Divergence of Cash Flow and Voting Rights, Opacity, and Stock Price Crash Risk: International Evidence

HYUN A HONG, JEONG-BON KIM AND MICHAEL WELKER

Voluntary Compliance with Journal of Accounting Research Data Policy

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

Data handling and programming were performed by Hyun A Hong. All the finance and accounting data are initially downloaded in 2008 and are subsequently supplemented between 2013 and 2017 during the Journal of Accounting Research (JAR) revision process.

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). We recommend that more than one author is able to vouch for the stated source of the raw data.

The data sources and the years in which data were downloaded or obtained are as follows:

1) We begin the sample collection procedure by identifying firms that have at least two classes of shares according to the Datastream country lists during the period of 1995 to 2007. Hyun A Hong downloaded the list of firms' securities across countries from Datastream in 2008.
2) For each share class, she extracted Friday-to-Friday weekly finance data from Datastream for the following variables. If the value of turnover was missing from Datastream, it was manually collected from Bloomberg (Doidge [2004]). All these data downloads occurred in 2009. In addition, she collected lagged annual financial statement variables from Worldscope in 2009.
3) Because data on the number of voting rights attached to the superior and inferior voting shares for each firm are required, these data are hand-collected from Datastream Manuals, Moody's International Manuals, filings with the national stock exchanges, firms' annual reports, and the firm lists compiled by Doidge [2004]. In conducting this research she received significant support from Craig Doidge with respect to data collection. In addition, if the data ware not clear or not available, they were requested from each firm through faxes, emails, and phone calls. She collected this list of supplementary dual-class firms between 2008 and 2009.

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 agreement, any restrictions imposed by the organization on the authors with respect to publishing certain results).

The list of dual-class firms is manually collected across countries, and all the other data are collected from the public databases as described above.

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 papers, we require information about subject eligibility and/or selection, as well as any exclusion criteria.

We follow the steps detailed below to obtain the firm-level annual sample employed to test the relation between stock price crash risk and the ownership-control wedge.

Step 1: Hyun A Hong began the sample collection procedure by identifying firms that have at least two classes of shares according to the Datastream country lists during the period of 1995 to 2007. After identifying these firms, she followed prior studies and impose the following sample selection criteria: (1) firms must have at least two classes of shares with distinct voting and cash flow rights; (2) both classes of shares must be publicly traded on a domestic stock exchange; (3) the inferior voting shares cannot be convertible into the superior voting shares (though the opposite direction is allowed); (4) neither share class can receive a fixed dividend; and (5) neither share class can be redeemable or callable at the option of the firm at a pre-arranged price.

Step 2: The data collection procedures followed the guidelines established in Nenova [2003] and Doidge [2004]. First, for each share class, she extracted Friday-to-Friday weekly data from Datastream for the following variables: closing stock price, market value of all equity outstanding, weekly return, dividends paid during the week, number of shares outstanding, and turnover. If the value of turnover is missing from Datastream, it is obtained from Bloomberg (Doidge [2004]). In addition, she collected lagged annual financial statement variables from Worldscope.

Step 3: For each fiscal year, a firm was included in the sample only if it had at least 26 weekly stock return observations.

Step 4: Because data on the number of voting rights attached to the superior and inferior voting shares for each firm are required, these data are hand-collected from Datastream Manuals, Moody's International Manuals, filings with the national stock exchanges, firms' annual reports, and the firm lists compiled by Doidge [2004].

Step 5: The final sample consists of 3,350 firm-year observations from 449 firms across 20 countries.

Appendix A in the paper provides detailed definitions of the variables in our analysis.

5. Prior to final acceptance of the paper, the computer program used to convert the raw data into the dataset used in the analysis plus a brief description that enables other researchers to use this program. Instead of the program, researchers can provide a detailed step-by-step description that enables other researchers to arrive at the same dataset used in the analysis. The purpose of this requirement is to facilitate replication and to help other researchers understand in detail how the sample was formed, including the treatment of outliers, Winsorization, truncation, etc. This programming is in most circumstances not proprietary. However, we recognize that some parts of the data generation process may indeed be proprietary or otherwise cannot be made publicly available. In such cases, the authors should inform the editors upon submission, so that the editors can consider an exemption from this requirement.

We use SAS to convert the raw data into the datasets used in the analyses and use SAS to run the analyses. The detailed step-by-step description on how we arrive at the final datasets is included in part 4 above. The SAS code and a list of the identifiers of sample firms are stored on the JAR

website.

6. Data and programs should be maintained by at least one author (usually the corresponding author) for at least six years, consistent with National Science Foundation guidelines.

We will maintain the data and programs for at least six years.