Given the unique nature of the study, all choices made in constructing the field data (e.g. sample selection, how inquiries were created and sent to companies, etc.) are explained in considerable detail in the paper. Additional variables (e.g. revenue, company age, etc.) are defined in Appendix B of the paper- with no addition truncation or alteration- and merged into the field dataset.

To highlight three key areas of the data construction process:

1) Firm Selection:

The sample of 250 companies was randomly selected from the possible set of 685 (see criteria described in the paper). Within the paper, the set of S&P 900 was as of October 2017 (i.e. S&P set of firms changes over time).

2) Data Collection:

Inquiries for each of the sample companies were sent via their website or e-mail hotline as described in the paper. In particular, after identifying the location of the helpline (as described in the paper), research associates sent in the inquiry and waited (according to the methodology in the paper) for a response.

3) Data Processing:

As readers of the paper will see, the analysis in the paper is econometrically very basic (e.g. t-stats, OLS, and logistic regressions). Moreover, the choices around coding variables (e.g. definition of closing an inquiry) are described in detail in the paper’s appendix. Following the aforementioned data collection, there was no additional modification, selection, or truncation of any variables used in the dataset (field collection data or control variables collected from archival sources noted in the paper in Appendix B).

For each company in the sample, the following data fields were recorded: text around anonymous reporting, availability of phone/e-mail/web reporting, all obstacles encountered during reporting process (as displayed in chart), provider, date and text of the response to each inquiry. Each of these variables was coded according to the specific methodology described in the paper (with no data points deviating from this process). For each company in the sample, the control variables were collected and merged with the final dataset as described in the paper (e.g. revenue from Compustat) with no further modification. There was no transformation or truncation of any control variables or other data collected from the field research.