# Pre-analysis plan outline:

# Impacts of access to free news subscriptions of traditional news outlets on news consumption behavior, political attitudes, and democratic preferences

Caio Castro[*]  and  João F. Pugliese[†]

September 6, 2022

## Contents

[*]Sao Paulo School of Economics, caio.castro@fgv.br

[†]Stanford University, jfpugli@stanford.edu

# 1 Introduction

This project implements a field experiment aimed at providing access to major news outlets during the period before and after the 2022 Brazilian general elections. We partner with five leading outlets in the news industry, all of which employ paywalls, to randomize free three-month digital subscriptions to their content for about 1600 Brazilian voters.

Our experimental design has two distinctive features. First, we do not randomize *which* news outlet participants receive, but rather ask participants to choose their preferred outlet (among the five we offer) before randomizing the treatment. This feature of our experiment allows us to estimate the effects of consuming news outlets on the population that would willingly choose to consume such outlets. We consider this preference-aligned intervention to be an attractive empirical target since it is an intervention that more closely resembles organic changes in consumption. In addition, given previous research establishing a taste for like-minded news, we expect consumers to choose news outlets that fit their ideological positions. Thus, our design aims to increase news consumption, both in general and of these outlets, without introducing major shifts in *which* news respondents are reading.

In a baseline survey, all participants have to choose one among five possible outlets: *O Globo*, *Valor Econômico*, *Estadão*, *Carta Capital*, and *Gazeta do Povo*. The first three are among the country's top four largest circulation newspapers (considering both print and digital paid subscriptions), largely representing the "mainstream media" in the press.[1] The last two, *Carta* and *Gazeta*, have smaller circulation numbers but are widely considered leading outlets on the left and right-wing, respectively. Thus, we allow for a choice set of newspapers with diverse ideological orientations. We also nudge participants to subscribe to newsletters, download the outlet's apps on their smartphones, and implement an encouragement treatment, providing small monetary incentives tied to participants' performance in news quizzes specific to the outlets they choose.

To investigate the impacts of access to these news outlets, we collect survey data on a range of outcomes, including news knowledge, polarization, political engagement, trust in the electoral system, and support for democratic principles. We also obtain Analytics administrative data from the newspapers with detailed consumption patterns, both in browsers and smartphone apps. We study if and how news consumption responds to the treatment and how the treatment impacts individuals' information knowledge, political attitudes and preferences, and democratic values. Our experiment allows us to document the causal effects of news consumption of each one of the outlets conditional on each individual's choice of outlet.

We investigate outcomes broadly divided into two groups: i) news consumption and information knowledge and ii) political attitudes and preferences. The first group reflects outcomes intrinsically tied to news consumption, while the second captures news consumption's "political spillovers".

Covering news consumption, we first explore whether people consume more news or substitute between different mediums by using self-reported data. We also aim to inspect how people's opinions about the press and the outlets react to the treatment. We then turn to informational knowledge, covering three groupings: news knowledge, political knowledge, and knowledge of notorious events happening in the country. All three involve quizzes asking participants to correctly identify pieces of information. In order to measure

---

[1]The other popular newspaper among the top four, *Folha de São Paulo*, was unable to participate in the experiment.

factual awareness of notorious events, we aim to rely on phone surveys through WhatsApp that will take place once we identify a relevant political event.[2]

After establishing changes in the news diet and its consequences on information knowledge, we analyze the impacts on political attitudes and preferences. We study affective polarization between the two main candidates in the 2022 elections, and between right-wing and left-wing politicians and supporters; approval of the current government; turnout and voting decisions in the 2022 elections; and confidence in institutions, such as Congress and the Supreme Court. We also analyze political engagement and participation by considering attitudes taken by respondents during and after the electoral period, such as public displays of support or explicit campaigning for candidates. Finally, we investigate whether the treatment instills democratic values in the readers. We try to measure democratic values by considering hypothetical actions taken by politicians (such as contesting electoral results) and gauging whether the willingness to vote decreases in response to these actions.

The main contribution of this project is to provide an experimental design to investigate the impacts of access to major news outlets on individual's preferences, focusing on consumption changes along the preference curve of customers. By letting participants choose their most preferred newspaper among the five, we can explore a relatively natural change in consumption.

We anticipate that we may write multiple papers on different outcomes. However, all results may be combined into a single paper.

## 2 Experimental design, sample recruitment, setting, and flowchart

**Intervention overview.** The core of our design is as follows: in an online survey, we ask participants to choose between five news outlets to receive a free three-month subscription possibly. Then, we randomize participants into three groups: control $C$, treatment $T$, and treatment+incentives $TI$. $T$ and $TI$ are offered coupon discounts to claim a free three-month digital subscription to one newspaper's website (we describe the difference between groups $T$ and $TI$ below). Participants are offered five news outlets to choose from. The outlets will cover different ideological positions, allowing consumers a range of options. The control group will not receive any subscription but will also make a choice, allowing us to compare differences conditional on newspaper choice. The newspapers independently follow up on account registration by sending promotional emails inviting subscribers to register in the newsletters, download the smartphone app, and follow the outlets on social media as part of their consumer retention strategies. To facilitate usage, we also send instructions directing to the smartphone app download page and newsletter registration,

In group $TI$, compared to group $T$, we perform one extra intervention: we encourage group $TI$ to read the news by providing extra incentives in the form of news quizzes every two weeks. These news quizzes are short and tailored to each newspaper, and as such, participants only answer the quiz about the news outlet they chose and were randomized to receive. The quizzes are composed of questions about the newspapers' content covered in the past two weeks. Participants who perform well on the quizzes have extra chances to win sweepstakes of around U$20.

---

[2]We do not commit ex-ante to any rule as to how to choose such events as they happen; they will however, focus on the most relevant political stories circulating on the country. We will commit, however, to reporting results of any event we decide to cover, to tie our hands against cherry-picking certain events with particular results.

**Mechanisms underlying the intervention.** The direct mechanism behind our intervention is budget constraint alleviation. Some consumers may perceive news from these outlets as valuable but may not purchase the subscription because of price constraints. Indeed, the monthly prices of these subscriptions hover between R$15 to R$20, about 1.2% to 1.65% of the minimum monthly wage. However, consuming news is not as straightforward as consuming a commodity. Apart from prices, news consumption can depend on random exposition to headlines, "push factors" such as significant events, spillovers between topics (such as cultural content), habit formation, and of course, time constraints (Chen and Yang, 2019). Thus, we view providing links to app download pages and newsletters sign-ups as a different mechanism, trying to increase the exposure to news of the chosen newspaper.

**Randomization and group size.** Finally, we discuss the size of each group. We start by discussing the relative sizes of the treatment arms $TI$ and $T$. Many people do not actively consume news – i.e., opening an app or website, scanning through headlines, clicking on news articles and reading through them. News consumption can sometimes be daunting when much of the current news has a negative outlook, and can be a product of habit formation. Thus, we worry that simply giving out free subscriptions may not be enough to move individuals toward higher consumption. Our pilot data shows that only about 55-60% of consumers redeemed a coupon providing a free one-month subscription, and the average usage time from administrative data from our partner newspapers hovered around 1.5 minutes per day (about 10 minutes per week), with considerable heterogeneity between heavy users and users who did not access any content at all.

With this in mind, we will allocate a much larger fraction of participants to the treatment $TI$. We will keep the treatment $T$ as part of the experiment to document differences in consumption between the two groups and to gauge the effectiveness of the simple treatment package (coupons + newsletter/app message at the beginning of treatment).

To decide between the exact sizes of the control $C$, treatment $T$, and treatment $TI$ groups, we use pilot data from 150 respondents to run a power analysis. We use the self-reported news consumption in minutes as the main outcomes to compare $C$ vs. $TI$ and the administrative consumption data to compare power between $T$ and $TI$ treatments. We performed a simple simulation exercise considering different proportions of $C$, $T$, and $TI$, using bootstrap to draw from the pilot sample and a post-double selection method to select controls and estimate the standard error. After estimating the standard error, we scale it appropriately by the sample size. Importantly, we use average costs per individual in each group along with our budget of around U$25,000 to determine the sample size in each simulation.

Figure 1 displays the experimental design along with sample sizes. We allocate 55% of participants to the Control $C$ group, 40% to the Treatment+Incentives $TI$ group, and the remaining 5% to the Treatment $T$ group. Randomization was performed within the survey platform and was done through indepedendent draws, meaning that we did not impose the percentages of each group. Thus, the actual proportions of participante in each group are 54.6%, 4.7%, and 40.6%.

## 2.1 Sample recruitment, eligibility, and stratification

**Recruitment strategy.** Our recruitment strategy employs online advertisements across different Meta platforms – Facebook, Instagram, and Messenger. The ads are planned to run over a whole week. The

main set of ads is general, covering the Brazilian population over 18 years old. Recruiting online has the benefits of being cheaper, thus allowing for larger sample sizes and providing opportunities to interact with participants that professional survey companies would not be able to provide.[3] The downside of online recruitment through Meta platforms is, of course, a lack of representativeness.

The first point is that not everyone uses Instagram or Facebook. Thus, we are selecting social media users. The second issue is that the algorithm will try to maximize "conversions" – i.e., display the ad to those most likely to engage with it. Finally, in addition to those two issues, ads are not text-neutral; for instance, our ad, displayed in Figure 2, mentions "news consumption" in its body, which makes it more likely that people who are interested in the news click on it. Taken together, pilot data shows that these three issues result in a skewed sample toward left-wing, news-aware, richer, and higher-education participants. There is also a larger share of female participants in the pilot samples.

To try and correct for some of these imbalances, in addition to the main ad set, we also run targeted ads. Our targets include: i) individuals declaring having complete high-school education or lower, men, and those interested in husbandry and agriculture. We also ran targets that exclude people with higher-education (undergraduate level or higher). We note that all targets are relatively broad; for instance, the ad which targets people interested in husbandry and agriculture was roughly half the size of the whole population aged over 18 years. Our main goal with the targeting is to ensure we have enough variation in income, political views, and baseline news consumption to study heterogeneity according to these characteristics.

**Eligibility criteria.** Our eligibility criteria consist of: i) being a Brazilian citizen; ii) being 18 years or older; iii) having WhatsApp installed on their phone; iv) not currently paying a subscription for any news services; v) declaring interest in receiving the free 3-month subscription. The first two criteria are straightforward. The third criterion is operational: we will conduct part of our communications and the phone surveys through WhatsApp. We chose WhatsApp over SMS to increase engagement: in Brazil, WhatsApp is used by virtually all individuals with smartphones, and it is much more usual to check your WhatsApp than your SMS inbox.[4]. The fourth and fifth criteria are implemented to increase the odds that our sample is responsive to the treatment.[5]

**Stratification.** To "tie our hands" concerning heterogeneous treatment effects and ensure balance in these subgroups, we will also implement five strata in the randomization process. The first one concerns the political position reported by the individual (center, left, or right). The second one is the level of news knowledge at the baseline, calculated as described in Section 4.2 and using above and below-median as the groupings. Similarly, our third and fourth strata will consider self-reported news consumption and self-reported social media usage at baseline, both measured in minutes and also split by the median. Finally, our fifth stratum is the individual's newspaper of choice.

Importantly, we use data from the pilot to determine the strata cut-offs for continuous variables that

---

[3]For instance, many professional survey companies have a contract with their usual panelists that forbade them from sharing phone numbers with third parties, which would limit our ability to field phone surveys and send nudges through WhatsApp.

[4]*DataFolha*, one of the leading pollsters in Brazil, conducted a representative survey of app usage in March 2022. WhatsApp was used by 92% of those interviewed and is by far the most used social media among Brazilians. The numbers vary by years of schooling: among those who have completed only elementary school, 82% use WhatsApp; for high school, it is 95% and for superior education is 98%. Source: https://www1.folha.uol.com.br/tec/2022/07/94-tem-conta-em-alguma-rede-social-whatsapp-ldera-com-92.shtml

[5]In pilot data, about 80% of individuals expressed interest in a free one-month subscription.

depend on median splits – i.e., news knowledge, self-reported news consumption, and social media usage.

## 2.2 Setting

We consider Brazil an exciting environment to test news consumption's effects for political and economic reasons. Politically, Brazil is the fourth-largest democracy in the world and, since 2016, has been undergoing a democratic erosion process, with ongoing fears of further autocratization. The 2022 General Election is poised to be a tight race between left-wing former president Lula and the right-wing incumbent Jair Bolsonaro. Bolsonaro has been notorious for his attacks on the electoral system, in unique the electronic voting system, much in the spirit of Donald Trump's remarks about electoral fraud in the 2020 United States election. The level of concern with misinformation during the elections is reflected in a request by the Brazilian Federal Public Ministry (which is analogous to the Public Prosecution Office) for the delay of WhatsApp's implementation of the "Communities" feature, which would allow some users to send messages to up to 2,560 people at once. The 2018 elections were similarly mired in controversy over the use of WhatsApp and Telegram to quasi-organically spread misinformation, triggering in 2019 a parliamentary commission to investigate the usage of fake profiles to influence the 2018 elections.

Economically, Brazil is a highly unequal country. For example, the price of a monthly news subscription (among the most significant news organizations) is about 1.2% to 1.65% of the minimum wage.[6] This implies three desirable features for our experiment. First, since our primary treatment effectively reduces prices to zero, we expect this shift to be more significant in the Brazilian context. Second, we can investigate whether income inequality translates into information inequality. Third, in October 2022, Brazil will hold presidential elections, making for an interesting setting to evaluate how information changes voters' decisions and opinions.

Both economic and political dimensions suggest Brazil is an excellent opportunity to implement our treatment and investigate the effects of reading top outlets on news knowledge and political attitudes.

## 2.3 Experiment flow

This section describes the "user journey" throughout the experiment. Figure 3 presents the flowchart of the experiment, from the pre-treatment period, consisting of the screening and baseline surveys, to the treatment period up until the endline survey.

**Screening survey.** First, by clicking on the advertisement, the respondent opens a screening survey. On this survey, we check the eligibility, which consists of:

1. Being a Brazilian citizen;

2. Being 18 years old or more;

3. Having WhatsApp installed on phone;

4. Not currently paying for news subscriptions;

---

[6]Comparatively, the standard price for the *New York Times* subscription is around $17, costing the median American about 0.5% of their monthly gross income.

5. Declare interest in having the free 3-month subscription.

The screening survey ends if the respondent does not meet one of the conditions. On the other hand, if the respondent meets all the conditions, we proceed with a brief explanation of the study tasks, which we group into three different tasks: i) the possible coupon redemption and account creation, in case they are chosen to receive one; ii) the two main surveys, a baseline survey to be distributed in two days and an endline survey to follow in three months; and iii) the phone surveys that will happen regularly through WhatsApp. We then present the consent form, obtain consent from the participants, and collect contact information (email, telephone number) and demographic information (income, race, sex, and education).

After the end of the screening survey, we send a "welcome to the study" message through WhatsApp to confirm that the telephone number is reachable.

**Baseline survey.** Two days after the screening survey, we send the baseline survey to the individual by email and WhatsApp. Contacts who do not answer the baseline survey are excluded from the participant pool.

By the end of the baseline survey, we randomize the treatment: Control participants receive an end-of-survey message. In contrast, those in treatment receive an extra module with their coupon and instructions on redeeming it for each newspaper.

**Treatment implementation.** In addition to the survey instructions, we also send an email and WhatsApp message containing a document with instructions for coupon redemption and account creation on the newspaper websites. One day after first sending the baseline survey message, we follow up with a reminder. A few days later, we also remind those individuals who have not yet redeemed their coupon.[7]

**Phone surveys.** We run three types of phone surveys: i) news consumption survey, ii) incentives treatment quiz, and iii) notorious events survey. The news consumption survey is answered by all participants and repeats questions about news consumption quantity and mediums (TV, radio, etc.). As mentioned earlier in the design, the incentives treatment quiz is only sent to the Treatment Incentives (*TI*) group. It also happens biweekly and intersperses with the news consumption survey.

Finally, the notorious events survey is triggered at the researchers' discretion. Its objective is to measure participants' awareness of ongoing political events being covered extensively by the media in the country. Possible examples include the arrest of a former minister due to corruption, the opening of an inquiry to investigate one of the candidates, or the start of a scandal due to the leakage of some private information.

**Endline surveys.** After three months, all participants receive the endline survey via email and WhatsApp. We also send two reminders, one and two days after the first message, to try and reduce attrition.

# 3 Data and Benchmark Models

**Data.** For most of the outcomes, we will repeat the question from the baseline survey in the endline survey, and as such, we will have both pre and post-treatment data. The pre-treatment data will be used as controls for the post-treatment outcomes. Thus, we will have a cross-section of treated and control

---

[7]This information is given to us by the newspapers, and we aim to check redemption and send the coupon reminder no longer than one week after the baseline survey.

participants for most of the data.

There are a few exceptions to this. First, for self-reported news consumption outcomes, the phone surveys will repeat the questions every two weeks, allowing us to build a panel with multiple responses over the treatment period. A similar procedure is also feasible for administrative data from the newspapers and the browsing data since we obtain time-stamped Analytics data that allows for a complete time series for each participant.

The standard procedure for all scale-based categorical variables and numerical variables will be to standardize the variable by the control group mean and standard deviation. Standardized coefficients will be presented along with the raw control mean and standard deviation. In some situations, however, we may want to fully explore how the categorical variable has changed in response to the treatment. In such cases, we will transform categorical variables into multiple indicator variables and perform inference on each. Relatedly, categorical variables that are not scale-based will be transformed into multiple indicator variables for each possible response. Treatment effects on indicator variables will always be presented as percentage points.

**Benchmark model.** The main regression specification is

$$y_i = \alpha + \beta_T \cdot \text{Treatment}_i + \beta_{TI} \cdot \text{TreatmentIncentives}_i + X_i'\delta + \epsilon_i, \tag{1}$$

where $i$ indexes a participant and $\beta_T$ and $\beta_{TI}$ are the coefficients of interest. We may also perform double machine learning corrections on Equation (1). This can include, for instance, a method to choose the vector of controls, $X_i$, such as the post-double-selection method (Belloni et al., 2014). We do not commit to a particular double correction because the literature on the best ways to perform such corrections is still evolving. Generally, the vector of controls $X_i$ will include covariates such as strata indicators and the baseline measurement of $y_i$ whenever it is available.

**Panel extension.** We can construct a panel dataset for the outcomes where we can establish time series, such as self-reported news consumption from biweekly phone surveys or analytics data. As such, we can also use participant-fixed effects. In this case, we will also track results over time, as in the following specification:

$$y_{it} = \alpha + \sum_{t=1}^{T} \lambda_t \cdot \left( \beta_{T,t} \cdot \text{Treatment}_i + \beta_{TI,t} \cdot \text{TreatmentIncentives}_i \right) + \mu_i + \lambda_t + \epsilon_i, \tag{2}$$

In Equation (2), standard errors will be clustered at the participant level. Again, we may perform double machine learning corrections on the equation.

**Heterogeneous Effects Regressions.** To study heterogeneous effects, we will employ a simple specification where we interact the treatment assignments with the stratification indicators described above. The regression specification is:

$$\begin{aligned} y_i = \alpha + \beta_T \cdot \text{Treatment}_i + \beta_{TI} \cdot \text{TreatmentIncentives}_i + \\ S_i \cdot \left( \beta_{T,s} \cdot \text{Treatment}_i + \beta_{TI,s} \cdot \text{TreatmentIncentives}_i \right) + \\ \gamma S_i + X_i'\delta + \epsilon_i. \end{aligned} \tag{3}$$

Here, $S_i$ indicates whether or not the individual is part of the stratum being analyzed. The strata are chosen *ex-ante* to ensure balance in the cells, so we can reliably estimate heterogeneous treatment effects. In Equation (3) $S_i$ can represent different strata, which include:

1. $Outlets_i$: news outlet the individual choice as part of the treatment,

2. $AboveMedConsump_i$: whether the participant's self-declared news consumption is above or below the median baseline in our study population,

3. $AboveMedSocial_i$: whether the participant's self-declared social media usage is above or below the median baseline in our study population,

4. $AboveMedNewsKnowledge_i$: whether the participant's news knowledge is above or below the median baseline in our study population,

5. $PoliticalPosition_i$: whether the participant's political position is center, right, or left.

In addition to these heterogeneous effects built-in by the stratification, we also plan to explore the role of age, gender, and income as moderators. Finally, following our strategy above, we will re-estimate our specifications by interacting the treatment indicator with an indicator variable of the above and below median value on our sample.

# 4   Outcome Variables

To investigate the treatment-induced news consumption of the chosen outlet, we use administrative data provided by the five partner newspapers to track news consumption. The control group is coded as having zero for all observations, by default, since they declare having no digital news subscription. In practice, members of the control group could have bought a news subscription while the experiment was running. However, since we cannot observe their data due to data regulations, we must rely on the assumption that they did not purchase a subscription.

We consider the outcome variables in the five families described below for the impact evaluation. We divide outcomes into *primary outcomes*, which are the primary goal of our analysis, and *secondary outcomes*, which complement the analysis of the experiment. Each primary outcome and secondary outcomes belong to a family. Table 1 describes these family-wide indices and their components, which can be found in more detail below.

To summarize the families, we will construct indices that combine the primary outcome variables within each family, weighting by the inverse of the covariance between variables, as described in Anderson (2008). In constructing these indices, we orient the outcome variables so that more positive values have the same qualitative meaning.

| | Primary outcomes |
|---|---|
| **0. Manipulation check** | News consumption of chosen outlet (minutes), administrative data |
| **Family I. News consumption and opinions** | 1. News consumption (minutes), self-reported<br>2. News diet index<br>3. Positive opinions about press index<br>4. Willingness to continue subscribing |
| **Family II. Information knowledge** | 5. News knowledge<br>6. Event knowledge and awareness (measured by phone surveys)<br>7. Political knowledge |
| **Family III. Political attitudes** | 8. Polarization Lula-Bolsonaro<br>9. Political engagement index |
| **Family IV. Democratic preferences** | 10. Trust in electoral system index<br>11. Willingness-to-vote for democratic politicians index |

Table 1: Families of Outcomes

**Multiple hypothesis testing.** One of the main purposes of a pre-analysis plan is to tie hands against data mining. As such, we commit ex-ante to performing multiple hypothesis adjustments on the main primary outcomes listed in Table 1. As Anderson (2008) argues, summary index tests are methods to adjust for multiple hypothesis testing when we try to infer the overall impact of the treatment, and we make extensive use of these indices to summarize outcomes within families. We use them for each family and construct the primary outcomes within each family. In addition, we plan to inform False Discovery Rates (FDR) adjustments on the $p$-values for the 11 null hypotheses being tested – namely, that the treatment had effects on the 10 primary outcomes listed above. Our standard commitment is to perform the two-stage linear step-up procedure outlined in Benjamini et al. (2006). However, we appreciate that the multiple hypotheses testing literature is still evolving. Thus we also recognize that we may employ other FDR-adjustment techniques closer to the literature frontier, provided these procedures are clear improvements upon earlier methods. In such case, this will be explicitly noted in the final paper. Examples include Barber and Candès (2015) and Fithian and Lei (2020).

We now proceed to list and describe the primary and secondary outcomes. We note that the list of secondary outcomes presented here is not exhaustive. We also discuss some outcomes as needed.

## 4.0   Manipulation check

**Primary outcome:** News consumption of chosen outlet measured by administrative data.

- Our main objective with the manipulation check is to gauge whether the treatment was successfully implemented.

- Data comes from Google Analytics and contains a timestamp for any URL accessed in the partner newspapers' website or app.

- To measure minutes of usage, we use the difference between timestamps in two subsequent entries.

- This data covers both browsers and apps, both on computers and smartphones.

- We winsorize any entry with a duration above 5 minutes, and we discard the last entry of the day as we cannot measure its duration.

- Control participants are assumed to have zero minutes consumed.

## 4.1   Consumption, opinions and valuation of news

### 4.1.1   Total minutes of news consumption

**Primary outcome:** Self-reported minutes per day consuming news: ranges from 0 to 120 minutes per day (winsorized at 120 minutes).

**Secondary outcomes:** News consumption from browsing data. Calculated using timestamps similar to the administrative data, but only counts data from Google Chrome users who opt-in to providing this data.

- Browsing data is richer but has limitations since it can only capture browser data.

- With this data, we can measure the precise time spent reading news across different news websites.

### 4.1.2   Qualitative measure of amount of news consumption

**Primary outcome:** None.

**Secondary outcome:** Answer to question: *Thinking back over the last 4 weeks, how closely did you follow news about Brazil's politics?*

- 4-point scale from "Not at all closely" to "very closely."

### 4.1.3 News consumption medius (diet)

**Primary outcome:** News diet index.

- We ask individuals the following question for each medium of consumption: *"Over the past 4 weeks, how often did you watch/read/listen/get news from...?.* Possible answers are a 5-point scale from "Never" to "Very often".

- The nine mediums are: cable TV news, local TV news, print newspapers, national evening network TV news, social media websites (examples given are Facebook, Instagram, and Twitter), radio/podcasts, websites or apps other than social media, video platforms (examples given are TikTok, Kwai, and Youtube).

- We standardize responses into an index measuring *absolute* deviations from the control group. We calculate absolute deviations since we want to measure overall substitution between mediums.

**Secondary outcome:** Each of the eight individual components of the news diet index.

### 4.1.4 Opinions about press and news

**Primary outcome:** "Positive opinions about press" index.

- We group as index variables asking: level of confidence in the press, whether journalists provide fair coverage of elections, and whether journalists provide fair coverage of politics.

- Confidence in the press is measured as a 4-point scale from "Not at all" to "A lot."

- Fair coverage questions are measured as a 4-point scale from "Not at all often" to "Very often".

**Secondary outcomes:** Each of the three individual components of the index above, trust in the five partner newspapers, and changes in the perceived ideological position of newspapers.

- We ask respondents whether they trust the news reported by each of the partner newspapers. This is measured as a 5-point scale between "None at all" and "A great deal".

- We also ask respondents to classify the newspapers into ideological positions, from a 7-point scale between "Extreme left" and "Extreme right". We can thus measure the change in their perceptions from baseline.

### 4.1.5 Willingness to continue subscribing to news outlet

**Primary outcome:** Take-up of an offer to renew subscription

- At the end of the experiment, we will offer participants in both the control and treatments group an offer to renew their subscription to keep access to the content from the newspapers.

- We will measure differential take-up between treatment and control groups as a measure of the "willingness to continue subscribing."

**Secondary outcome:** None.

## 4.2 Information knowledge

### 4.2.1 News knowledge

**Primary outcome**: Score on 8-statement news knowledge quiz. Three will be "true news," three will be "incorrect news," and two will be "fake news."

- Scoring: Participants pick "True," "False," or "Not familiar with this information" for each of the 8 statements.

    1. One point for getting it right.
    2. Zero points for getting it wrong.
    3. Half-a-point for choosing "not familiar."

- The statement types included in the quiz are defined as:

    1. "True news": a set of relevant news that ran in the press at most two weeks before the quiz launch.
    2. "Incorrect news": set of incorrect news based on "true news" but slightly modified to be factually incorrect.
    3. "Fake news": set of actual, "malicious" fake news circulating; aimed to be chosen given their relevance to political events or the campaign (selected in partnership with a co-editor of Estadão Verifica). [8]

**Secondary outcomes:** scoring for each of the types of news – "true", "incorrect", and "fake."

### 4.2.2 Notorious events knowledge quiz

As notorious news happens during the study period, we can ask people whether they have been exposed to these events and are aware of them. These are similar to the Informational Knowledge quizzes above but with specific events. The number of exact events covered is undetermined ex-ante.

**Primary outcomes:** Score in the quiz of the relevant event being contemporaneously covered in the news (indicator variable). The number of items in the quiz may vary according to the event.

**Secondary outcomes:** None.

---

[8]A fact-checking division of Estadão, one of the partner newspapers.

### 4.2.3 Political knowledge

Next, we will investigate individuals' political knowledge with questions regarding political facts. We will keep the same format as the previous quizzes, with "True", "False", or "I am not familiar with this information". The difference in this quiz is that the aim is to capture political rather than news knowledge. Examples of questions asked may include electoral rules for the Senate/Congress, the composition of presidential tickets, and the name of Supreme Court Judges/Speaker of the House.

**Primary outcomes:** Score in the quiz of political knowledge.

**Secondary outcomes:** None.

## 4.3 Political attitudes

### 4.3.1 Polarization

**Primary outcome:** Polarization between Bolsonaro-Lula and Left-Right wing supporters.

- We use the "feeling thermometer" to measure affective polarization for different individuals/groups:

    1. President Bolsonaro.
    2. Former president Lula.
    3. Right-wing supporters.
    4. Left-wing supporters.
    5. Right-wing politicians.
    6. Left-wing politicians.
    7. Supreme Court.

- Polarization between Bolsonaro and Lula and between right and left-wing supporters serve as measures of the polarization between the two main camps of Brazilian politics.

**Secondary outcomes**: Polarization between Left and Right-wing politicians, and polarization between Supreme Court and Bolsonaro.

### 4.3.2 Political engagement

**Primary outcome**: Political engagement index.

- We try to measure how intensively individuals participate in political activities, among the following:

    – Actively sharing content for a candidate in WhatsApp/Telegram.
    – Actively sharing content for a candidate on other social media.
    – Helping a candidate out by trying to convert voters.
    – Helping a candidate out by actively joining the campaign as a volunteer for them.

- – Helping a candidate out by donating money to their campaign.

- – Watching debates for governor, president.

- – Watching interviews with candidates.

- – Changing profile picture in social media to support a candidate.

- – Distributing pamphlets for a candidate.

- – Attaching stickers to their cars/motorcycles.

- – Attending rallies.

- Each of these outcomes will be measured according to a 4-point scale to try and capture the intensity.

**Secondary outcomes:** Turnout and voting decision.

- We measure self-reported turnout and voting decisions in the presidential elections.

- This is considered a secondary measure because we believe we are underpowered to identify effects on such a latent outcome.

## 4.4  Democratic opinions and preferences

We aim to measure democratic preferences. Nevertheless, there are many challenges in measuring the level of support for democracy.[9] In our case, the main concern is not the level itself but rather the sensitivity of the measures to the treatment in our setting. For instance, most broad measures of democracy try to capture a general feeling for democratic government systems.[10] Given that Brazil has been a democratic country since 1985, variables capturing general tastes for democracy generally have small variation and thus are likely to be unresponsive to the treatment.

Our main goal is to gauge whether our treatment can alleviate contemporaneous concerns of democratic backsliding in modern democracies. We focus on two main aspects: distrust in the electoral system and support for anti-democratic, "hardball" measures. The latter can include actions such as expanding the Supreme Court (or reducing retirement ages), attacking the free press, or abusing executive orders.

Measuring trust in the electoral system is relatively straightforward. However, measuring support for anti-democratic actions can be more delicate. Ideally, we want to capture voters' willingness to vote-out politicians who take or support undemocratic actions. Inspired by candidate-choice experiments (such as Graham and Svolik (2020)), we implement questions capturing "willingness-to-vote" where we describe hypothetical actions that can be taken by politicians (or specific public offices, such as the president) and ask participants whether these actions made them more or less prone to vote for the hypothetical politician. In addition, we also ask a batch of "traditional" questions measuring broader support for democracy. Finally, we also include a question about military rule, given Brazil's recent military dictatorship (1964-1985).

---

[9]See, for instance, the International Institute for Democracy and Electoral Assistance (IDEA) resource guide, "Measuring Public. Support for Democracy"

[10]As an example, the World Value Surveys asks, among other questions, "how important is it for you to live in a country that is governed democratically?".

### 4.4.1 Electoral dimension

**Primary outcome:** Trust in electoral system index.

- We ask individuals the following questions regarding their trust in the electoral system:

  1. Free and fair elections: *In your view, are federal elections held in Brazil free and fair?*. Possible answers are "Yes" or "No."
  2. Trust in the electronic voting system: a 3-point scale from "Do not trust" to "Trust a lot."
  3. Confidence in elections.

- We standardize responses into an index measuring *absolute* deviations from the control group.

**Secondary outcome:** Each of the three individuals components of the trust in electoral system index. We also include a question about support for an alternative voting system called the "auditable printable vote", which was pushed by supporters of incumbent Jair Bolsonaro but widely discredited by the Electoral Court.

### 4.4.2 Willingness-to-vote for democratic politicians

**Primary outcome:** Willingness-to-vote for democratic politicians index.

- We present statements regarding the opinion of their candidate, and we present a 5-point scale from "much more willing" to "much less willing" so that the respondent express their opinion. The statements include:

  1. Respecting electoral results.
  2. Overusing executive orders if Congress is not cooperating.
  3. Prosecuting journalists who reveal misconduct without revealing sources.
  4. Not fulfilling judicial orders, if politician believes judge was partial.
  5. Not authorizing protests from opposition leaders in public spaces.

- We then join these statements as an index similarly to other outcomes.

**Secondary outcome:** Each of the individual components of the "willingness-to-vote" preference index.

### 4.4.3 Confidence in institutions (following World Value Surveys)

**Primary outcome:** None.

**Secondary outcome:** Individual's self-reported confidence in the following institutions: television, government, courts, and congress.

- 4-point scale from "Not at all" to "A great deal" [11].

- We may show results jointly in an index, as we did for other variables.

---

[11] Following the World Value Surveys

### 4.4.4 Other democratic outcomes

**Primary outcomes:** None.

**Secondary outcomes:** Reported support for military rule, the importance of democracy, and opinions on how democratically ruled the country is today. These outcomes are all based on World Value Survey questions.

1. Support for military rule is measured by asking participants whether they believe military rule is a good way to rule the country. It is measured on a 4-point scale from "Very bad" to "Very good."

2. The importance of democracy is measured on a 10-point scale, from "Not important at all" to "Absolutely important."

3. Subjective opinion about how democratically the country is ruled today is measured on a 10-point scale from "Not at all democratic" to "Completely democratic."

# 5 Figures and Tables

**Total unique ad views:**
**1,000,966**

**Screening Survey**
*(N = 5,394)*

**Eligible respondents:**
*(N = 2,055)*

**Baseline Survey**
*(N = 1,664)*

**Choice of News Outlet**

| Newspaper 1 *Carta Capital* *(N = 388)* | Newspaper 2 *Estadão* *(N = 438)* | Newspaper 3 *Gazeta do Povo* *(N = 155)* | Newspaper 4 *O Globo* *(N = 433)* | Newspaper 5 *Valor Econômico* *(N = 250)* |

**Control (C)**
*(N = 909)*

**Treatment (T)**
*(N = 78)*

**Treatment+Incentives (TI)**
*(N = 677)*

- Coupon for free 3-month subscription of chosen news outlet

- Coupon for free 3-month subscription of chosen news outlet

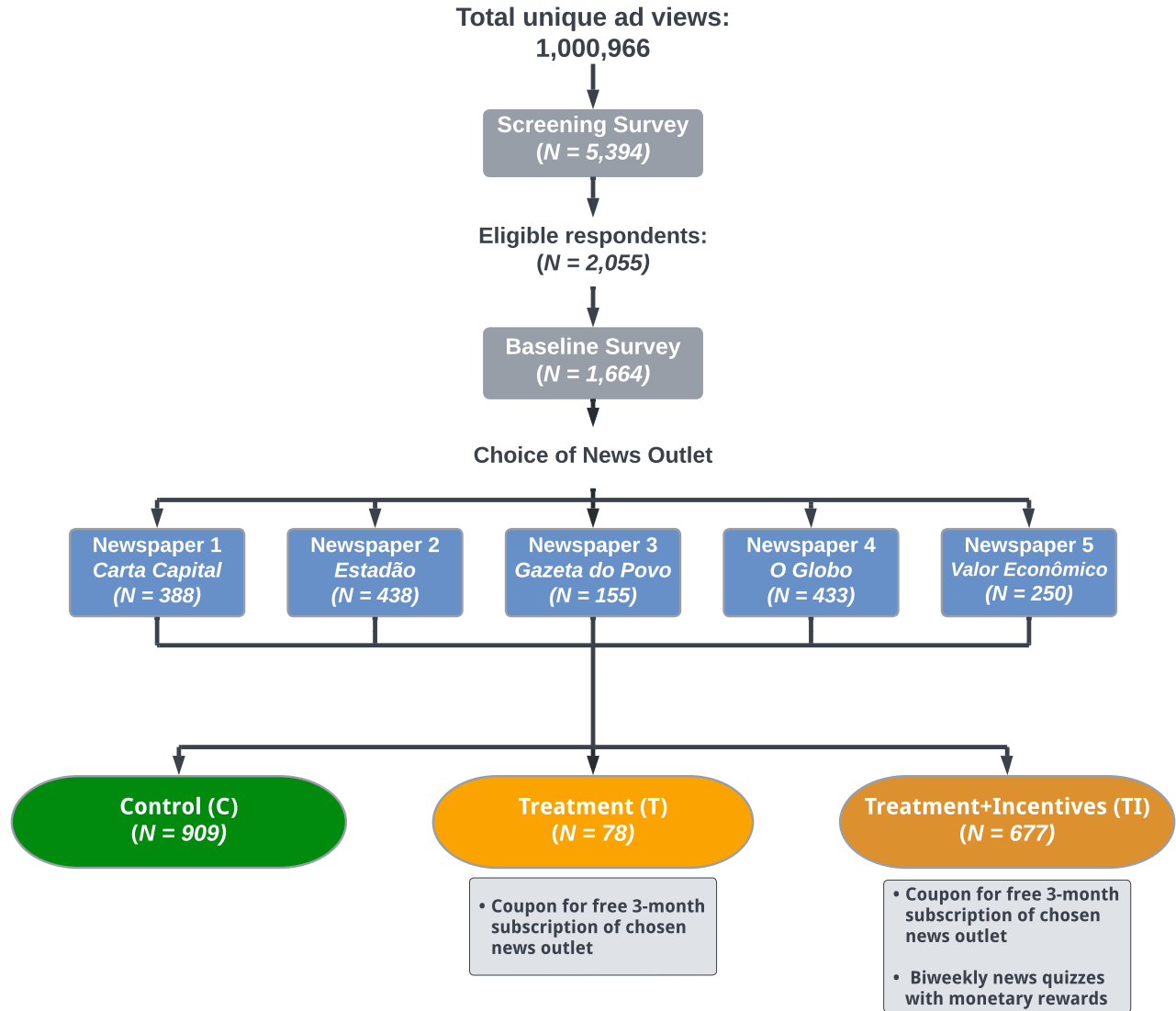- Biweekly news quizzes with monetary rewards

Figure 1: Experimental design and sample sizes

*Notes:* This figure presents the experimental design. Participants are recruited through online ads and then fill out a screening survey to determine eligibility. They are then sent a link to the Baseline Survey, where they choose their preferred news outlet. After that, they are randomized into three arms, Control (C), Treatment (T), and Treatment+Incentives (TI). These arms are stratified by newspaper. A brief description of the treatment procedures is included in the gray boxes below the treatments.

Figure 2: Advertising to encourage participation in the study

*Notes:* This figure presents an example of an ad used in the recruitment phase of the experiment. On the main text, the ad reads: "Take part in the Academic News Study, conducted by researchers at the Economics Department of Stanford University and *Fundação Getúlio Vargas*". The image's main text reads: "Academic News Study". On the gray box, the text is "Take part!". Finally, at the bottom, we write "Conducted by researchers from the universities:" along with the logos of the Department of Economics at Stanford and the Sao Paulo School of Economics, a school associated with *Fundação Getúlio Vargas*.
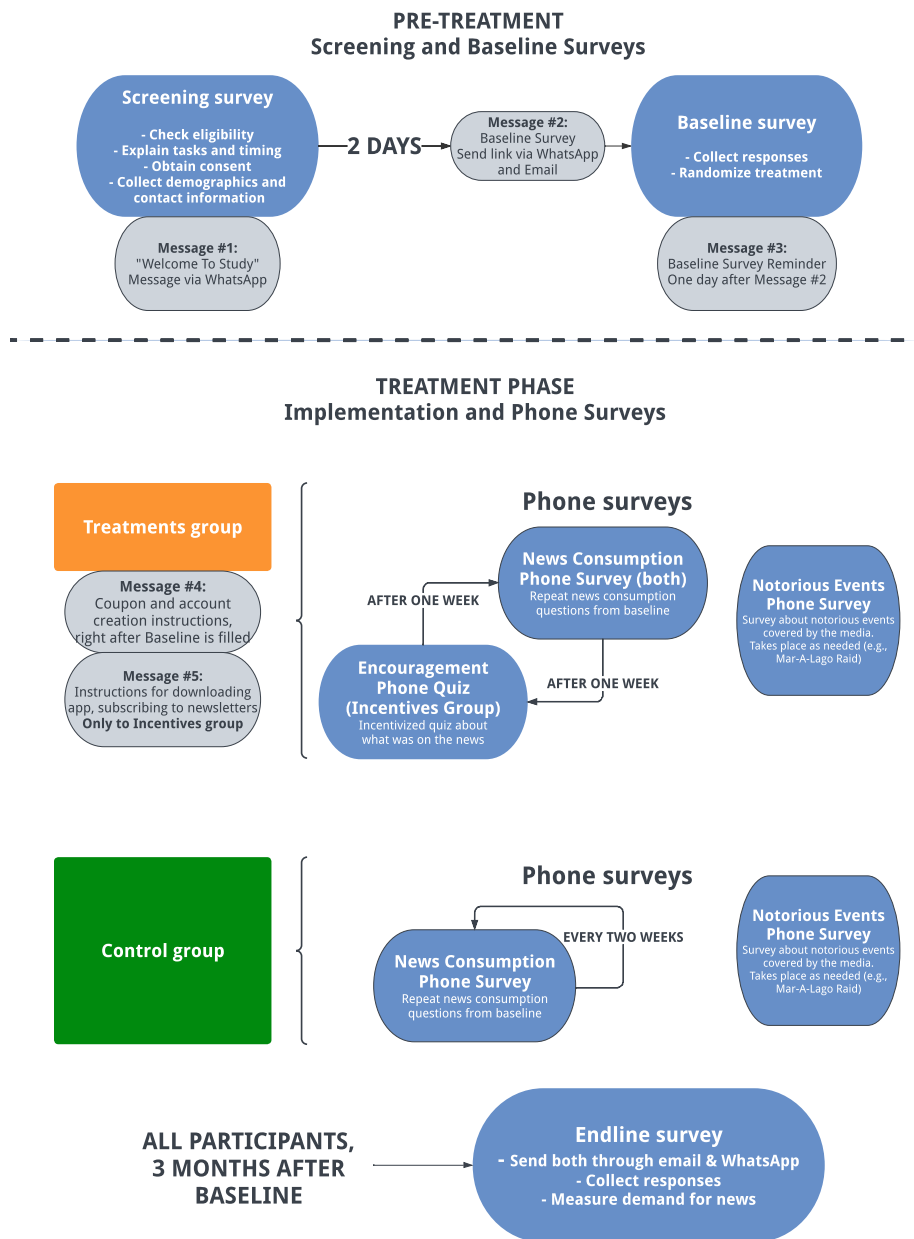
Figure 3: Flowchart of the period, from screening survey to endline survey.

*Notes:* This figure presents the flowchart of the experiment.

- During the **Pre-Treatment** phase, participants will:

  1. Fill out the screening survey, where we explain tasks, collect consent, demographic data, and contact information.

  2. Receive messages with links to the Baseline survey two days after filling out the Screening survey and a reminder.

  3. Be randomized into Control ($C$), Treatment ($T$), or Treatment+Incentives ($TI$) groups.

- Treatment $T$ and Treatment+Incentives $TI$ groups will receive a coupon to redeem a free three-month subscription in one of the partner newspapers, along with instructions on account creation and coupon redemption for their three-month subscriptions.

- During the **Treatment** phase, participants will:

  1. $TI$ group only: receive instructions for downloading the app, subscribing to newsletters, and answering biweekly encouragement phone quizzes covering news reports. In addition, each participant answers a survey tailored to the outlets they chose to receive the subscription.

  2. All groups: answer biweekly phone surveys about news consumption, including minutes spent and mediums of consumption and answer notorious events phone surveys about ongoing relevant events in the news.

# References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Barber, R. F. and Candès, E. J. (2015). Controlling the false discovery rate via knockoffs. *The Annals of Statistics*, 43(5):2055 − 2085.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3):491–507.

Chen, Y. and Yang, D. Y. (2019). The impact of media censorship: 1984 Or brave new world? *American Economic Review*, 109(6):2294–2332.

Fithian, W. and Lei, L. (2020). Conditional calibration for false discovery rate control under dependence.

Graham, M. H. and Svolik, M. W. (2020). Democracy in america? partisanship, polarization, and the robustness of support for democracy in the united states. *American Political Science Review*, 114(2):392–409.