README Chiou and Klingler

Rule Significance and Interbranch Competition in Rulemaking Processes

AMERICAN POLITICAL SCIENCE REVIEW

Fang-Yi Chiou

Jonathan Klingler

Content

README Chiou and Klingler.docx

codebook_dataverse.pdf

SupplementalMaterials.pdf

Regression Results for Table 5, Appendix G, and robustness checks.pdf

MCMC results for the IRT model.Rda

R script for IRT model and related Tables and Figure.R

raters_30March22.txt

supplemental.txt

tableA1data.csv

tableA2data.csv

regression.do

Main results.dta

Table 5.dta

replication.zip

codebook_dataverse.pdf

The document ``**codebook_dataverse.pdf''** explains all datasets and variables used in the study, including those used as raters to generate significance scores for RINs, those used to generate descriptive tables and figures for our rules universe and significance scores, and the agency-Congress data and variables used to conduct our empirical analysis. This document also explains differences between the variables available on RINs only observed in the Spring 1995 issue of the Unified Agenda and the variables available for all other RINs.

SupplementalMaterials.pdf

The document **``SupplementalMaterials.pdf''** is the online appendix for the paper. It includes additional information on how the RIN data were obtained for our rules universe, and description of the rater variables used in the IRT. The appendix also includes information on our item response model and validation of our estimates through evaluation of construct validity and examination of extreme cases. The appendix also includes summary statistics of the variables used in our empirical analyses and a proof of Proposition 1 in the paper. Finally, the appendix includes the full results for Model CP and a discussion of examining the judicial branch as a veto player.

Regression Results for Table 5 and Appendix G, and robustness checks.pdf

The document **``Regression Results for Table 5 and Appendix G, and robustness checks.pdf''** contains additional results referred to or summarized in the main paper or the supplemental materials. Specifically, these include the full results for our robustness check in the main paper in which we consider rule productivity in terms of all rules, with and without NPRMs summarized in Table 5 in the main paper; the full results for our additional models including the Supreme Court referred to in Appendix G; and the full results of our robustness checks in which other legislative actors were used as the bicameral legislative veto pivots which were referred to in the main paper.

Replicating the IRT and Related Figures

The R script **``R script for IRT model and related Tables and Figure.R"** is run with the data file **``raters_30March22.txt''** to replicate the IRT model from our set of raters. This script is run with the file **``MCMC results for the IRT model.Rda''** to reproduce Table 1 in the main paper and Table A.5 in the Supplemental Materials/Appendix. The same R script is run with the data file **``supplemental.txt''** to generate Figures 1 and 2 in the main paper, as well as Figures A.1 and A.2 and Table A.4 in the Supplemental Materials/Appendix. This R script is run with the data files **``tableA1data.csv''** and **``tableA2data.csv''** to generate Tables A.1 and A.2 in the Supplemental Materials/Appendix, respectively.

Replicating the Empirical Results and Related Figures

The script in the Stata do-file **``regression.do''** is run with the data file **``Main results.dta**'' and **``table 5.dta**'' to reproduce the results presented in Tables 2, 3, 4, and 5 in the main paper. This do file also reproduces Tables A.8 and A.9 in the Supplemental Materials/Appendix, and Tables D.1, D.2, D.3, D.4, D.5, D.6, D.7, D.8, D.9, D.10, D.11, D.12, D.13, D.14, D.15, and D.16 in **``Regression Results for Table 5 and Appendix G, and robustness checks.pdf''**

Replicating the Data Used in the Main Study

The file, "**replication.zip**" is an archive containing all of the source data and scripts required to replicate the datafiles used in the main study and described above. All Python scripts were run using Python 2.

The files in "replication.zip" should be extracted to the user's Desktop in a folder called "Replication" and following this, in each R script in this archive, the path to the Desktop only needs to be redefined for the user's machine for data to be processed appropriately.

Replication Instructions for raters_30March22.txt

The XML files **UA1995F.xml, UA1996S.xml, UA1996F.xml, UA1997S.xml, UA1997F.xml, UA1998S.xml, UA1998F.xml, UA1999S.xml, UA1999F.xml, UA2000S.xml, UA2000F.xml, UA2001S.xml, UA2001F.xml, UA2002S.xml, UA2002F.xml, UA2003S.xml, UA2003F.xml, UA2004S.xml, UA2004F.xml, UA2005S.xml, UA2005F.xml, UA2006S.xml, UA2006F.xml, UA2007S.xml, UA2007F.xml, UA2008S.xml, UA2008F.xml, UA2009S.xml, UA2009F.xml, UA2010S.xml, UA2010F.xml, UA2011S.xml, UA2011F.xml, UA2012.xml, UA2013S.xml, UA2013F.xml, UA2014S.xml, UA2014F.xml, UA2015S.xml, UA2015F.xml, UA2016S.xml, UA2016F.xml, UA2017S.xml, UA2017F.xml, UA2018S.xml, UA2018F.xml,** and **UA2019S.xml** were downloaded from https://www.reginfo.gov/public/do/eAgendaXmlReport. The Python script **Initial UA Parser.ipynb** processes each of these 47 XML files, each of which comprises a single issue of the Unified Agenda (note that there is a single yearly issue for 2012), and produces a set of 47 CSV files of the format UA+YYYY+(S/F)+U.csv, where YYYY denotes the year and S/F denotes the Spring or Fall issue. These files are **UA1995FU.csv, UA1996SU.csv, UA1996FU.csv, UA1997SU.csv, UA1997FU.csv, UA1998SU.csv, UA1998FU.csv, UA1999SU.csv, UA1999FU.csv, UA2000SU.csv, UA2000FU.csv, UA2001SU.csv, UA2001FU.csv, UA2002SU.csv, UA2002FU.csv, UA2003SU.csv, UA2003FU.csv, UA2004SU.csv, UA2004FU.csv, UA2005SU.csv, UA2005FU.csv, UA2006SU.csv, UA2006FU.csv, UA2007SU.csv, UA2007FU.csv, UA2008SU.csv, UA2008FU.csv, UA2009SU.csv, UA2009FU.csv, UA2010SU.csv, UA2010FU.csv, UA2011SU.csv, UA2011FU.csv, UA2012U.csv, UA2013SU.csv, UA2013FU.csv, UA2014SU.csv, UA2014FU.csv, UA2015SU.csv, UA2015FU.csv, UA2016SU.csv, UA2016FU.csv, UA2017SU.csv, UA2017FU.csv, UA2018SU.csv, UA2018FU.csv,** and **UA2019SU.csv.**

The CSV file **slowroll.csv** is a dataset created by Rachel Potter used in the study in her 2017 JOP article. The CSV file **FYC.pca.scores.csv** is a dataset provided by Rachel Potter with information on her New York Times data on rules.

The Python script **Compartmentalized Docket Scrape.ipynb** draws on **rins.csv**, which is a handcoded list of RINs from our rules universe, to obtain data from the Regulations.gov API on the dockets associated with the RINs in our study. This script takes a very long time to run (more than 24 hours) and can easily be interrupted, so the output was saved in compartmentalized CSV files of roughly 500 RINs each. If the script is interrupted, the existing number of processed RINs may be added to object “rinlistindex1” on line 24, with the number of existing CSV files added to the object “compartmentupdated” on line 31, and the script may be run to continue calling and saving data from the point of the interruption. The files **docketids0.csv** through **docketids78.csv** (note that there is no “docketids28.csv” file) are the result of this process.

The Python script **Unified Rater Parsing.ipynb** draws on **rins.csv,** to obtain data from the Regulations.gov API on the public comments associated with the RINs in our study. This script takes a very long time to run and can be easily interrupted, so the output was saved in compartmentalized CSV files of roughly 500 RINs each. If the script is interrupted, the existing number of processed RINs may be added to object “rinlistindex1” on line 25, with the number of existing CSV files added to the object “compartmentupdated” on line 32, and the script may be run to continue calling and saving data from the point of the interruption. The files **pcraters0.csv** through **pcraters9999.csv** are the result of this process.

The CSV files **nprms0.csv, nprms1.csv, nprms2.csv, nprms3.csv, nprms4.csv, nprms5.csv, nprms6.csv, nprms7.csv,** and **nprms8.csv** are handcoded datasets verifying the existence of NPRMs in the data.

The Python script **Unified Agenda Parser March 2020.ipynb** draws on the file **rin_uadates.csv**, which is a handcoded dataset consisting of information about the first issue of the Unified Agenda in which a RIN appeared to parse the Unified Agenda XML files and create a set of 47 CSVs with additional information extracted from the Unified Agenda data. These files are **UA1995Fmar_20.csv, UA1996Smar_20.csv, UA1996Fmar_20.csv, UA1997Smar_20.csv, UA1997Fmar_20.csv, UA1998Smar_20.csv, UA1998Fmar_20.csv, UA1999Smar_20.csv, UA1999Fmar_20.csv, UA2000Smar_20.csv, UA2000Fmar_20.csv, UA2001Smar_20.csv, UA2001Fmar_20.csv, UA2002Smar_20.csv, UA2002Fmar_20.csv, UA2003Smar_20.csv, UA2003Fmar_20.csv, UA2004Smar_20.csv, UA2004Fmar_20.csv, UA2005Smar_20.csv, UA2005Fmar_20.csv, UA2006Smar_20.csv, UA2006Fmar_20.csv, UA2007Smar_20.csv, UA2007Fmar_20.csv, UA2008Smar_20.csv, UA2008Fmar_20.csv, UA2009Smar_20.csv, UA2009Fmar_20.csv, UA2010Smar_20.csv, UA2010Fmar_20.csv, UA2011Smar_20.csv, UA2011Fmar_20.csv, UA2012mar_20.csv, UA2013Smar_20.csv, UA2013Fmar_20.csv, UA2014Smar_20.csv, UA2014Fmar_20.csv, UA2015Smar_20.csv, UA2015Fmar_20.csv, UA2016Smar_20.csv, UA2016Fmar_20.csv, UA2017Smar_20.csv, UA2017Fmar_20.csv, UA2018Smar_20.csv, UA2018Fmar_20.csv,** and **UA2019Smar_20.csv.**

The Python script **Federal Register API for NPRM and ANPRM Counts.ipynb** draws on **rins.csv** to obtain data from the Federal Register API on the number of NPRMs and ANPRMs existing for a RIN in that data source. This script takes a very long time to run and can be easily interrupted, so the output was saved in compartmentalized CSV files of roughly 1000 RINs each. If the script is interrupted, the existing number of processed RINs may be added to object “rincountindex1” on line 26, with the number of existing CSV files added to the object “compartmentupdated” on line 33, and the script may be run to continue calling and saving data from the point of the interruption. The files **recode_fr0.csv** through **recode_fr38.csv** are the result of this process.

The Python script **Federal Register API for NPRM and ANPRM Counts and Dates.ipynb** draws on the file **rin_uadates.csv** (only to get a list of NPRMs in our rules universe), to obtain data from the Federal Register API on the first NPRM existing for a RIN in that data source. This script takes a very long time to run and can be easily interrupted, so the output was saved in compartmentalized CSV files of roughly 1000 RINs each. If the script is interrupted, the existing number of processed RINs may be added to object “rinlistindex1” on line 27, with the number of existing CSV files added to the object “compartmentupdated” on line 32, and the script may be run to continue calling and saving data from the point of the interruption. The files **fr_nprms0.csv** through **fr_nprms38.csv** are the result of this process.

The CSV file **spring1995_cleaned_v2.csv** is a handcoded dataset consisting of several raters obtained from PDF copies of the Spring 1995 issue of the Unified Agenda. These data are discussed further in the Codebook. The file **spring95rins.csv** is a handcoded list of the RINs identified in the Spring 1995 issues of the Unified Agenda.

The Python script **Separate Statutory and Judicial Deadline Raters Parser.ipynb** draws on the file **rin_uadates.csv** (only to get a list of NPRMs in our rules universe), to parse the Unified

Agenda XML files and create a set of 47 CSVs with specialized information on statutory and judicial deadlines for the RINs in our study. These files are **UA1995F_deadlines.csv, UA1996S_deadlines.csv, UA1996F_deadlines.csv, UA1997S_deadlines.csv, UA1997F_deadlines.csv, UA1998S_deadlines.csv, UA1998F_deadlines.csv, UA1999S_deadlines.csv, UA1999F_deadlines.csv, UA2000S_deadlines.csv, UA2000F_deadlines.csv, UA2001S_deadlines.csv, UA2001F_deadlines.csv, UA2002S_deadlines.csv, UA2002F_deadlines.csv, UA2003S_deadlines.csv, UA2003F_deadlines.csv, UA2004S_deadlines.csv, UA2004F_deadlines.csv, UA2005S_deadlines.csv, UA2005F_deadlines.csv, UA2006S_deadlines.csv, UA2006F_deadlines.csv, UA2007S_deadlines.csv, UA2007F_deadlines.csv, UA2008S_deadlines.csv, UA2008F_deadlines.csv, UA2009S_deadlines.csv, UA2009F_deadlines.csv, UA2010S_deadlines.csv, UA2010F_deadlines.csv, UA2011S_deadlines.csv, UA2011F_deadlines.csv, UA2012_deadlines.csv, UA2013S_deadlines.csv, UA2013F_deadlines.csv, UA2014S_deadlines.csv, UA2014F_deadlines.csv, UA2015S_deadlines.csv, UA2015F_deadlines.csv, UA2016S_deadlines.csv, UA2016F_deadlines.csv, UA2017S_deadlines.csv, UA2017F_deadlines.csv, UA2018S_deadlines.csv, UA2018F_deadlines.csv,** and **UA2019S_deadlines.csv.**

The Python script **Page Length API Scrape.ipynb** draws on the file **fullrinlist.csv,** which is a handcoded list of all the RINs in our rules universe (including Spring 1995), to obtain data from the Federal Register API on the page length of the first Federal Register entry for an RIN in that data source. This script takes a very long time to run and can be easily interrupted, so the output was saved in compartmentalized CSV files of 1000 RINs each. If the script is interrupted, the existing number of output files (all of which must consist of the full set of planned 1000 RINs) may be added to object "existingfiles" on line 26, and the script may be run to continue calling and saving data from the point of the interruption. The 40 files **fr_pagenum20.csv** through **fr_pagenum239.csv** are the result of this process.

The CSV file **wapo.csv** is a handcoded dataset consisting of the Washington Post raters described in our codebook. **RIN_order APSR.csv** is a handcoded dataset consisting of an ordered list of the RINs in our dataset to ensure consistency across the data in our rules universe.

The R script **unified_rater_merge.R** draws from all of the files listed up until this point in the "*Replication Instructions for raters_30March22.txt"* section to produce **rules_universe_march24_22 APSR.csv**, which is used to continue the replication procedure for "raters_30March22.txt." The script also produces **supplemental_source.csv** which is used in the replication of the "supplemental.txt" file.

The R script **R script for generating rater data file.R** draws from several files. The file **meetingRater.xls** is a handcoded datafile consisting of our lobbying rater, and the file **All RINs After correcting errors (complete).xls** is a handcoded datafile consisting of our Hein Online citation rater. These two files are processed along with "rules_universe_march24_22 APSR.csv," described in the prior paragraph, to generate the output file, **raters_30March22.txt**, used in the main study.

Replication Instructions for Main results.dta

Part I: Replicating agency_vars_yearly_apr1222 APSR.csv

The DTA file **HANDSL01113C20_BSSE.dta** is a dataset downloaded from https://legacy.voteview.com/k7ftp/junkord/HANDSL01113C20_BSSE.DTA, and is the original source of our ideal point estimates. The Excel files **house_assignments_103_115_3.xls** and **senate_assignments_103_115_3.xls** are datasets of House and Senate committee assignments, respectively, created by Charles Stewart and Jonathan Woon and downloaded from web.mit.edu/17.251/www/data_page.html.

The R script **committee_medians_113 APSR.R** processes these files to generate two CSV files which contain the House and Senate committee median ideal points used in our study, respectively, **committee_medians_house_113 APSR.csv**, and **committee_medians_senate_113 APSR.csv.**

The R script **calculate_committee_ip_113 APSR.R** processes these two CSV files, along with **agencynames_numbers APSR.csv** (a handcoded dataset of every agency in our study's oversight committees in the House and Senate), and **department_order APSR.csv** (a handcoded dataset of the standard order of department IDs used to ensure order consistency throughout our data) to generate the output CSV file **agency_vars_yearly_apr1222 APSR.csv.**

Part II: Replicating agency_vars_outlays APSR.csv

The CSV files **fedscope1998.csv** through **fedscope2019.csv** contain employment data on federal agencies from the years 1999 through 2019, and were downloaded from https://www.fedscope.opm.gov/employment.asp by navigating to the September link for the relevant year. In the upper left, under Employment, the folder "Agency – All" was highlighted but not opened. In the bottom left, the "Insert Before Rows" image button was clicked. Then in the very bottom to the left of the question mark, the File image button was clicked and the data were obtained by exporting a CSV. The file **employment95.csv** was obtained through an email request to fedstats@opm.gov requesting September total employment count data for all agencies for 1995, 1996, and 1997. The CSV file **CJ.csv** is a handcoded dataset containing the OPM codes for each department in our study.

The R script **fedscopemerger APSR.R** processes the CSV files described in the prior paragraph, along with **department_order APSR.csv** (a handcoded dataset of the standard order of department IDs used to ensure order consistency throughout our data) to generate the Excel output file **agency_vars_outlays APSR.xls.**

Part III: Replicating rules_universe_may20_v2 APSR.xls

Several files created during the replication procedure for "raters_30March22.txt" are used in this section. These files are **UA1995Fmar_20.csv, UA1996Smar_20.csv, UA1996Fmar_20.csv, UA1997Smar_20.csv, UA1997Fmar_20.csv, UA1998Smar_20.csv, UA1998Fmar_20.csv, UA1999Smar_20.csv, UA1999Fmar_20.csv, UA2000Smar_20.csv, UA2000Fmar_20.csv, UA2001Smar_20.csv, UA2001Fmar_20.csv, UA2002Smar_20.csv, UA2002Fmar_20.csv, UA2003Smar_20.csv, UA2003Fmar_20.csv, UA2004Smar_20.csv, UA2004Fmar_20.csv, UA2005Smar_20.csv, UA2005Fmar_20.csv, UA2006Smar_20.csv, UA2006Fmar_20.csv, UA2007Smar_20.csv, UA2007Fmar_20.csv, UA2008Smar_20.csv, UA2008Fmar_20.csv,**

UA2009Smar_20.csv, UA2009Fmar_20.csv, UA2010Smar_20.csv, UA2010Fmar_20.csv, UA2011Smar_20.csv, UA2011Fmar_20.csv, UA2012mar_20.csv, UA2013Smar_20.csv, UA2013Fmar_20.csv, UA2014Smar_20.csv, UA2014Fmar_20.csv, UA2015Smar_20.csv, UA2015Fmar_20.csv, UA2016Smar_20.csv, UA2016Fmar_20.csv, UA2017Smar_20.csv, UA2017Fmar_20.csv, UA2018Smar_20.csv, UA2018Fmar_20.csv, and **UA2019Smar_20.csv.** These files created in the replication of "raters_30March22.txt" also include **recode_fr0.csv** through **recode_fr38.csv** and **fr_nprms0.csv** through **fr_nprms38.csv.** The CSV file **agencynames.csv** is a handcoded dataset of the agency and department names and acronyms associated with the RIN prefixes used in the study.

The Python script **Unified Agenda Final Rule Date.ipynb** draws on **rin_uadates.csv**, a handcoded list of RINs in our study, described earlier, to parse the Unified Agenda XML files and create a set of 47 CSVs with specialized information on the final rule dates of the RINs in our study. These files are **UA1995Ffd.csv, UA1996Sfd.csv, UA1996Ffd.csv, UA1997Sfd.csv, UA1997Ffd.csv, UA1998Sfd.csv, UA1998Ffd.csv, UA1999Sfd.csv, UA1999Ffd.csv, UA2000Sfd.csv, UA2000Ffd.csv, UA2001Sfd.csv, UA2001Ffd.csv, UA2002Sfd.csv, UA2002Ffd.csv, UA2003Sfd.csv, UA2003Ffd.csv, UA2004Sfd.csv, UA2004Ffd.csv, UA2005Sfd.csv, UA2005Ffd.csv, UA2006Sfd.csv, UA2006Ffd.csv, UA2007Sfd.csv, UA2007Ffd.csv, UA2008Sfd.csv, UA2008Ffd.csv, UA2009Sfd.csv, UA2009Ffd.csv, UA2010Sfd.csv, UA2010Ffd.csv, UA2011Sfd.csv, UA2011Ffd.csv, UA2012fd.csv, UA2013Sfd.csv, UA2013Ffd.csv, UA2014Sfd.csv, UA2014Ffd.csv, UA2015Sfd.csv, UA2015Ffd.csv, UA2016Sfd.csv, UA2016Ffd.csv, UA2017Sfd.csv, UA2017Ffd.csv, UA2018Sfd.csv, UA2018Ffd.csv,** and **UA2019Sfd.csv.**

The Python script **Federal Register API for Promulgation Dates (Final Not Needed; Interim and Direct).ipynb** draws on **rin_uadates.csv** to obtain data from the Federal Register API on the promulgation dates of all RIN from our study in that data source. This script takes a very long time to run and can be easily interrupted, so the output was saved in compartmentalized CSV files of 1000 RINs each. If the script is interrupted, the existing number of processed RINs may be added to object "rincountindex" on line 24, and the existing number of existing output files may be added to object "compartment" on line 31, and the script may be run to continue calling and saving data from the point of the interruption. The files **fr_finalcomp20.csv** through **fr_finalcomp239.csv** are the result of this process.

The CSV file **spring1995_cleaned_v2.csv** is a handcoded dataset consisting of several raters obtained from PDF copies of the Spring 1995 issue of the Unified Agenda. These data are discussed further in the Codebook. The file **spring95rins.csv** is a handcoded list of the RINs identified in the Spring 1995 issues of the Unified Agenda. **RIN_order APSR.csv** is a handcoded dataset consisting of an ordered list of the RINs in our dataset to ensure consistency across the data in our rules universe.

The R script **replication_dv APSR.R** draws on all of the CSV files listed in this subsection to produce the output file **rules_universe_may20_v2 APSR.xls.**

Part IV: Replicating Main_results.dta

The R script **Code for main empirical results APSR 20220410.R** draws on several CSV files already described, "agency_vars_yearly_apr1222 APSR.csv" from Part I in this section, "agency_vars_outlays APSR.xls" from Part II in this section, "rules_universe_may20_v2 APSR.xls" from Part III in this section, "X (ORIGINAL FOR COMPARISON).txt" from the next section. This script also draws on the handcoded file, **complete_ivs APSR.xls**, along with **Chen_Johnson_Agency_IdealPointsJTP_2_DeptNum.xls**, obtained from Jowei Chen and Tim Johnson, **HANDSL01114A20_STAND_ALONE.xls,** obtained from voteview.com in DAT format and saved as an Excel file, and **jcs_supreme_court_2020_long.RData**. obtained from Lee Epstein at https://epstein.usc.edu/jcs. The output of this R script is the file **main results.txt.** The content of this file was copied and pasted into STATA's Data editor and saved as **main results.dta.**

Replication Instructions for Table 5.dta

The R script **Code for Table 5 APSR 20220410.R** draws on the same files already described in the prior paragraph, "agency_vars_yearly_apr1222 APSR.csv," "agency_vars_outlays APSR.xls," "rules_universe_may20_v2 APSR.xls," "X (ORIGINAL FOR COMPARISON).txt," "complete_ivs APSR.xls," "Chen_Johnson_Agency_IdealPointsJTP_2_DeptNum.xls," "HANDSL01114A20_STAND_ALONE.xls," and "jcs_supreme_court_2020_long.RData." The output of this R script is the file **Table 5.txt.** The content of this file was copied and pasted into STATA's Data editor and saved as **Table 5.dta.**

Replication Instructions for supplemental.txt

The CSV file **rules_universe_march24_22.csv** is a dataset of information on the RINs in our rules universe which is produced by the R script "unified_rater_merge.R" in the process of replicating "raters_30March22.txt," and described above. The CSV file **agencynames.csv** is a handcoded dataset of the agency and department names and acronyms associated with the RIN prefixes used in the study. The CSV file **slowroll.csv** is a dataset created by Rachel Potter used in the study in her 2017 JOP article. The CSV file **supplemental_source.csv** is a dataset of additional information on the RINs in our rules universe, and this file is also produced by the R script "unified_rater_merge.R" in the course of replicating "raters_30March22.txt," and described above.

Finally, **X (ORIGINAL FOR COMPARISON).txt** is a dataset of the significance estimates produced by running the R script, **R script for generating significance scores.R,** which processes the data file "MCMC results for the IRT model.Rda" from the files used in the main study, and the CSV file "rules_universe_march24_22 APSR.csv" to produce the output.

The R script **supplemental APSR.R** processes the five bolded data files in the paragraph above to produce the output file **supplemental.txt.**

Replication Instructions for tableA1data.csv and tableA2data.csv

The XML files, **UA1999S.xml, UA1999F.xml, UA2000S.xml, UA2000F.xml, UA2001S.xml, UA2001F.xml, UA2002S.xml, and UA2002F.xml** were downloaded from https://www.reginfo.gov/public/do/eAgendaXmlReport. The file **rin_uadates.csv** is a handcoded dataset consisting of information about the first issue of the Unified Agenda in which a RIN appeared. The Python script **Table UA Parser.ipynb** parses the eight aforementioned XML files,

drawing on rin_uadates.csv, to produce eight datafiles with information on the RINs in each Unified Agenda issue, which are **UA1999SUCD2.csv, UA1999FUCD2.csv, UA2000SUCD2.csv, UA2000FUCD2.csv, UA2001SUCD2.csv, UA2001FUCD2.csv, UA2002SUCD2.csv, and UA2002FUCD2.csv.**

The Python script **Table FR Parser.ipynb** calls the Federal Register API to collect data on every proposed rule and final rule published in the Federal Register in 2000, saving the proposed rule data in **fr_audit1.csv, fr_audit2.csv,** and **fr_audit3.csv**, and saving the final rule data in **fr_audit1_rule.csv, fr_audit2_rule.csv, fr_audit3_rule.csv, fr_audit4_rule.csv,** and **fr_audit5_rule.csv.**

The R script **tablesA1A2 APSR.R** processes the eight Unified Agenda CSV files, along with the eight Federal Register CSV files listed above, and produces **tableA1data.csv** and **tableA2data.csv.**

Finally, the file "**LaTex for Figure 3.tex"** is the tex file required to construct Figure 3 in our paper.

For any questions on this APSR Dataverse entry, please reach out to Jonathan Klingler (jklingle@olemiss.edu) or Fang-Yi Chiou (fchiou@gate.sinica.edu.tw).