

Historical FRED-MD Vintages

Overview

This zip file contains historical vintages of FRED-MD for every month from August 1999 to December 2014. We have constructed these vintages using historical Haver Analytics data. Below are a description of the data, some guidance for determining the difference between these historical vintages and the real-time vintages, and a brief account of two comparisons we perform to check the accuracy of the historical vintages.

Data

The St. Louis Fed has regularly backed up the Haver databases to which it subscribes since August 1999.¹ This set of archived databases gives us access to historical versions of various economic and financial time series. By matching each FRED-MD series to a relevant Haver series (or combination of Haver series), we are able to use these archived Haver databases to construct historical FRED-MD vintages. Specifically, we construct a particular FRED-MD vintage YYYY-MM by using the last version of each series that was available by the end of month MM of year YYYY.

Each version of a particular Haver series is associated with a “date last modified” characteristic, which is the date that the version was updated by Haver. We use this as a proxy for the version’s actual release date. Because we only care about the version of a series at the end of the month when constructing a particular FRED-MD vintage, this is only problematic if a version was released before the end of the month but was not updated by Haver until after the start of the following month. Given that FRED-MD consists of relatively popular national series, it seems unlikely that Haver would ever update one of these series a day or more after it was released.

Before using the Haver data to construct the FRED-MD vintages, we perform tests on each Haver series to check for missing versions of the series and for any changes in the availability of historical observations. Specifically, we flag a series if (i) the number of observations increases by more than one between consecutive versions of the series or (ii) a non-missing observation switches to a missing observation between consecutive versions of the series. Note that the only instance where these two flags would not identify a missing version of a series is when the missing version contains revisions to historical data but introduces no new observations. If a series is flagged with one of the above issues, we investigate the cause and resolve it if possible.

While we have access to archived Haver databases starting in August 1999, there are some series that do not have versions going back that far. For each of these series, this is due to one of the following reasons:

1. The series did not exist prior to the date of its earliest version in our archived Haver databases.
2. Haver did not carry the series prior to the date of the series’ earliest version in our archived Haver databases.

There is nothing we can do to resolve these issues. Thus, when constructing a particular historical FRED-MD vintage, we simply remove any series for which we do not have access to historical versions prior to the FRED-MD vintage date.

Note that some series in FRED-MD do not get revised over time. Rather than matching them with a Haver series, we simply construct historical versions of these series by removing observations from the current version.

¹ The databases were saved weekly through August 2005. Since then they have been saved daily.

Differences from the Real-Time Vintages

We construct the historical vintages of FRED-MD to match as closely as possible the 128 series in the most recent version of the dataset at the time (vintage 2017-04). However, there are some minor differences. Differences between the last historical vintage (2014-12) and the first real-time vintage (2015-01) can be found in the PDF titled “Changes to Historical FRED-MD Vintages” available in this zip file. This PDF also documents all changes to the historical FRED-MD vintages in reverse chronological order (from 2014-12 to 1999-08). For changes to the real-time FRED-MD vintages (2015-01 to current), please see the PDF titled “Changes to FRED-MD” found on the FRED-MD and FRED-QD homepage (<https://research.stlouisfed.org/econ/mccracken/fred-databases/>).

Checks for Accuracy

We perform two comparisons to check the accuracy of the procedure used to construct the historical FRED-MD vintages: first, a comparison to vintages 2015-01 to 2016-12 that we constructed in real time (i.e. the ones we published online); second, a comparison to the dataset used in the paper “Implications of Dynamic Factor Models for VAR Analysis” by James H. Stock and Mark W. Watson (2005).

Comparison to Real-Time Vintages

Using the same procedure implemented to construct the historical FRED-MD vintages from 1999-08 to 2014-12, we construct FRED-MD vintages from 2015-01 to 2016-12. We then compare them to the same vintages that we constructed in real time and published online. We find no significant differences.

Comparison to Stock and Watson Dataset

We first identify all series in the Stock and Watson dataset (SW henceforth) that have a comparable series in FRED-MD. For each of these series, we then compare the SW series to its counterpart in FRED-MD for all FRED-MD vintages between 2004-01 and 2004-12. We use a range of FRED-MD vintages in the comparison because, while all SW series end in 2003:12, there is some ambiguity as to when each SW series was updated. For example, it is possible that one series in SW was updated in January 2004 while another was updated in March 2004. Differences in precision (e.g. rounding to nearest tenth vs. rounding to nearest hundredth) and units (e.g. persons vs. thousands of persons) are taken into account when comparing series. We find no significant differences.