

Made in the U.S.A.? A Study of Firm Responses to Domestic Production Incentives

Data Description

Rebecca Lester
Stanford Graduate School of Business

Below is a description of the data used in the study “Made in the U.S.A.? A Study of Firm Responses to Domestic Production Incentives” pursuant to the *JAR* Data and Code Sharing Policy.

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

As the sole author of this paper, I handled the data and conducted all analyses.

I had assistance from two outside parties with obtaining some of the raw data (see more details below under Item #2). First, Jeff Hoopes (Assistant Professor of Accounting, UNC) created a Perl program for me in April 2014 to search financial statements for search terms I provided to him. Second, the Stanford Research Support Team helped me obtain additional data related to firms’ tax net operating losses, positive performance shocks, and other tax incentives claimed. I was involved with these data at all stages, ranging from providing the initial search terms to reviewing the excerpts to cleaning the data that were ultimately used in the project.

Finally, I created all of the code used to create the final output.

2. *A detailed description of how the raw data were obtained or generated, including data sources, the date(s) on which data were downloaded or obtained, and the instrument used to generate the data (e.g., for surveys or experiments).*

The data I use are primarily from the following three sources:

Dataset identifying DPAD Firms: To identify the DPAD firms studied in this paper, I created a list of search terms (listed in Appendix A of the manuscript) and provided this list to Jeff Hoopes, who created a Perl program for me in April 2014 to search annual financial statements. I personally reviewed each excerpt and coded each match so as to correctly identify firms that discuss or disclose the benefit and the first financial statement year in which the terms are identified. Excerpts range from a general discussion of the benefit in any section of a firm’s financial statements to disclosure of the specific amount of DPAD benefit in a firm’s effective tax rate reconciliation. I disclose Jeff’s assistance in the Acknowledgements section of the paper.

Historical Financial Information: I obtained firm financial information from the Compustat (Annual, Quarterly, and Segments) databases via WRDS in April 2014, June 2014, October 2014, April 2016, November 2018, and January 2019.

Bureau of Economic Analysis Data: I obtained jurisdiction-specific investment and employment data from the Bureau of Economic Analysis in June 2014, October 2014, April 2016, January 2018, September 2019, and January 2019.

Additional analyses also use data from these other sources:

Identification of Extraterritorial Income Exclusion (ETI) firms and firms with positive shocks: To identify ETI firms and firms with positive performance shocks, I provided lists of search terms to the Stanford Research Support Team; see footnote 18 and Online Appendix F (footnote 32), respectively,

in the manuscript for a list of these search terms. The Stanford Research Support Team used a Python script to search financial statements for these search terms. I reviewed these excerpts and then hired research assistants through the Stanford Research Support Team to further review and clean these data. In collaboration with the Stanford Research Support Team, I reviewed the research assistants' work daily and completed a final check of the excerpts prior to merging these data with the main dataset. I disclose this assistance by the Stanford Research Support Team and the RAs in the Acknowledgements section of the paper.

Measurement of tax net operating losses: Details on the amount of firms' net operating losses were hand-collected from firm financial statements. These data were obtained by providing a list of search terms to the Stanford Research Support Team, who used a Python script to search financial statements. I reviewed these excerpts and then hired research assistants through the Stanford Research Support Team to further review and clean these data. A subset of these data were also used in another project co-authored with Shane Heitzman (*Net Operating Loss Carryforwards and Corporate Savings Policies*). I disclose this assistance by the Stanford Research Support Team and the RAs in the Acknowledgements section of the paper.

Repatriators under the American Jobs Creation Act of 2004: I obtained a list of repatriating firms from Michelle Hanlon, who compiled this list for the paper "Real Effects of Accounting Rules: Evidence from Multinational Firms' Investment Location and Profit Repatriation Decisions" co-authored with John Graham and Terry Shevlin (*JAR* 2011). I disclose the use of these data in the Acknowledgements section and in Online Appendix B, footnote 31.

Foreign subsidiary details from Exhibit 21: I downloaded a list of foreign subsidiaries disclosed in Exhibit 21 to firm financial statements from Scott Dyreng's website in June 2014 and again in April 2016. I disclose use of these data in the Acknowledgements section and also cite the paper that introduced this measure to the academic literature throughout the manuscript (Dyreng and Lindsay, 2009).

R&D Tax Credit Data: I estimate a firm's R&D Tax Credit in two ways. First, I obtain hand-collected data used by Jeff Hoopes in his paper *The Effect of Temporary Tax Laws on Understanding and Predicting Corporate Earnings* (2018 working paper). These data cover the years 1997 through 2010. I further estimating a firm's R&D tax credit for the period through 2013 using publicly available financial statement data from Compustat. I disclose the use of the data provided by Jeff Hoopes in the Acknowledgements section and in footnote 7.

No survey instrument or experiment was used.

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, any restrictions imposed by the organization on the authors, such as restrictions to publish certain results.)*

I used confidential data on U.S. multinational firms from the U.S. Bureau of Economic Analysis. These data can only be accessed by researchers who have received Special Sworn Employee status. Contact information for the BEA employees that oversee the Special Sworn Employee program has been provided to the editors at the same time as this final data description.

4. *A complete description of the steps necessary to collect and 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.

I describe the process to collect and process the data used in the final analyses in Section 4, 5.3, 6.1, 6.2, Appendix B, and Appendix F of the paper. Additional details are included in Item #2 above.

5. *The computer programs or code used to convert the raw data into the final dataset used in the analysis, plus a brief description that enables other researchers to use this program. 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, 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 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 final dataset used in the analysis. In such cases, the authors should inform the editors upon initial submission, so that the editors can consider an exemption from the code sharing requirement. Whenever feasible, authors should also provide the identifiers (e.g., CIK, CUSIP) for their final sample. Authors should consult our FAQ Sheet on the JAR website for further details.*

See Code Description for explanations of the content and source of main tables, detailed processing descriptions, and relevant sections of code. Also see the file Sample Firm Identifiers; note that this list corresponds to the sample after requiring requisite data for the test of H1 and thus includes all DPAD firms and possible matched control firms.

6. *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 the National Science Foundation guidelines.*

I agree to maintain data and programs used in this paper for the six-year time period suggested by the National Science Foundation.