

Data Description

Paper: “Accounting for Goodwill” by Stefan Huber, and Charles McClure

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

Stefan Huber conducted all analyses and data cleaning. All coauthors have access to the data and code.

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 can vouch for the stated *source* of the raw data.

The raw sources of data, access methods, and access dates are as follows:

- We obtained corporate takeover auctions bidding data for the period 2001 - 2012 from Alexander Gorbenko and Andreij Malenko, who collected this data in two published papers:
Gorbenko, Alexander S., and Andrey Malenko. "Strategic and financial bidders in takeover auctions." The Journal of Finance 69.6 (2014): 2513-2555.
Gorbenko, Alexander S. "How do valuations impact outcomes of asset sales with heterogeneous bidders?." Journal of Financial Economics 131.1 (2019): 88-117.
- We hand collect corporate takeover auctions bidding data for the period 2012 - 2022 from the deal background sections of SEC merger filings
- We obtain data on purchase price allocation from BVR DealStats (most recent data pull: November 10, 2022). These data were purchased by Rice University. We supplement it with handcollected data on purchase price allocation from acquirers' 10Ks, whenever data is missing in the DealStats database.
- We obtained data on firm fundamentals from Compustat (most recent data pull: October 3, 2023). These data were purchased by Rice University and accessed through Wharton Research Data Services (WRDS).
- We obtained data on executive tenure and compensation from Execucomp (most recent data pull: October 3, 2023). These data were purchased by the University of Chicago and accessed through Wharton Research Data Services (WRDS).

- We obtained data on M&A transactions from Thomson Reuters SDC (now called Refinitiv) (most recent data pull: September 2022). These data were purchased by the Rice University.
- We obtained data on market-adjusted announcement returns from the WRDS Event-Study Suite (most recent data pull: August 27, 2024). These data were purchased by the University of Chicago and accessed through Wharton Research Data Services (WRDS).

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 paper does not use proprietary data.

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.

A complete description of the steps used to create the sample is described in Section 3. Calculation of all variables are as described in the body of the paper.

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.

We attest that all data manipulations are done via computer programs, the code of which is provided in our replication package.

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.

Details of how to run our code are in the file `README.md` in our replication package. Firm identifiers are Deal Numbers by SDC Platinum and are in our replication package under `identifiers/identifiers.csv`.

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 are provided in our replication packet in the folder `logfiles`. This folder includes the log files for our data cleaning, estimation, bootstrapping, and counterfactual analyses.

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.

Data and programs used in this study will be maintained by Stefan Huber by at least 6 years, consistent with National Science Foundation guidelines.