

# Data Description Sheet - Syndicated Lending Relationships, Information Asymmetry and Market Making in the Secondary Loan Market

Matthew A. Phillips

To be provided upon initial submission on a separate data description sheet:

1. *A description of which author(s) handled the data and conducted the analyses.*

Matthew Phillips handled the data and conducted the analyses. Predoctoral researcher Oliver Sun ran the WRDS earnings surprise code and provided the data output but did not perform any analysis.

2. *A detailed description of how the raw data were obtained or generated, including data sources, the specific date(s) on which data were downloaded or obtained, and the instrument used to generate the data (e.g., for surveys or experiments). We recommend that more than one author is able to vouch for the stated source of the raw data.*

The raw data for secondary loan pricing and market maker identities were obtained from Thomson Reuters Refinitiv. The data files were received April 12, 2023 (pricing data) and the market maker identities on July 29, 2019. The LIN-FacilityID mapping from Refinitiv was received on April 11, 2023. CLO trading data was obtained from Refinitiv on April 11, 2023. Dealscan new vintage data was obtained on July 12, 2022 and old vintage on April 11, 2024. Compustat data was obtained on April 10, 2024. IBES data was obtained on April 10, 2024. CapitalIQ events data was received on April 29, 2024. The WRDS earnings surprise code was run by predoctoral researcher Oliver Sun on April 29, 2024.

3. *If the data are obtained from an organization on a proprietary basis, the authors should privately provide the editors with contact information for a representative of the organization who can confirm data were obtained by the authors. The editors would not make this information publicly available. The authors should also provide information to the editors about the data sharing agreement with the organization (e.g., non-disclosure agreements, and any restrictions imposed by the organization on the authors). In particular, the authors should indicate if an organization or data provider imposes restrictions on the publication of the results, has not given the authors full control of the relevant data, requires that the results must be reviewed or approved prior to public release of the paper or publication.*

The Refinitiv representative I received the data from was Eric Trujillo and order forms have been sent to JAR. No person requires that the results must be reviewed or approved prior to the public release of the paper or publication.

To be provided in the paper or the online appendix:

4. *A complete description of the steps necessary to download, obtain or collect as well as process the data used in the final analyses reported in the paper. For experimental and survey papers, we require information about the instructions and instruments used to generate the data, subject eligibility and/or selection, as well as any exclusion criteria. The full set of instructions and instruments can be provided in the online appendix.*

Sample selection criteria, data processing and analysis information are provided in the paper.

Raw files for loan quote data are referred to as “Historical Loan Price File” from LPC Collateral and will be downloaded as monthly csv files. The CLO trading data are referred to as LPC\_Collateral\_TradeHistory and will be downloaded as annual excel files. The market maker identities must be requested directly from Refinitiv and come as a csv file. All other data are available via WRDS.

The processing from the raw data files to the final analysis is available in the internet appendix code. The code must be run in this order:

- The first file to run is the A1\_DS\_LA\_DEDUP SAS file that creates a unique record for each LIN for the new vintage of Dealscan. It retains a unique lead arranger but this is not used directly in the analysis as the lead arranger identities are based on the old vintage DS designation.
- The second chunk of code to run is the processing of the raw loan quote files which is in the first 174 lines of the A2\_Stata\_2025 do file.
- The third file to run is the esurprises (WRDS) canned code.
- The fourth file to run is the A2\_SAS\_2025 SAS file.
- The final file to run is the A2\_Stata\_2025 do file from lines 175 through the end of the document.

The purpose and procedure of the code is embedded in the code files. The user must manually map the market maker identities from Refinitiv to the ultimateparentid (and name the variable ds\_upid) from Dealscan in addition to having access to the raw data from all the data providers listed above in a SAS data file for the code to run. I have provided the mapping SAS data files I performed in this appendix.

5. *After downloading or obtaining the raw data, all manipulations of the data should be done via computer programs. The code for these manipulations should be included in the code submitted upon acceptance (see below). No manipulations of raw data can take place manually or outside the computer code provided. If compliance with this requirement is not feasible, the authors need to explain and disclose any manipulations of the raw data (e.g., manually created variables or file conversions). When feasible, we also encourage the authors to share the code that downloads the data.*

As discussed above, the manual mapping of market makers to ultimateparentid is the sole manual creation of a variable that occurs outside of the program file. I have provided the files with the manual mapping in the internet appendix for users to review. Modifications of this based on the updated market maker file received are reflected in the A2\_SAS\_2025 SAS file.

To be provided upon acceptance of the paper and prior to publication:

6. *The computer programs (i.e., code) used to (1) convert the raw data into the final dataset used in the analysis, (2) to execute the statistical or econometric analysis, and (3) to generate the tables or to produce the output used in constructing tables of the manuscript. A brief description that enables other researchers to understand and run the code should be provided. The purpose of this requirement is to facilitate replication and to help other researchers understand in detail how the raw data were processed, the final sample was formed, variables were defined, outliers were treated, and which commands were used in the analysis, etc. This code or programming is in most circumstances not proprietary. However, we recognize that some parts of the code or data generation process may be proprietary, including from the authors' perspective. Therefore, instead of disclosing the proprietary portion of the code or program, researchers can provide a detailed step-by-step description of the code or the relevant parts of the code such that it enables other researchers to arrive at the same results that the authors obtained and presented in their manuscript. In such cases, the authors should inform the editors upon initial submission, so that the editors can consider an exemption allowing the step-by-step description. Whenever feasible, authors are required to provide the identifiers (e.g., CIK, CUSIP) for their final sample. Authors should consult our FAQ Sheet on the JAR website for further details.*

The code files are annotated and include all steps from raw files through analysis output to LaTeX. All facilityids used in the main analysis are included in the jar\_facilityids.dta file. A description of the order of code to run is available in the response to point 4 above.

7. *A comprehensive log file that shows the execution of the entire code. This log file should cover all the steps that convert the raw data into a final dataset and the execution of all statistical and econometric analyses presented in the tables of the manuscript. The portion of the log file that shows proprietary code or data may be masked. In this case, the reader should be referred to the step-by-step description provided as per the requirements in Item 6.*

The log files have been provided to JAR. They are referenced as follows:

- A1\_DS\_LA\_DEDUP.txt
- A2\_Stata\_2025\_pt1.smcl
- esurprises.txt
- A2\_SAS\_2025.txt

- A2\_Stata\_2025\_pt2.smcl
8. *An assurance that the data and programs will be maintained by at least one author (usually the corresponding author) for at least six years, consistent with National Science Foundation guidelines.*

All files will be maintained for at least six years following the NSF guidelines.