DATA AND CODE SHARING POLICY SHEET
for
"Do managers of U.S. defined benefit pension plan sponsors use regulatory freedom strategically?"

Michael Kisser* John Kiff† Mauricio Soto‡

July, 2017

*Norwegian School of Economics (michael.kisser@nhh.no)
†International Monetary Fund (jkiff@imf.org)
‡International Monetary Fund (msoto@imf.org)

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

Michael Kisser handled the data and conducted the 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).

The paper works with different sources of data. We categorize the raw data according to the different sections of the paper.

Section 4: Raw data on corporate defined benefit pension plans from 1999 to 2007 were downloaded from the Boston Center for Retirement Research, Subsection Form 5500 Annual Reports. The download took place on November 4, 2010.[1] Because the Form 5500 only contains information on which mortality table has been used (by a given pension plan at a specific point in time), we further hand collected information on age specific death rates for all mortality tables of our sample period.[2]

Section 5: Raw data including financial statement information of U.S. corporations were downloaded from Compustat (WRDS) on December 7, 2016. Additional raw data on pension accounting information of U.S. corporations were downloaded from Compustat (WRDS) on October 2, 2014. For robustness checks, we also employ data on marginal tax rates (raw data were obtained by email from Prof. John Graham on December 16, 2016 (Graham and Mills, 2008)) as well as on the Gompers-Ishii-Metrick (GIM) corporate governance index (raw data were downloaded from Prof. Metrick's webpage on December 16, 2016; (Gompers et al., 2003)).[3]

Section 6: Raw data on corporate defined benefit pension plans for 2012 were downloaded from the homepage of the U.S. Department of Labor. The download to place on September 3, 2014.[4] Because the Form 5500 does not contain data on segement interest rates, we downloaded this information from the webpage of the Internal Revenue Service and manually incorporated this information in our computer programming code.[5]

All files have been downloaded and stored by Michael Kisser. They are are available upon request.

3. If the data are obtained from an organization on a proprietary basis...

Not relevant.

4. A complete description of the steps necessary to collect and process the data used in the final analyses reported in the paper

The corresponding information can be found at the beginning of Section 4 and 5 of the main paper, Tables 1 and 2 (of the main paper) and in the Online Appendix.

[1] The original download link was http://crr.bc.edu/data/form-5500-annual-reports/

[2] The pdf files summarizing the death rates under the different mortality tables were downloaded from the Society of Actuaries between October 18, 2011 and October 26, 2011. The files are available upon request.

[3] The original download link was http://faculty.som.yale.edu/andrewmetrick/data.html

[4] The original download link was https://www.dol.gov/ebsa/5500main.html

[5] The original download date is unknown but the information on segment interest rates can be downloaded from http://www.irs.gov/Retirement-PLans/Funding-Yield-Curve-Segment-Rates and the file is available upon request.

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 data has been handled in STATA and we have made all code files available. In total, there are seven different do files which are numbered chronologically. Files starting with "0" simply make the downloaded raw data user friendly (e.g. by aggregating individual files into one large stata file). The file starting with "1" reads the aggregated raw data relevant for Section 4 and performs the data cleaning steps that are explained in the paper and the Online Appendix. The file starting with "2" performs the analysis relevant for Sections 4.1 to 4.3. The file starting with "3" performs the analysis relevant for Section 4.4. The files starting with "4" merge the data relevant for Section 5. The file starting with "5" performs the analysis relevant for Section 5. The file starting with "6" cleans the raw data relevant for Section 6. The file starting with "7" performs the analysis relevant for Section 6.

We also provide separately three Excel files which include the respective identifiers for the final samples used in Sections 4, 5 and 6 of the paper. The file *Ident_Sec4.xls* includes the employment identification number (EIN; from Form 5500) and the year for which the Form 5500 was submitted (year). Consistent with Panel A of Table 1, the final sample consists of 48,800 plan-years. The file *Ident_Sec5.xls* includes the Global Company Key (gkvey; from CCM) and the fiscal year for which the annual report was submitted (fyear). Consistent with Panel B of Table 1, the final sample consists of 2,566 plan-sponsor years. The file *Ident_Sec6.xls* includes the employment identification number (EIN; from Form 5500) corresponding to the year 2012 of the Form 5550 submission. Consistent with Online Appendix Table 9, the final sample consists of 5,204 plans.

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

Michael Kisser will maintain all data and programs for at least six years.

References

Gompers, Paul, Joy Ishii, and Andrew Metrick, 2003, Corporate governance and equity prices, *Quarterly Journal of Economics* 118, 107–156.

Graham, John, and Lilian Mills, 2008, Simulating marginal tax rates using tax return data, *Journal of Accounting and Economics* 46, 366 – 388.