Keep Winning with WinRed? Online Fundraising Platform as the Party's Public Good

Overview of Replication Materials and Codebook

Seo-young Silvia Kim and Zhao Li

Contents

Steps to Replicate **2**

Operating System, Software Version, and Computing Power **2**

Replication Outputs **2**

Data Sets **3**

R Scripts **4**

Codebook for Data Sets **5**

panel_filtered.Rda 5
actblue_blog_content_list.Rda 9
winred_blog_content_list.Rda 9
portfolio.Rda 10
corp_itemized.Rda 11
pnas_refunds_by_day.csv 12

Steps to Replicate

In this document, datasets and other such global environment objects are indicated by LaTeX syntax `\textbf` (bold), package names, data frame's column names, folder structure, and file formats are indicated by `\texttt` (typewriter), and file names (including `R` scripts) are indicated by `\textit` (italicized).

1. Download the replication .zip file.
2. Unzip the .zip file.
3. Run *main.R*. If the relevant packages are not yet installed, uncomment the installation steps outlined in the same script and install the required `R` packages.

Operating System, Software Version, and Computing Power

All replication work was conducted on a Linux-based as well as a Windows-based RStudio, using R version 4.4.1. The packages used are recorded with the `renv` package.

Replication Outputs

All replicated figures, in `.pdf` format are found under `/fig`. Similarly, all replicated `LaTeX` tables, in `.tex` format, are found under `/tab` subfolder. All model estimates, in `.Rda` format, generated from the `R` scripts described the preceding section are found under `/output`. However, note that the following figures and tables are not accompanied by replication code as they do not report original analytical results:

- Tables 1 and F.3 present our choices of keywords for the purpose of training keyword-assisted topic models.
- Figure 1 contains the screenshots of top documents for the party topic analyzed through `keyATM` package.
- Figure 2 contains the screenshot of WinRed's launch announcement.
- Table B.1 summarizes pricing structures by ActBlue and WinRed as reported in sources that we reference.
- Table E.1 summarizes the candidate payoff matrix used in the public good game.
- Table E.2 describes the pure-strategy equilibria without party intervention.
- Figure G.1 contains screenshots of fundraisers by Marjorie Taylor Greene.

Data Sets

All data sets used for replication are stored under `/data/raw`. The following table shows descriptions of individual files.

panel_filtered.Rda	This file contains a single list object, **df_ls**, which itself contains four quarterly panel data sets of congressional Republican candidates from Q1 of 2015 to Q4 of 2020. The data includes each candidate's quarterly fundraising receipts, WinRed adoption and usage (or lack thereof), and a host of other covariates. The first element of this list object contains said panel data set for all Congressional Republican candidates, the second for House Republican candidates, the third for Senate Republican candidates, and the fourth for Republican incumbents. It is used for replicating tables and figures related to `PanelMath` analyses.
actblue_blog_content_list.Rda	This file contains all ActBlue's blog entries, scraped on Nov 15, 2022. There are total 420 blog post texts.
winred_blog_content_list.Rda	This file contains all WinRed's blog entries, scraped on Nov 13, 2022. There are total 108 blog post texts.
portfolio.Rda	This file contains what amounts of money congressional candidates asked for when soliciting donations from individual campaign contributors through their official campaign website throughout early 2022 election cycle.
corp_itemized.Rda	This file contains a panel dataset of campaign contributions made from committees sponsored by interest groups (such as corporations) to congressional candidates from Q1 of 2015 to Q4 of 2020.
pnas_refunds_by_day.csv	This file contains select variables from the dataset *refunds_by_day.csv* as a part of the replication materials of Posner et al. (2023). This is a campaign-by-day-by-refund type panel dataset. For each of the campaigns studied Posner et al. (2023) and each day during August 1–November 3, 2020, this dataset summarizes the donations that said campaigns received through WinRed by donation source and refund status.

R Scripts

All R scripts are found under /R, except for *init.R*, which installs the R packages to be used and lives in the root directory. Given that package versions may change over time, at the point of replication, it's best to use the installation steps outlined in *main.R*, but we leave *init.R* as a trace of how we installed the packages and set up the directory.

A description of each individual script is below.

main.R	Executes the replication in a single step
utilities.R	Loads R packages and user-defined functions, and sets the working directory using the `here` package
01_descriptives.R	Replicates Figures H.1 and H.2
02_prediction_exercise.R	Replicates Tables H.1 and H.2
03_platform_adjustment.R	Replicates Table G.1 and Figure 4
04_blog_text_analysis.R	Replicates Tables 2 and F.4, and Figures 3, F.1, F.2, F.3, F.4, F.5
05_panel_benchmark.R	Replicates Figure I.1 and Table I.1
06_synthetic_panel.R	Replicates Tables I.2, I.3, and I.4, and Figures 5, 6, I.2, I.3, I.5, I.6, I.7, I.8, I.9, and I.10
07_heterogeneous_effects.R	Replicates Figures 7 and 8
08_placebo_corporate.R	Replicates Figure I.4
09_recurring_donations.R	Replicates Tables I.5, I.6, and I.7

Other figures and tables are not targets of replication. See explanation under an earlier section titled "Replication Outputs." on page 2.

Codebook for Data Sets

All these datasets are stored under `/R` subfolder, as previously stated.

panel_filtered.Rda

This is the main file for replicating tables and figures related to `PanelMatch` analyses. It contains a list object named **df_ls**, which itself includes four data sets:

full	A quarterly panel data set of all congressional Republican candidates from Q1 of 2015 to Q4 of 2020, which includes each candidate's quarterly fundraising receipts, WinRed adoption and usage (or lack thereof), and a host of other covariates.
house	A quarterly panel data set of House Republican candidates from Q1 of 2015 to Q4 of 2020, containing the same set of variables as those in ***full***
senate	A quarterly panel data set of Senate Republican candidates from Q1 of 2015 to Q4 of 2020, containing the same set of variables as those in ***full***.
inc	A quarterly panel data set of Republican incumbents from Q1 of 2015 to Q4 of 2020, containing the same set of variables as those in ***full***.

Related scripts. This file is used in

- *01_descriptives.R* to replicate Figures H.1 and H.2.
- *02_prediction_exercise.R* to replicate Tables H.1 and H.2.
- *05_panel_benchmark.R* to replicate Figure I.1 and Table I.1.
- *06_synthetic_panel.R* to replicate Tables I.2, I.3, and I.4, and Figures 5, 6, I.2, I.3, I.5, I.6, I.7, I.8, I.9, and I.10.
- *07_heterogeneous_effect.R* to replicate Figures 7 and 8.
- *08_placebo_corporate.R* to replicate Figure I.4.
- *09_recurring_donations.R* to replicate Tables I.5, I.6, and I.7.

Variable names and definitions. Please see below for names and definitions of variables found in each element of the list object **df_ls** that *panel_filtered.Rda* contains:

Column name	Description
`office`	The office for which a candidate was running. "H" for House and "S" for Senate.
`state`	Abbreviation for the state in which a candidate was running.

`state_cd`	The congressional district in which a candidate is running. For House candidates, this would be the state abbreviation concatenated with the district number (including a leading 0 if single digit). For Senate candidates, this would simply be the state abbreviation.
`cand_id`	Candidates' Federal Election Commission (FEC) ID numbers.
`treated`	An indicator of whether a candidate had joined WinRed by a given year-quarter.
`treated_sum`	The total number of year-quarters for which a candidate had appeared on WinRed by the end of 2020.
`rpt`	A factor variable indicating year-quarters ranging from Q1 of 2015 to Q4 of 2020.
`last_name`	Candidate last name.
`first_name`	Candidate first name.
`inc`	The incumbency status of a candidate: incumbent, challenger, or open-seat.
`first_date`	The first date by which a candidate appeared on WinRed.
`incumbent`	An indicator of whether a candidate was an incumbent (1) or not (0).
`gender`	An indicator of whether a candidate is male (1) or female (0).
`seniority`	The number of years for which a candidate had served in Congress.
`PVI`	The Cook Partisan Voter Index (PVI) score for the congressional race that a candidate was in, scaled in terms of the Democrats' net advantage and ranging from -33 to 43. For example, -33 corresponds to a congressional district or state with a 33-point advantage for recent Republican presidential candidates.
`recipient_cfscore`	The recipient CFscore of a candidate (Bonica 2014).
`dwnom1`	The 1^{st}-dimension DW-NOMINATE score of a candidate.
`indv_ttl`	The dollar amount of a candidate's total fundraising from individual donors in each year-quarter.
`ttl`	The dollar amount of a candidate's total fundraising in each year-quarter.
`open`	An indicator of whether a candidate was running in an open-seat race (1) or not (0).
`senate`	An indicator of whether a candidate was running for the Senate (1) or not (0).
`first_quarter`	The first year-quarter during which a candidate appeared on WinRed.
`rpt_int`	An integer index corresponding to each year-quarter, ranging from 1 (representing 2015 Q1) to 24 (representing 2020 Q4).
`cand_id_int`	An integer index corresponding to `cand_id`.
`no_election`	An indicator of whether a candidate was not up for election (1) or otherwise (0).

`office_election`	An indicator of whether a candidate was running for the House, was a Senate candidate up for re-election, or was a Senate candidate not up for re-election.
`indv_ttl_log`	The total dollar amount ($\log(\cdot + 1)$ transformed) of individual campaign contributions raised by a candidate in each year-quarter.
`ttl_log`	The total dollar amount ($\log(\cdot + 1)$ transformed) of candidate fundraising in each year-quarter.
`oth_log`	The total dollar amount ($\log(\cdot + 1)$ transformed) of campaign contributions raised from non-individual sources by a candidate in each year-quarter.
`diff_ttl_log`	The difference between `ttl_log` in the present quarter and that in the preceding quarter.
`diff_indv_ttl_log`	The difference between `indv_ttl_log` in the present quarter and that in the preceding quarter.
`diff_oth_log`	The difference between `oth_log` in the present quarter and that in the preceding quarter.
`diff_item_log`	The difference between the total dollar amount ($\log(\cdot + 1)$ transformed) of a candidate's fundraising from itemized donations in the present quarter and that in the preceding quarter.
`diff_unitem_log`	he difference between the total dollar amount ($\log(\cdot + 1)$ transformed) of a candidate's fundraising from unitemized donations in the present quarter and that in the preceding quarter.
`state_club_size`	The proportion of same-state candidates who joined WinRed during the same year-quarter as a given candidate.
`avg_first_date_state`	The average date on which same-state candidates adopted WinRed.
`proximity_recipient_cfscore`	The inverse of the average squared distance between a candidate's recipient CFscore and those of WinRed adopters (Bonica 2014).
`proximity_dwnom1`	The inverse of the average squared distance between a candidate's 1^{st}-dimension DW-NOMINATE score and those of WinRed adopters.
`avg_first_date_proximity_cfscore`	The weighted average date of WinRed adopters, weighted by each adopter's distance in recipient CFscores to a given candidate (Bonica 2014).
`avg_past_ttl`	The average pre-2020 dollar amount of quarterly total fundraising for a candidate.
`avg_past_ttl_log`	The average pre-2020 dollar amount ($\log(\cdot + 1)$ transformed) of quarterly total fundraising for a candidate.
`ttl_opp_log`	The total dollar amount ($\log(\cdot + 1)$ transformed) of campaign funds raised by the opponent of a candidate in a year-quarter.
`avg_ttl_opp`	The average quarterly total dollar amount of fundraising by opponent(s) of a candidate.
`avg_indv_opp`	The average quarterly total dollar amount of campaign contributions raised from individual donors by opponent(s) of a candidate.

`avg_ttl_opp_log`	The average quarterly total dollar amount ($\log(\cdot + 1)$ transformed) of fundraising by opponent(s) of a candidate.
`avg_indv_opp_log`	The average quarterly total dollar amount ($\log(\cdot + 1)$ transformed) of campaign contributions raised from individual donors by opponent(s) of a candidate.
`avg_past_ttl_opp_log`	The average pre-2020 dollar amount ($\log(\cdot + 1)$ transformed) of quarterly total fundraising for opponent(s) of a candidate.
`region`	The U.S. region a candidate is from: North Central, Northeast, South, or West.
`pct_indv`	The ratio between `indv_ttl` and `ttl`.
`pct_unitem`	The ratio between the total dollar amount of candidate fundraising from unitemized campaign contributions and `ttl`.
`item_log`	The itemized individual donations' dollar amount ($\log(\cdot + 1)$ transformed) of candidate fundraising in each year-quarter.
`unitem_log`	The unitemized individual donations' dollar amount ($\log(\cdot + 1)$ transformed) of candidate fundraising in each year-quarter.
`indv_opp_log`	The individual donations' dollar amount ($\log(\cdot + 1)$ transformed) of campaign funds raised by the opponent of a candidate in a year-quarter.

actblue_blog_content_list.Rda

Related scripts. This file is used in *04_blog_text_analysis.R* to replicate Tables 2 and F.4, as well as Figures 3, F.1, F.2, F.3, F.4, and F.5.

Variable names and definitions. This file loads a list of 420 entries named `blog_content_list`. Each element is a blog post from ActBlue's blog, from its beginning to the scrape date. Each list element is again a list with six elements: `title`, `date`, `tag`, `content`, `social`, and `modified`. While others are straightforward, `modified` is a cleaned-up version of `content`, such as deleting extra spaces and tabs, and is the target of analyses such as semantic coherence, topic modeling, and so on.

winred_blog_content_list.Rda

Related scripts. This file is used in *04_blog_text_analysis.R* to replicate Tables 2 and F.4, as well as Figures 3, F.1, F.2, F.3, F.4, and F.5.

Variable names and definitions. This file loads a list of 108 entries named `blog_content_list`. Each element is a blog post from WinRed's blog, from its beginning to the scrape date. Each list element is again a list with eight elements: `title`, `date`, `content`, `header`, `subheader`, `body`, `itemize`, and `modified`. As with ActBlue's blog entries, `modified` is a cleaned-up version of `content`, such as deleting extra spaces and tabs, and is the target of analyses such as semantic coherence, topic modeling, and so on.

portfolio.Rda

This is the main file for replicating tables and figures related to adjustment to the increased FEC donation limit. It contains a list object named **df_ls**, which itself includes two data sets: one for Senate candidates and one for House candidates.

Related scripts. This file is used in *03_platform_adjustment.R* to replicate Table G.1 and Figure 4.

Variable names and definitions.

Column name	Description
`last_name`	Last name of candidates.
`state` or `state_cd`	The state (for Senate candidates) or state-district of the candidate (for House candidates).
`year`	The election cycle.
`date`	The date in which the fundraising portfolio was scraped.
`portfolio`	The amount suggested for fundraising to potential donors.
`url`	The main fundraising URL associated with the official campaign website of the candidate (linked to `portfolio`).
`party`	The partisan affiliation of the candidate.
`FEC_ID_cand`	FEC's unique candidate ID.
`class`	This is only contained in the `senate` data, indicating the class of Senate candidates, or when whose term of service expires (if elected).
`first_name`	First name of candidates.
`incumbent`	Incumbency status of candidates.

corp_itemized.Rda

This file is used to replicate the placebo test for our baseline `PanelMatch` analyses using corporate PAC contributions (see Appendix I.3). It contains a data frame object titled **corp_itemized**, which is a panel dataset of campaign contributions made from committees sponsored by interest groups (such as corporations) to congressional candidates.

Related scripts. This file is used in *08_placebo_corporate.R* to replicate Figure I.4.

Variable names and definitions.

Column name	Description
`org_tp`	Organization types for committees registered with the Federal Election Commission (FEC). "C" represents corporations as classified by the FEC. "T" represents trade associations as classified by the FEC. "W" represents corporations without capital stock as classified by the FEC. "CRP" represents other business organizations as classified by the Center for Responsive Politics. Finally, "L" represents labor unions as classified by the FEC.
`transaction_dt`	Transaction dates (in mdy format) of campaign contributions made by committees.
`transaction_amt`	Dollar amounts of campaign contributions made by committees.
`cand_id`	Candidates' Federal Election Commission (FEC) ID numbers.

pnas_refunds_by_day.csv

This is a campaign-by-day-by-refund type panel dataset that summarizes the donations received by campaigns studied in campaigns studied in Posner et al. (2023) through WinRed by donation source and refund status. It is used to analyze how congressional Republican candidates' historical reliance on unitemized Donors predicted their use of pre-checked recurring donations on WinRed (see Appendix I.6).

Related scripts. This file is used in *09_recurring_donations.R* to replicate Tables I.5, I.6, and I.7.

Variable names and definitions.

Column name	Description
`date`	The date (in y-m-d format) by which a campaign's donations raised through WinRed is summarized.
`source`	The source of donations raised by a campaign through WinRed. "Individual Donations" sum up all donations received through WinRed from individual contributors (separated by refund status; see the next variable). "Part of a weekly chain" denotes donations given in a weekly recurring fashion through WinRed. "Part of a monthly chain" denotes donations given in a monthly recurring fashion through WinRed. And "Precipitating donation" denotes the first donations as a part of either type of chains described earlier.
`recipient`	Abbreviated names of campaigns studied in Posner et al. (2023). "cg" is Cory Gardner. "djt" is Donald J. Trump. "dp" is David Purdue. "je" is Joni Ernst. "jj" is John James. "kk" is Kim Klacik. "km" is Kevin McCarthy. "lg" is Lindsey Graham. "lj" is Lacy Johnson. "maga" is the Trump Make America Great Again Committee. "mmc" is Mitch McConnell. "mms" is Martha McSally. "rnc" is the Republican National Committee. "nrcc" is the National Republican Congressional Committee. "nrsc" is the National Republican Senatorial Committee. "sc" is Susan Collins. "sd" is Steve Daines. "ss" is Steve Scalise. "ts" is Tim Scott. "tt" is Thom Tillis.
`refunded`	An indicator of whether donations are summarized for a given source (see description above) that were refunded (1) or not (0).
`donation_amount`	The total dollar amount of donations received by a given campaign on a given day by source and refund status.
`treat`	An indicator of whether a campaign ever added a weekly default option on their WinRed homepage between August 1, 2020 and November 3, 2020 (1) or not (0).
`weeks_to_treat`	The number of weeks since a campaign added a weekly default option on their WinRed homepage between August 1, 2020 and November 3, 2020 (1) or not (0).

References

Bonica, Adam. 2014. "Mapping the Ideological Marketplace". *American Journal of Political Science* 58 (2): 367–386.

Posner, Nathaniel, Andrey Simonov, Kellen Mrkva, and Eric J. Johnson. 2023, October. "Dark defaults: How choice architecture steers political campaign donations". *Proceedings of the National Academy of Sciences* 120 (40): e2218385120. Publisher: Proceedings of the National Academy of Sciences.