Observation Driven Mixed-Measurement Dynamic Factor Models with an Application to Credit Risk

By: Drew Creal, Bernd Schwab, Siem Jan Koopman, Andre Lucas

The default, credit ratings, and loss-given-default data comes from Moody's and is unfortunately not publically available. Here, we describe the databases and transformations taken for the results in the paper.

We match information from four files in the Moody's default and recovery database (DRD). "**Senrratg_new**" contains issuer identifier, rating transition date, and new rating; "**mastdflt**" contains issuer identifier, default date, and default number; "**mastissr**" contains issuer identifier, industry sector, and country information, and "**dfltissu**" contains a default number and the post-default price of a reference bond. The database is proprietary and available to subscribers of www.moodys.com default risk services (DRS).

We discard non-U.S. based firms. Ratings are initially aggregated into the usual seven broad rating groups, and default (D). These are eventually further re-grouped into four final groups: Investment Grade

(IG) that contains ratings Aaa down to Baa3; double B (BB) that contains Ba1 to Ba3; single B (B) that contains B1 to B3; triple C (CCC) that contains Caa1-C3. A company that defaults is marked as a transition into D.

We apply the following standard filters when cleaning default data: If there are multiple defaults of the same firm, we discard the firm history after the first default. If a firm history begins with a default, we exclude the firm. If a firm defaults, we ignore a previous rating withdrawal. If a firm defaults jointly with other firms on the same business day, we check that the firms are not linked through a parent-subsidiary relationship; this is done on a case-by-case basis, based on the ultimate parent number in "**senrratg_new**" and the qualitative default description in "**mastdflt**". If multiple defaults are due to a parent-subsidiary relationship, we count one default for the firm group.

Loss-given default is constructed as one minus a recovery rate. The recovery rate is the post-default price of a reference bond, and is in percentages of face value. The recovery rate and defaults are matched across files through their default number. We observed 14 recovery rates larger than 1 in the data. This can happen with market based recovery rates. We censor these recovery rates at 99.99%, such that the corresponding LGDs are put at 1 basis point.

The macroeconomic time series are available from the FRED database at the Federal Reserve Bank of St. Louis, and are publicly available. We provide the macroeconomic series used in the paper in Ox files.