

CRSP SURVIVOR-BIAS-FREE US MUTUAL FUND DATABASE

DECEMBER 2019 QUARTERLY UPDATE

These release notes accompany the January 2020 quarterly release of the CRSP Survivor-Bias-Free US Mutual Fund Database, and contain data through December 2019.

MUTUAL FUND RETURNS RERUN

The December release of the CRSP Mutual Funds database, roughly 20,000 (one quarter of one percent) of daily and monthly returns were presented as price appreciation rather than total returns. A new presentation of the underlying reinvest_nav values from our data source resulted in our populating the reinv_nav field with zeros. These returns have been corrected to reflect total returns. A new database is available for download from our MOVEit site and is available from WRDS as well.

NOTES ABOUT THIS RELEASE

- Several clean-up tasks have been completed and applied in this release:
- We have removed duplicate entries from the fund_flows table. These duplicate rows were the result of revisions reported in the NSAR/A filings that were included in the 201909 release. In this 201912 release, original data was replaced with the revised figures to avoid duplicate primary keys.
- Historical fund fees data that was excluded from the 201909 release pending further research, has now been added back to the fund_fees table. This is data with enddt prior to Jan 31, 1998.
- Several corrections were applied to fund_summary table. Per_com values were edited for crsp_fundnos 31294, 31295 and 31296. The corrections were based on data sourced from the SEC N-Q filings.
- Over 3400 daily_nav values were removed from the daily_nav table. These nav values were duplicate prices with incorrect caldt that did not fall on trading days. This caused invalid daily returns 0.0 values for some funds in the 201909 release.
- We have removed non-linking crsp_portno keys from the crsp_portno_map table. These crsp_portno keys linked to neither the holdings nor the fund_flows data. We are continuing to work with our data vendors to source the missing holdings data. Once available, the holdings data will be added in future releases.

- We have corrected over 1000 maturity_dt values in the holdings_co_info table. Entries with invalid maturity_dt were either replaced with correct dates or null where no data was available.
- We have made over 30,000 corrections to the monthly return values by adding the previously missing dividends.

FILE VERSION SPECIFICS

The database has 67,417 total funds. Of these 32,174 are active and 35,243 are delisted.

The 32,174 active mutual fund classes are assets within 12,044 unique mutual funds (determined by unique crsp_cl_grp)

The 35,243 inactive mutual fund classes are assets within 17,676 unique mutual funds (determined by unique crsp_cl_grp)

The biggest table is HOLDINGS with 224,950,946 rows.

The mutual fund database is available in SAS and ASCII formats and as a CRSPAccess database for access through CRSPSift for Windows.

NUMBER OF ROWS SUMMARY

TABLE	ROWS
CONTACT_INFO	259,443
CRSP_CIK_MAP	38,660
CRSP_PORTNO_MAP	71,961
DAILY_NAV	135,524,176
DAILY_RETURNS	136,246,118
DIVIDENDS	3,519,076
FRONT_LOAD_DET	105,040
FRONT_LOAD_GRP	33,558
FUND_FEES	515,455
FUND_FLOWS	1,826,023
FUND_HDR	67,417
FUND_HDR_HIST	489,626
FUND_STYLE	177,905
FUND_SUMMARY	2,220,130
HOLDINGS	224,950,946
HOLDINGS_CO_INFO	1,831,268
MONTHLY_NAV	7,657,751
MONTHLY_RETURNS	7,665,202
MONTHLY_TNA	7,630,133
REAR_LOAD_DET	153,777
REAR_LOAD_GRP	50,049

SYSTEM REQUIREMENTS

- SAS version 8, SAS version 9.1, or a relational database of your choice.
- Sufficient disk space to load and manipulate the database:

DATABASE FORMAT	DISK SPACE REQUIRED
SAS V 9.1	21.7GB
ASCII DATA	21.6GB
CRSPACCESS	13.8GB
R	3.29GB

GENERAL NOTES:

- A blank field means either that data was not applicable to a fund or that data is not available.
- A field with a “?” in it is a flag to indicate further research is required.
- A field has a zero in it when 1) The fund did not exist; 2) No value could be found; or 3) A calculated value could not be calculated because information was missing.

INSTALLS AND DATA FORMATS

ZIP INSTALLS

CRSP uses a zip process for data installation. ZIP utilities that are standard on all platforms should work to unzip the databases. On Windows computers, CRSP recommends using 7-ZIP, which can be downloaded from this site: <http://www.7-zip.org/>

Several important items to note follow:

- All directories are extracted to lower case letters. For example: C:\crspdata\miz201308
- All Windows Users: Extracting a zipped file does not create an installed program in Windows. It will not show in the Windows Control Panel. When uninstalling the data, just delete the directories that contain the detailed files
- Red Hat Linux Users: If your version of Red Hat is earlier than 6.0, you will need to install an updated version of the “File Roller” program in order to install the data

STEPS TO INSTALL

- From the directory containing your downloaded zip file, right click (in Windows) or run File Roller (in Red Hat)
- Set the extract destination to the current higher level directory where data is normally stored and unzip will add the rest of the path (default is C:\ in Windows or /home in Red Hat) For example, in Windows, extracting to C:\ will install the data under C:\crspdata

Contact CRSP Client Services for assistance at: support@crsp.org or call 312-263-6400, Option 2.