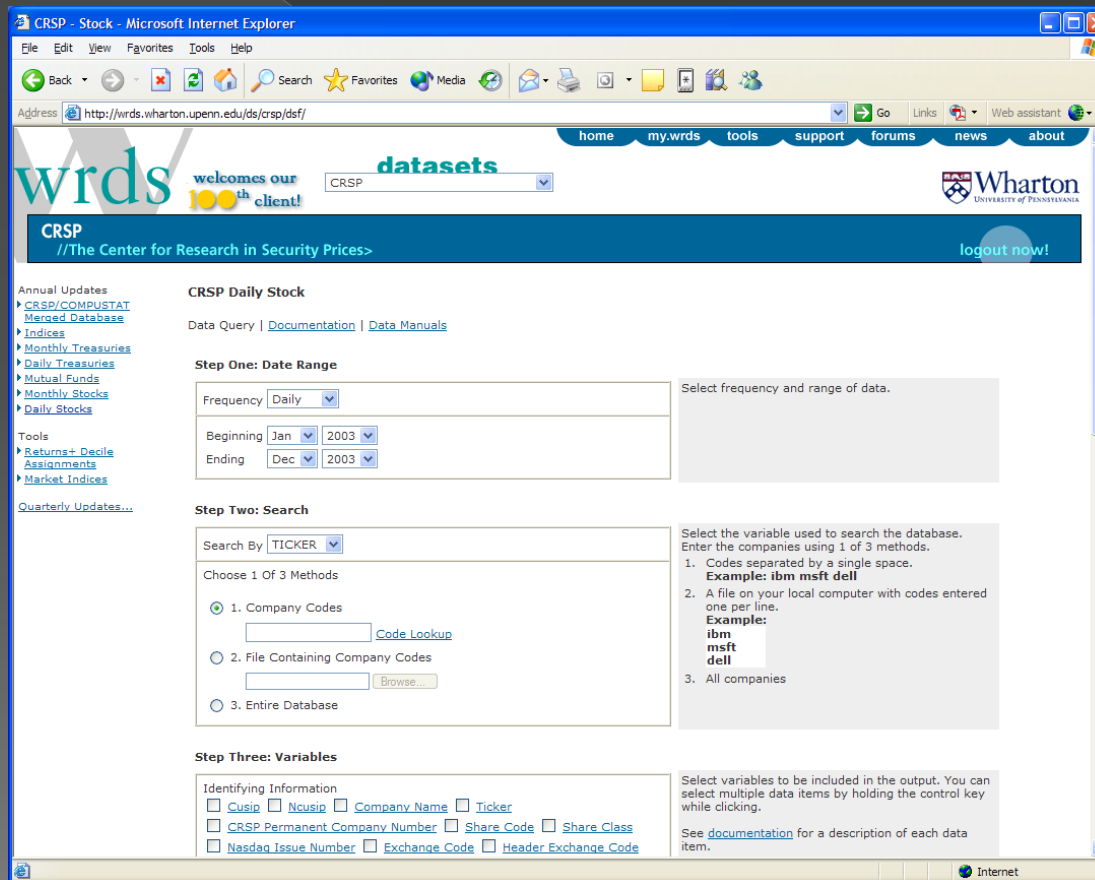
Outline

- Introduction
 - > Accessing CRSP/Compustat data through WRDS
 - > CRSP_DP.exe goals
- General Presentation
 - > Outputs
 - > Coverage
 - > Installation
 - > Interface
- Usage
 - > Let's go
 - > Options
- Event Study



Accessing CRSP/Compustat data through WRDS

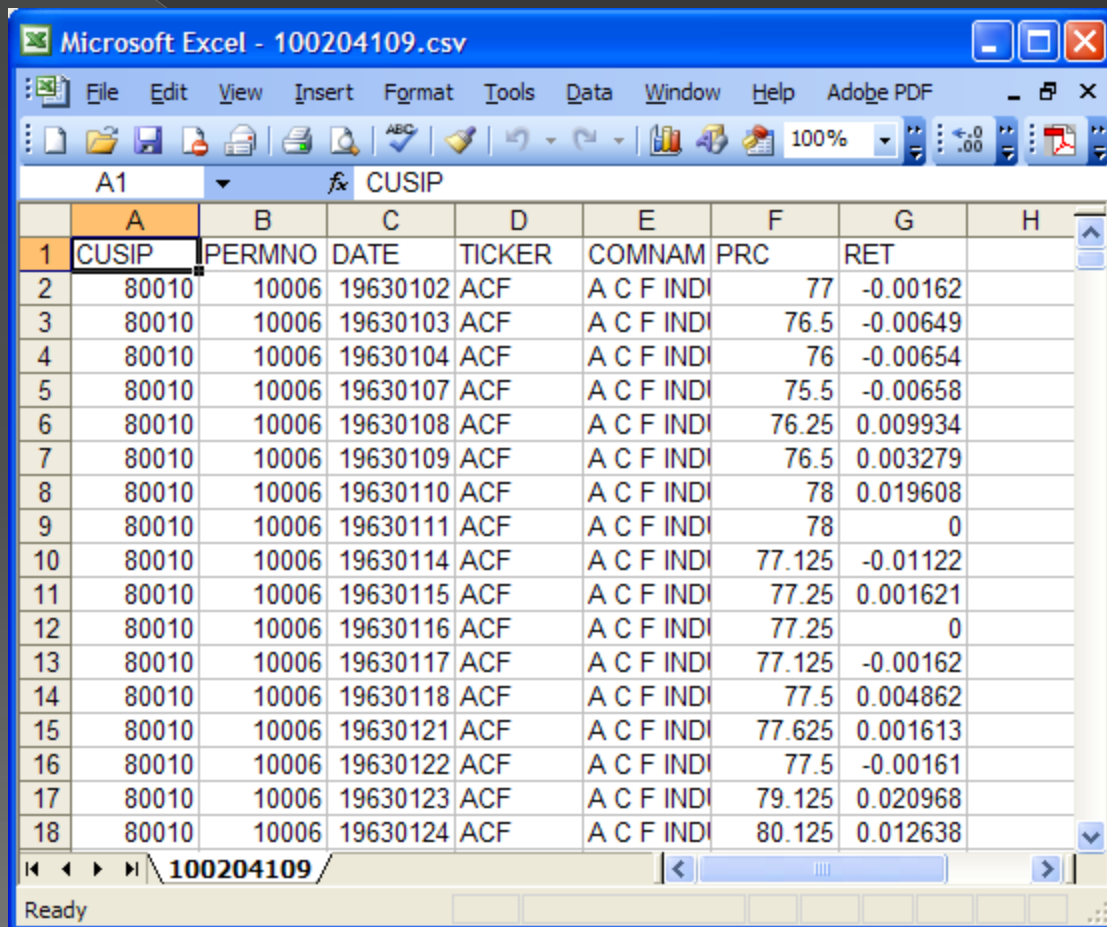
- The WRDS Interface is user-friendly



Accessing CRSP/Compustat data through WRDS

- But for large data extractions (typical for research projects):
 - > The output is quickly very large. This is due to the output file organization :
 - Common fields are replicated for each date
 - For 1,000 assets, 10 years, and 2 fields, we receive more than 5,000,000 records in one single flat file.
 - For the year 2003 and in daily frequency, extracting all fields for all issues generate a file of more 400 Mo
 - > Alpha-numeric error codes are difficult to handle under most econometric software
 - > Date matching between CRSP and Computstat was a tricky process
 - > Many “by hand” data cleaning remains to be done (duplicate records, ...), every one doing it on his own way
- Compustat and CRSP data are merged but not all automatically ...

Accessing CRSP/Compustat data through WRDS



Microsoft Excel - 100204109.csv

File Edit View Insert Format Tools Data Window Help Adobe PDF

100%

A1 CUSIP

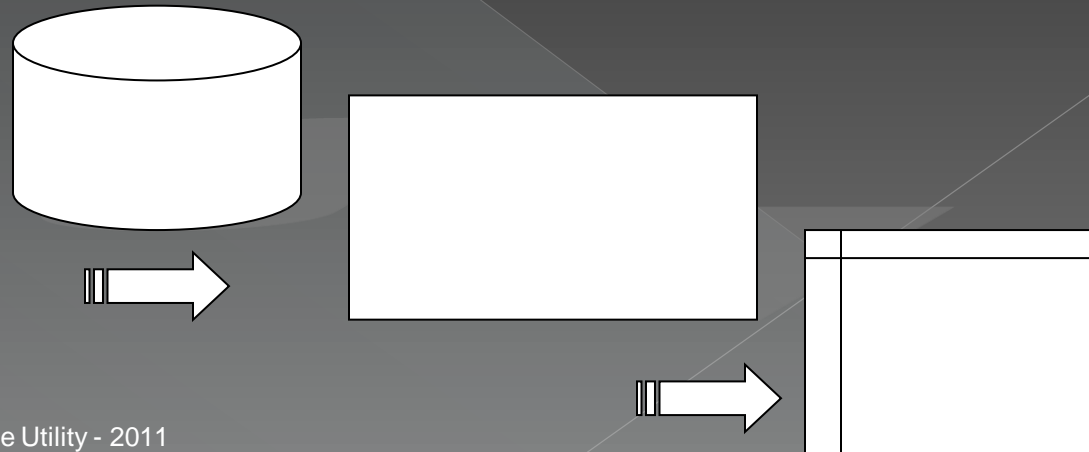
	A	B	C	D	E	F	G	H
1	CUSIP	PERMNO	DATE	TICKER	COMNAM	PRC	RET	
2	80010	10006	19630102	ACF	A C F INDI	77	-0.00162	
3	80010	10006	19630103	ACF	A C F INDI	76.5	-0.00649	
4	80010	10006	19630104	ACF	A C F INDI	76	-0.00654	
5	80010	10006	19630107	ACF	A C F INDI	75.5	-0.00658	
6	80010	10006	19630108	ACF	A C F INDI	76.25	0.009934	
7	80010	10006	19630109	ACF	A C F INDI	76.5	0.003279	
8	80010	10006	19630110	ACF	A C F INDI	78	0.019608	
9	80010	10006	19630111	ACF	A C F INDI	78	0	
10	80010	10006	19630114	ACF	A C F INDI	77.125	-0.01122	
11	80010	10006	19630115	ACF	A C F INDI	77.25	0.001621	
12	80010	10006	19630116	ACF	A C F INDI	77.25	0	
13	80010	10006	19630117	ACF	A C F INDI	77.125	-0.00162	
14	80010	10006	19630118	ACF	A C F INDI	77.5	0.004862	
15	80010	10006	19630121	ACF	A C F INDI	77.625	0.001613	
16	80010	10006	19630122	ACF	A C F INDI	77.5	-0.00161	
17	80010	10006	19630123	ACF	A C F INDI	79.125	0.020968	
18	80010	10006	19630124	ACF	A C F INDI	80.125	0.012638	

100204109

Ready

CRSP_DP.exe Goals

- Automatic Data management :
 - > “clean” data matrices generation
 - > error code handling
 - > inclusion/exclusion lists handling (Permnos, Sic codes, Share Codes)
 - > date matching between CRSP data and Compustat data (look ahead bias, ...)
 - > calendar time – event time switching for event studies
- Collective productivity improvement



Outline

- Introduction
 - > Accessing CRSP/Compustat data through WRDS
 - > CRSP_DP.exe goals
- General Presentation
 - > Outputs
 - > Coverage
 - > Installation
 - > Interface
- Usage
 - > Let's go
 - > Options
- Utilities
 - > File Copier
 - > Multiple Issue Cleaner
 - > Event Study
- Conclusion



Outputs

- **Matrix Output**

- > A cube with, for each field, the same plane:

PERMNO	2/1/2001	3/1/2001	...
100001			
100002			
...			

Field Values

- **Flat File Output** : each row is transposed and the resulting column vectors are stacked. The processing is very efficient under Gauss using the load and reshape commands
- **Other output format** : everything possible (eg: xml, ...)

Coverage

- CRSP
 - > Daily database 1962-2007
 - > Monthly database 1925-2007
- Datastream (France, Germany, UK, Italy, Spain)
 - > Daily database 1970-2006
 - > Monthly database 1970-2006
- Compustat Merge Database
 - > Old format: Bank Annual, Bank Quarterly, Industrial Annual, Industrial Quarterly, Price/Dividends
 - > New format: Fundamentals Annual, Fundamentals Quarterly, Security Monthly
- CRSP Indices
- REITS Data daily/monthly
- REITS Indices daily/monthly

Installation

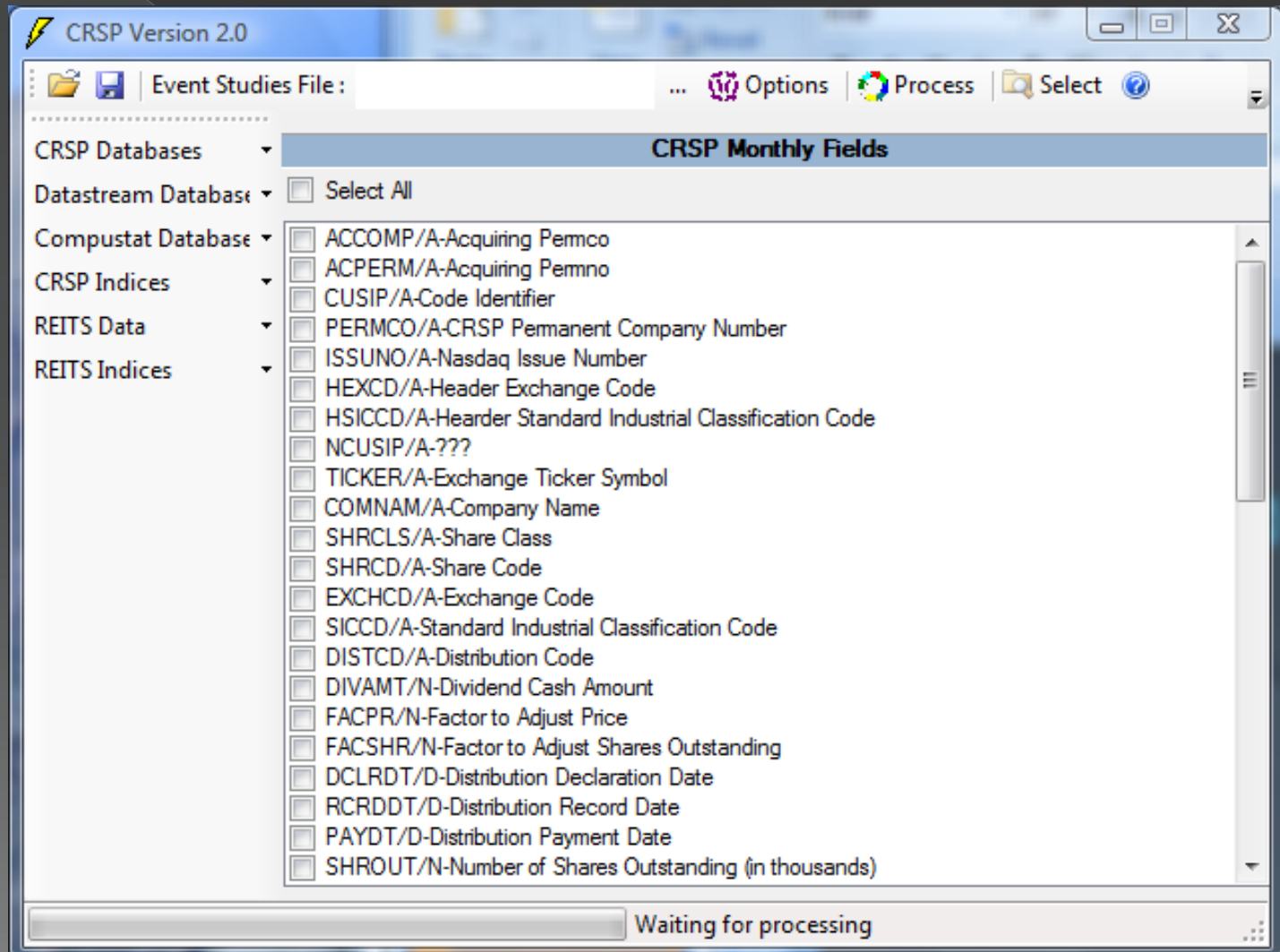
⦿ Requirements

- > A good hard drive (external USB powered HD for laptop are ok)
- > 2GoRAM, a good Dual Core processor
- > At least Windows XP + SP2

⦿ Installation

- > Program setup can be done through Internet
- > The database (> 100 Gb) must be copied from disk to disk

Interface



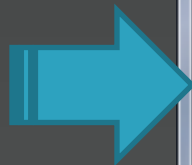
Outline

- Introduction
 - > Accessing CRSP/Compustat data through WRDS
 - > CRSP_DP.exe goals
- General Presentation
 - > Outputs
 - > Coverage
 - > Installation
 - > Interface
- Usage
 - > Let's go
 - > Options
- Event Study



Usage

- 1. Select a Database



CRSP Version 2.0

Event Studies File : ... Options Process Select

CRSP Databases ▼ **CRSP Monthly Fields**

Datastream Database ▼ Select All

Compustat Database ▼

CRSP Indices ▼

REITS Data ▼

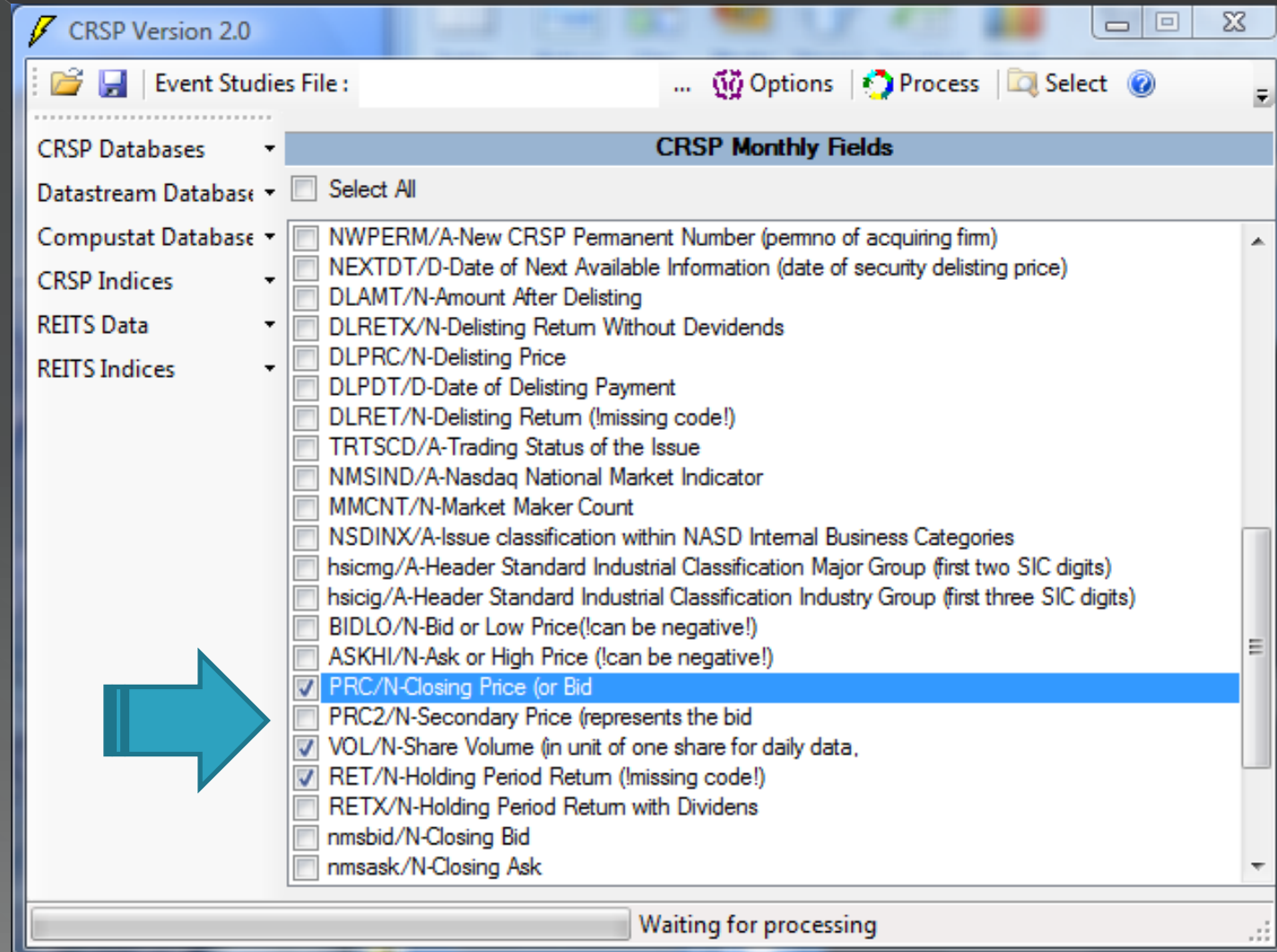
REITS Indices ▼

- ACCOMP/A-Acquiring Permco
- ACPERM/A-Acquiring Permno
- CUSIP/A-Code Identifier
- PERMCO/A-CRSP Permanent Company Number
- ISSUNO/A-Nasdaq Issue Number
- HEXCD/A-Header Exchange Code
- HSICCD/A-Header Standard Industrial Classification Code
- NCUSIP/A-???
- TICKER/A-Exchange Ticker Symbol
- COMNAM/A-Company Name
- SHRCLS/A-Share Class
- SHRCD/A-Share Code
- EXCHCD/A-Exchange Code
- SICCD/A-Standard Industrial Classification Code
- DISTCD/A-Distribution Code
- DIVAMT/N-Dividend Cash Amount
- FACPR/N-Factor to Adjust Price
- FACSHR/N-Factor to Adjust Shares Outstanding
- DCLRDT/D-Distribution Declaration Date
- RCRDDT/D-Distribution Record Date
- PAYDT/D-Distribution Payment Date
- SHROUT/N-Number of Shares Outstanding (in thousands)

Waiting for processing

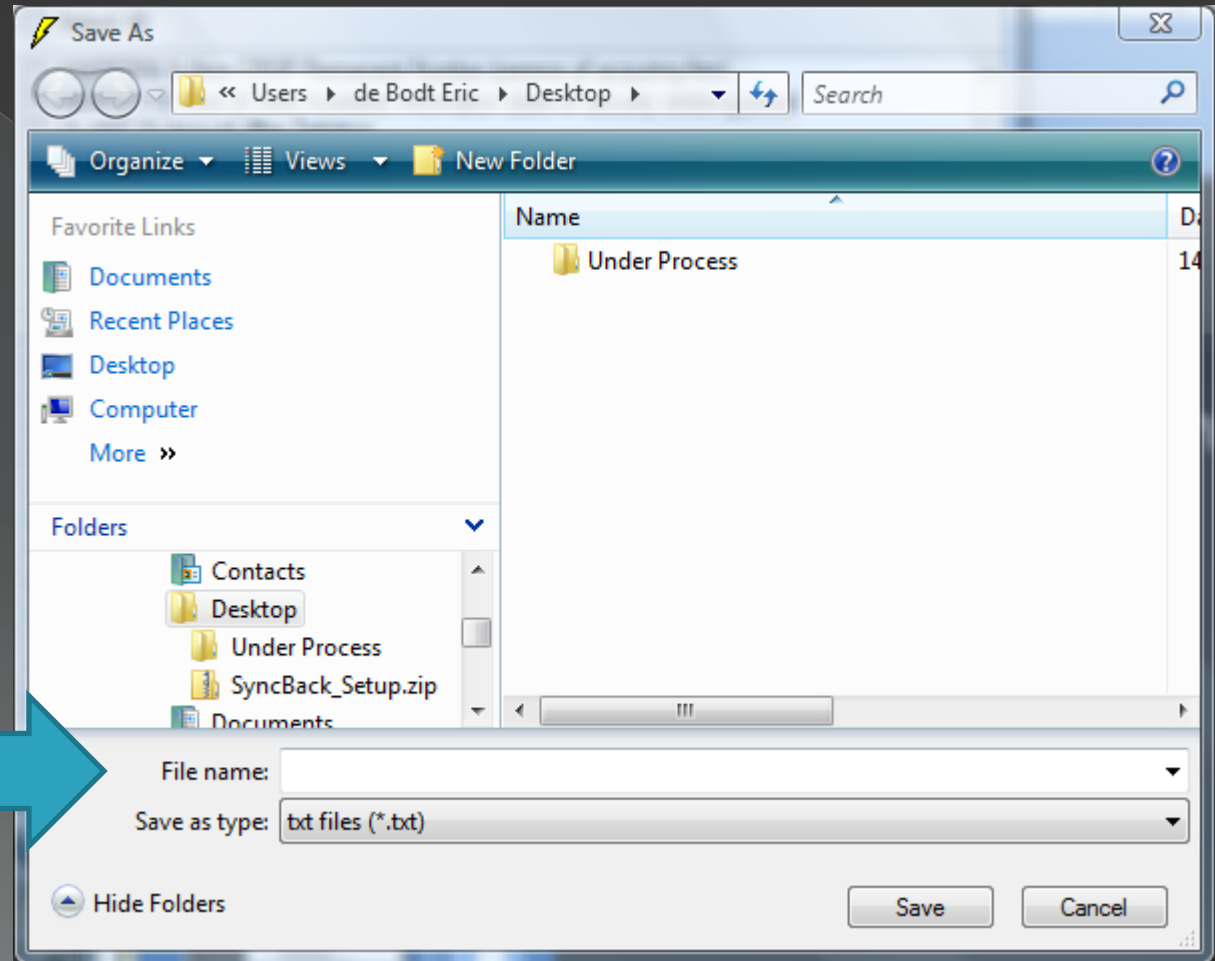
Usage

- 2. Select fields



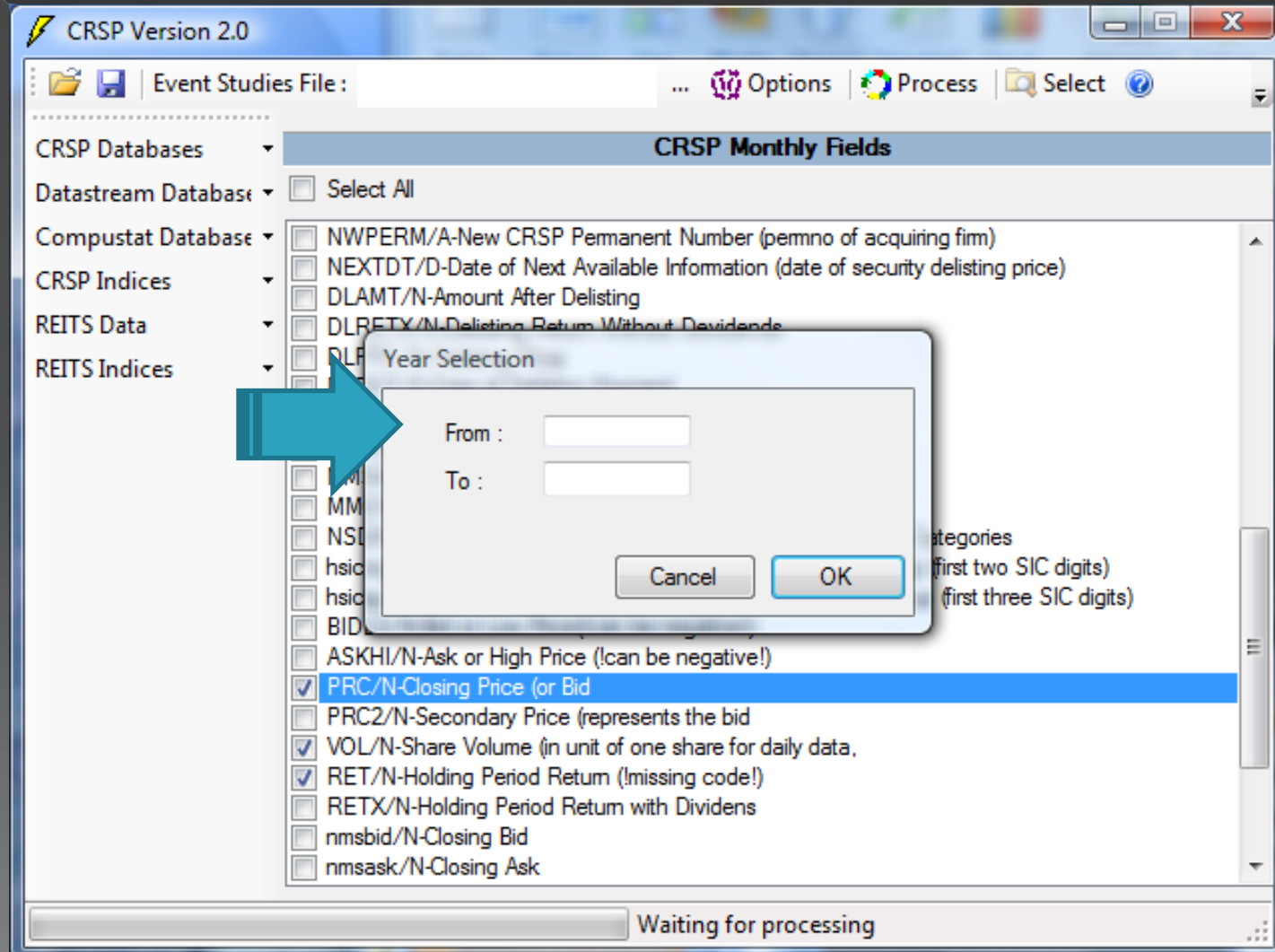
Usage

- 3. Save your configuration



Usage

- 4. Select a period



Usage

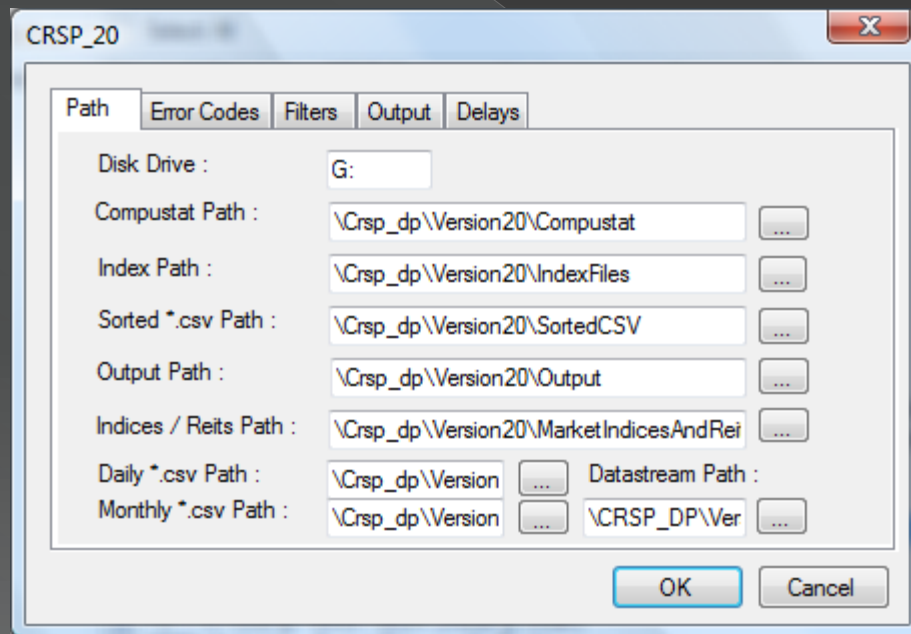
5. Get back your data

The image shows a Windows Explorer window on the left and a Microsoft Excel window on the right. The Explorer window displays the 'Output' folder containing files: CRSP_M_RET.csv, dates.csv, size.csv, and sorted_permnos.csv. A blue arrow points from the Explorer window to the Excel window. The Excel window shows a spreadsheet with columns A through L and rows 1 through 23. The first row (row 1) is the header row with the following values: PERMNO, 19900131, 19900228, 19900330, 19900430, 19900531, 19900629, 19900731, 19900831, 19900928, 19901031, 19901130, 19901231. The subsequent rows contain numerical data for each PERMNO.

PERMNO	19900131	19900228	19900330	19900430	19900531	19900629	19900731	19900831	19900928	19901031	19901130	19901231
10001	18.565	18.565	18.565	18.565	18.565	18.565	18.881	18.881	18.881	18.881	18.881	18.881
10002	0	0	0	0	0	0	0	0	0	0	0	0
10003	0	0	0	0	0	0	0	0	0	0	0	0
10005	1.072	1.072	1.072	1.072	1.072	1.072	1.072	1.072	1.072	1.072	1.072	1.072
10007	3.25	3.25	3.25	3.25	3.25	3.25	3.25	0	0	0	0	0
10009	0	0	0	0	0	0	0	0	0	0	0	0
10010	13.177	13.177	13.177	13.177	13.177	13.177	13.177	13.177	13.177	13.177	13.177	17.788
10011	5.468	5.468	5.962	5.962	5.962	5.962	5.962	5.962	5.962	5.962	5.962	5.962
10012	4.253	4.253	6.228	6.228	6.228	6.228	6.228	6.228	6.228	6.228	6.228	6.228
10016	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206	1027.206
10018	2.221	2.221	2.221	2.221	2.221	2.221	2.221	2.221	2.221	2.221	2.221	6.073
10019	33.645	33.645	33.645	33.645	33.645	33.645	33.645	40.409	40.409	40.409	40.409	40.409
10020	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786	2386.786
10021	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09	2.09
10024	1.256	1.256	1.256	1.256	1.256	1.168	1.168	1.168	1.168	1.168	1.168	1.168
10025	65.573	65.573	65.573	65.573	65.573	65.573	65.573	65.573	65.573	65.573	65.573	73.365
10026	66.594	66.594	66.594	66.594	66.594	66.594	66.594	66.594	66.594	66.594	75.132	75.132
10027	22.154	22.154	22.154	0	0	0	0	0	0	0	0	0
10028	3.741	3.741	3.741	3.741	3.741	3.741	3.741	3.741	3.741	3.741	3.741	3.741
10029	60.695	60.695	60.695	60.695	60.695	60.695	60.695	60.695	60.695	60.695	65.607	65.607
10032	36.464	36.464	36.464	36.464	36.464	36.464	36.464	36.464	36.464	36.464	43.419	43.419
10034	-98	-98	-98	-98	-98	-98	-98	-98	-98	-98	-98	-98

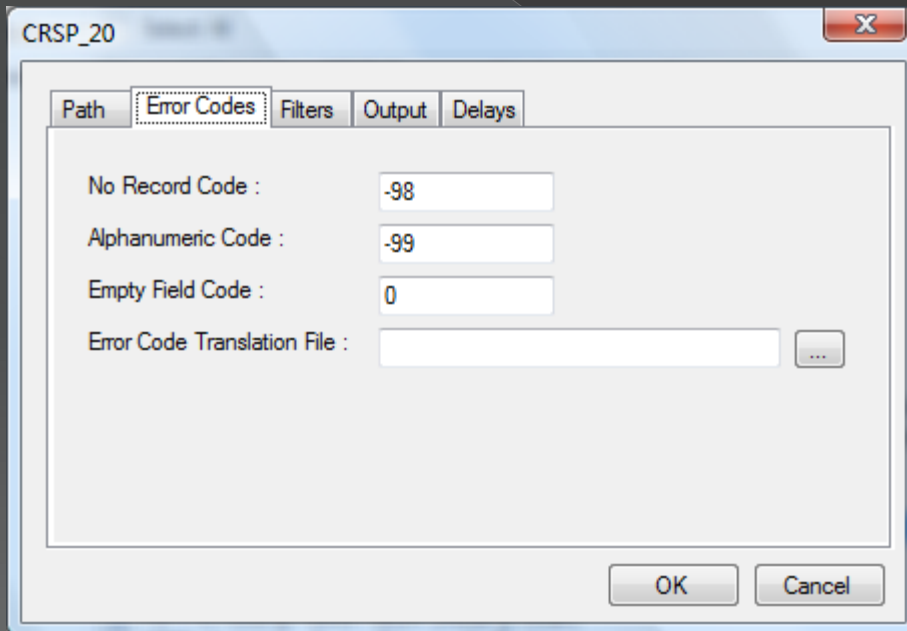
Usage

- Options: Paths



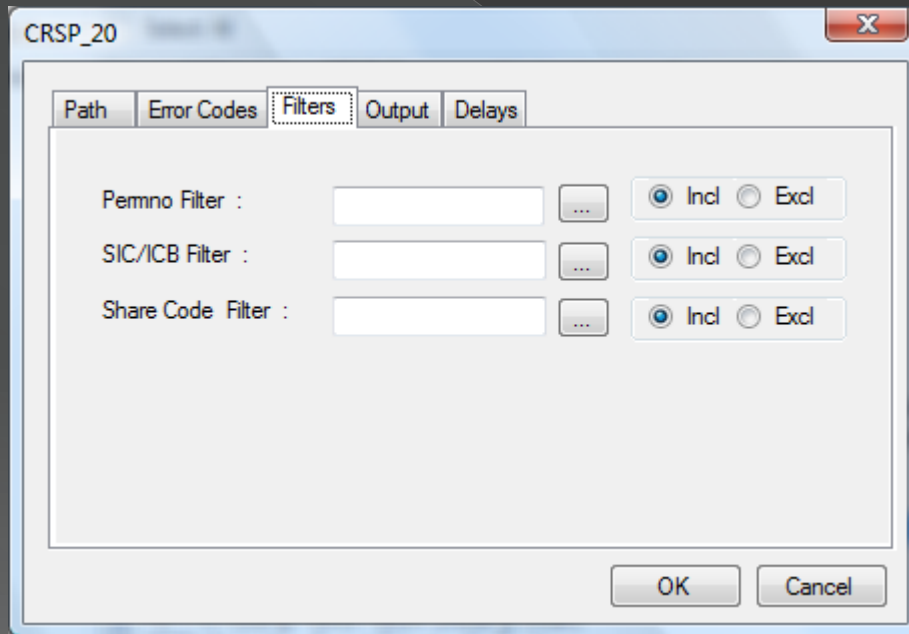
Usage

- Options: Error Codes



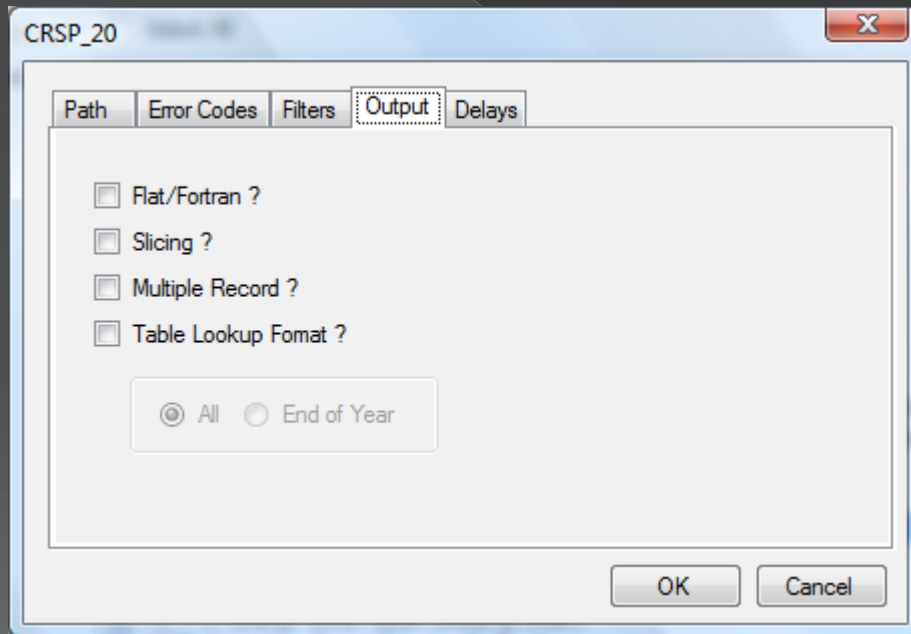
Usage

- Options: Filters



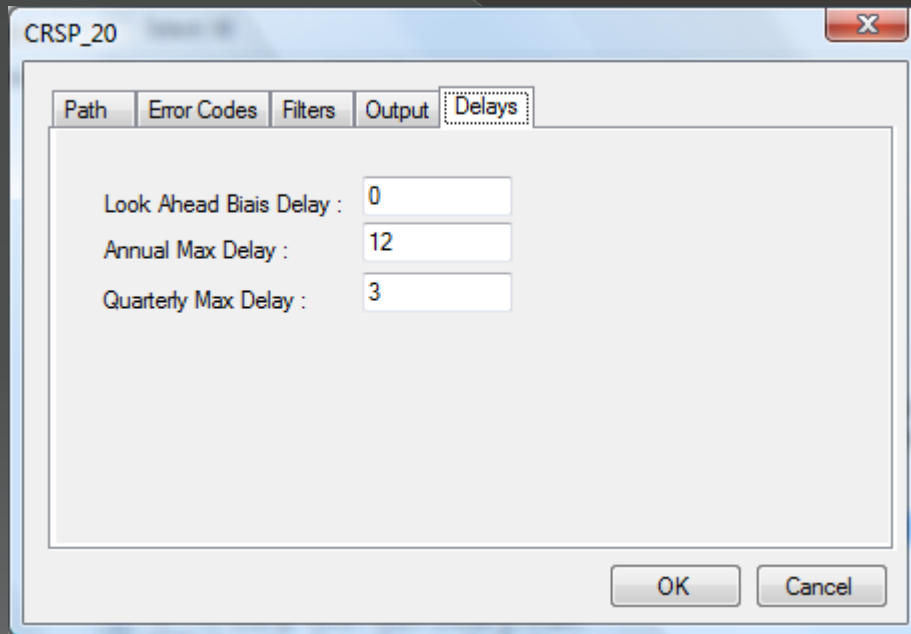
Usage

- Options: Output



Usage

- Options: Delays



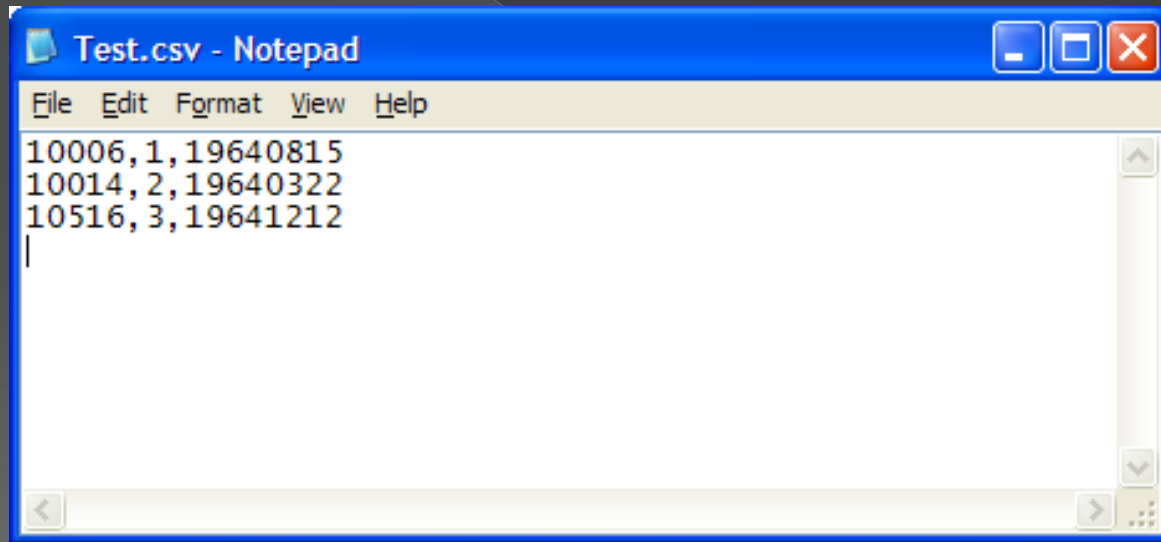
Outline

- Introduction
 - > Accessing CRSP/Compustat data through WRDS
 - > CRSP_DP.exe goals
- General Presentation
 - > Outputs
 - > Coverage
 - > Installation
 - > Interface
- Usage
 - > Let's go
 - > Options
- Event Studies



Event Study

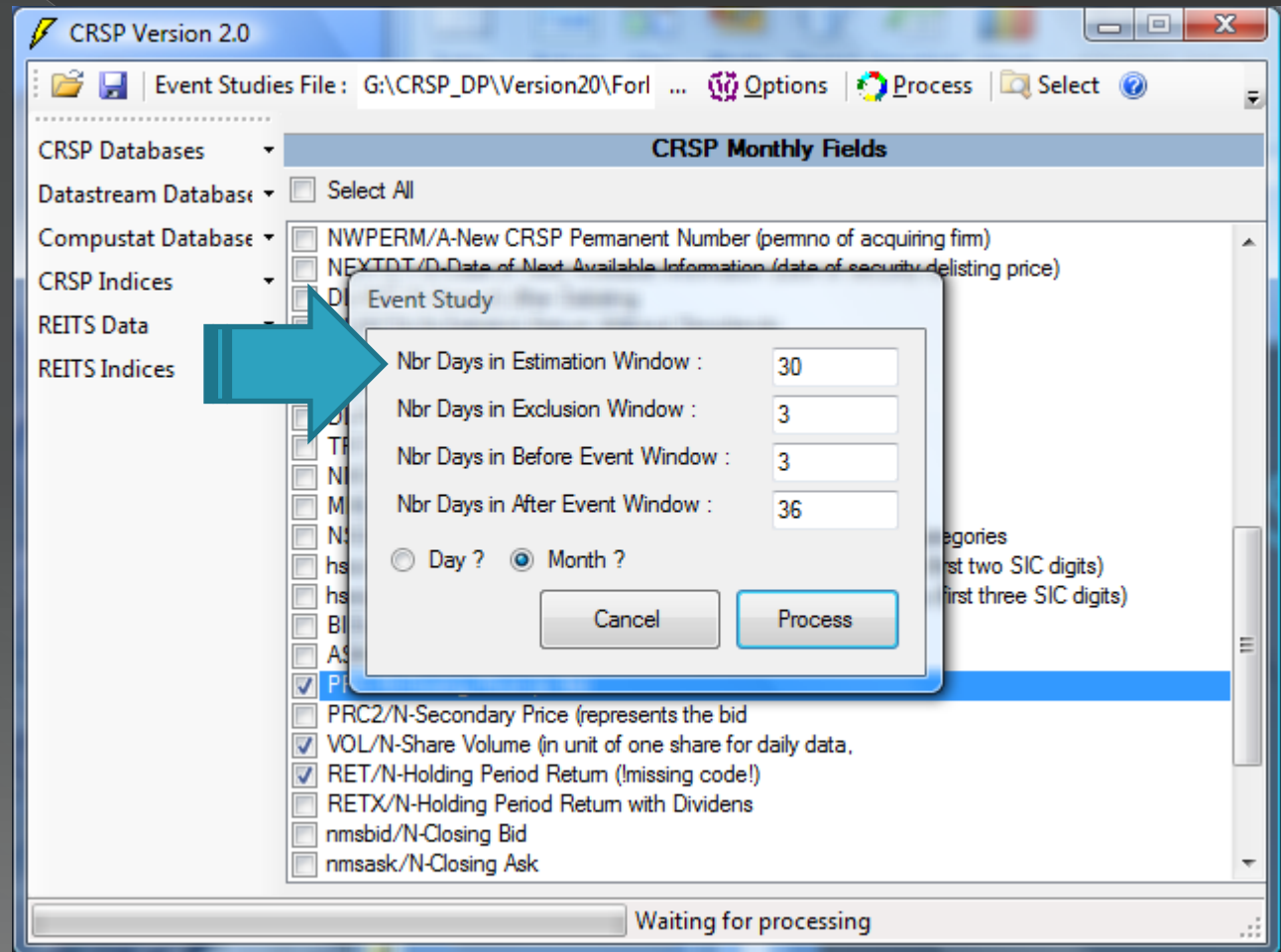
- Switch from calendar dates to event dates
- Inputs : A sample description file



```
Test.csv - Notepad
File Edit Format View Help
10006,1,19640815
10014,2,19640322
10516,3,19641212
```

Event Studies

- Options



Event Studies

- Output

CRSP_M_RET.csv - Microsoft Excel

Home Insert Page Layout Formulas Data Review View

Clipboard Font Alignment Number Cells Editing

A1 PERMNO

	A	B	C	D	E	F	G	H	I	J	K
1	PERMNO		t-36	t-35	t-34	t-33	t-32	t-31	t-30	t-29	t-28
2	10051	2939	-0.05357	-0.01887	-0.01923	-0.01961	0.08	-0.07407	0	0	
3	10108	258	0.04386	0.151261	0	-0.0292	0.135338	0.062914	0.018692	0.045872	0.0
4	10108	1400	0.071429	-0.05185	0.240234	-0.09606	0.10453	0.009464	-0.20156	0.09589	-0.0
5	10108	1208	-0.17029	-0.12664	0.0525	-0.07126	-0.08951	0.067416	0.221053	0.034483	0.2
6	10108	723	-0.03949	0.059947	0.115115	-0.11616	0.122689	0.076172	0.009079	-0.0913	-0.
7	10123	2168	0.392857	-0.02564	0.105263	0.380952	0	0.103448	-0.0625	0.166667	0.0
8	10123	3539	0.033898	-0.10656	0.119266	0	0.204918	0.088435	-0.00425	0.100629	0.0
9	10123	2218	0.207317	-0.12121	0.206897	0.152381	0.194215	-0.02422	-0.0922	0.058594	-0.
10	10126	1684	0.00885	0	0.017699	0.021739	0.042918	0.004115	0.020492	-0.00405	
11	10137	3600	-0.00692	0.030172	0.016736	0.014156	0.00823	0	-0.01861	-0.11392	-0.0
12	10137	1655	0.073009	0.086351	-0.05769	-0.00204	0.112229	-0.08752	-0.08367	0.08882	-0.0
13	10145	1079	0.042146	0.051471	0.064895	0.036304	0.009554	0.023849	0.102167	0.050562	-0.0
14	10145	354	0.043379	0.028446	0.054894	0.066802	-0.00569	0.121756	-0.08532	0.050373	0.0

Ready CRSP_M_RET 100%