Department of Economics Discussion Paper Series

Visual Bias

Giulia Caprini

Number 1016
May, 2023

Department of Economics
Manor Road Building, Manor Road
Oxford, OX1 3UQ

ISSN 1471-0498

Visual Bias*

Giulia Caprini†

May 4, 2023

Abstract

I study the non-verbal language of leading pictures in online news and its influence on readers' opinions. I develop a visual vocabulary and use a dictionary approach to analyze around 300,000 photos published in US news in 2020. I document that the visual language of US media is politically partisan and significantly polarised. I then demonstrate experimentally that the news' partisan visual language is not merely distinctive of outlets' ideological positions, but also promotes them among readers. In a survey experiment, identical articles with images of opposing partisanships induce different opinions, tilted towards the pictures' ideological poles. Moreover, as readers react more to images aligned with their viewpoint, the news' *visual bias* causes issue polarization to increase. Finally, I find that media can effectively slant readers using neutral texts and partisan pictures: this result calls for the inclusion of image scrutiny in news assessments and fact checking, today largely text-based.

Keywords: Media bias, polarization, non-verbal language, news photography

L82 D91 D83

*I am especially grateful to Andrea Mattozzi and Riccardo Salerno for their continued support throughout this project. This work also greatly benefited from comments by Zeinab Aboutalebi, Giacomo Calzolari, Flavia Cavallini, Thomas Crossley, David Yanagizawa-Drott, Ruben Durante, Andrea Ichino, Mustafa Kaba, David K. Levine, David Martinez de Lafuente, Ro'ee Levy, Patrizia Mealli, Massimo Morelli, Salvatore Nunnari, Nancy Qian, David Stromberg, Junze Sun, Ekaterina Zhuravskaya, and seminar participants at Alghero (Ibeo), CEPR workshop on Media (EIEF/Tor Vergata), Bocconi (LEAP), Collegio Carlo Alberto, EEA, EUI, Oxford University, Yale, University of Milan, Italian Economists Society, and the Royal Economics Society. I gratefully acknowledge financial support from the 2019 EUI Research Council grant, and from the 2020 Early Stage Researchers grant at the European University Institute, where this work was largely developed. The experiment was approved by the IRB ("Ethics Committee") of the European University institute, and pre-registered in June 2021 in the AEA RCT Registry at: https://doi.org/10.1257/rct.7904-2.0

†Oxford University and Nuffield College. Email: Giulia.caprini@economics.ox.ac.uk

I INTRODUCTION

"*We don't see things as they are, we see things as we are*"

–Anais Nin (writer), 1961

Four facts characterize today's access to written news: first, people increasingly read their news online, finding news pieces through social media platforms or news apps (Shearer, 2021). Second, news readers first encounter these news pieces through short previews, a format consisting of a headline, a short summary text, and a leading image (Figure I): this format confers leading pictures a prominent position over other news elements. Third, people heavily rely on the content of news previews, for instance when they share the news pieces in their social media feeds without reading the full text (Gabielkov et al., 2016). Fourth, most initiatives tackling online misinformation and news quality assessments are concerned with the analysis of written contents, and pictures largely escape systematic scrutiny. Taken together, these dynamics suggest that leading images already gained – and can keep gaining – strategic importance in the communication strategy of news media, and in particular of ideologically-partisan outlets: not only the unspoken content can reach a broad audience, but the ambiguity of intended meaning in pictures allows to provide controversial hints and cues while limiting the potential backfire. This power will be further amplified with the introduction of generative AI tools that lower the cost and skills required to fabricate a photo-realistic illustration.

This study explores the role of leading images in the communication strategy of politically-slanted news producers, investigating the extent of their influence over and above text. The paper documents two complementary broad instances: on the one hand, news providers with different political leanings employ systematically different visual language; on the other hand, partisan visual narratives are effective in slanting readers' opinion towards the news outlets' ideological poles. Jointly, these dynamics indicate that the news' *visual bias* is a tangible expression of media bias (De Vreese, 2004; Groseclose and Milyo, 2005; McCombs and Reynolds, 2009; DellaVigna and Gentzkow, 2010; Prat and Strömberg,2013; Strömberg, 2015; Prat,2018).

The analysis is organized in two Sections. I first study the partisanship of the news' visual language, namely the extent to which the characteristics of the leading images are distinctive of their news outlets' political leaning. I collect about 300'000 leading images from news published between December 2019 and December 2020 by the main US news outlets, and I exploit computer vision tools to extract information on the images' content (such as the subjects and objects depicted, their characteristics, and contextual aspects of the image). Drawing from existing studies in photography, linguistics, semiotics, psychology, and political science, I combine this information into key measures to decode meaning from visual contents. I thus construct a "visual vocabulary" of interpretable tokens, which pertain several dimensions relevant to convey political cues through graphic elements. Borrowing from text-analysis methods, I map the images in my dataset to the vector of tokens in my visual vocabulary, using this representation to analyse systematically the portrayal of subjects and compare it across pictures. To measure the partisanship of the news' visual language I employ the leave-out estimator of phrase partisanship developed by Gentzkow, Shapiro, and Taddy (2019). The results of this initial analysis demonstrate that the lead images chosen by liberal and conservative news outlets are systematically different. Both lexically, in terms of the visual words adopted to characterize a subject, and syntactically, in terms of the subjects, characterizations, and contexts depicted. These differences contribute to making partisan narratives hard to untangle for readers. Overall, the news' visual polarization appears close to the verbal polarization that Gentzkow, Shapiro, and Taddy (ibid.) document in Congress sessions of recent years.

The second Section of the paper presents a survey experiment conducted on a nationally-representative sample of 2'000 US residents to examine the effect of visual partisanship on news readers' opinion. I test two hypotheses: whether partisan leading images distinctive of Republican/Democrat outlets effectively slant the audience towards their respective party, and whether partisan images increase the polarization of public opinion. The results indicate that individuals exposed to identical news previews but leading pictures with opposite partisan loadings formulate significantly different opinions, with the slant following that of the news outlet. Moreover, I find that the news' visual bias causes a significant increase of issue polarization in the general public: the slanting effect of images interacts with readers' prior, and audiences on both sides of the political spectrum react more distinctly to pictures aligned with their viewpoint. This pattern implies that the polarizing effect of visual bias is further exacerbated if readers' source their news exclusively from like-minded outlets. Finally, and equally importantly, the experiment demonstrates that by creating news pieces with politically neutral texts led by visually partisan images, newscasts can bypass text-based fact checking and still slant readers' opinion effectively. This result calls for an inclusion of image scrutiny in the quality assessments of news.

This study seeks to contribute to the understanding of media language in three ways. From a methodological viewpoint, I design a visual vocabulary for the systematic interpretation of pictures, adapting to the study of images a NLP framework common in text analysis. Several studies explore the graphic tools and elements relevant to political visual framing (among recent ones, see Peng,2018; Boxell, 2021; Haim and Jungblut, 2021; Ash et al. 2021). Like in these works, my approach enables the study of a wide range of image characteristics, examining the relative incidence of distinct *visual words* across sources with opposite political stances. Additionally, the method here adopted allows to investigate both the lexicon and the syntax structure of visual language. By including syntactically-coherent combinations of visual words, the visual vocabulary models the interdependency of distinct pictorial elements, allowing the extraction of deeper-level symbolisms in the images. Symbolic semantics makes use of logical relationships between visual words, which in this paper are encoded following linguistic intuitions inherited from verbal language (syntactic classes and relationships).

Second, this study relates to numerous analyses performing automatic detection of bias and language polarization (see, e.g. Greene and Resnik, 2009; Yano, Resnik, and Smith, 2010; Recasens, Danescu-Niculescu-Mizil, and Jurafsky, 2013; Gentzkow, Shapiro, and Taddy, 2019). I draw in particular from Demszky et al. (2019), who use the measure of phrase partisanship originally developed by Gentzkow, Shapiro, and Taddy (2019) to study the political polarization in the text of tweets related to mass shootings events in the US. I employ their partisanship estimator to document for the first time the systematic visual partisanship of US news.

Third, this paper presents novel evidence on the impact of leading images on news readers' opinion. I show that US media convey their political bias through news leading pictures, and I document a significant causal effect of *visual bias* on the polarization of public opinion. In this respect, this paper relates to several works that identified the correlation between increasing polarization of media and the general population's political stance, underscoring the imperative to accurately detect news bias and to understand its nature (e.g. Gentzkow and Shapiro, 2010,2011; Prior, 2013). A large literature documents that recently –and in particular during the first months of the Covid-19 pandemic– partisan divisions significantly shaped health behavior, support to specific policies, attributions of responsibility, and general beliefs (e.g., Allcott et al., 2020a, 2020b; Druckman et al., 2020; Gadarian, Goodman, and Pepinsky, 2021); Gollwitzer et al., 2020; Romer and Jamieson, 2020). There is additional evidence suggesting that issue polarization is rising and documenting its possible causes (see e.g. Doherty, Kiley, and Asheer,

FIGURE (I)
News Preview on Social Media

Notes: The Figure shows the format of a news preview on Twitter. Its key elements are: the name of the news source (A), the news' leading text (B), the news' leading image (C), and the news' header (D). In this format, lead images occupy the largest area share. Photo by Brooks Kraft for Getty Images ("The two-story Board Room in the Eccles Building, Washington, DC"). Image registered and available at `shorturl.at/hnuAC`.

2019, Levy, 2021), to which the present paper adds by demonstrating the causal effect of news visual bias in this direction.

The remainder of the paper is organized as follows: I first explore the non-verbal language of US news, and I quantify the visual partisanship (Section II). Then, I test the effect of partisan images on general public opinion (Section III). I close with a summary of the findings and a general discussion of their implications (Section IV).

II VISUAL PARTISANSHIP IN US NEWS

This first Section documents the extent of visual partisanship in US news between December 2019 and December 2020. I find a high degree of polarization across the visual narratives adopted by news sources across the political spectrum. To estimate visual partisanship I begin by applying a dictionary method, which entails creating a visual vocabulary and expressing images as vectors of dictionary entries; I then use a partisanship estimator borrowed from text-analysis to measure language distance.

II.A Method: A dictionary-based approach to the study of pictures

To perform a comprehensive analysis of the leading pictures collected from US news I adapt a dictionary-based methodology originally developed to study texts. This approach transforms the pictures in a convenient format, allowing me to then exploit the existing measures of language distance from text analysis (the partisanship estimator illustrated in Subsection *II.C*).

Dictionary methods entail counting words from a predefined lexicon (the dictionary) in a big corpus, with the intent to explore or test hypotheses about the corpus itself. The essence of the method consists of transforming a document in a vector of counts or indicators for the presence of given language elements. The reference vocabularies are generally composed of *unigrams*, *bigrams*, and/or *trigrams*, namely series of one or two/three consecutive words (or word roots) that, once combined (and before the removal of stopwords and word suffixes/prefixes) compose the phrases of a text; these elements are commonly referred to as *tokens*.

I adapt this procedure to study the news' visual language and to extract computationally the meaning of the large number of leading images in my dataset. By processing the raw information described in previous section, I draft a vocabulary of graphic and content-related *features* which, once combined, result in the pictures' backbone. Following the parallel with text analysis, these can be considered as my set of "visual tokens".

Describing dictionary methods for text analysis, Gentzkow, Kelly, and Taddy (2019) illustrate the simplifications that help reduce raw text to a representation suitable for statistical analysis:

> "*We typically make three kinds of simplifications: dividing the text into individual documents, reducing the number of language elements we consider, and limiting the extent to which we encode dependence among elements within documents. The result is a mapping from raw text Δ to a numerical array C. A row c_i of C is a numerical vector with each element indicating the presence or count of a particular language token in document i.*"

I reduce the pictures in my set to simpler representations through three steps. The first entails dividing the corpus into single documents; in my application, since the attributes of interest are defined at the single image level, I consider each picture as an individual document. The second step entails adapting the number of language elements that are considered. The purpose of my analysis is to study how the visual narrative differs among sources with different political leanings. To this extent, I consider both general graphic elements and politically-relevant cues in the pictures, as described in details in the next Section.

The third step entails encoding the dependence among elements within a document. In text analysis, this is aided by including consecutive words/stems (bigrams and trigrams) in the vocabulary. The study of consecutive words helps the extraction of meaning from a text because words' contiguity proxies pertinence to the same textual object. In images, instead, I model the pertinence of multiple characteristics to the same portrayed object using combinations of co-occurring features of each depicted element (this "features engineering" is described in Subsection *II.B*.3). The unigrams, bigrams and trigrams in my visual vocabulary are thus represented by single graphic features, feature-pairs, and features-triplets. To create meaningful combinations I exploit the features' semantic categories, organizing them in a syntax as in verbal language. To this extent, I extract image features passing the pictures to the computer vision algorithms from the Azure suite by Microsoft; for each image, these algorithms provide me with an array of *words* describing each detected elements, as well as the pixel-coordinates of the "bounding boxes" the element falls in. Hence, this type of features extraction not only detects the features' presence, but also directly provides annotated meanings.[1] This *annotated output* is crucial to later produce syntactically-meaningful combinations of features, and it also implies that all entries of the vocabulary are interpretable by design.

II.B Creating a Visual Vocabulary

This Subsection describes the three phases of the vocabulary creation: the collection of pictures (*II.B*.1), the features extraction (*II.B*.2), and the features engineering (*II.B*.3 and *II.B*.4).

II.B.1 Retrieving Pictures

News sources. I begin by constructing a comprehensive list of the relevant news outlets from a list of the top 50 US news media by digital circulation from Similarweb.com. The circulation metric is based on the number of Unique Visitors per Month (UVM), and it indicates how many people in the U.S. market visit a website in a month.[2] I discard sources that do not cover political news and are exclusively focused on entertainment, celebrity news, fashion, beauty

[1] In this way, the cognition (and re-cognition") phase of image reading coincides with features extraction. Conversely, other methods for mapping pictures to vectors of features express visual tokens as patterns of pixels, which may or may not mark an identifiable object and whose meaning is not annotated (as, for instance, the "Bag of Visual Words" method).

[2] UVM data by SimilarWeb.com accessed on October 27, 2020. See `https://www.similarweb.com`.

news, or local news.[3] I derive the sources' party affiliation through the political bias ratings from Adfontesmedia.com and Allsides.com, keeping the sources with concordant partisanship attribution.[4] The final sample consists of 22 sources, evenly divided on the two sides of the political spectrum in terms of news pieces produced. Appendix Section A.1.1 lists the sources as well as their partisanship scores documented by the sources described.

News data. From the Twitter accounts of the selected sources, I obtain all the news articles shared on the social media between December 1, 2019 and Dec 13, 2020. I focus exclusively on tweets sharing written news pieces (discarding links to video, voice recordings etc.), and I filter out all news pieces written by an outlet but tweeted by a different source. As sources commonly share their pieces multiple times to maximize audience, I keep only the latest version of each piece. The resulting dataset counts 284'294 unique valid news pieces.

From the articles' metadata I retrieve and store the headline, description, publication outlet, publication date, and leading image. An article's leading image is the main picture accompanying a news piece, the one displayed in the preview when news are shared on social media.[5]

*II.B.*2 Features Extraction

As mentioned in Subsection *II.A*, the visual vocabulary in this paper builds on image features expressed in terms of their annotated meanings. This grants that features can later be combined to decode contents of higher semantic level, and that all vocabulary entries are interpretable by design. I pass each picture to image analysis algorithms from the computer vision API by Microsoft (Azure cognitive services).[6] The following paragraphs illustrate the algorithms used and their output.

Image analysis, Face detection, Face verification, and emotion recognition. The image analysis algorithm detects the presence of faces and assigns tags to the picture based on the depiction of "iconic", recognizable items (e.g. clothing pieces, natural elements, animals, etc.). I pass images that contain at least one human face to the face detection, description (age, gender, hair colour, eye-nose-mouth landmarks etc) and emotion recognition algorithms. The latter classifies the emotions expressed by a face into happiness, sadness, anger, fear, contempt,

[3] The labels correspond to the tags in the descriptions by Similarweb.com; discarded outlets are mainly small local outlets, with the notable exception of the "Los Angeles Times", the "Chicago Tribune" and the "Arizona Republic".

[4] `https://www.adfontesmedia.com`, `https://www.allsides.com`

[5] See Figure I for an illustration of the news previews on social media.

[6] For a list, see `https://learn.microsoft.com/en-us/azure/cognitive-services/computer-vision`

disgust and surprise. Using the subset of images containing a human face, I check whether the depicted persons are members of the US congress or prominent figures of the US recent public scene. To this purpose, I first train the face-verification algorithm on a comprehensive set of images created by manually selecting 9 pictures of each congressmen and congresswomen sitting in the 114, 115, 116, and 117th US congresses.[7] Then, to record the presence of prominent public figures outside the setting of Congress (e.g. Governors, Supreme Court judges, athletes, actors etc.) I pass the pictures to Microsoft's "celebrities" API, a face-verification algorithm pre-trained to recognize a wide set of celebrities. In addition, whenever the picture contains a congressperson, I record her relative political leaning as measured through the first dimension of the Common Space DW-NOMINATE score from McCarty, Poole, and Rosenthal (2015).[8]For each of the above-mentioned extracted elements, a confidence score is returned along with the bounding box coordinates of each detected element; this allows to determine the element's position within an image.

Finally, the Image Analysis algorithm returns general information on each image. For instance, it identifies and categorizes the pictures using a category taxonomy with parent/child hierarchies (e.g. "indoor_marketstore"); it describes the "type" of image, such as whether it is a drawing or clip art, whether an image is black and white or color and, for color images, dominant and accent colors. The algorithm also produces a list of image tags from a set of thousands of recognizable objects, living things, scenery, and actions. These tags are indicators for noteworthy contextual elements of a picture, such as natural elements (e.g. fire, water, etc.), transportation means (e.g. cars, ambulances, etc.), architectural elements (e.g. skyscrapers, castles, etc.), or text content (e.g. banners, signals, etc.).[9] Importantly, in the case of general image information, the API does not return bounding box coordinates. Hence, these elements can be used as general image descriptors but do not possess information on location (this implies that they can hardly be "attributed" to other elements to characterise their depiction).

Overall, the final information set extracted from each image through the computer vision suite includes the following: detection of people and recognition of politicians and celebrities;

[7]When a person's portraits did not cover a wide range of angles, I added a 10th picture to her set. Portraits were chosen so to include different camera angles for each person.

[8]I attribute Donald Trump (who didn't seat in congress before the presidency) the same DW-NOMINATE score as the most partisan Republican congressperson (Tommy Tuberville, with score 0.916). I attribute Joe Biden the same partisanship he had as congressman in 2008 (-0.314).

[9]For a complete list of the tags classes available, see `https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/category-taxonomy`.

details on people position (coordinates), head poses (pitch, yaw, roll), facial expressions, position of landmarks (nose, eyes, mouth, etc); detection and recognition of objects, details on colors; detection of places and background elements through tags; details on the image category, type, color scheme, tags.

In the rest of the paper, I refer to this set as the "raw information" on leading pictures given by the algorithms.

II.B.3 Features Engineering: Single Features

This subsection describes how the "raw information" is then processed to obtain a vocabulary of valid tokens to decode visual language. I do so in two steps: first, I derive meaningful individual features (for instance, using the coordinates of a face to derive its size) and I organize them in syntactic classes; second, I combine individual features in couples and triplets. Jointly, single features and combinations compose the visual vocabulary.

Features Class "Subjects" (S):

"Subject features" are attributes capturing characteristics that are constant for a subject across all pictures in the sample (such as a given person's name or political party affiliation). This syntax class encompasses indicators for whether or not a person is well-know to the public (a "celebrity" status), whether the depicted person is a man or a woman, the subjects' names, and the subjects' relative position in the political partisanship distribution (measured through the first dimension of the Common Space DW-NOMINATE score from Poole and Rosenthal, 1985).[10] This partisanship score originally ranges from -1 to +1 and is scaled so that the lowest scores are those of liberal Democrats and the highest scores are those of conservative Republicans; dividing the index domain into 10 equally spaced bins, I produce four vocabulary indicators to mark pro-Democratic partisanship, four to mark the relative pro-Republican leaning of a congress member, and one indicator to mark the central bin. The unique identification of people within an image is essential to correctly attribute the adjectives features through features' combinations. For this reason, the "Subjects" class also includes within-picture unique identifiers for all the persons portrayed. Those are indicators for their saliency rank within a picture, obtained from the weighted average of their face area share (70%) and centrality in the picture (30%) (the higher the rank, the more salient the Subject).[11] Appendix Section A.1.2 provides a summary of all *Subject* subclasses.

[10] Data from `voteview.com`, by Lewis et al., 2021

[11] The maximum number of individuals I contemplate in the data is 10 persons in the same image.

Features Class "Adjectives" (A):

With the term "Adjective features", I refer to other features defined at face-level, like Subjects. However, Adjectives indicate variable attributes that subjects may possess in given pictures and not in others. I organize Adjective features by their pertinence to three dimensions: *Size*, *Centrality*, and *Kinesics*. The first two pertain the pictures' proxemics, i.e. the way the space is used in the portrayal, while the third concerns the dynamics and gestures portrayed.

-Size. Subjects' size is relevant to image analysis primarily because higher graphic dimension induces higher visibility. Humans do not receive a picture's content through a single glance, but rather via separate scans. Hence, the longer a person looks at a picture, the higher the chances marginal details will be "seen". Bigger objects are always more likely to be grasped by viewers. In this sense, we can interpret the relative size of depicted objects as informative of the illustrative intent behind the choice of a picture: if an element occupies a large portion of the image, the person who chose the illustration likely meant to highlight the given element to the viewers. Therefore, objects' size proxies a criterion of precedence among the objects portrayed in the picture. The visual vocabulary includes three individual features for subjects' size: a "close-up" indicator for faces whose area is equal or greater than 1/6 of the total image area, a "mid range" indicator, for faces from 1/6 to 1/24 of the total image area, and a "long shot" indicator for faces with size below 1/24 of the total image area.

-Centrality. I define an object's *centrality* in a picture as its ability to attract the viewer's attention. It is measured in terms of proximity to the two vertical and two horizontal parallel lines that divide a picture in three equal sections, vertically and horizontally, following the "rule of thirds" (Figure II). Such "attention lines" have been shown to attract and guide viewers' attention within a picture (see, e.g. Koliska and Oh, 2021) and are often marked in cameras' viewfinders to aid photographers' frame choice.[12] For every face in a picture, its centrality is inversely proportional to the distance between the eyes-midpoint and the closest of 5 focal points (either the center of the image, or one of the four intersections of the attention lines determined through the rule of thirds). Formally, it is expressed as:

$$c^{ROTC}(x,y) = argmax_i e^{-\left(\sqrt{(\frac{x-x_i}{W})^2+(\frac{y-y_i}{H})^2}\right)} \tag{1}$$

[12]This is only one of the ways to measure viewer's attention towards a subject. Other aspects that are relevant to this extent include color vidvidness, body posture etc.; however, they do not form part of my "raw information" set on the images, and thus nevertheless necessarily fall outside the scope of the present work.

(a) *Horizontal lines* (b) *Vertical lines* (c) *Rule of thirds grid*

FIGURE (II)

The rule of thirds in "Hand in hand" by Steve McCurry

Notes: The Figure illustrates an application of the "rule of thirds", which the photographer uses to guide the viewer's attention towards the elements of interest (placed along the two vertical and two horizontal lines that divide the image in equal thirds). Picture: *India*, "Hand in hand" gallery by Steve McCurry.

where i indicates the focal point, x and y are the coordinates of the eyes' midpoint, x_i and y_i are the coordinates of point i, and W and H express the total width and height of the image. The distances in (1) are measured in pixels, with the top-left angle of the images marking the (0,0) coordinate. The distance between focal point i and the eyes-midpoint is normalized with respect to the image dimensions to ensure cross-pictures comparability. The centrality index therefore ranges in 0-1, with higher values indicating higher proximity to a focal point.

The visual vocabulary includes four main individual features related to subjects' centrality: an indicator for high centrality ($C \geq 0.95$), medium-high centrality ($0.85 \leq C < 0.95$), medium-low centrality ($0.75 \leq C < 0.85$), and low centrality ($C < 0.75$).

-Kinesics. Kinesics is broadly defined as the study of body movements (Bowden, 2015; Furnham and Petrova, 2010; Walters, 2011). It entails body dynamics such as gestures, facial expressions, eye behavior, or touching. Some gestures, such as facial expressions or eye movements, have been recognized as important markers of the emotional and cognitive inner state of a person. The particular look on a person's face, for instance, provides reliable cues as to approval, disapproval, or disbelief (Bailenson et al., 2008; Grabe and Bucy, 2009; Lunenburg, 2010). Importantly, those elements are also relevant political cues in visual narratives: during the 2016 US election campaign, for instance, news websites of varying ideologies portrayed the two candidates displaying more positive and less negative emotions of the candidate they supported (Peng, 2018; Boxell, 2021). In line with these findings, I include in the visual vocabulary a feature for each of the seven emotions detected by the emotion recognition algorithm (happiness, sadness, anger, fear, contempt, disgust and surprise).[13] I additionally create a "negative emotion" indicator taking value 1 whenever the subject expresses either sadness, anger, fear, contempt or disgust. Correspondingly, I create a "positive emotion" indicator taking value 1 when the subject's face expresses happiness. The negative- and positive-emotion indicators exclude surprise: lead pictures often portray subjects during speeches or public appearances, and the algorithm often mistakenly associates the depicted persons' open mouths to surprise. If none of the emotion variables (including surprise) takes value 1, I code emotion as "neutral". While single-subject, close-up portraits focus the viewer's attention on the portrayed person's characteristics (and in this context, her emotion), shots with many subjects tend to shift the attention from an individual to a group. For this reason, the interaction of multiple persons shapes the overall emotion expressed in a picture. For the visual vocabulary to capture this aspect, and to deepen the understanding of images' emotional loading, I include a "triggered emotion" feature: for each portrayed person, it is the average emotion of the subjects whose glance is directed towards the person.[14] For each subject in a picture I also include an indicator for the number of other people who are looking towards her, as well as one indicating whether she is looking towards someone. The vocabulary also captures a number of other kinesics dimensions. It encompasses indicators for each of the individual head pose angles used to construct the triggered emotionality (namely positive, neutral, or negative yaw/pitch/roll), as such angles by themselves can affect a viewer's judgement of the social relations among the depicted subjects (Ekman, 2009); Gawronski and Payne, 2011). It captures two types of color-related factors that can be used to augment or decrease the salience of a subject in a picture: the light exposure of a subject's face (measured through indicators for "normal exposure", "bright exposure", and "dark exposure") and its blur level (measured through indicators for low blurring, medium blurring, and high blurring).[15] The vocabulary also contains indicators for non-verbal cues related to the mode of dress: those capture whether the portrayed subjects are wearing clothes or accessories with particular distinctive features, such as uniforms, formal

[13] I consider the emotion as correctly identified when the algorithm expresses a confidence level of 80% or higher (as in Peng, 2018).

[14] If persons X, Y and Z are portrayed in a picture, with X and Y looking towards Z, the latter's triggered emotion is the average of X and Y's expressions. I use the subjects' head pose angles as proxy for their glance direction and to compute their sight regions, borrowing this approach from studies on intelligent vehicle systems. Appendix Section A.2.1 describes the details of the method.

[15] Importantly, there exist a number of other color-related dimensions that are known to be relevant devices in visual communication (such as the overall color dominance in a picture, or the presence of evocative color patterns); unfortunately, those dimensions are outside the scope of the algorithms employed.

dresses, suits, ties, hats, face masks, scarves, or glasses. From a visual communication perspective, individual clothing pieces –e.g. a suit– can be used to prime specific evaluations – such as success and wealth (Ekman, 2009). In addition, the presence of specific clothing or accessories –such as face masks– can be *per se* politically and socially important signals.

Appendix Section A.1.2 provides a summary of all *Adjective* subclasses, with *AS* labelling features pertaining size, *AC* indicating centrality, and *AK* indicating kinesics.

Features Class "Context elements" (C):

The third and last Syntax class of features encompasses indicators for the presence of specific contextual elements, varying at the image level. Previous research on political candidates' imagery has shown the communicative relevance of contextual features –such as the portrayal of many individuals together– or of structural characteristics –such as the camera angle– (see, e.g., Sutherland et al., 2013; Abele et al., 2016; Haim and Jungblut 2021). In light of the existing evidence, the visual vocabulary includes indicators for the following elements varying at image level and defining its context: the presence of any human face; the total number of faces; the presence of any well-know person and the total number of recognized individuals; the joint presence of any two individuals from a "common subjects" lists encompassing all well-known individuals who appear in at least 50 images; the presence and total number of members of the Republican (Democratic) party sitting in at least one of the 114th-117th Congresses; the presence and number of politicians in each range of the politicians' partisanship distribution; the presence and total number of men or women in images featuring at least one human face; the relative presence of the two sexes (majority women/men/equal number); the average emotionality –neutral, positive or negative– for pictures featuring at least one human face, measured through the average emotion expressed among all faces; the number of faces expressing each of the seven emotions, in pictures with at list one recognized individual; the presence, total number, and relative frequency (majority/minority/equal number) of individuals wearing a facial mask; the picture camera angle, derived from the head poses of portrayed people.

Other individual features pertaining context are conveyed (without transformations) by image tags. As mentioned, tags are indicators for the general presence (that is, without coordinates details) of particular elements or objects within the image. These mark the presence of natural elements (e.g. fire, water, etc.), transportation means (e.g. cars, ambulances, etc.), architectural elements (e.g. skyscrapers, castles, etc.), or text content (e.g. banners, signals, etc.).[16]

[16] An example of image tagging is available at `https://rb.gy/421`.

Because tags information doesn't include location details (i.e., there are no bounding boxes, as mentioned), from a list of all possible tags in my dataset I manually classify them into subclasses, to expand the list of tags through meaningful tag mixes. Appendix Section A.1.2 provides a summary of *Context* tokens, with *CNtagmix* indicating features derived from tags, and *CNtxt* indicating other general context features.

*II.B.*4 Features Engineering: Combinations

As mentioned, text-analysis vocabularies include bigrams and trigrams (couples and triplets of consecutive words). This allows to decode the pertinence of two language elements to the same textual object through words proximity. As noted, visual language doesn't have a clear order of words; for this reason, I rely on features-combinations to model the pertinence of multiple characteristics to a portrayed object, exploiting overlapping locations or simple co-occurrence. The bigrams and trigrams in my visual vocabulary are represented by features pairs and triplets. To combine features in a meaningful way I exploit their syntactic roles, distinguishing among represented subjects, characteristics of their depiction, and characteristics of the context and interactions between subjects. This structure is intended to help decode the meaning of images the same way words syntax helps to analyse a text: individual features convey the lexical composition of an image, while features' combinations inform about their "visual syntax". In the remainder of this work, I refer to vocabulary entries as *tokens*, a term that indicates either an individual feature or a combination of features.

The structure of syntax classes combinations is summarized in Table I. Table II presents summary statistics for the three syntax classes within the visual vocabulary, distinguishing between single-feature tokens (upper panel) and feature-combinations tokens (lower panel).[17] The vocabulary encompasses a total of about 50k different subjects, each referring to either individual persons or couples (row 1), 40 different individual adjectives (row 2), and 0.5M individual elements of context (row 3), for a total of about 0.55M distinct tokens derived from single features (column 3 of upper panel). This adds to more than 3.8M distinct combinations, of which about 2.6M involve a Subject (column 1, row 4), 2M involve an Adjective (row 5), and 1.7M involve a Context element (row 6). Column 6 of Table II summarizes the occurrences of vocabulary tokens in the images analyzed images in this study. Overall, the pictures display single individual Subjects 0.6M times, individual Adjectives 1.3M, and individual context elements 3.2M

[17]These statistics only encompass visual words contained in analysed images. For instance, a token for the presence of Johnny Cash is only included if the singer is present at least once across the images analysed.

times. In combinations, Subjects score an additional 26.6M appearances, Adjectives 24.6M, and Context elements 8.8M. Appendix Tables A.1.2, A.1.3, and A.1.4 provide additional summary statistics for each of the subclasses within the Subjects, Adjective, and Context groups, respectively.

TABLE (I)

Features Engineering: Summary of combinations.

		-	SR	SN	SC	SP	SG	SRxSC
S	SR	✓						
	SN	✓	✓					
	SC	✓	✓					
	SP	✓	✓		✓			✓
	SG	✓	✓		✓			✓
A	AK	✓	✓	✓	✓	✓	✓	
	AC	✓	✓	✓	✓	✓	✓	
	AS	✓	✓	✓	✓	✓	✓	
	ASxAC	✓	✓	✓	✓	✓	✓	
	AKxACxAS	✓	✓	✓	✓	✓	✓	
	AKxAS	✓	✓	✓	✓	✓	✓	
	AKxAC	✓	✓	✓	✓	✓	✓	
C	CNtagmix	✓	✓	✓				
	CNtxt	✓	✓	✓				
	CNtxt×CNtagmix	✓						

Notes: The Table shows the features combinations included in the vocabulary, marked by "✓". The syntax classes are: **Group S**= Subject features: either defining subjects' constant characteristics across images (*SN*= person name, *SP* = political partisanship, *SG*= sex, *SC* celebrity status) or varying ones (*SR* = Saliency Rank). **Group A**= Adjective features: those related to Kinesics (AK), Size (AS), and Centrality (AC). **Group C**= Context features: i.e. general context features (*CNtxt*), and those derived from tags and their mix (*CNtagmix*).

TABLE (II)

Vocabulary summary statistics by syntax class

Syntax classes:	**N distinct tokens in syntax class**	**Share of total distinct tokens**	**Total distinct tokens across classes**	**Total class presence in all pictures**
Single features:				
Subject ("S")	49'951	.0918	543'840	633'159
Adjective ("A")	40	.0000735	543'840	1'261'862
Context ("C")	493'849	.908	543'840	3'221'858
Combinations of features:				
Subject ("S")	2'582'232	.677	3'816'127	26'611'909
Adjective ("A")	2'031'314	.532	3'816'127	24'677'817
Context ("C")	1'717'958	.45	3'816'127	8'844'042
of which:				
Subject×Adjectives	2'031'314			1'271'981
Subject×Context	484'063			311'031
Subject×Subject	66'855			41'166
Context×Context	1'233'895			586'414

Notes: The Table presents summary statistics for the three syntax classes within the Visual Vocabulary, separately for tokens that consist of single features (upper panel) and for feature-combinations (lower panel). Column 3 presents the number of distinct tokens for each class, i.e., the distinct values a class can assume. Column 4 indicates the proportion of tokens in the class relative to the total number of (single or combined) vocabulary features, reported in Column 5. Column 6 lists the total occurrences of tokens in the class across images. Note: all statistics encompass only tokens that appear at least once in the images analyzed.

II.C Measuring Visual Partisanship in US news

*II.C.*1 Pre-processing

I restrict my attention to features used at least 10 times in at least one of the 2-week periods, used in at least 10 different periods, and used at least 50 times across all periods.[18] Similarly, I remove features that appear too frequently because their use is likely not informative about the inter-party differences I wish to measure, while I remove infrequently used features to economize on computation. The resulting vocabulary contains about 4.3M unique visual tokens used 39.96M times in 280'310 leading images.

I analyze data at the image level and within time periods of two weeks, for a total of 26 periods between Dec 2019 and December 2020.

*II.C.*2 Estimating Visual Partisanship

I study the *visual* partisanship in leading images by adapting the leave-out estimator of phrase partisanship introduced in Gentzkow, Shapiro, and Taddy (2019). Like these authors, I define partisanship as the expected posterior probability that an observer with a neutral prior would correctly guess a picture's political leaning (i.e. whether it was published by a Republican-leaning or Democrat-leaning source) after observing a single token randomly drawn from the image. If the token is used equally in images by Republican- and Democrat-leaning news sources, this probability is .5 and the token is uninformative of the image's political leaning.

This leave-out estimator solves the problem of finite-sample bias, which arises because the features an image could contain are many more than those present in any image leading the news. As a consequence, many pictures' features are used mostly by one party or the other purely by chance; however, naive estimators interpret such differences as evidence of partisanship, leading to a bias estimate that is much larger than the true signal in the data.

The leave-out estimate of partisanship π^{LO} between images from Democrat-leaning sources, $i \in D$, and images from Republican leaning sources, $i \in R$, is

$$\pi^{LO} = \frac{1}{2}\left(\frac{1}{|D|} \sum_{i \in D} \hat{\mathbf{q}}_i \hat{\boldsymbol{\rho}}_{-i} + \frac{1}{|R|} \sum_{i \in R} \hat{\mathbf{q}}_i (1 - \hat{\boldsymbol{\rho}}_{-i}) \right) \tag{2}$$

where $\hat{\mathbf{q}}_i = \mathbf{c}_i / m_i$ is the vector of empirical token frequency for image i, with $\mathbf{c}_i$ being

[18]Following the text analysis by Demszky et al. (2019), whose programming code I used as basis for the analysis in this subsection. See code available at `https://github.com/ddemszky/framing-twitter`.

the vector of token counts for image i and m_i being the sum of token counts for image i; $\hat{\boldsymbol{\rho}}_{-i} = \left(\hat{\mathbf{q}}^{D\setminus i} \oslash \left(\hat{\mathbf{q}}^{D\setminus i} + \hat{\mathbf{q}}^{R\setminus i}\right)\right)$ is a vector of posterior probabilities, excluding image i and any token that is not present in least two images. Here, $\oslash$ denotes element-wise division and $\hat{\mathbf{q}}^{G} = \sum_{i \in G} \mathbf{c}_i / \sum_{i \in G} m_i$ denotes the empirical token frequency of images in group G. This LO estimator captures two components of image partisanship: the difference between groups (posterior probability for each feature) and the similarity within a group (dot-product between the feature vector of each speaker and that of their group).

*II.C.*3 Overall polarization

Figure III shows that in the entire period between December 2019 and Dec 2020 the visual language of leading images was highly polarized, with estimates ranging between .518 and .535, and a mean level around .525. Following Demszky et al. (2019), I quantify the noise by calculating the leave-out estimates after randomly assigning images to parties: the values resulting from random assignment are close to .5, suggesting that the actual values capture a true signal in the data.

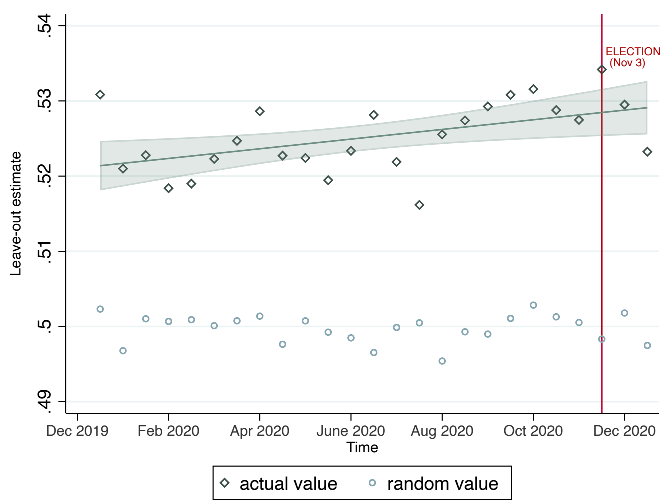


FIGURE (III)
Visual polarization

Notes: The Figure shows the partisanship of leading images estimated through the leave-out estimator (Gentzkow, Shapiro, and Taddy, 2019). The shaded region represents the 95% confidence interval of the linear regression fit to the actual values. I quantify noise by calculating the leave-out estimates after randomly assigning images to parties (Demszky et al. (2019): the values resulting from random assignment are all close to .5, suggesting that the actual values are not a result of noise.

As term of comparison for the magnitude of the estimated visual polarization, Gentzkow, Shapiro, and Taddy (2019) estimated the verbal polarization of US congress members from

1800s to present time, estimating verbal polarization at .52 around the year 2000. In another study, Demszky et al. (2019) use the same polarization measure to estimate the verbal polarization of Twitter discussions on mass-shootings in the US, finding a verbal polarization range between .517 and .547, with a mean of about .53.

*II.C.*4 Polarization by topic

In this subsection, I explore the extent of visual polarization dividing the images by the topic of the news pieces they lead. I model the news' topics by analysing the text of the tweets describing (and linking to) the articles.[19] For this I use BERTopic, a topic modelling approach that operates through sentence-trasformers to create embeddings, and exploits a class-based tf-idf for clustering.[20] The algorithm creates the tweets embeddings using a pre-trained BERT-based model for tasks of semantic similarity in English ("Paraphrase-MiniLM-L6-v2"). It lowers the dimensionality of the tweets embeddings with UMAP, to then cluster the reduced embeddings in semantically similar groups to define topics.[21] To get a sense of the words composing each topic, I extract the most important words in each cluster through their within-cluster tf-idf score ("class-based tf-idf", or c-tf-idf). The c-tf-idf score of a word is a proxy of information density: the higher the score of a word, the more representative it should be of its topic. Hence, the list of words with the highest scores provide for each topic an easily interpretable description. The unsupervised model identifies 75 granular topics[22] I manually inspect their descriptions to reduce their number to 8 macro-topics. The granular topics, their descriptions, and this hierarchical clustering are summarized in Appendix Section A.1.1. The 7 macro topics roughly pertain to the following categories: environment (grouping news related to natural events, animals, and climate), politics (grouping news on domestic or foreign politics), health and covid (grouping news on healthcare, and those related to the pandemic from a medical perspective), economy (grouping news pertaining to finance, economic policy, businesses and management), security (including news related to reform, social movements/protests, and crime), society (grouping news pertaining to education, the judicial system, and lawmaking), and entertainment (including movies, sports, and celebrity news), which I discard from further analyses. About half of the tweets eligible for the analysis by topic are assigned to a mixed category: those are news

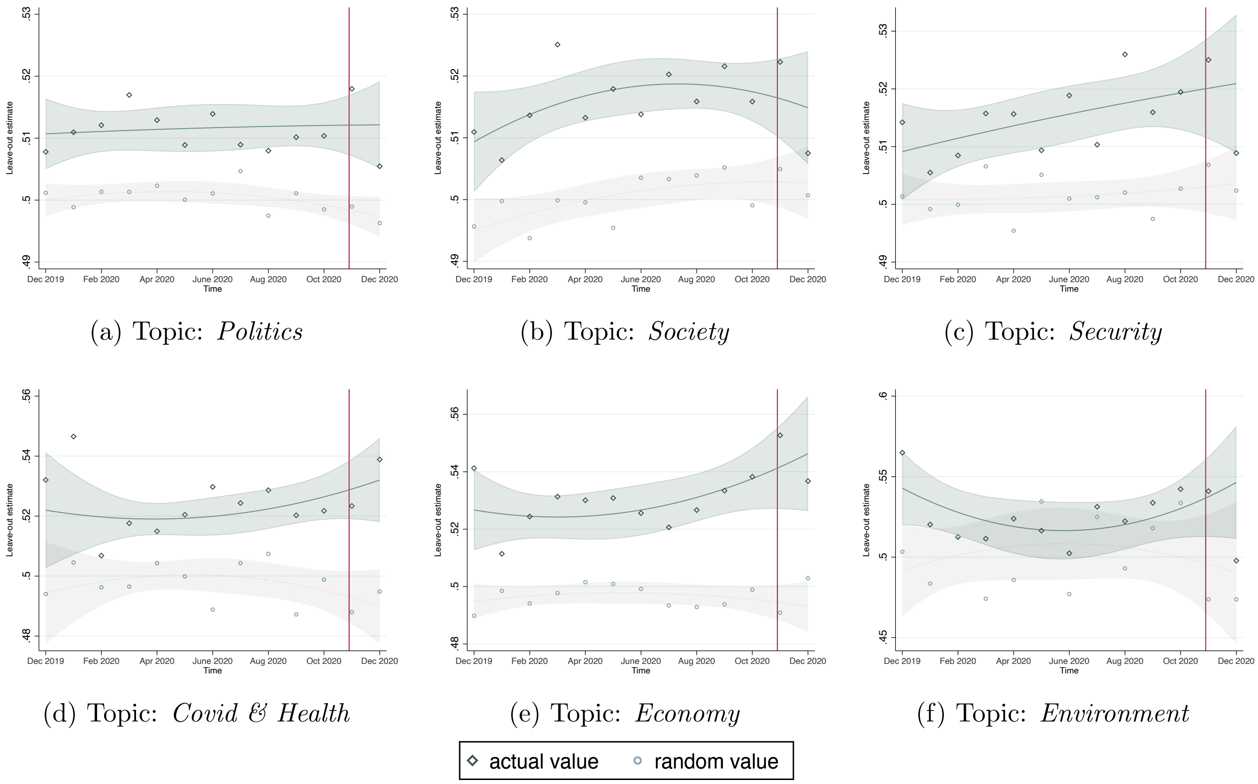
[19]This excludes all tweets that contain no other text than the url of the shared articles.

[20]For more details on BERTopic, see `https://maartengr.github.io/BERTopic`.

[21]Using HDBSCAN for clustering.

[22]I run a sensitivity analysis with a topic number ranging from 5 to 150 clusters. The optimal number of topics was identified both by comparing the BERTopic coherence score and manual inspection of topics' descriptions.

pieces that pertain multiple topics equally, or whose topic is otherwise difficult to assign.[23] To preserve the internal coherence of other topics I separate the miscellaneous category.



(a) Topic: *Politics*

(b) Topic: *Society*

(c) Topic: *Security*

(d) Topic: *Covid & Health*

(e) Topic: *Economy*

(f) Topic: *Environment*

FIGURE (IV)
Visual Polarization By News Topic

Notes: The Figure shows the monthly partisanship of leading images, by news topic (in subtitles). The shaded regions represents the 95% confidence interval of a second order polynomial fit to the values. Random values correspond to leave-out estimates obtained after randomly assigning images to parties. The vertical red line marks the date of the 2020 Presidential election (Nov 3).

Figure IV plots monthly polarization estimates within each topic.[24] These results indicate that the documented visual polarization during the entire period under analysis *across* news topics originates –at least in part– in the different visual language used *within* topics.
The estimates appear noisier for news on "environment" than for others.[25] The overlapping confidence bounds should not be interpreted as evidence of a non-partisan visual language by

[23]For example, the following news piece is equally relevant to the "security" and "entertainment" topics: *"Police shot tear gas and rubber bullets into a massive crowd that lined the streets of Argentina's capital city to pay their respects to soccer legend Diego Maradona, who died at the age of 60"*.

[24]To aid precision, I cut tokens used less than 8 times (i.e. 2 per week) within each event-topic combination. As in the overall analysis, I cut tokens at the bottom 0.0001 of the tf-idf score distribution. Similar results are obtained using a 0.00001 threshold (i.e. virtually no cut).

[25]As above, each panel portrays two series: one of real estimates and another obtained from random assignment of news sources to parties; hence, the more the latter departs from .5, the noisier the estimates within that time frame/topic.

news sources. Indeed, random values frequently depart from .5, suggesting the poor reliability of the estimates in this domain and that no conclusive inference can be drawn. This pattern likely indicates that the visual vocabulary contains too few of the elements relevant to characterize the partisan narratives of media sources for environment-related news. I discuss these aspects more in details in Subsection *II.D*.

II.C.5 Lexical and Syntactic Analysis of Visual Partisanship

I explore further the dimensions in which the visual languages of Republican-leaning and Democrat-leaning sources differ, as well as which visual features types are distinctive ideological markers. I study the partisanship of the visual tokens via the tokens' log-odds ratio of Democrats relative to Republicans.

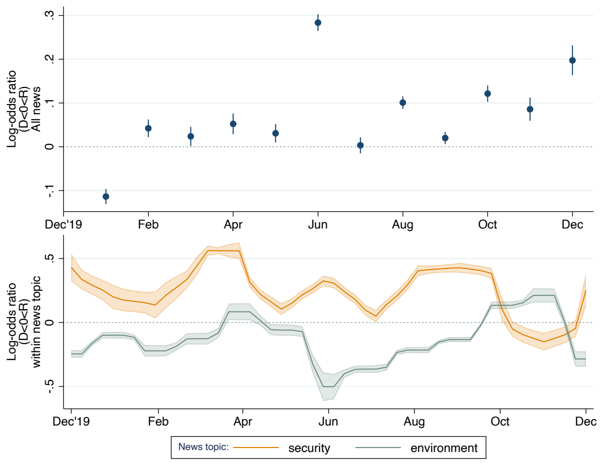
Lexicon

In text documents, the analysis of lexicon proved useful to detect clear partisan markers, namely expressions whose partisanship was constant across periods and topics (e.g. "*death tax*", see Gentzkow, Shapiro, and Taddy, 2019). The "stability" of partisanship (as in the case of clear partisan markers) is relevant because repeated exposure to stably partisan depictions is likely to develop readers' intuition, and enable them to become aware of latent slant. Similarly, unambiguous signals help developing and training tools for the automated detection of partisanship.[26] In the context of US news visual language, two types of patterns complicate the interpretation of visual words as partisan markers. On the one hand, visual words that are partisan markers within topic but not across topics: their partisanship varies with news context. On the other hand, visual words whose partisanship varies *both* over time and across news topics.

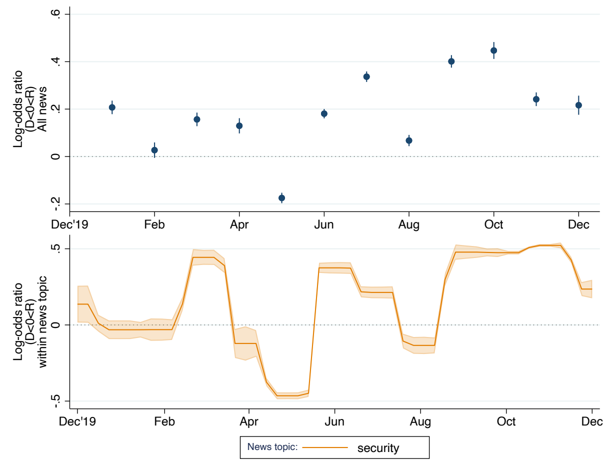
Figure V presents an example from the two above mentioned patterns. The figure displays the log odds ratios of Republican vs. Democrat token use, where positive values hence reveal higher frequency of use by Republican-leaning news sources, and vice versa. Panel (a) presents an example of a "partisan marker" whose partisanship changes across news topics: the context tokens "*fire*" and "*firefighters*", respectively indicating the presence in a picture of either fire or firefighters. As displayed in the upper half of the panel, the overall use of either token was generally more frequent in images from Republican-leaning sources (see the positive log odds ratios). The bottom half of the panel repeats the analysis focusing on two topics where these tokens exhibit different partisan leanings; the depiction of either fire or firefighters appears to

[26]This section documents the existence of the different lexical patterns, and calls for further exploration on their prevalence.

be a Republican marker in news pertaining to "security", but a relatively Democratic marker in news pertaining to "environment".



(a) Token: *"Fire & /or Firefighters"*



(b) Token: *"Police & Weapon"*

FIGURE (V)

Lexical Analysis: Log Odds-Ratios of Two Visual tokens

Notes: The Figure shows the log-odds ratios for two different tokens (indicated in panel headers). Ratios are presented for news overall (upper half of panels) and within topic (bottom half).

Panel (b) of Figure V plots instead the log odds ratios of a token resulting from a feature combination of a human subject (*police*) with a context element (*weapons*). The depiction of armed police forces is overall more frequent among Republican media sources. However, in the month of May 2020, the token is more likely to be part of the visual narratives presented by Democrat sources. This shift in partisanship coincides with the period following the homicide of George Floyd and the resulting protests. The coverage of these events by the Democrats involved the frequent representation of armed police forces, as confirmed in the lower half of the panel that narrows the analysis to news pertaining to the relevant topic.

Overall, the evidence presented in Figure V indicates that the partisanship of the visual lexicon of US news may change over time and across news topics. This complexity could make the detection of partisan narratives more challenging for images than for text. As such, the identification of partisanship in visual language requires a nuanced understanding of both the specific temporal reference of the images and the associated news topics. This is relevant from a policy viewpoint: while personal intuition can sometimes be adequate in revealing overtly partisan characterizations, detecting subject-driven visual partisanship in news requires readers to actively seek out coverage from multiple sources. Policy efforts geared towards enhancing readers' ability to identify partisan news will likely be more effective when they hinge on facilitating the comparison of diverse visual narratives.

Syntax

To explore the syntax of US news' partisan visual language, I distinguish between three syntax categories that offer distinct framing devices: depiction of human subjects ("subjects"), characterizations of such subjects ("subject+adjectives"), and context depiction ("context").[27]

Figure VI shows the evolution over time of the monthly log-odds ratios in absolute value for each syntax group. Loosely speaking, the more these values deviate from 0, the higher the mean partisanship of the tokens in the syntax group (disregarding the direction of partisanship, hence the absolute value). The Figure shows that over the entire period under analysis, tokens in the "subject" and "subject+adjectives" classes had the highest differential use by news sources of opposite political leaning. As a partisan device, subject choice dominates subjects' characterisation in most periods. This partial dominance of subject choice suggests that the syntax of visual partisanship may in itself complicate the identification of partisan images by readers. As mentioned, in fact, audiences are plausibly less cognizant of "subject-driven" partisan visual narratives than of "adjective-driven" and "context-driven" ones.

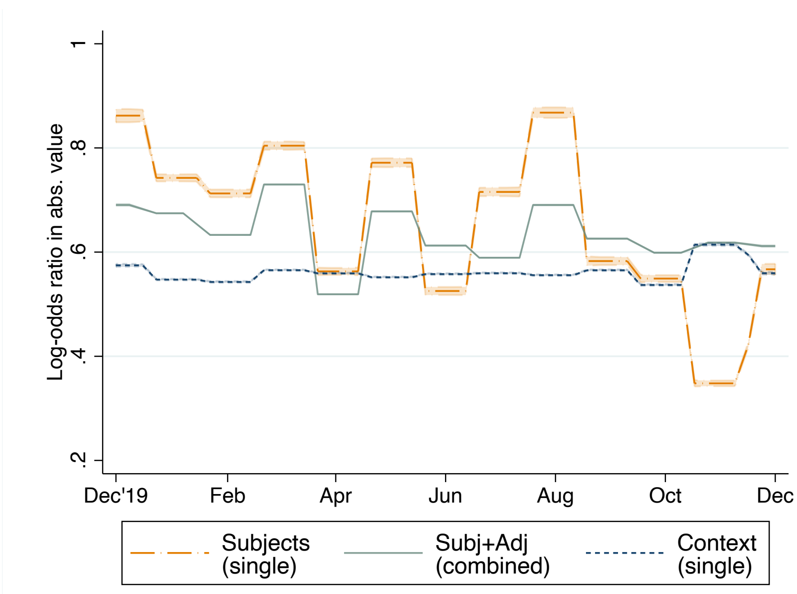


FIGURE (VI)
Visual Syntax Analysis

Notes: The Figure shows the absolute value of log-odds ratios over time for three syntax groups. The higher the average partisanship (disregarding the direction of partisanship), the more relevant the syntax class in visually partisan narratives for a given period.

[27] This set includes inanimate objects, animals, physical spaces or social events as specified in the previous subsection and in Appendix Section A.1.2.

In addition, Figure VI shows that the rank of syntax classes evolves over time, suggesting that different visual syntaxes emerge in partisan pictures over time. This pattern remarks that in studying visual contents one should avoid abstracting from the context and period in which contents emerge. This prescription has tangible implications for instance in initiatives to automate bias detection, and similar others.

II.D Evaluating the Method: Perks and Pitfalls

I briefly discuss some pitfalls and perks of the method proposed in this Section, that is, of adopting a dictionary approach to study the language of pictures. The ability of dictionary-based procedures to extract meaning from a corpus is always tightly linked to the design of the underlying vocabulary. This is particularly relevant in the study of pictures, where the link between symbols and meaning is less express and more context-dependent than in words, as a vast semiotics literature suggests (Peirce,1931 Cassirer,1944; Morris,1946; Knowlton,1964 and 1966 Veltrusky,1976; Eco,1979; Hołowka,1981; Cassidy,1982; Sebeok,1985; Langer,2009). It is thus worth noting three key observations regarding the methodology proposed in this Section.

First, while the number of tokens that can be extracted from text documents is finite, the potential *tokenizations* of an image are infinite, and subject only to the limitations of computer vision progress. This means that the interpretation of results must be approached with caution, as inference is inherently constrained by the "terminology" embedded in the vocabulary, just as is the case for all dictionary-based approaches. For instance, the visual language of environment-related news could actually be very polarized, just in dimensions not captured by the dictionary used in this study (see Figure IV).

Second, this Section presents a metric for identifying differences in visual language across the political spectrum. To this end, the analysis employs a uniform vocabulary for all images and sources, thereby generating "internally valid" results within the framework of the employed lexicon, regardless of its comprehensiveness. This is particularly pertinent to discussions on algorithmic bias: as the methodology used in this paper relies on identical algorithms to extract features from all images, any detected differences in visual language are plausibly net of algorithmic biases.

Finally, but no less importantly, compiling a visual vocabulary of tokens with annotated meaning achieves features interpretability by design. Interpretable visual tokens enable the capture of both "lexical" and "semantic" language traits, a goal otherwise challenging to achieve.

III THE EFFECT OF VISUAL PARTISANSHIP ON OPINION

The evidence presented so far indicates that US news' visual language is significantly partisan. In this Section I test and verify that distant partisan visual narratives also have significantly different impacts on news readers' opinion. In fact, to consider visual partisanship as an expression of media bias (a "*visual bias*"), leading images need not only be distinctive of Republican- or Democrat-leaning sources, but also to promote the corresponding ideological positions.

As pre-registered, I introduce a survey experiment to test two main hypotheses: first, whether partisan leading images distinctive of Republican/Democrat outlets slant the audience towards their respective party; second, whether partisan images interact with the readers' priors and increase polarization of public opinion. Following the pre-analysis plan I additionally explore the heterogeneity of the estimates along three dimensions: the baseline opinion on the issue, the perceived issue salience, and the respondents' self-reported prior knowledge on the issue. Neither of these dimensions appears to be a strong predictor of respondent' sensibility to lead images, and no neat patterns arise. I discuss the heterogeneity of results (or lack thereof) in the Online Appendix.

III.A Experimental strategy

I conduct a survey experiment on a nationally representative sample of the US population consisting of 2'000 respondents. The eligible population of the study consists of US citizens between 18 and 65 years of age. I recruit survey respondents on IPSOS's survey panel between July 2, 2021 and July 22, 2021.[28] Each respondent is exposed to news on five news issues, displayed sequentially. An issue is introduced through the following steps:

1. The respondent reads a short summary of the news.

2. She is asked to evaluate her knowledge of the issue and express the issue's relevance to her personal life/experience (i.e. its perceived salience), and her viewpoint on the issue. Some general questions (e.g. on demographics) follow the end of this section.

[28]The experiment was approved by the European University Institute's Ethics Committee (the EUI's IRS board), and informed consent was obtained from the respondents at the beginning of the survey. The experiment was pre-registered in the AEA RCT Registry with digital object identifier (DOIs): https://doi.org/10.1257/rct.7904-2.0

3. In the second part of the survey, the respondent accesses a page containing one piece of news on the issue. The piece appears in the same compact format of news previews when pieces are shared on social media, as described in the introductory Section. This preview's main elements are a brief summary-text on top (what is widely called "lead statement", an introductory text that summarizes the key details of a news piece), a leading image, a header (i.e. the title), and a byline (the line of text below the title generally providing context details). Respondents can read or skim the news pieces through a scroll-down movement, as in social media. The overall news look is that of the news previews featured on Facebook, mimicking the one illustrated in the right panel of Figure I.

4. After being exposed to the news, respondents express their opinion on the issue shifting a graphic slider to provide a numeric answer.

Steps 1-2 and 3-4 repeat five times, one for each issue.[29] The experiment consists of exogenously varying the images leading the news in step 3 among three alternatives: non partisan, distinctive of Democrat-leaning sources (hereinafter: "Democrat-leaning") or distinctive of Republican-leaning sources (hereinafter: "Republican-leaning"). All other aspects of the news previews (texts, headlines, bylines and graphic look) are held constant. Treatment assignment is randomized at individual level, and respondents are equally likely exposed to either treatment branch (with treatment status for each issue being orthogonal to the status in others).

The text in the news pieces is non-partisan, depicting facts covered by both liberal and conservative news sources without using partisan narrative frames or language.[30] Democratic- (Republican-) leaning images contain Democratic- (Republican-) visual features with high partisanship score (measured following the method described in Section II), hence they depict issues in a manner that is distinctive of Democrats (Republicans) news outlets. Vice versa, non-partisan images (hereinafter "neutral" images) contain features with low partisanship scores. Images are congruent with the true coverage on the same issues from outlets on both political sides (news pieces sourced from `www.allsides.com`). Appendix section A.2.1 displays and describes the chosen images.

The treatment news issues pertain to five news topics characterised by visually partisan language as described in Section *II.C*.4, i.e. Security, Politics, Economy, Covid & Health, and

[29]The order of issues is randomized.

[30]For all the issues the text is based on news pieces on rated "non-partisan" on `www.allsides.com`.

Society. I select one recent news issue from each of those broader topics, and respectively: the debate on police budget cuts (hereinafter: "Police funds" issue); Biden's efforts to renew the 2015 US-Iran nuclear deal ("Iran deal"); the FED forecasts on inflation ("Inflation"); the anti-Covid measures implemented in March 2020 in the US ("Covid measures"); the institution of Juneteenth as Federal holiday ("Juneteenth").

I collect respondents' opinions on these issues through the following questions:

- Police funds: *The total state and local government spending on police is currently about \$119 billion a year. If you were to decide the police budget, how much would you set it to?* [Answers readjusted to range in -100%/+100%]
- Iran deal: *From 0 to 100, in your opinion what is the probability for Biden to succeed in reviving the 2015 nuclear deal with Iran?*
- Inflation: *From 0 to 100, in your opinion what is the probability of inflation returning to pre-pandemic levels by July 2022?*
- Covid measures: *From 0 to 100, how much do you approve of the pandemic handling by public health experts in March last year?*
- Juneteenth: *From 0 to 100, how much do you support the creation of a new federal holiday for Juneteenth?*

The outcome variable of interest is the respondent's opinion on an issue after being exposed to the news. For each news issue, I estimate the following specification through OLS:

$$Y_i = \beta_0 N_i + \beta_1 D_i + \beta_2 R_i + \beta_3 X_i + \epsilon_i \tag{3}$$

where Y_i is the post-treatment opinion expressed by respondent i on a given issue, N_i is an indicator for exposure to news led by neutral-leaning images, and D_i and R_i are similar indicators for exposure to news led by Democrat-leaning or Republican-leaning leading images, for the given issue. Finally X is a vector of demeaned control variables uncorrelated with the treatment indicators, to aid the precision of the estimates.[31]

As expected by virtue of randomization, for all the treatment news issues the respondents in the three treatment branches are balanced in terms of observable characteristics, and the

[31] List of treatment-independent controls: 4 age groups, ethnicity (White, Black, Latinx, Asian, Native American), literacy, political opinion (liberal-conservative), level of interest for politics, party preference, previous knowledge on the issue, perceived salience of the issue, baseline opinion on the issue, main type of information outlet (Radio, TV, Social networks, Newspapers), frequency of use for 6 media outlets (Fox News, Breitbart, New York Post, MSNBC, New York Times, CNN), technical aspects of the survey filling (indicator for low screen resolution, total number of clicks in the survey introduction), and State of residence fixed effects. Appendix Table A.2.2 reports the estimates omitting all controls other than baseline opinion.

standardized difference is always below the critical threshold of 0.25 (see Imbens and Rubin, 2015).[32] Appendix Table A.2.1 summarizes the main variables of the study.

III.B Do partisan images affect public opinion?

I investigate the extent to which Democrat- and Republican- leaning images shift public opinion by testing, for each news issue, the significance and the equality of treatment coefficients β_1 and β_2 in (3). The analysis presented in this Section excludes survey respondents who do not pass an attention check placed at half survey; it also discards single answers given after treatment exposures of strictly less than 5 seconds.[33] Both exclusion criteria are pre-registered.

Table III reports the estimated treatment effects on the respondents' opinion on each issue. News issues are ordered by the distinctiveness of Democrats' and Republicans' baseline ideological positions on the issues, measured at the beginning of the survey with a general question (opinions are most similar for Police funding, and least similar for Juneteenth,as inferred from the distribution of respondents' opinions on the issues measured at baseline. (Appendix Figure A.2.1). All dependent variables are readjusted to range between -50 and +50, with the exception of the "Defund Police" issue, whose opinion ranges between -100% and +100% of the true Police budget.[34] All the dependent variables have been adjusted so that higher and lower values correspond respectively to Democrats' and Republicans' ideological positions relative to an intermediate position ("indifference", marked with value 0); hence positive coefficients indicate a relatively pro-Democratic opinion stance, and vice versa. Round parentheses contain robust standard errors, while square brackets contain the p-values for two-sided tests of equality (with tested coefficients pairs indicated on the left) using heteroskedasticity-robust standard errors.

Police funding. The dependent variable is the answer to the question: "*If you were to decide the police budget, how much would you set it to? [Relative to the current budget of $119 billion]*". Coefficients represent the deviation, in percentage points, from the indifference position (0, indicating no budget variation). To ease the comparison with other issues, the answers to this question (ranging between -100 and +100) have been adapted so that positive values indicate a budget decrease (i.e. a positive budget cut). Relative to the news piece led by a

[32]See the Online Appendix for the five Tables (A.2.1 to A.2.5) displaying the balance of observables characteristics across treatment branches for the 5 treatment news issues.

[33]This is the time just sufficient to load the news page and immediately scroll down to the "next page" button.

[34]Respondents are given as reference the State and Local total Police expenditure in 2018 (119 billion). Data accessed on January 29, 2021 from: `https://state-local-finance-data.taxpolicycenter.org/pages.cfm`

TABLE (III)
Impact of Leading Images On News-Readers' Opinion

Dependent variable:	(1) Opinion on "Defund Police" (Budget cut in -100 +100)	(2) Opinion on "Iran deal" (Confidence, in -50+50)	(3) Opinion on "Inflation" (Confidence, in -50+50)	(4) Opinion on "Covid measures" (Blame in -50+50)	(5) Opinion on "Juneteenth" (Policy support, in -50+50)
Neutral images (N)	-0.594	-0.421	1.084	-7.260	8.687
	(0.745)	(0.692)	(0.414)	(0.223)	(0.457)
Democrat images (D-N)	1.376	-0.283	-0.548	1.364	-0.495
	(1.097)	(0.987)	(1.071)	(1.291)	(0.836)
	[0.210]	[0.775]	[0.609]	[0.291]	[0.554]
Republican images (R-N)	-2.394	-2.329	-2.941	-0.903	-0.229
	(1.309)	(0.898)	(1.118)	(1.352)	(0.830)
	[0.068]	[0.010]	[0.009]	[0.505]	[0.782]
Democrat-Republican (D-R)	3.771	2.047	2.392	2.267	-0.266
	(1.240)	(0.910)	(1.161)	(1.366)	(0.840)
	[0.002]	[0.025]	[0.039]	[0.097]	[0.751]
Observations	1565	1599	1615	1584	1542
Controls:	Y	Y	Y	Y	Y

Notes: The Table presents OLS estimates of the effect of the Democrat-leaning (D), neutral (N), and Republican-leaning (R) news-leading images on respondents' opinion after exposure to the news. Column headers indicate the relevant news issue. The dependent variable for the "Defund Police" issue ranges in [-100,+100], while all others range in [-50+50]. Dependent variables are adjusted so that the maximum value corresponds to Democrats' ideological position (thus, positive coefficients indicate a pro-Democratic opinion, and vice versa). In the Table, round parentheses present robust standard errors and square brackets contain the p-values for two-sided tests of equality between coefficients (tested pairs are noted on the left). Treatment-independent controls are indicators for: 4 age groups, ethnicity (White, Black, Latinx, Asian, Native American), literacy, political opinion (liberal-conservative), level of interest for politics, party preference, previous knowledge on the issue, perceived salience of the issue, baseline opinion on the issue before treatment exposure, main type of information outlet (Radio, TV, Social networks, Newspapers), frequency of use for 6 media outlets (Fox News, Breitbart, New York Post, MSNBC, New York Times, CNN), technical aspects of the survey filling (indicator for low screen resolution, total number of clicks in the survey introduction), and State of residence fixed effects.

Republican-leaning image, the same news piece led by a Democratic-leaning picture significantly increases the desired budget cut by an additional 3.77 percentage points – equivalent to about $ 4.5 billion in monetary terms (st. error = 1.240, p-value = .002). A comparison of the maximum opinion spread produced by image variation (that is, the difference between the largest and the smallest treatment coefficients) and the smallest effect exerted by news exposure (that is, the smallest coefficient in absolute value) provides an indication of the effect of visual partisanship relative to the more general effect of news previews. The rationale is the following: as all treatment branches display the same text content, all coefficients capture the effect of exposure to the constant elements (headline, summary, byline, etc.). Given this, any difference in opinion across treatment branches identifies the additional effect that image partisanship can exert on top of the overall effect of news previews.[35] In this first news issue, image variation can increase the desired Police budget cuts by up to 3.77 percentage points, (from a minimum of -2.39 to a maximum of 1.37), that is more than 6 times the increase produced from overall exposure to news previews (amounting to .54 percentage points, as indicated by the smallest coefficient in absolute value, that of the neutral treatment).[36]

Iran deal. For this issue, the dependent variable is the answer to the question: "*From 0 to 100, in your opinion what is the probability for Biden to succeed in reviving the 2015 nuclear deal with Iran?*". The answers to this question have been adapted to range in -50 + 50 and the coefficients can once again be interpreted as deviation from the indifference position (moved from "50" to 0). All the coefficients are negative, indicating a perceived likelihood of deal success lower than 50%. Compared to respondents exposed to Republican-leaning leading images, those exposed to Democratic-leaning images judge the deal success as significantly more likely, with a margin of 2.05 percentage points (st. error= .910, p-value = .025). Similarly, respondents' exposed to neutral images report a higher perceived likelihood of the deal success (with a margin of 2.33 percentage points, estimated with st. error = .898 and p-value = .010).

I compare again the coefficient range to the smallest treatment coefficient in absolute value: while the deal news always produce a loss in confidence of Biden's success, the variation in images can produce an additional confidence loss, 5.54 times as big.[37]

Inflation. For this issue, the dependent variable is the answer to the question: "*From 0 to 100, in your opinion what is the probability of inflation returning to pre-pandemic levels by July 2022?*". Once again, the answers to this question have been adapted to range in -50 + 50, and coefficients represent deviation from an indifference stance (moved from "50" to 0). Compared to respondents exposed to Republican-leaning images, those who see Democratic-leaning images report a higher perceived likelihood of Biden's success, with a 2.39 percentage points difference (st. error = 1.161, p-value = .039); the smallest effect is obtained by news exposure with Democrat-leaning images, with a .54 p.p. coefficient. Hence, the variation in images attains about 4.4 times the opinion change of the news preview overall.

Covid measures. For this issue, the dependent variable is the answer to the question: "*From*

[35] Note: experimental images lead only neutral (non politically partisan) text elements. The experiment does not speak to the impact of partisan text.

[36] Image variation produces an opinion change 634% that produced by the news preview with a neutral image.

[37] In this issue, image variation produces a change in opinion up to 2.33 p.p.; exposure to the news attains a minimum opinion change of .42 p.p. (negative). The former is 554 % the latter.

0 to 100, how much do you approve of the pandemic handling by public health experts in March last year?". Also in this case answers have been adapted to range in -50 + 50, and coefficients represent deviation from the indifferent opinion (moved from "50" to 0) while the +50 value corresponds to Democrats' ideological pole (i.e. the strongest blame for the covid handling). Compared to Republican-leaning leading images, Democratic-leaning ones increase (i.e. decrease by less) the dissatisfaction for the pandemic management, with a 2.27 percentage points gap (st. error = 1.366, p-value = .097). As above, I compare the coefficients' range to the smallest treatment coefficient (-5.89); image variation produces an additional approval increase of more than a third the increase from overall exposure to the news (+38%).

Juneteenth. For this issue, the dependent variable is the answer to the question "*From 0 to 100, how much do you support the creation of a new federal holiday for Juneteenth?*". Again, the answers have been adapted to range in -50 + 50, and coefficients represent deviation from the indifferent opinion (moved from "50" to 0). For this issue treatments lead to negligible differences and imprecise estimates: the opinion margin between Republican-leaning and Democratic-leaning images amounts to .266 percentage points, and the effect is statistically indistinguishable from 0 (st. error= .840, p-value .751).

Overall, the results in Table III indicate that leading images have a non-negligible impact on news readers' opinion. Pictures distinctive of Democrats/Republican news outlets pull the audience towards their respective parties' ideological poles. These results indicate that the visual partisanship documented in Section II promotes news outlets' ideological positions among readers, hence the "visual bias" is a tangible expression of political bias in the media.

In the experiment, when images produce a significant impact on opinion its magnitude ranges between 38% and 634% of the overall effect from exposure to news previews; in 3 of the 4 precisely estimated impacts, the "slanting effect" of pictures dominates that of other elements of the news previews, and notably of written content.[38] One implication of these findings is that a news piece rated as "non partisan" through a text-based analysis could still exert a partisan influence on readers. Hence, any measure of political bias in news shall take into account both text and images to prevent slanted media from using pictures strategically.

Another pattern highlighted by the results in Table III concerns the decrease of the effect of images relative to text in the distance between parties' ideological positions. As above mentioned, parties' stances are most similar for the "Defund police" issue, and least similar for

[38]This refers to the short text appearing in news previews. Survey participants did not read a full article.

the "Juneteenth" issue (Appendix Figure A.2.1). The effect of images relative to text shrinks for two reasons: first, due to a decrease in the numerator (i.e. the maximum distance across treatment branches, which becomes smaller across columns from left to right); second, and more evidently, due to an increase in the denominator (the smallest coefficient in absolute value, which in the last column is more than 15 times bigger than in the first). Independently of the images leading the news, readers seem to react more to news previews covering issues for which the ideological positions across parties are more distinct. These patterns are suggestive but should not be taken as conclusive evidence: a formal assessment of these relationships requires testing a large number of issues, hence falls outside the scope of the present work.

III.C Does visual partisanship cause opinion polarization?

I test whether the exposure to partisan images causes polarization to increase in the general public, exploring the treatment effects across political affiliations. I find that exposure to partisan images increases overall issue polarization.

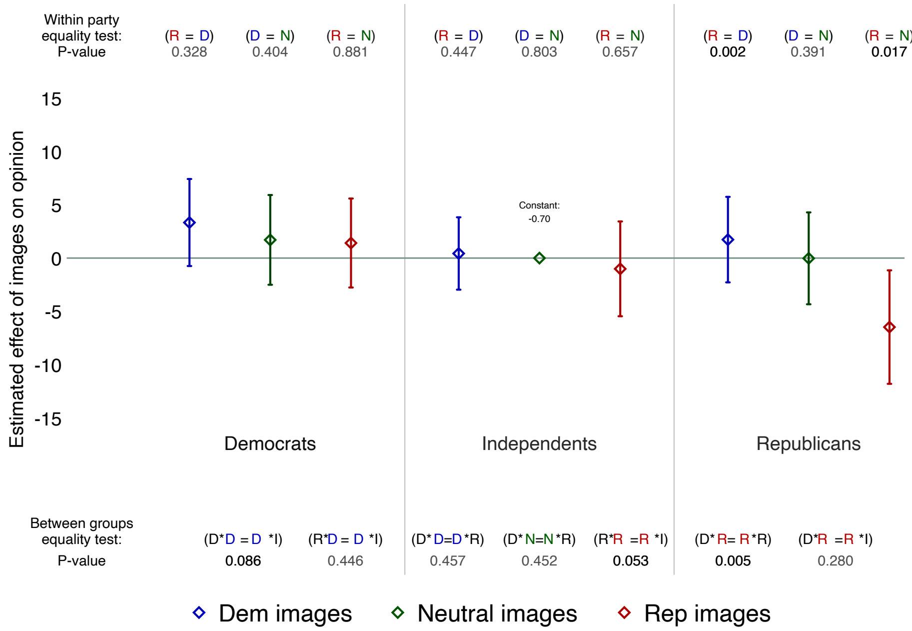
III.C.1 Polarization within-party

Within-party polarization occurs if the exposure to different leading images brings members of the same group apart. The results indicate that visual partisanship increases polarization *within* parties, but within each group/topic the effect is statistically indistinguishable from 0.

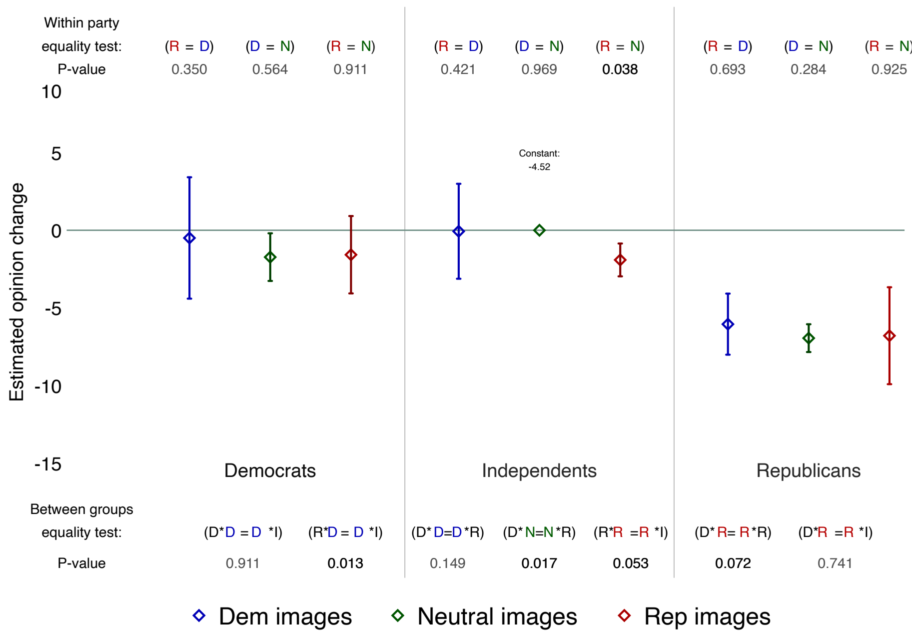
Figures VII, and VIII show the heterogeneous impact of leading images on individuals from different political affiliations separately for the *Defund Police*, *Covid measures*, *Iran deal*, and *Inflation* news issues.[39] The equality tests on top of each panel compare the treatment coefficients within the same party (e.g. whether Dem-leaning and Rep-leaning images have the same effect on Democrats); the p-values therefore measure the within-group polarization induced by visual bias for a given news issue. The equality tests at the bottom of each Figure instead compare the effect of exposure to the same leading image across different parties (e.g. whether Dem-leading images have the same effect on Democrats and on Republicans); the corresponding p-values indicate the extent to which one's party affiliation affects the impact of leading images.[40] The relative distance among the three treatment coefficients varies with the political party, indicating that the effect of an image interacts with readers' political stance. For

[39] Appendix Table A.2.3 reports the point estimates for all the issues, including *Juneteenth*.

[40] The test label indicates respondents' party affiliation through the external letter, and the images while the internal letters. For instance, test (D*R=R*I) compares coefficients for Democrats and Independents both exposed to Republican-leaning images.



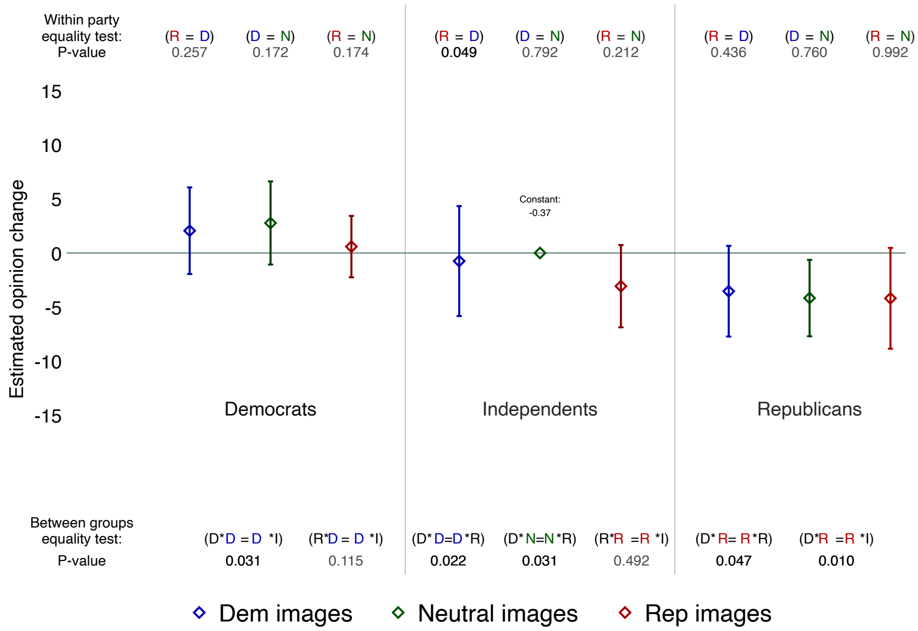
(a) News issue: *Defund Police*



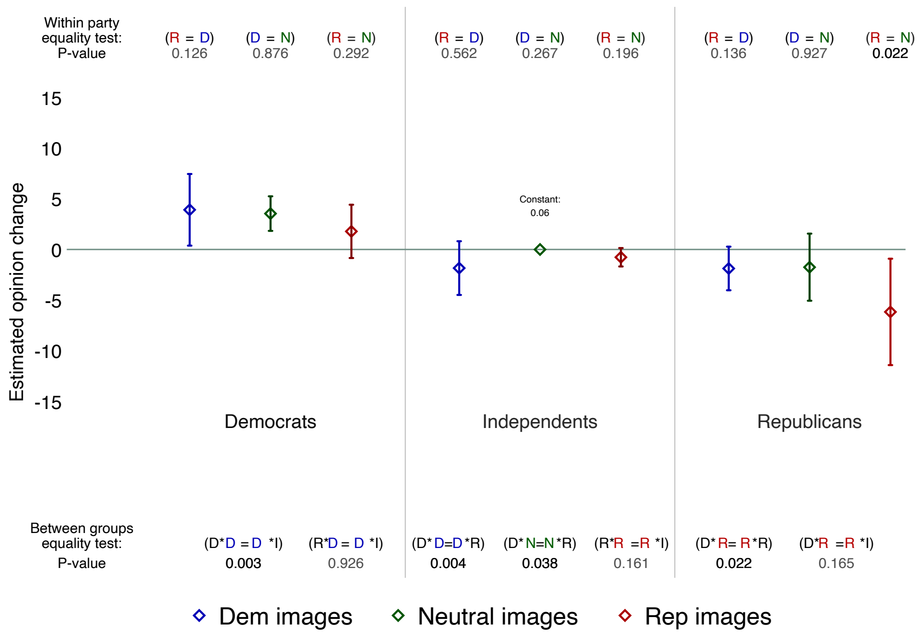
(b) News issue: *Covid Measures*

FIGURE (VII)
Heterogeneous effects of images on opinion,
by respondents' political party affiliation

Notes: The Figure shows OLS estimates of opinion changes after news exposure (news issues indicated below each panel). Treatments are interacted with respondent's party affiliation. Omitted regression category: Republicans exposed to Rep-leaning images. Lines indicate 95% CI (heteroskedasticity-robust st. errors). Equality tests on top of each Figure compare coefficients within each party; those at the bottom compare coefficients across parties (tested coefficients indicated in parentheses). All p-values are for two-sided tests of equality, with bold font marking statistical significance at 10 percent level or higher.



(a) News issue: *Nuclear deal with Iran*



(b) News issue: *Inflation*

FIGURE (VIII)
Heterogeneous effects of images on opinion,
by respondents' political party affiliation

Notes: The Figure shows OLS estimates of opinion changes after news exposure (news issues indicated below each panel). Treatments are interacted with respondent's party affiliation. Omitted regression category: Republicans exposed to Rep-leaning images. Lines indicate 95% CI (heteroskedasticity-robust st. errors). Equality tests on top of each Figure compare coefficients within each party; those at the bottom compare coefficients across parties (tested coefficients indicated in parentheses). All p-values are for two-sided tests of equality, with bold font marking statistical significance at 10 percent level or higher.

Democrats, partisan and non-partisan images never have statistically distinguishable effects, with the exception of the difference between Republican-leaning and neutral images leading the news on Iran (p-value: 0.065). For Independents, the only significant difference is once again between Republican-leaning and neutral images leading the news on Iran (p-value: 0.011). Among Republicans, the effect of Republican-leaning vs. Democrat leaning images differs for news on police funding (p-value: 0.002) and on inflation (p-value: 0.061); similarly, the effect of Republican-leaning images differs from that of neutral images for police funding (p-value: 0.016) and inflation news (p-value: 0.050).

Does the exposure to the same partisan image produce equal effect on Democrats and Republicans? Overall, partisan images seem to increase polarization more within the Republican party. On the one hand, the coefficient of Democrats exposed to Rep-leaning images is always statistically different from that of Republicans exposed to the same images (see the third coefficient from the left, marked by a red line, in each Figure). Equality is rejected with at least 90% confidence in all four news issues. On the other hand, equality of effects for Democrats and Republicans exposed to Dem-leaning images is rejected with $<$ 90% confidence in two of the four news issues, as indicated by the tests at the bottom of Figures VII and VIII (test label: D*D=D*R; p-values: 0.004 for Covid measures; 0.098 for Iran deal). Similarly, the equality test for Democrats and Republicans exposed to neutral images (test label: D*N=N*R) rejects the equality in the same two issues (p-values: 0.037 for Covid measures; 0.003 for Iran deal).

*III.C.*2 Polarization across parties

The polarizing effects *within* party could "cancel out" in society as a whole if partisan visual narratives made Republicans exposed to Dem images and Democrats exposed to Rep images sufficiently close to one another. As shown in Figures VII and VIII, this does not generally hold true (with the exception of the "Defund police" issue, where opinions are less dispersed). In addition, the Figures highlight that a partisan image slants opinions always towards a given direction, independently of whether the reader is a Democrat or a Republican. In other words, the rank of coefficients for Republican-leaning, neutral, and Democrat- leaning images is nearly constant. This means that the decoding process of the visual contents is homogeneous in the population studied (different people do not interpret the images differently).[41] Jointly, these patterns imply that in this multi-party society, the –always opposite– slants of Dem-leaning and

[41] This is not a given, and it rather depends on the culutural homogeneity of the population under analysis.

Rep-leaning images are added up, and albeit not significant within each party, they produce a statistically significant polarization in the general public. Moreover, if readers are only exposed to partisan images aligned with their party's viewpoint (hereinafter: "partisan aligned images"), as in information echo-chambers, the effect of visual bias on issue polarization is exacerbated.

The polarizing effects of partisan aligned images are formally verified by testing the equality between the maximum and minimum opinion distance images can instil between Democrats and Republicans. This amounts to testing whether the distance between Democrats and Republicans exposed to aligned images is greater than the distance between the same groups exposed to images that do not reinforce their prior.[42] If members of different parties had homogeneous reactions to the partisan images, the tested quantities would be equal. Appendix Table A.2.3 reports the p-values of the corresponding one sided-tests (lower Panel, first line), showing that the null hypothesis is rejected in all four issues.

In synthesis, the direction and the magnitude of the effects of visually partisan narratives are such that they cause polarization to increase in the general public. Moreover, the presence of information echo chambers can exacerbate this effect by increasing exposure to partisan-aligned images and reducing that to opposite-leaning pictures.

III.D Inference from the Experiment: Within and Beyond

I here summarize the conclusions that can be drawn from the experiment, its limits of inference, and the patterns that remain as speculation for future studies.

The experiment described in this paper was designed to answer the question of whether an effect of lead images on opinion and issue polarization exists; the main experimental results confirm both dynamics. The estimates' external validity must however be considered with care. For instance, testing different images or news issues could lead to higher/lower estimates. Further research is needed to explore the sensitivity of the estimates to variation of different news elements, including graphic rendering (testing different formats of online news), news topics, or slant in text.

The experiment also finds that in 3 out of the 4 news issues in which leading images had a significant impact, the opinion variation accruing to images was larger than the general effect of news previews, including the (politically neutral) text elements. This proves that images *can* be more relevant than text in affecting the opinion of online-news readers; however, it would be

[42] This nets out the gap accruing to the political stance and independent of the images.

factually incorrect to generalize the result and conclude, for instance, that visual bias dominates text bias.

More generally, the experiment measured the opinion of a large sample of individuals on a limited number of news issues (5): hence, any pattern arising from comparisons across issues offers, at best, suggestive evidence. Further research is needed to test simultaneously several issues and allow solid inference *across* issues. Nevertheless, even the limited number of news tested suffices to prove an important point: newscasts can bypass text-based fact checking and still effectively slant readers' opinion in relevant domains such as politics, economy, or security.

IV CONCLUSION

This study explores how pictures leading online news pieces in the US exert a political influence on news readers. The first part of this paper explores the non-verbal language of US news, and it documents a high degree of partisanship in the visual narratives adopted by news sources across the political spectrum. The analysis of visual language highlights that both lexical and syntactic aspects concur to making partisan narratives difficult to disentangle, potentially more in images than in text. This analysis suggests that policy efforts aimed at reinforcing readers' ability to detect partisan coverage should concentrate on facilitating the comparison of diverse visual narratives by readers.

The second part of the paper tests the direct effect of visually-partisan images on public opinion. It finds that partisan visual narratives slant readers' opinion towards the outlets' ideological poles, hence that visual partisanship is an expression of political media bias. The experimental results also show that news visual bias has a positive causal effect on issue polarization, as readers on both sides of the political spectrum react more distinctly to pictures aligned with their political stance. This pattern implies that the polarizing effect of visual bias is further exacerbated if readers' source their news exclusively from like-minded outlets. Finally, the experiment demonstrates that newscasts can bypass text-based fact checking and still be effective in slanting readers' opinion, by writing politically neutral text and conveying their bias through partisan images. This result calls for an inclusion of image scrutiny in the quality assessments of news.

Nuffield College and Oxford University

References

Abele, A. E., N. Hauke, K. Peters, E. Louvet, A. Szymkow, and Y. Duan (2016). “Facets of the fundamental content dimensions: Agency with competence and assertiveness-Communion with warmth and morality”. *Frontiers in psychology* 7, 1810.

Allcott, H., L. Boxell, J. C. Conway, B. A. Ferguson, M. Gentzkow, and B. Goldman (2020a). *What explains temporal and geographic variation in the early US coronavirus pandemic?* Tech. rep. National Bureau of Economic Research.

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Yang (2020b). “Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic”. *Journal of Public Economics* 191, 104254.

Ash, E., R. Durante, M. Grebenschikova, and C. Schwarz (2021). “Visual representation and stereotypes in news media”.

Bailenson, J. N., S. Iyengar, N. Yee, and N. A. Collins (2008). “Facial similarity between voters and candidates causes influence”. *Public opinion quarterly* 72.5, 935–961.

Bowden, M. (2015). *Winning body language: Control the conversation, command attention, and convey the right message without saying a word.*

Boxell, L. (2021). “Slanted images: Measuring nonverbal media bias during the 2016 election”. *Political Science Research and Methods.*

Cassidy, M. F. (1982). “Toward integration: Education, instructional technology, and semiotics”. *ECTJ* 30.2, 75–89.

Cassirer, E. (1944). “An Essay on Man (New Haven and London)”. *Yale University.*

Dari, S., N. Kadrileev, and E. Hullermeier (2020). “A Neural Network-Based Driver Gaze Classification System with Vehicle Signals”, 1–7.

De Vreese, C. H. (2004). “The effects of frames in political television news on issue interpretation and frame salience”. *Journalism & Mass Communication Quarterly* 81.1, 36–52.

DellaVigna, S. and M. Gentzkow (2010). “Persuasion: empirical evidence”. *Annu. Rev. Econ.* 2.1, 643–669.

Demszky, D., N. Garg, R. Voigt, J. Zou, M. Gentzkow, J. Shapiro, and D. Jurafsky (2019). “Analyzing polarization in social media: Method and application to tweets on 21 mass shootings”. *arXiv preprint arXiv:1904.01596.*

Doherty, C., J. Kiley, and N. Asheer (2019). “Partisan antipathy: More intense, more personal”. *Pew Research Center*.

Druckman, J. N., S. Klar, Y. Krupnikov, M. Levendusky, and J. B. Ryan (2020). “How affective polarization shapes Americans’ political beliefs: A study of response to the COVID-19 pandemic”. *Journal of Experimental Political Science*, 1–12.

Eco, U. (1979). *A theory of semiotics*. Vol. 217. Indiana University Press.

Ekman, P. (2009). *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*. WW Norton & Company.

Furnham, A. and E. Petrova (2010). *Body language in business: Decoding the signals*. Palgrave Macmillan.

Gabielkov, M., A. Ramachandran, A. Chaintreau, and A. Legout (2016). “Social clicks: What and who gets read on Twitter?” In: *Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science*, 179–192.

Gadarian, S. K., S. W. Goodman, and T. B. Pepinsky (2021). “Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic”. *Plos one* 16.4, e0249596.

Gawronski, B. and B. K. Payne (2011). *Handbook of implicit social cognition: Measurement, theory, and applications*. Guilford Press.

Gentzkow, M., B. Kelly, and M. Taddy (2019). “Text as data”. *Journal of Economic Literature* 57.3, 535–74.

Gentzkow, M. and J. M. Shapiro (2010). “What drives media slant? Evidence from US daily newspapers”. *Econometrica* 78.1, 35–71.

— (2011). “Ideological segregation online and offline”. *The Quarterly Journal of Economics* 126.4, 1799–1839.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). “Measuring group differences in high-dimensional choices: method and application to congressional speech”. *Econometrica* 87.4, 1307–1340.

Gollwitzer, A., C. Martel, W. J. Brady, P. Pärnamets, I. G. Freedman, E. D. Knowles, and J. J. Van Bavel (2020). “Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic”. *Nature human behaviour* 4.11, 1186–1197.

Grabe, M. E. and E. P. Bucy (2009). *Image bite politics: News and the visual framing of elections*. Oxford University Press.

Greene, S. and P. Resnik (2009). “More than words: Syntactic packaging and implicit sentiment”. In: *Proceedings of human language technologies: The 2009 annual conference of the north american chapter of the association for computational linguistics*, 503–511.

Groseclose, T. and J. Milyo (2005). “A measure of media bias”. *The Quarterly Journal of Economics* 120.4, 1191–1237.

Haim, M. and M. Jungblut (2021). “Politicians’ self-depiction and their news portrayal: Evidence from 28 countries using visual computational analysis”. *Political Communication* 38.1-2, 55–74.

Hołowka, T. (1981). “On conventionality of signs”.

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jha, S. and C. Busso (2020). “Estimation of Driver’s Gaze Region from Head Position and Orientation using Probabilistic Confidence Regions”. *arXiv preprint arXiv:2012.12754*.

Knowlton, J. Q. (1964). “A socio-and psycho-linguistic theory of pictorial communication.”

— (1966). “On the definition of “picture””. *AV communication review* 14.2, 157–183.

Koliska, M. and K. Oh (2021). “Guided by the Grid: Raising Attention with the Rule of Thirds”. *Journalism Practice*, 1–20.

Langer, S. K. (2009). *Philosophy in a new key: A study in the symbolism of reason, rite, and art*. Harvard University Press.

Lee, J., M. Muñoz, L. Fridman, T. Victor, B. Reimer, and B. Mehler (2018). “Investigating the correspondence between driver head position and glance location”. *PeerJ Computer Science* 4, e146.

Levy, R. (2021). “Social media, news consumption, and polarization: Evidence from a field experiment”. *American economic review* 111.3, 831–870.

Lewis, J. B., K. Poole, H. Rosenthal, A. Boche, A. Rudkin, and L. Sonnet (2021). “Voteview: Congressional Roll-Call Votes Database”. *See https://voteview.com/*.

Lunenburg, F. C. (2010). “Louder than words: The hidden power of nonverbal communication in the workplace”. *International Journal of Scholarly Academic Intellectual Diversity* 12.1, 1–5.

McCombs, M. and A. Reynolds (2009). “How the news shapes our civic agenda”. In: *Media effects*. Routledge, 17–32.

Morris, C. (1946). “Signs, language and behavior.”

Parks, D., A. Borji, and L. Itti (2015). "Augmented saliency model using automatic 3d head pose detection and learned gaze following in natural scenes". *Vision research* 116, 113–126.

Peirce, C. S. (1931). "Collected Papers, Cambridge". *Harvard University Press* 1958.2, 643.

Peng, Y. (2018). "Same candidates, different faces: Uncovering media bias in visual portrayals of presidential candidates with computer vision". *Journal of Communication* 68.5, 920–941.

Poole, K. T. and H. Rosenthal (1985). "A spatial model for legislative roll call analysis". *American journal of political science*, 357–384.

Prat, A. (2018). "Media power". *Journal of Political Economy* 126.4, 1747–1783.

Prat, A. and D. Strömberg (2013). "The political economy of mass media". *Advances in economics and econometrics* 2, 135.

Prior, M. (2013). "Media and political polarization". *Annual Review of Political Science* 16, 101–127.

Recasens, M., C. Danescu-Niculescu-Mizil, and D. Jurafsky (2013). "Linguistic models for analyzing and detecting biased language". In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1650–1659.

Romer, D. and K. H. Jamieson (2020). "Conspiracy theories as barriers to controlling the spread of COVID-19 in the US". *Social science & medicine* 263, 113356.

Sebeok, T. A. (1985). *Contributions to the Doctrine of Signs*. University Press of America.

Shearer, E. (2021). "More than eight-in-ten Americans get news from digital devices". *Pew Research Center* 12.

Strömberg, D. (2015). "Media and politics". *economics* 7.1, 173–205.

Sutherland, C. A., J. A. Oldmeadow, I. M. Santos, J. Towler, D. M. Burt, and A. W. Young (2013). "Social inferences from faces: Ambient images generate a three-dimensional model". *Cognition* 127.1, 105–118.

Veltrusky, J. (1976). "Some aspects of the pictorial sign". *Semiotics of Art. Ladislav.*

Walters, S. B. (2011). *Principles of kinesic interview and interrogation*. Boca Raton, FL: CRC Press.

Yano, T., P. Resnik, and N. A. Smith (2010). "Shedding (a thousand points of) light on biased language". In: *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, 152–158.

A.1 Appendix 1 (for the Analysis of Visual Partisanship)

A.1.1 List of US News Sources:

The following is the list of the main News sources extrapolated from *Similarweb.com*. The partisanship scores are listed after each media Twitter handle (first the rating by Allsides.com, then that by Adfontesmedia –where positive numbers mark Republican leaning):

> AlterNet (-30.33, LL); TheAtlantic (-19.66, L); Salon (-19.35, LL); politicususa (-16.62, LL); theintercept (-16.5, LL); MSNBC (-13.76, LL); CNN (-12.15, LL); voxdotcom (-11.93, LL); GuardianUS (-10.35, L); TIME (-10.22, L); NYTimes (-8.71,LL); NBC News (-8.61, L); Politico (-7.98, L); PittsburghPG (4.8, R); TPInsidr (7.67, R); RealClearNews (13.07, R); nypost (14.2, RR); FreeBeacon (15.9, RR); WashTimes (16.12, R); FoxNews (17.19, R); realDailyWire (18.63, RR); BreitbartNews (25.67, RR).

TABLE (A.1.1)
Topics-reduction Scheme:

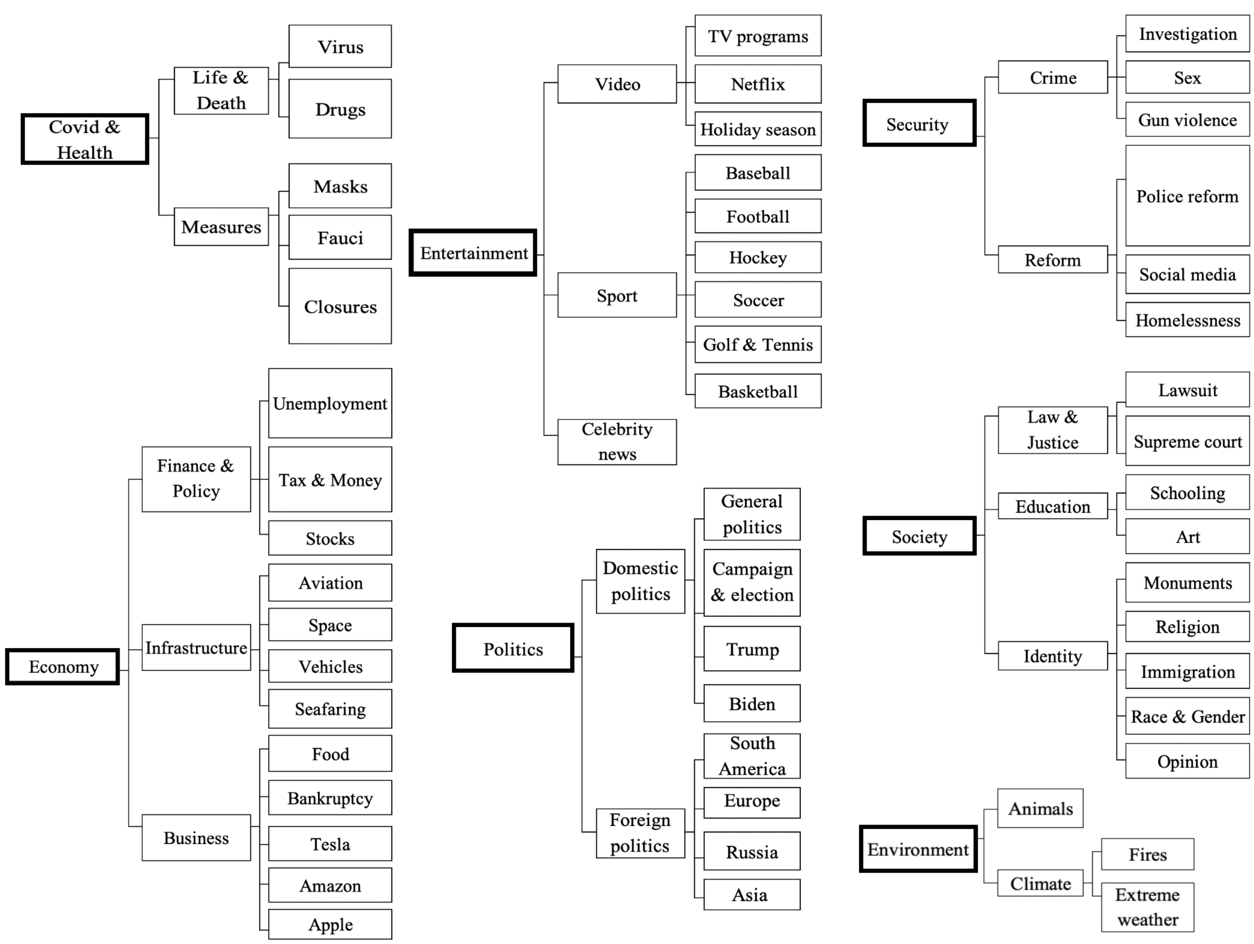


TABLE (A.1.2)

VOCABULARY SUMMARY STATISTICS: "SUBJECT" FEATURES AND FEATURE-COMBINATIONS INVOLVING "SUBJECTS"

"Subjects" subclasses:	Label in syntax index	N distinct tokens in subclass	Share of total "Subject" tokens	Total "Subject" tokens	Total presence in pictures
Single features:					
Celebrity status (y/n)	"SCele"	2	0.00004	49'951	1'231
Name	"SNam"	49'928	0.9995	49'951	96'091
Rank in terms of centrality	"SRnk"	10	0.00022	49'951	278'873
Political leaning decile	"SPoL"	9	0.00018	49'951	48'218
Gender	"SGen"	2	0.00004	49'951	208'746
Combinations of features:					
Celebrity status (y/n)	"SCele & ..."	1165	0.000451	2'582'232	6'924'932
Name	"SNam & ..."	2'573'099	0.996	2'582'232	3'515'178
Rank in terms of centrality	"SRnk & ..."	70'046	0.0271	2'582'232	8'899'051
Political leaning decile	"SPoL & ..."	3'689	0.00143	2'582'232	1'510'694
Gender	"SGen & ..."	1'168	.000452	2'582'232	7'429'796

Notes: The table provides summary statistics for the "Subjects" syntax class. The upper panel refers to single features only (excluding combinations). The bottom panel refers to features combinations. All statistics encompass only tokens that appear at least once in analysed images. Column 1 lists the subclasses in the "Subjects" class, marked in the data using the labels in Column 2. Column 3 indicates the number of distinct tokens for each subclass (denoting the different values a subclass can take). Column 4 indicates the percentage of tokens in the subclass out of all "Subject" tokens, whose total is listed in Column 5. Column 6 lists the total occurrence of subclass tokens across images.

TABLE (A.1.3)

VOCABULARY SUMMARY STATISTICS: "ADJECTIVE" FEATURES AND FEATURE-COMBINATIONS INVOLVING "ADJECTIVES"

"Adjetives" subclasses:	Label in syntax index	N distinct tokens in subclass	Share of total "Adjective" tokens	Total "Adjective" tokens	Total presence in pictures
Single features:					
Centrality	"ACent"	4	.1	40	222'771
Face size	"ASize"	3	.075	40	190'373
Use of Face Mask	"AKmsk"	2	.05	40	181'998
Blur level	"AKblr"	3	.075	40	220'737
Light exposure	"AKexp"	3	.075	40	194'887
Head yaw	"AKyaw"	3	.075	40	2'386
Head pitch	"AKpit"	3	.075	40	197'197
Facial emotion	"AKem"	9	.225	40	5'384
Triggered emotion	"AKtrem"	3	.075	40	43'187
Observed by # people	"AKseen"	7	.175	40	2'942
Combinations of features:					
Centrality	"ACent"	1'156'201	.569	2'031'314	13'414'075
Face size	"ASize"	1'109'596	.546	2'031'314	13'023'993
Use of Face Mask	"AKmsk"	251'437	.118	2'031'314	3'433'893
Blur level	"AKblr"	305'424	.15	2'031'314	3'686'366
Light exposure	"AKexp"	269'820	.133	2'031'314	3'540'878
Head yaw	"AKyaw"	283'479	.14	2'031'314	3'366'994
Head pitch	"AKpit"	272'911	.134	2'031'314	3'555'286
Facial emotion	"AKem"	314'368	.155	2'031'314	3'559'047
Triggered emotion	"AKtrem"	112'330	.0553	2'031'314	777'954
Observed by # people	"AKseen"	20'340	.01	2'031'314	113'546

Notes: The table provides summary statistics for the "Adjectives" syntax class. The upper panel refers to single features only (excluding combinations). The bottom panel provides refers to features combinations. All statistics encompass only tokens that appear at least once in analysed images. Column 1 lists the subclasses in the "Adjectives" class, marked in the data using the labels in Column 2. Column 3 indicates the number of distinct tokens for each subclass (denoting the different values a subclass can take). Column 4 indicates the percentage of tokens in the subclass out of all "Adjective" tokens, whose total is listed in Column 5. Column 6 lists the total occurrence of subclass tokens across images.

TABLE (A.1.4)
VOCABULARY SUMMARY STATISTICS: "CONTEXT" FEATURES AND FEATURE-COMBINATIONS INVOLVING "CONTEXT"

"Context" subclasses:	Label in syntax index	N distinct tokens in subclass	Share of total "Context" tokens	Total "Context" tokens	Total presence in pictures
Single features:					
Tags and tags combinations	"CNtagmix"	479'327	.971	493'849	3'141'234
Context elements	"CNtxt"	14'522	.0294	493'849	81'924
Combinations of features:					
Tags and tags combinations	"CNtagmix"	1'632'796	.95	1'717'958	8'714'560
Context elements	"CNtxt"	1'319'057	.768	1'717'958	8'366'586

Notes: The table provides summary statistics for the "Context" syntax class. The upper panel refers to single features only (excluding combinations). The bottom panel refers to features combinations. All statistics encompass only tokens that appear at least once in analysed images. Column 1 lists the subclasses in the "Context" class, marked in the data using the labels in Column 2. Column 3 indicates the number of distinct tokens for each subclass (denoting the different values a subclass can take). Column 4 indicates the percentage of tokens in the subclass out of all "Context" tokens, whose total is listed in Column 5. Column 6 lists the total occurrence of subclass tokens across images.

A.1.2 Visual tokens in the Vocabulary

This section summarizes the visual vocabulary structure, listing individual features (not combinations) contained in syntax classes and subclasses (categories in bold, token names in italic).

- Class S: **Subject** identity traits (Features constant across portrayals of the same individual, but varying across different individuals).
  - **Name (SN)** (es. *"Kate Blanchett"*, *"Johnny Cash & Kate Blanchett"* etc.: a token per each known single person or couple ever portrayed or joinly portrayed);
  - **Celebrity Status (SC)** (token indicating that a person is a celebrity);
  - **Sex (SG)** (*"Male"* or *"Female"*);
  - **Party bin (SP)** (9 tokens for politicians' partisanship decile; data by Poole and Rosenthal, 1985);
  - **Face Salience Rank (SR)** (i.e. *"Person 1", "Person 2",..., "Person 10"*) Face unique identifier within a picture. The number is a rank based on a person's relative salience within the image, measured as a weighted average of face size and centrality (weighting respectively 70% and 30%); rank=1 indicates the most salient person in the picture.
- Class A: **Adjectives**, modality of a person's representation (*Features that vary across individuals and across representations of a given individual*).
  - **Face Size (AS)** (*"Large"*: face area share F $>$ 1/6 image; *"Medium"*: F $\in [1/6, 1/24]$; *"Small"*: F $< 1/24$);
  - **Face Centrality (AC)** (*"Very High"*: centrality C $\in [.95, 1]$; *"M-High"*: $C \in [.85, .95)$; *"M-Low"*: $C \in [.75, .85)$; *"Very Low"*: $C < .75$);
  - **Facial emotion (AKem)** (*"Anger"; "Contempt"; "Disgust"; "Fear"; "Happiness"; "Sadness"; "Surprise"*);

  - **Emotion triggered in portrayed observers (AKtrem)** (Mean emotion: *"Positive"* if happiness; *"Neutral"* if no prominent emotion among observers; *"Negative"* if Anger, Contempt, Disgust, Fear, or Sadness);
  - **Head pitch (AKpit)** (*"Negative"*: pitch P $< -15°$; *"Neutral"*: P $\in [-15°, +15°]$; *"Positive"*: $P > 15°$);
  - **Head yaw (AKyaw)** (*"Right profile"*: yaw Y $< -30°$; *"Frontal"*: Y $\in [-30°, +30°]$; *"Left profile"*: $Y > 30°$);
  - **Mask (AKmsk)** (Indicator for person wearing a mask);
  - **Face blur level (AKblr)** (*"High", "Medium", "Low"*);
  - **Face light exposure (AKexp)** (*"Overexposed"*- bright, *"Regular exposure"*, *"Underexposed"*- dark);
  - **Number of observers (AKseen)** (Indicator for people observing a person summing to *1-9*);

- <u>Class C</u> (describing **Context** attributes):

  - **Within class C, General image descriptors focused on persons** (**CNtxt**):

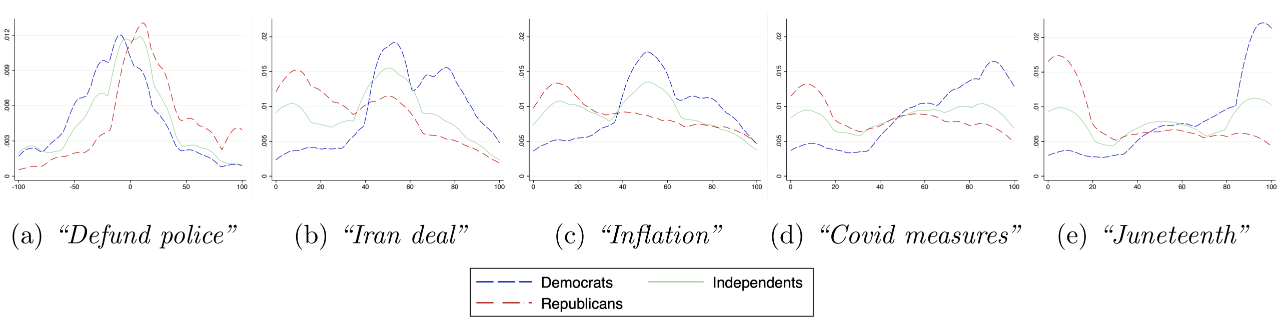
    * **Presence of persons, celebrities and congresspeople** (indicators for number of persons, 0 to 10; *Celebrity*: presence of at least one well-known person; *Congresspeople*: presence of at least one congress member; *People but no celebrity*: presence of people but no celebrities; *Celebrity but no congressperson*: presence of celebrities but no congresspeople);
    * **Triplets of names** (es. *Donald Trump & Kate Blanchett & Johnny Cash*: a token per each triplet of well-known persons ever portrayed jointly);
    * **Men and women representation patterns** (Tokens indicating the presence of: *men*; *women*; *men only*; *women only*; *majority of men*; *majority of women*; *same number of women and men*);
    * **Republicans, Democrats, and Independents representation patterns** (Tokens indicating the presence of:*Democrats*; *Republicans*; *Independents*; *Dems only*; *Reps only*; *Majority of Dems*; *Majority of Reps*; *N. Dems = N. Reps*);
    * **Mask wearing patterns** (indicators for image portraying wearing a mask:*At least one person*, *All the persons*; *Majority of people*; *Minority of people*);
    * **Facial emotion patterns** (indicators for average emotion: *Positive*; *Neutral*; *negative*);
    * **Image shot angle** (indicators for shot angle: *From above, From below, From front*; it is derived from average camera angle of person's face portrayals);

  - **Within class C, General image descriptors from image tags, and tags mix** (**CNtagmix**):

    * **Tags for Person** (tags characterizing persons in the picture, e.g. "*Policeman*");
    * **Tags for Animals** (indicators for animals or animal groups, e.g. "*Flock*");
    * **Tags for Things** (tags indicating the presence of specific objects, e.g. "*Knife*");

∗ **Tags for Verbs** (indicators for actions a subject is executing, e.g. "*Walking*");
∗ **Tag for Qualifiers** (tags for characteristic of subjects/objects, e.g. "*Yellow*");
∗ **Tag for Bodyparts** (tags for visible body parts of humans/animals, e.g. "*Waist*");
∗ **Tag for Places** (tags for the place where the picture is taken, e.g. a "*Station*")
∗ **Tag: Background** (Single background tags, couple and triplets; Background tags indicate items in the image background, e.g. "*Tower*");
∗ **Tag: Setting** (single setting tags, and in combination with background tags; setting tags refer to the pictures setting, situation or theme, e.g. "*Volleyball*");
∗ **Tag mix: Thing(wearable)** ×2 **or** 3
∗ **Tag mix: Person + Thing(wearable) or Thing(wearable)** ×2
∗ **Tag mix: Verb(drinking) + Thing(beverage)**
∗ **Tag mix: Person + Verb(transitive; with object:animals) + Animal**
∗ **Tag mix: Person + Verb(transitive; with object:animals) + Animal**
∗ **Tag mix: Person + Verb(transitive; with object:things) + Thing**
∗ **Tag mix: Person + Verb(intransitive; with subject: person)**
∗ **Tag mix: Animal + Verb(with subject: animal)**
∗ **Tag mix: Thing + Qualifier(color)**
∗ **Tag mix: Thing + Qualifier(material)**
∗ **Tag mix: Person + Verb(with subject: person; with object: transports) + Thing(transportation mean)**
∗ **Tag mix: Thing** ×2 **or** 3
∗ **Tag mix: Place + Background + Background + Background**
∗ **Tag mix: Background** ×2 **or** 3
∗ **Tag mix: Setting** ×2 **or** 3
∗ **Tag mix: Setting(event) + Place**

A.2 Appendix 2: Survey Experiment



(a) *"Defund police"* (b) *"Iran deal"* (c) *"Inflation"* (d) *"Covid measures"* (e) *"Juneteenth"*

FIGURE (A.2.1)
Densities of opinions at Baseline on News Issues, by Respondents' Affiliation

Notes: The Figure displays the densities of opinions on the five news issues at baseline (before treatment), dividing respondents by party affiliation. Opinion modes suggest the ideological distance between Democrats and Republicans in the sample is smaller in the "Police funds" issue and wider for the "Juneteenth" issue.

A.2.1 News Issues, Leading text, and Leading images

This Section illustrates the choice of treatments' news issues, texts, and pictures. Topics As illustrated in Section *II.C*.4, significant visual partisanship characterizes US news in Politics, Covid & Health, Economy, Security, and Society. To identify news issues pertaining each of these topics, I rely on the list of relevant news issues drafted by `allsides.com`, which tags issues by their topic. This website compares news issues from sources with opposite slants, and it periodically publishes "Headline Roundups" (syntheses of the main news issues within a given period) to highlight the different takes of the Democrat-leaning and Republican-leaning news sources.[43] I use these roundups to identify valid issues within each news topic (as listed in the following paragraphs) and to ensure the coherence between treatments' framing and actual media coverage. Based on the roundups, for each news issue I draft a headline, a byline, and a leading text coherent with the neutral tone of non-partisan coverage (i.e. that of news sources rated as "Centre", neither Democrat- nor Republican- leaning, on `allsides.com`).

I select three treatment leading images (Dem-leaning, Rep-leaning and neutral) for each issue. A fundamental aspect to consider in the choice of images is that the partisanship scores of tokens determined through the method in Section II pertained to a period preceding that of the experiment (Dec 2019- Dec 2020, vs. July 2021). Some notable events intervened in between, such as the switch from a Republican to a Democrat US presidency. This implies that the partisan news outlets' narratives (including the visual one) in use during the experiment could potentially differ from the ones tested in Section II.[44] I therefore adopt a two-steps process to select experimental pictures that can be both "partisan" according to the method introduced in Section II and aligned with the partisan visual narratives in place at the time of the experiment (2021). I first identify a set of leading pictures actually published in news pieces on the same issue and rated as "strongly Democratic","Strongly Republican", or "center" on allsides.com. Then, I select the ones whose visual features have "partisan loading" in the event-topic vocabularies.

[43]Roundups are available at: `https://www.allsides.com/story/admin`.

[44]For instance, Republican outlets previously framing the governments' economic policy through positive affect could later have adopted the opposite stance (to criticize the new government), and the same for Democrat-leaning outlets.

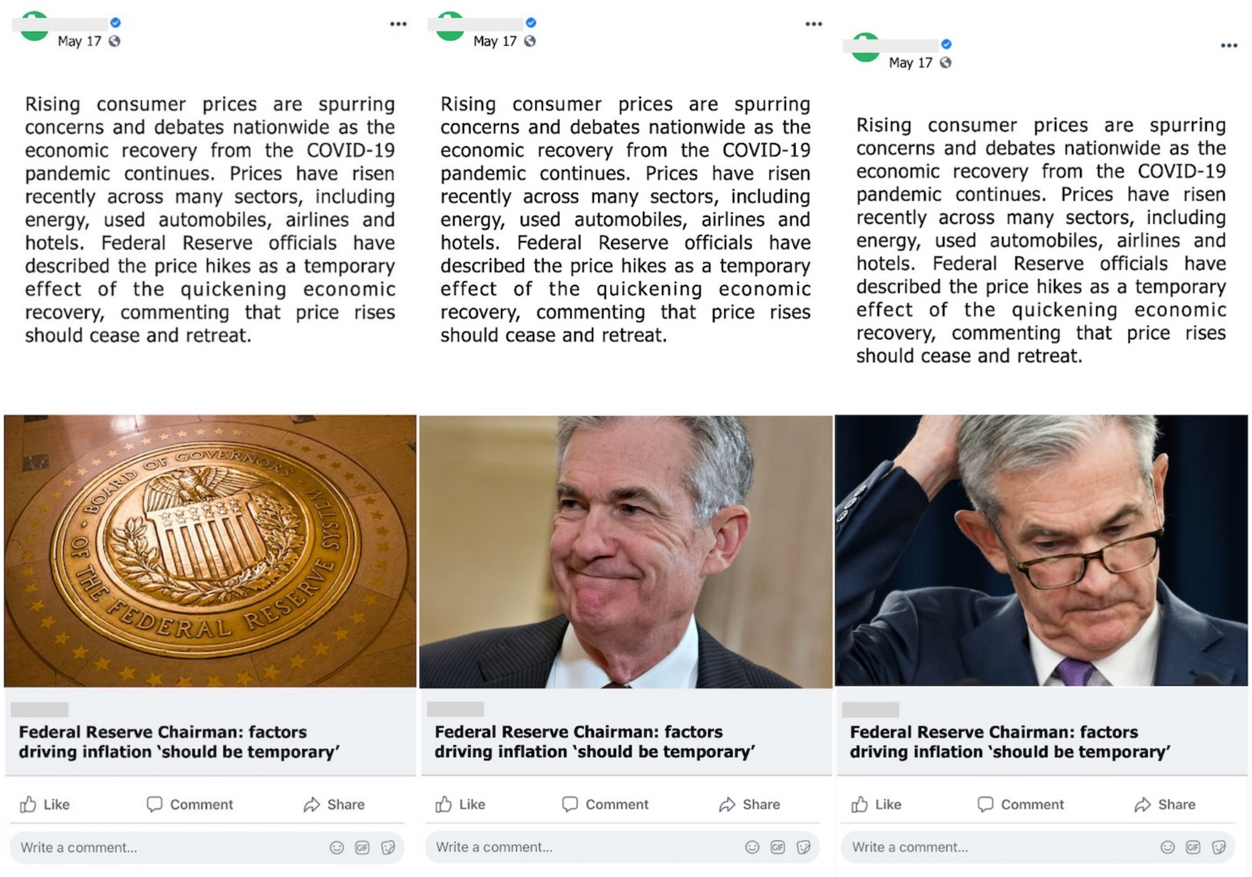
A.2.1.1 Topic: ECONOMY.

Issue: FED's forecasts on inflation.

Headlines Roundup:[45]

> *"The Federal Reserve maintains that current inflation will only be temporary, a stance that President Joe Biden and other prominent Democrats have echoed while advocating for spending packages they say will better the lives of average Americans. Right-rated voices have covered inflation fears more prominently, with some accusing Democrats of dismissing inflation fears while supporting harmful economic policy. Left- and center-rated voices have been less accusatory, often exploring the likelihood of inflation worsening and financially-sustainable legislation being agreed upon in Congress."*

Treatments:



(a) Leading image: *Neutral* (b) Leading image: *Dem-leaning* (c) Leading image: *Rep-leaning*

FIGURE (A.2.2)
Treatments for topic "Economy"

Notes: The Figure shows the treatments (news previews) for "Inflation" issue, in the "Economy" news topic.

[45] From Allsides.com's Headline Roundup "*The Politicization of Inflation*", available at: `https://www.allsides.com/story/perspectives-politicization-inflation`

A.2.1.2 Topic: COVID & HEALTH.

Issue: The effectiveness of anti-Covid measures.

Headlines Roundup:[46]

> *"Opinions range far and wide on the Trump administration's response to the COVID-19 outbreak. Many voices, particularly on the left, criticized the U.S. and White House responses. Others, especially on the right, tended to focus more on China's response to the virus as being worthy of stricter scrutiny. Some minimized the role that the administration was playing, focusing instead on other key actors and decisions."*

Treatments:

(a) Leading image: *Neutral* (b) Leading image: *Dem-leaning* (c) Leading image: *Rep-leaning*

FIGURE (A.2.3)
Treatments for topic "Covid & Health"

Notes: The Figure shows the treatments (news previews) for "Covid Management" issue, in the "Covid & Health" news topic.

---

[46] From Allsides.com's Headline Roundup "*Trump and the Politics of Coronavirus*", available at: `https://www.allsides.com/story/opinions-trump-and-politics-coronavirus`

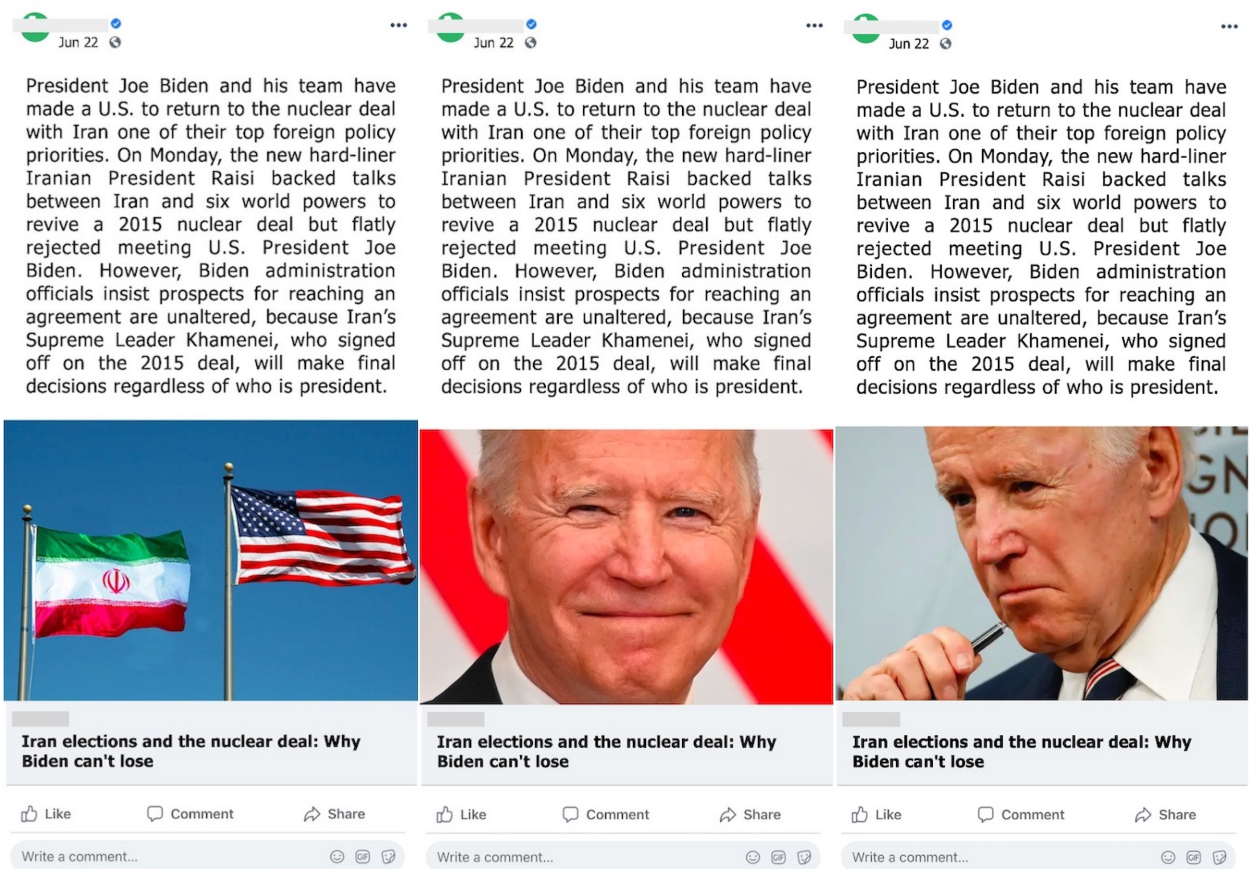
A.2.1.3 Topic: POLITICS.

Issue: Renewal of the US-Iran nuclear deal.

Headlines Roundup:[47]

> *"U.S. President Joe Biden is intent on restoring the 2015 nuclear agreement with Iran [...]. According to sources close to European and U.S. negotiators,* [his chief negotiator] *Malley is expected to offer Tehran a Goldilocks-style deal: just enough sanctions relief so Iran will return to the pact but not so much that it would leave Biden vulnerable to attacks from hard-liners at home "*

Treatments:



(a) Leading image: *Neutral* (b) Leading image: *Dem-leaning* (c) Leading image: *Rep-leaning*

FIGURE (A.2.4)
Treatments for "Politics" topic.

Notes: The Figure shows the treatments (news previews) for "Iran deal" issue, in the "Politics" news topic.

[47] From Allsides.com's "*U.S. Mounts All-Out Effort to Save Iran Nuclear Deal*", available at: `https://www.allsides.com/news/2021-04-15-1349/us-mounts-all-out-effort-save-iran-nuclear-deal`

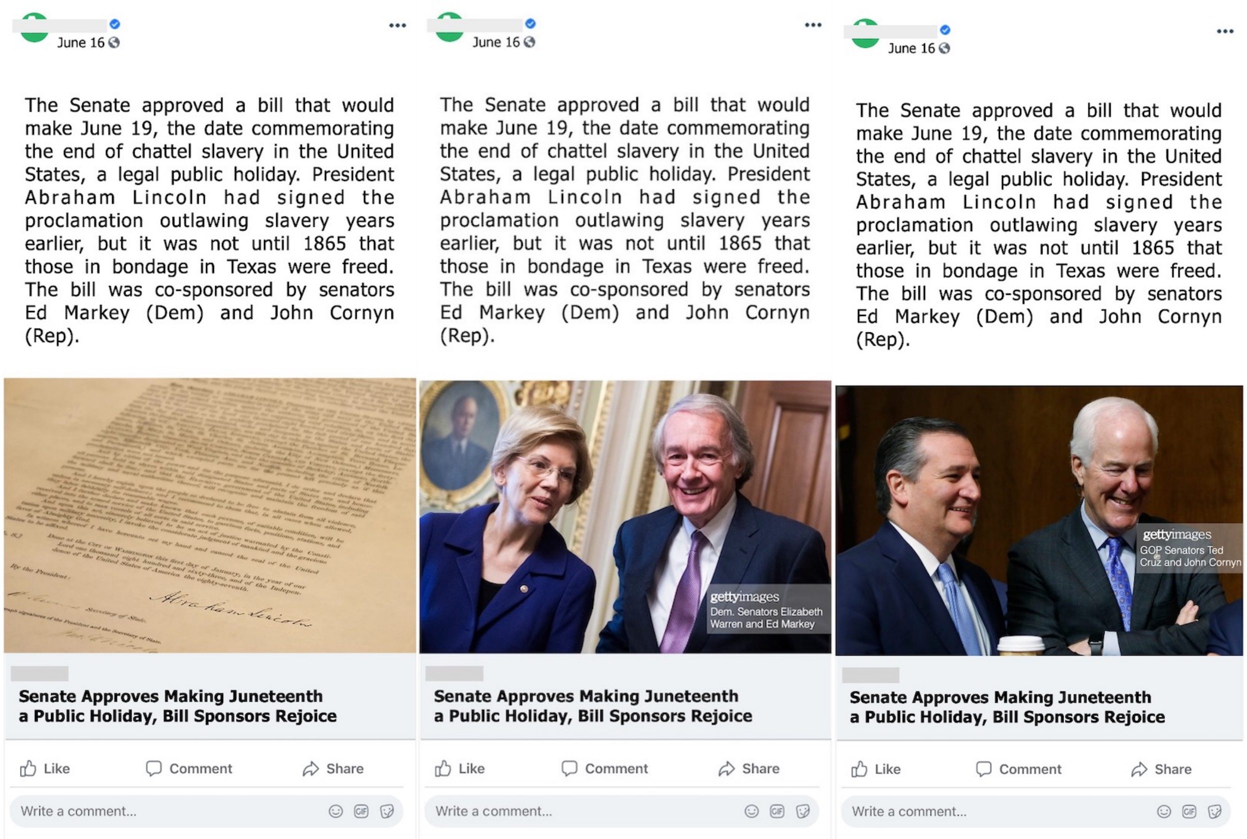
A.2.1.4 Topic: SOCIETY.

Issue: Juneteenth becomes a Federal holiday.

Headlines Roundup:[48]

> *"Most of the opinions about Juneteenth this year were framed around the day becoming an official holiday. Opinions were more common from left- and center-rated outlets. Many left-rated voices celebrated the decision; many also called it a "hollow victory" and grouped it with other "symbolic gestures that are presented as progress without any accompanying economic or structural change." Some right-rated voices criticized that narrative and its proponents, arguing that "there is no concession or show of good faith that will ever placate their ever-increasing litany of demands."*

Treatments:



(a) Leading image: *Neutral* (b) Leading image: *Dem-leaning* (c) Leading image: *Rep-leaning*

FIGURE (A.2.5)
Treatments for "Society" topic.

Notes: The Figure shows the treatments (news previews) for "Juneteenth" issue, in the "Society" news topic.

[48] From Allsides.com's "*Juneteenth 2021*", available at: `https://www.allsides.com/story/perspectives-juneteenth-2021`

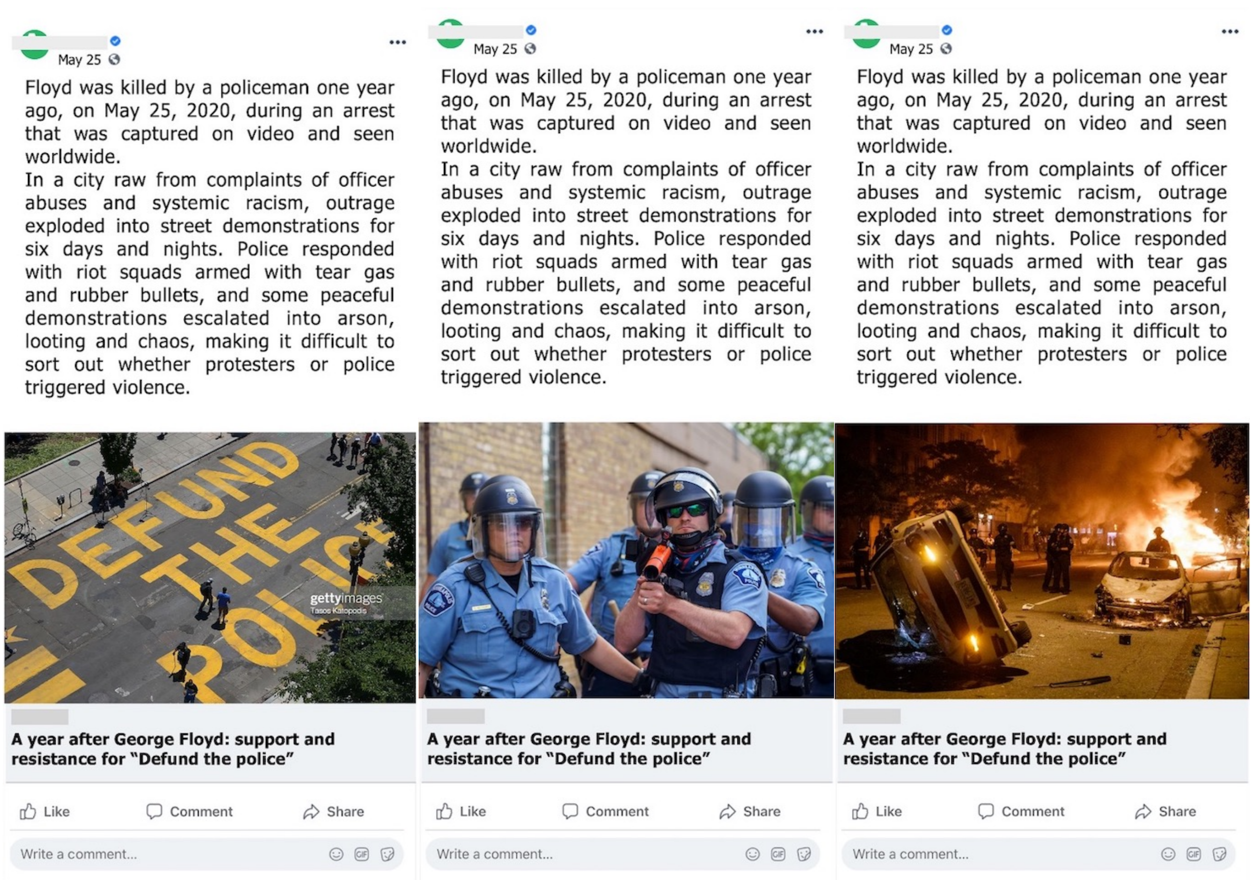
A.2.1.5 Topic: SECURITY.

Issue: Police budget cuts.

Headlines Roundup:[49]

> *"Some left-rated voices advocated for addressing systemic issues and reforming communities by reallocating significant funds from law enforcement to housing and education budgets. Several also called for an end to mass incarceration, police militarization, and police in schools. Some voices from the right argued that police systems should remain intact, pointing to possible correlations between cities with progressive law enforcement policies and rising crime rates. Many voices from all sides of the spectrum advocated for some form of police reform or reduced funding."*

Treatments:



(a) Leading image: *Neutral* (b) Leading image: *Dem-leaning* (c) Leading image: *Rep-leaning*

FIGURE (A.2.6)
Treatments for "Security" topic.

Notes: The Figure shows the treatments (news previews) for "Police budget" issue, in the "Security" topic.

[49] From Allsides.com's "*Defunding the Police*", available at: `https://www.allsides.com/story/perspectives-defunding-police`

TABLE (A.2.1)
Survey Experiment Summary Statistics

		Mean	Sd	Min	Max
Age bracket:	– 18-34	.2273185	.4192057	0	1
	– 35-44	.2565524	.4368403	0	1
	– 45-54	.2525202	.4345675	0	1
	– 55-65	.2636089	.4407007	0	1
Ethnicity:	– Caucasic	.8089718	.3932103	0	1
	– African-American	.0927419	.2901436	0	1
	– Latin American	.0645161	.245732	0	1
	– Asiatic	.0579637	.2337337	0	1
	– Native American	.015625	.1240509	0	1
Schooling < 8 yrs.		.0095766	.097415	0	1
Party affiliation:	– Democrat	.3886089	.487557	0	1
	– Independent	.3069556	.4613471	0	1
	– Republican	.3044355	.4602839	0	1
Politics interest:	– Very low	.0922379	.2894344	0	1
	– Low	.1673387	.3733721	0	1
	– Medium	.3447581	.4754091	0	1
	– High	.2620968	.4398859	0	1
	–Very high	.1335685	.3402739	0	1
Political opinion	(Liberal/Conservative)	4.048387	1.723639	1	7
Gets news from:	– Fox News	1.144153	1.151808	0	3
	– CNN	1.316028	1.143477	0	3
	– Breitbart	.3513105	.7434829	0	3
	– NYT	1.012097	1.058479	0	3
	– MSNBC	1.020665	1.041294	0	3
	– NYPost	.7923387	.9480834	0	3
Main info. source:	– Newspapers	.1789315	.3833915	0	1
	– Radio	.0453629	.2081513	0	1
	– Socials	.1355847	.3424333	0	1
	– TV	.5146169	.4999123	0	1
Clicks in introduction		1.628024	1.339399	1	28
Low screen resolution		.2011089	.4009303	0	1
Defund Police	(baseline opinion)	.1334203	44.36863	-100	100
”	(Post treatment opinion)	1.272364	44.13409	-100	100
Iran deal	(baseline opinion)	-3.100806	28.06095	-50	50
”	(Post treatment opinion)	-1.830141	27.76713	-50	50
Inflation	(baseline opinion)	-2.071069	28.96654	-50	50
”	(Post treatment opinion)	-.4188508	28.7	-50	50
Covid measures	(baseline opinion)	4.081653	33.6061	-50	50
”	(Post treatment opinion)	6.431452	32.87022	-50	50
Juneteenth	(baseline opinion)	5.163306	37.67628	-50	50
”	(Post treatment opinion)	7.272177	37.56104	-50	50
Defund police issue has low salience		.2520161	.4342799	0	1
Iran deal issue has low salience		.2535282	.4351403	0	1
Inflation issue has low salience		.2510081	.4337025	0	1
Covid measures issue has low salience		.2530242	.4348543	0	1
Juneteenth issue has low salience		.2681452	.4431053	0	1
Familiarity with *Defund police* issue		2.389113	.7378551	0	3
Familiarity with *Inflation* issue		1.537802	.9781178	0	3
Familiarity with *Iran deal* issue		1.307964	1.000462	0	3
Familiarity with *Covid measures* issue		2.160786	.8902761	0	3
Familiarity with *Juneteenth* issue		1.995968	.8743466	0	3
Observations	1984				

TABLE (A.2.2)
Impact of Leading Images On News-Readers' Opinion
(Only control: baseline opinion on issue)

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Opinion on "Defund Police"	Opinion on "Iran deal"	Opinion on "Inflation"	Opinion on "Covid measures"	Opinion on "Juneteenth"
	(Budget cut in -100 +100)	(Confidence, in -50+50)	(Confidence, in -50+50)	(Dissatisfaction, in -50+50)	(Policy support, in -50+50)
Neutral images (N)	-0.517	-0.541	0.903	-7.475	8.740
	(0.672)	(0.769)	(0.486)	(0.317)	(0.653)
Democrat images (D-N)	1.379	-0.257	-0.494	0.866	-0.737
	(0.465)	(1.246)	(1.359)	(0.314)	(1.109)
	[0.059]	[0.850]	[0.740]	[0.070]	[0.554]
Republican images (R-N)	-1.516	-1.832	-2.239	-0.159	-0.426
	(0.519)	(1.180)	(0.804)	(0.631)	(0.448)
	[0.06]	[0.22]	[0.07]	[0.82]	[0.41]
Democrat-Republican (D-R)	2.895	1.575	1.745	1.025	-0.311
	(0.422)	(0.845)	(1.014)	(0.799)	(1.461)
	[0.006]	[0.159]	[0.184]	[0.290]	[0.845]
Observations	1574	1608	1625	1595	1551
"Baseline opinion" control:	Y	Y	Y	Y	Y
Other controls:	N	N	N	N	N

Notes: The Table presents OLS estimates of the effect of the Democrat-leaning (D), neutral (N), and Republican-leaning (R) news-leading images on respondents' opinion after exposure to the news (column headers indicate the relevant news issue). The dependent variable for the "Defund Police" issue ranges in [-100,+100], while all others range in [-50+50]. Variables are adjusted so that the higest value in the range always corresponds to Democrats' ideological position (hence positive coefficients indicate a pro-Democratic opinion shift, and vice versa). The specifications only control for the baseline opinion expressed on the issue before treatment exposure, and no other covariates. Round parentheses contain robust standard errors; square brackets contain the p-values for two-sided tests of equality between coefficients (tested pairs indicated on the left) using robust standard errors.

TABLE (A.2.3)
Impact of Leading Images by Readers' Political Party Affiliation

	(1) Defund Police	(2) Iran deal	(3) Inflation	(4) Covid measures	(5) Juneteenth
Dependent variable: Post-treatment opinion on topic					
Democrats x Dem-leaning images (D)	3.880	0.984	2.466	-0.341	1.242
	(1.172)	(1.936)	(1.735)	(2.134)	(2.878)
Democrats x neutral images (N)	2.197	1.939	2.236	-2.073	0.680
	(1.360)	(2.006)	(0.962)	(0.386)	(0.537)
Democrats x Rep-leaning images (R)	2.449	-0.555	-0.103	-2.390	1.116
	(1.113)	(1.621)	(1.106)	(1.349)	(0.419)
Independents x Dem-leaning images (D)	0.632	-0.923	-1.865	-0.184	-1.017
	(1.107)	(2.505)	(1.055)	(1.031)	(0.229)
Independents x Rep-leaning images (R)	-0.835	-3.764	-1.626	-1.895	-0.781
	(1.493)	(1.913)	(0.880)	(0.769)	(1.042)
Republicans x Dem-leaning images (D)	1.249	-3.456	-1.685	-4.874	-1.983
	(2.303)	(1.964)	(0.673)	(0.897)	(0.848)
Republicans x neutral images (N)	-1.279	-4.721	-1.277	-5.555	0.022
	(2.621)	(2.115)	(0.842)	(0.835)	(1.018)
Republicans x Rep-leaning images (R)	-5.956	-4.939	-5.571	-5.577	-1.155
	(1.369)	(1.976)	(1.956)	(1.454)	(1.602)
H0 for equality tests:	P value:	P value:	P value:	P value:	P value:
Dem*(D) - Rep*(R) $\leq$ Dem*(R) - Rep*(D):	0.002	0.030	0.027	0.086	0.564
Dem*(R) = Rep*(D):	0.424	0.091	0.118	0.275	0.010
Dem*(D) - Rep*(R) $\leq$ Dem*(N) - Rep*(N):	0.027	0.613	0.062	0.290	0.315
Observations	1574	1608	1625	1595	1551
Treatment-independent controls	Y	Y	Y	Y	Y

Notes: The Table presents OLS estimates of the effect of the Democrat-leaning (D), neutral (N) and Republican-leaning (R) news-leading images. The dependent variable is respondents' opinion after exposure to the news (column headers indicate the relevant news issue). Treatments are interacted with indicators of the respondent's political affiliation (Democratic, Independent, or Republican), which is measured before treatment. The dependent variable for the "Defund Police" issue ranges between -100 and 100, while all others range in -50+50. Variables are adjusted so that the highest value in the range always corresponds to the Democrats' ideological position (hence the largest of any two coefficients indicates a relatively more pro-Democratic opinion, and vice versa). Treatment-independent controls are the same as in the main specification (here excluding controls for political opinion and party preference). The panel below the regression coefficients reports the P-values for one-sided and two-sided tests of equality between coefficients (null hypotheses are indicated on the left) using robust standard errors. Heteroskedasticity-robust standard errors in parentheses.

A.3 Online Appendix

A.2.1 On measuring visual partisanship: gaze regions in pictures

This section describes the details of the methods used to determine the gaze regions of the subjects in a picture. I borrow this approach from studies on Intelligent Vehicle Systems, in which a driver's head pose is used to predict the attention patterns to the road. See, among others, Parks, Borji, and Itti, 2015; Lee et al., 2018; Dari, Kadrileev, and Hullermeier, 2020; Jha and Busso, 2020).

Given two subjects in a picture, A and B, I determine subject A's "gaze region", and measure whether B falls in that gaze region; if so, then I consider B as seen by A. I use this measure to construct the triggered emotionality measure described in the main text. The raw data from Microsoft's API include the measurement of the following head poses: pitch (ie. whether the chin is up or down), yaw (i.e. the horizontal rotation of the head, towards the left or the right), and roll (i.e. the head's inclination to the sides, namely bringing the ear closer to the shoulder). First, I determine an area of the picture that is "compatible" with an individual's gaze region, approximating this region through the information on the head's yaw and pitch. The accurate determination of a subject's gaze region in a 2-dimensional picture presents two main challenges. First, the head's position is expressed in degrees (yaw, pitch, roll), and the conversion of an angle to a length requires knowledge of the distance between the viewer and an object. In fact, the sight region flattened in a 2-dimensional space appears as a triangle whose base (i.e. the side most distant from the viewer) is proportional to the triangle's "height" (namely, the distance between viewer and object). This implies that for a given angle of a visual region, its section is wider the furthest is the observer. Actual distances between subjects in a picture can hardly be measured[50] and are thus often approximated. Another problem originates in the fact that the sight angle γ between A and B can result from multiple combinations of A's yaw and pitch, as we ignore the distance between subject and cannot exactly determine the relative contributions of a head's pitch and yaw in producing γ. To illustrate, imagine a viewer in the center of the picture and consider the picture's bottom-left corner: such point could be visible both if the person had yaw$= -90$ (i.e. her head was completely turned to the right) and pitch < 0 (i.e. looking downward), and if the person had pitch=0 (gaze at own eyes' level) and head

[50] This would requiring information on focal lengths and the presence of a know object whose dimension is known (e.g. a 1 euro coin).

turned more toward the camera (e.g. yaw 45). In particular, the more distant the person from the camera, the closer to 0 could be her head's yaw while maintaining sight of the point at the bottom-left corner of the picture. The ambiguity is once again due to the lack of knowledge of distances between subjects, and the flattening of the scene on a 2-dimensional surface.

To work around the difficulty, I determine each person's "plausible" sight region using rather ample criteria, and then imposing further requirements to increase the precision. First, I consider a margin to the left and to the right of the head's yaw. Now, the eyes' main focus region is 30 degrees to each side, but 30 degrees is much less than the actual natural sight region as we also have 30 more degrees of near-peripheral area. Objects in this area would be more comfortably seen by turning the head more, however pictures often capture moments in which the individuals are reacting quickly to a visual stimulus, to which eyes naturally respond before head movements. Therefore, I take an intermediate length between the focus region and the near-peripheral region, and consider a margin of 45 degrees to each side of the yaw. Then, I consider the sign of the head's pitch, to pin down in which direction (upper or lower) to orient the area determined by the yaw.
For every observer (A) and other subject (B) in the picture, I consider B as falling within A's sight region if both of the folowing conditions are verified:

1. The angle γ generated by the line connecting A and B falls within a range around A's yaw equal to $3 * \sqrt{|yaw|}$.
2. The vertical distance between A and B (i.e the distance in coordinates $y_b - y_a$) and A's pitch have the same sign. Formally, the product of the two shall be non-negative: this indicates that A's head vertical inclination (upwards or downwards) is in B's direction. Given that for sufficiently small vertical distances or for pitches close to 0 the product may happen to be negative even if B is visible to A, I include a tolerance level considering as 0 values between -15 and +15 for both vertical distance and pitch, so to obtain a vertical vision span of 30 degrees. I also set vertical distance to 0 any distance between -0.9 and +0.9 between Y_a and Y_b.

If two or more persons fall within A's gaze region, I consider A to be looking only at the person that is closest to her. In this sense, since in images with at least three individuals about 94% of the persons have head yaw between -45 and 45 degrees (indicating a relatively frontal head pose), I rank subjects in a person's gaze region considering first image depth (namely distance from the camera), then breaking potential ties using horizontal distances (to the left and to the right of the viewer). I establish the relative distance of subjects from the camera using the faces' dimensions, considering two subjects with the same face size as equally distant, and allowing for a 5% tolerance in face area differences. I then exclude from a person's gaze region all the subjects who are behind her (and hence cannot be in sight). Finally, I exclude all subjects from the gaze regions of persons whose eyes are occluded (either covered or closed).

Having so approximated the focus of the persons' gaze (i.e. what they "see"), I compute the triggered emotion of observed individuals as the weighted average of their observers' emotions. The weights are proportional to the depth-distance of the observer: as stated in the previous section, I assume the picture to confer more visibility to the subjects whose features are meant to matter more.

The triggered emotionality measure rests on the assumption that glances can be used to transfer the observer's emotion to the observed person, thus that a person's facial expression is informative of the emotional evaluation of what she sees. The method is clearly limited in cases such as when individuals glance away from an emotionally triggering sight (instance plausibly more frequent with negative emotions). Nevertheless, it allows to go beyond the mere emotion-labelling of single faces, and to capture the deeper emotional loading of images with multiple individuals. To limit the method's possible flaws, I only measure triggered emotions in images with up to 3 persons: this safeguards the accuracy of the method (the more people are portrayed, the higher number of possible glance-interactions and emotion attributions), while at the same time includes the vast majority of pictures in my sample.

A.4 Experiment: Balance of observable characteristics across treatment groups

TABLE (A.2.1)
Balance of observable characteristics across treatment branches, "Defund police" news issue

		Republican		Neutral		Democrat		Normalized difference:		
Variables:		Mean	St. err.	Mean	St. err.	Mean	St. err.	(R-N)	(N-D)	(D-R)
Age bracket:	– 18-34	0.006	(0.019)	-0.009	(0.018)	0.003	(0.018)	0.037	-0.028	-0.009
	– 35-44	0.007	(0.020)	0.015	(0.020)	-0.021	(0.018)	-0.019	0.084	-0.064
	– 45-54	-0.003	(0.019)	-0.010	(0.019)	0.013	(0.019)	0.014	-0.051	0.037
	– 55-65	-0.010	(0.019)	0.003	(0.020)	0.006	(0.019)	-0.029	-0.006	0.035
Ethnicity:	– Caucasic	0.011	(0.017)	-0.032	(0.018)	0.021	(0.016)	0.107	-0.132	0.025
	– African-American	-0.008	(0.013)	0.028	(0.014)	-0.019	(0.012)	-0.117	0.157	-0.041
	– Latin American	0.005	(0.011)	0.007	(0.011)	-0.013	(0.010)	-0.008	0.084	-0.075
	– Asiatic	0.001	(0.011)	-0.002	(0.010)	0.001	(0.010)	0.015	-0.012	-0.003
	– Native American	-0.006	(0.004)	0.010	(0.007)	-0.004	(0.004)	-0.123	0.107	0.017
Schooling < 8 yrs.		-0.003	(0.004)	-0.001	(0.004)	0.004	(0.005)	-0.018	-0.050	0.068
Party affiliation:	– Democrat	-0.007	(0.022)	0.054	(0.022)	-0.047	(0.021)	-0.125	0.209	-0.084
	– Independent	0.006	(0.020)	-0.048	(0.019)	0.042	(0.021)	0.119	-0.196	0.077
	– Republican	0.001	(0.020)	-0.007	(0.020)	0.005	(0.020)	0.017	-0.026	0.009
Politics interest:	–Very low	0.017	(0.014)	-0.009	(0.012)	-0.008	(0.012)	0.087	-0.002	-0.085
	–Low	0.011	(0.017)	-0.002	(0.017)	-0.008	(0.016)	0.032	0.018	-0.050
	–Medium	-0.013	(0.021)	-0.004	(0.021)	0.016	(0.021)	-0.019	-0.043	0.062
	–High	-0.018	(0.019)	0.017	(0.020)	0.001	(0.019)	-0.078	0.036	0.042
	–Very high	0.003	(0.015)	-0.002	(0.015)	-0.001	(0.014)	0.015	-0.005	-0.011
Conservative-Liberal score		-0.075	(0.076)	-0.066	(0.075)	0.137	(0.074)	-0.005	-0.119	0.124
Gets news from:	–Fox News	-0.063	(0.050)	0.005	(0.052)	0.055	(0.050)	-0.060	-0.043	0.104
	–CNN	0.062	(0.050)	-0.004	(0.051)	-0.056	(0.050)	0.058	0.045	-0.104
	–Breitbart	0.032	(0.032)	-0.029	(0.031)	-0.002	(0.031)	0.085	-0.039	-0.047
	–NYT	0.060	(0.046)	-0.015	(0.046)	-0.043	(0.046)	0.072	0.026	-0.099
	–MSNBC	0.033	(0.046)	-0.014	(0.045)	-0.018	(0.045)	0.045	0.003	-0.048
	–NYPost	0.015	(0.041)	-0.016	(0.041)	0.001	(0.041)	0.033	-0.019	-0.014
Main info. source:	–Newspapers	0.020	(0.018)	-0.006	(0.017)	-0.014	(0.016)	0.067	0.023	-0.090
	–Radio	0.006	(0.010)	-0.014	(0.007)	0.008	(0.010)	0.105	-0.113	0.008
	–Socials	0.014	(0.016)	-0.024	(0.014)	0.010	(0.015)	0.113	-0.101	-0.012
	–TV	-0.058	(0.022)	0.038	(0.022)	0.019	(0.022)	-0.192	0.038	0.154
Clicks in introduction		-0.033	(0.056)	0.013	(0.072)	0.019	(0.055)	-0.031	-0.004	0.041
Low screen resolution		-0.005	(0.017)	0.019	(0.018)	-0.013	(0.017)	-0.059	0.080	-0.020
Topic of low subjective salience		-0.015	(0.019)	-0.017	(0.018)	0.030	(0.019)	0.005	-0.108	0.103
Topic familiarity:	– Low	-0.007	(0.005)	0.004	(0.007)	0.002	(0.006)	-0.081	0.015	0.066
	– Mid-Low	0.004	(0.013)	0.013	(0.013)	-0.017	(0.011)	-0.032	0.108	-0.077
	– Mid-High	-0.002	(0.021)	-0.013	(0.021)	0.015	(0.021)	0.023	-0.058	0.035
	– High	0.005	(0.022)	-0.005	(0.022)	-0.000	(0.022)	0.019	-0.009	-0.010
Topic baseline opinion		0.904	(2.004)	0.662	(1.908)	-1.517	(1.785)	0.005	0.051	-0.056
N of observations:		510		523		532				

Notes: The table presents the means and standard errors for each covariate specified, and the standardized difference between treatment groups for the "Defund Police" news issue to assess balance. Treatment branches are marked in column headers, with "Republican" ("Democrat") indicating being exposed to news on the issue lead by Republican-leaning (Democrat-leaning) images, and "Neutral" indicating non-partisan leading images.

TABLE (A.2.2)
Balance of observable characteristics across treatment branches, "Iran deal" news issue

Variables:		Republican Mean	Republican St. err.	Neutral Mean	Neutral St. err.	Democrat Mean	Democrat St. err.	Normalized difference: (R-N)	(N-D)	(D-R)
Age bracket:	– 18-34	0.010	(0.018)	0.003	(0.018)	-0.013	(0.017)	0.016	0.039	-0.055
	– 35-44	0.004	(0.019)	0.001	(0.019)	-0.005	(0.019)	0.006	0.014	-0.020
	– 45-54	-0.008	(0.019)	0.012	(0.019)	-0.004	(0.019)	-0.045	0.035	0.010
	– 55-65	-0.006	(0.019)	-0.016	(0.019)	0.022	(0.020)	0.023	-0.083	0.060
Ethnicity:	– Caucasic	-0.020	(0.018)	0.004	(0.017)	0.016	(0.016)	-0.058	-0.033	0.091
	– African-American	-0.013	(0.012)	0.008	(0.013)	0.005	(0.013)	-0.070	0.008	0.062
	– Latin American	0.016	(0.012)	-0.005	(0.010)	-0.011	(0.010)	0.083	0.030	-0.112
	– Asiatic	0.017	(0.012)	-0.017	(0.009)	-0.001	(0.010)	0.143	-0.071	-0.072
	– Native American	-0.003	(0.005)	0.001	(0.006)	0.003	(0.006)	-0.030	-0.016	0.046
Schooling $<$ 8 yrs.		-0.009	(0.002)	0.001	(0.005)	0.008	(0.006)	-0.116	-0.063	0.168
Party affiliation:	– Democrat	-0.003	(0.021)	0.003	(0.021)	0.000	(0.021)	-0.012	0.005	0.008
	– Independent	-0.013	(0.020)	-0.003	(0.020)	0.016	(0.020)	-0.022	-0.041	0.063
	– Republican	0.016	(0.020)	0.000	(0.020)	-0.017	(0.020)	0.035	0.037	-0.072
Politics interest:	– Very low	0.008	(0.013)	0.000	(0.012)	-0.008	(0.012)	0.027	0.029	-0.056
	– Low	-0.031	(0.015)	0.025	(0.017)	0.006	(0.017)	-0.149	0.047	0.102
	– Medium	0.037	(0.021)	-0.039	(0.020)	0.003	(0.021)	0.160	-0.089	-0.071
	– High	0.015	(0.019)	-0.006	(0.019)	-0.009	(0.019)	0.049	0.005	-0.054
	– Very high	-0.029	(0.013)	0.021	(0.015)	0.008	(0.015)	-0.152	0.038	0.114
Conservative-Liberal score		0.115	(0.074)	0.015	(0.074)	-0.131	(0.074)	0.058	0.086	-0.144
Gets news from:	– Fox News	0.002	(0.050)	-0.006	(0.050)	0.005	(0.051)	0.007	-0.010	0.003
	– CNN	-0.019	(0.050)	-0.003	(0.049)	0.022	(0.049)	-0.014	-0.022	0.036
	– Breitbart	-0.006	(0.031)	0.036	(0.032)	-0.030	(0.029)	-0.057	0.092	-0.034
	– NYT	-0.031	(0.045)	-0.014	(0.045)	0.045	(0.046)	-0.016	-0.056	0.071
	– MSNBC	-0.017	(0.045)	-0.006	(0.046)	0.023	(0.044)	-0.011	-0.028	0.039
	– NYPost	-0.029	(0.039)	-0.009	(0.040)	0.039	(0.041)	-0.021	-0.051	0.073
Main info. source:	– Newspapers	-0.002	(0.017)	0.013	(0.017)	-0.011	(0.016)	-0.037	0.062	-0.026
	– Radio	-0.002	(0.009)	-0.000	(0.009)	0.002	(0.009)	-0.008	-0.012	0.020
	– Socials	0.002	(0.015)	-0.011	(0.014)	0.009	(0.015)	0.040	-0.060	0.020
	– TV	0.010	(0.022)	-0.023	(0.022)	0.013	(0.022)	0.067	-0.074	0.006
Clicks in introduction		-0.101	(0.048)	0.085	(0.074)	0.016	(0.058)	-0.130	0.046	0.095
Low screen resolution		0.005	(0.017)	-0.005	(0.017)	-0.000	(0.017)	0.025	-0.011	-0.014
Topic of low subjective salience		-0.016	(0.018)	-0.009	(0.018)	0.026	(0.019)	-0.016	-0.081	0.096
Topic familiarity:	– Low	0.009	(0.019)	-0.013	(0.019)	0.004	(0.019)	0.049	-0.038	-0.011
	– Mid-Low	-0.011	(0.020)	0.021	(0.021)	-0.010	(0.020)	-0.069	0.067	0.002
	– Mid-High	0.014	(0.020)	-0.026	(0.019)	0.012	(0.020)	0.090	-0.083	-0.006
	– High	-0.012	(0.014)	0.017	(0.015)	-0.005	(0.014)	-0.088	0.067	0.021
Topic baseline opinion		0.338	(1.189)	-0.202	(1.203)	-0.137	(1.192)	0.020	-0.002	-0.017
N of observations:		534		536		529				

Notes: The table presents the means and standard errors for each covariate specified, and the standardized difference between treatment groups for the "Iran deal" news issue to assess balance. Treatment branches are marked in column headers, with "Republican" ("Democrat") indicating being exposed to news on the issue lead by Republican-leaning (Democrat-leaning) images, and "Neutral" indicating non-partisan leading images.

TABLE (A.2.3)
Balance of observable characteristics across treatment branches, “Inflation” news issue

		Republican		Neutral		Democrat		Normalized difference:		
Variables:		Mean	St. err.	Mean	St. err.	Mean	St. err.	(R-N)	(N-D)	(D-R)
Age bracket:	– 18-34	-0.025	(0.017)	0.011	(0.018)	0.015	(0.018)	-0.089	-0.011	0.099
	– 35-44	-0.003	(0.019)	-0.004	(0.019)	0.007	(0.019)	0.001	-0.024	0.023
	– 45-54	0.021	(0.019)	-0.005	(0.019)	-0.017	(0.018)	0.059	0.028	-0.088
	– 55-65	0.007	(0.019)	-0.002	(0.019)	-0.005	(0.019)	0.021	0.006	-0.027
Ethnicity:	– Caucasic	-0.019	(0.017)	-0.003	(0.017)	0.022	(0.016)	-0.040	-0.063	0.103
	– African-American	0.016	(0.013)	-0.007	(0.012)	-0.010	(0.012)	0.078	0.010	-0.088
	– Latin American	0.013	(0.011)	0.001	(0.010)	-0.014	(0.009)	0.048	0.063	-0.111
	– Asiatic	0.007	(0.011)	0.006	(0.011)	-0.012	(0.009)	0.004	0.079	-0.083
	– Native American	0.005	(0.006)	-0.001	(0.005)	-0.004	(0.004)	0.043	0.034	-0.077
Schooling < 8 yrs.		0.004	(0.005)	-0.005	(0.003)	0.001	(0.005)	0.091	-0.062	-0.030
Party affiliation:	– Democrat	0.013	(0.021)	-0.013	(0.021)	0.000	(0.021)	0.054	-0.028	-0.026
	– Independent	-0.010	(0.020)	0.025	(0.020)	-0.015	(0.020)	-0.076	0.085	-0.009
	– Republican	-0.003	(0.020)	-0.012	(0.020)	0.014	(0.020)	0.020	-0.056	0.036
Politics interest:	–Very low	-0.009	(0.012)	0.009	(0.013)	0.000	(0.012)	-0.062	0.028	0.034
	–Low	-0.007	(0.016)	0.001	(0.016)	0.006	(0.016)	-0.021	-0.015	0.036
	–Medium	-0.015	(0.020)	0.041	(0.021)	-0.026	(0.020)	-0.117	0.140	-0.023
	–High	0.020	(0.019)	-0.038	(0.018)	0.018	(0.019)	0.133	-0.127	-0.006
	–Very high	0.011	(0.015)	-0.013	(0.014)	0.002	(0.015)	0.071	-0.043	-0.028
Conservative-Liberal score		0.001	(0.074)	-0.010	(0.073)	0.009	(0.074)	0.006	-0.011	0.005
Gets news from:	–Fox News	-0.010	(0.050)	-0.015	(0.049)	0.025	(0.050)	0.005	-0.036	0.030
	–CNN	0.014	(0.049)	-0.025	(0.049)	0.011	(0.050)	0.034	-0.031	-0.003
	–Breitbart	0.015	(0.032)	-0.018	(0.029)	0.003	(0.030)	0.046	-0.030	-0.017
	–NYT	0.011	(0.046)	-0.028	(0.046)	0.017	(0.045)	0.037	-0.043	0.006
	–MSNBC	0.047	(0.046)	-0.084	(0.044)	0.037	(0.045)	0.125	-0.118	-0.009
	–NYPost	0.002	(0.041)	-0.014	(0.040)	0.012	(0.040)	0.017	-0.028	0.010
Main info. source:	–Newspapers	0.013	(0.017)	0.006	(0.017)	-0.020	(0.016)	0.017	0.069	-0.086
	–Radio	0.007	(0.009)	-0.001	(0.009)	-0.006	(0.008)	0.041	0.026	-0.067
	–Socials	0.003	(0.015)	0.009	(0.015)	-0.012	(0.014)	-0.016	0.061	-0.045
	–TV	0.007	(0.021)	-0.039	(0.022)	0.032	(0.022)	0.090	-0.143	0.052
Clicks in introduction		-0.076	(0.049)	0.080	(0.075)	-0.004	(0.054)	-0.107	0.056	0.060
Low screen resolution		0.013	(0.017)	-0.010	(0.017)	-0.003	(0.017)	0.059	-0.020	-0.039
Topic of low subjective salience		0.009	(0.019)	0.014	(0.019)	-0.023	(0.018)	-0.012	0.086	-0.074
Topic familiarity:	– Low	-0.035	(0.015)	0.043	(0.018)	-0.009	(0.016)	-0.204	0.132	0.072
	– Mid-Low	0.018	(0.020)	-0.030	(0.019)	0.011	(0.020)	0.107	-0.090	-0.016
	– Mid-High	0.002	(0.021)	0.001	(0.021)	-0.002	(0.021)	0.002	0.007	-0.008
	– High	0.014	(0.016)	-0.015	(0.015)	0.000	(0.016)	0.079	-0.041	-0.038
Topic baseline opinion		-0.400	(1.230)	-1.787	(1.234)	2.217	(1.241)	0.048	-0.140	0.091
N of observations:		545		538		532				

Notes: The table presents the means and standard errors for each covariate specified, and the standardized difference between treatment groups for the “Inflation” news issue to assess balance. Treatment branches are marked in column headers, with “Republican” (“Democrat”) indicating being exposed to news on the issue lead by Republican-leaning (Democrat-leaning) images, and “Neutral” indicating non-partisan leading images.

TABLE (A.2.4)
Balance of observable characteristics across treatment branches, "Covid measures" news issue

Variables:		Republican Mean	Republican St. err.	Neutral Mean	Neutral St. err.	Democrat Mean	Democrat St. err.	Normalized difference: (R-N)	(N-D)	(D-R)
Age bracket:	– 18-34	0.013	(0.018)	0.000	(0.018)	-0.013	(0.017)	0.030	0.031	-0.062
	– 35-44	0.026	(0.020)	-0.032	(0.018)	0.006	(0.019)	0.132	-0.088	-0.044
	– 45-54	-0.011	(0.018)	0.017	(0.019)	-0.005	(0.019)	-0.065	0.049	0.016
	– 55-65	-0.027	(0.019)	0.015	(0.020)	0.012	(0.020)	-0.094	0.008	0.086
Ethnicity:	– Caucasic	-0.010	(0.017)	-0.006	(0.017)	0.016	(0.016)	-0.009	-0.056	0.065
	– African-American	0.007	(0.013)	0.005	(0.013)	-0.012	(0.012)	0.006	0.061	-0.066
	– Latin American	-0.014	(0.009)	0.016	(0.012)	-0.001	(0.010)	-0.126	0.068	0.059
	– Asiatic	0.002	(0.010)	0.003	(0.011)	-0.004	(0.010)	-0.005	0.030	-0.025
	– Native American	0.002	(0.006)	-0.003	(0.005)	0.000	(0.006)	0.043	-0.028	-0.015
Schooling $<$ 8 yrs.		-0.009	(0.002)	0.003	(0.005)	0.006	(0.006)	-0.133	-0.028	0.156
Party affiliation:	– Democrat	0.005	(0.021)	-0.018	(0.021)	0.013	(0.021)	0.046	-0.062	0.016
	– Independent	0.022	(0.020)	-0.006	(0.020)	-0.017	(0.020)	0.060	0.024	-0.084
	– Republican	-0.027	(0.019)	0.023	(0.021)	0.004	(0.020)	-0.109	0.041	0.068
Politics interest:	–Very low	0.021	(0.014)	-0.006	(0.012)	-0.015	(0.011)	0.089	0.035	-0.124
	–Low	-0.021	(0.015)	0.015	(0.017)	0.007	(0.016)	-0.097	0.021	0.075
	–Medium	0.006	(0.021)	-0.011	(0.021)	0.005	(0.021)	0.037	-0.034	-0.003
	–High	-0.008	(0.019)	-0.007	(0.019)	0.015	(0.019)	-0.004	-0.049	0.053
	–Very high	0.002	(0.015)	0.009	(0.015)	-0.011	(0.014)	-0.019	0.060	-0.040
Conservative-Liberal score		0.001	(0.074)	0.014	(0.075)	-0.014	(0.073)	-0.008	0.017	-0.009
Gets news from:	–Fox News	0.004	(0.050)	-0.051	(0.050)	0.045	(0.050)	0.048	-0.084	0.036
	–CNN	0.036	(0.049)	-0.050	(0.050)	0.013	(0.050)	0.076	-0.054	-0.021
	–Breitbart	0.013	(0.032)	-0.007	(0.031)	-0.007	(0.029)	0.028	0.000	-0.028
	–NYT	-0.001	(0.045)	-0.019	(0.047)	0.019	(0.046)	0.017	-0.036	0.020
	–MSNBC	0.019	(0.045)	-0.007	(0.047)	-0.011	(0.045)	0.025	0.004	-0.029
	–NYPost	0.028	(0.041)	-0.017	(0.040)	-0.011	(0.040)	0.048	-0.007	-0.042
Main info. source:	–Newspapers	-0.009	(0.016)	0.012	(0.017)	-0.002	(0.017)	-0.054	0.036	0.018
	–Radio	0.002	(0.009)	-0.001	(0.009)	-0.002	(0.008)	0.015	0.005	-0.020
	–Socials	0.006	(0.015)	0.003	(0.015)	-0.009	(0.014)	0.008	0.038	-0.045
	–TV	0.003	(0.022)	0.001	(0.022)	-0.004	(0.022)	0.005	0.011	-0.015
Clicks in introduction		-0.034	(0.051)	-0.030	(0.055)	0.063	(0.077)	-0.003	-0.061	0.065
Low screen resolution		-0.023	(0.016)	0.013	(0.018)	0.010	(0.017)	-0.093	0.008	0.086
Topic of low subjective salience		0.023	(0.019)	-0.009	(0.018)	-0.014	(0.018)	0.073	0.013	-0.086
Topic familiarity:	– Low	0.014	(0.011)	-0.006	(0.010)	-0.009	(0.009)	0.083	0.014	-0.097
	– Mid-Low	-0.005	(0.014)	0.000	(0.015)	0.004	(0.015)	-0.014	-0.013	0.026
	– Mid-High	0.001	(0.021)	-0.018	(0.021)	0.017	(0.021)	0.040	-0.071	0.032
	– High	-0.011	(0.022)	0.023	(0.022)	-0.012	(0.021)	-0.069	0.072	-0.003
Topic baseline opinion		0.945	(1.481)	-1.146	(1.478)	0.176	(1.447)	0.062	-0.039	-0.023
N of observations:		531		520		533				

Notes: The table presents the means and standard errors for each covariate specified, and the standardized difference between treatment groups for the "Covid measures" news issue to assess balance. Treatment branches are marked in column headers, with "Republican" ("Democrat") indicating being exposed to news on the issue lead by Republican-leaning (Democrat-leaning) images, and "Neutral" indicating non-partisan leading images.

TABLE (A.2.5)
Balance of observable characteristics across treatment branches, "Juneteenth" news issue

		Republican		Neutral		Democrat		Normalized difference:		
Variables:		Mean	St. err.	Mean	St. err.	Mean	St. err.	(R-N)	(N-D)	(D-R)
Age bracket:	– 18-34	-0.005	(0.018)	-0.011	(0.018)	0.015	(0.019)	0.016	-0.064	0.048
	– 35-44	0.019	(0.020)	-0.001	(0.020)	-0.018	(0.019)	0.045	0.041	-0.085
	– 45-54	-0.018	(0.018)	0.016	(0.020)	0.002	(0.019)	-0.079	0.032	0.047
	– 55-65	0.003	(0.020)	-0.004	(0.020)	0.001	(0.020)	0.017	-0.011	-0.006
Ethnicity:	– Caucasic	0.000	(0.017)	0.012	(0.017)	-0.012	(0.018)	-0.029	0.059	-0.031
	– African-American	-0.008	(0.013)	-0.016	(0.012)	0.023	(0.014)	0.028	-0.128	0.101
	– Latin American	0.001	(0.011)	-0.006	(0.010)	0.005	(0.011)	0.031	-0.048	0.017
	– Asiatic	0.002	(0.010)	-0.003	(0.010)	0.001	(0.010)	0.023	-0.016	-0.007
	– Native American	0.002	(0.006)	-0.001	(0.005)	-0.001	(0.005)	0.026	0.005	-0.031
Schooling < 8 yrs.		0.004	(0.005)	0.001	(0.005)	-0.005	(0.003)	0.029	0.066	-0.094
Party affiliation:	– Democrat	-0.005	(0.021)	0.026	(0.022)	-0.021	(0.021)	-0.063	0.096	-0.032
	– Independent	-0.005	(0.020)	-0.016	(0.020)	0.021	(0.020)	0.025	-0.081	0.057
	– Republican	0.010	(0.020)	-0.010	(0.020)	-0.000	(0.020)	0.044	-0.021	-0.022
Politics interest:	–Very low	-0.009	(0.012)	0.021	(0.014)	-0.011	(0.012)	-0.100	0.106	-0.006
	–Low	-0.026	(0.016)	0.002	(0.017)	0.024	(0.018)	-0.077	-0.056	0.133
	–Medium	-0.012	(0.021)	-0.001	(0.021)	0.013	(0.021)	-0.022	-0.029	0.051
	–High	0.018	(0.020)	-0.011	(0.019)	-0.007	(0.019)	0.068	-0.009	-0.059
	–Very high	0.029	(0.016)	-0.010	(0.015)	-0.019	(0.014)	0.112	0.026	-0.138
Conservative-Liberal score		0.054	(0.077)	-0.129	(0.074)	0.070	(0.074)	0.108	-0.119	0.009
Gets news from:	–Fox News	0.004	(0.052)	-0.096	(0.050)	0.089	(0.050)	0.086	-0.163	0.073
	–CNN	-0.015	(0.051)	0.001	(0.051)	0.013	(0.050)	-0.014	-0.011	0.024
	–Breitbart	0.066	(0.034)	-0.079	(0.027)	0.010	(0.033)	0.206	-0.129	-0.073
	–NYT	0.031	(0.047)	-0.007	(0.048)	-0.024	(0.046)	0.035	0.016	-0.052
	–MSNBC	0.009	(0.046)	-0.015	(0.046)	0.005	(0.045)	0.024	-0.020	-0.004
	–NYPost	0.079	(0.043)	-0.040	(0.040)	-0.041	(0.040)	0.125	0.002	-0.126
Main info. source:	–Newspapers	0.031	(0.018)	0.018	(0.018)	-0.049	(0.014)	0.033	0.185	-0.218
	–Radio	-0.001	(0.009)	0.006	(0.010)	-0.005	(0.008)	-0.038	0.057	-0.019
	–Socials	0.006	(0.016)	-0.007	(0.015)	0.001	(0.015)	0.039	-0.024	-0.015
	–TV	-0.012	(0.022)	-0.010	(0.022)	0.022	(0.022)	-0.005	-0.065	0.069
Clicks in introduction		0.020	(0.063)	-0.059	(0.049)	0.036	(0.071)	0.061	-0.069	0.010
Low screen resolution		0.007	(0.017)	0.007	(0.018)	-0.014	(0.017)	-0.002	0.054	-0.052
Topic of low subjective salience		-0.011	(0.019)	0.004	(0.020)	0.007	(0.020)	-0.034	-0.007	0.042
Topic familiarity:	– Low	-0.008	(0.010)	-0.005	(0.011)	0.013	(0.012)	-0.011	-0.074	0.085
	– Mid-Low	-0.025	(0.017)	0.011	(0.018)	0.014	(0.018)	-0.090	-0.008	0.099
	– Mid-High	0.015	(0.022)	0.006	(0.022)	-0.020	(0.021)	0.018	0.053	-0.071
	– High	0.018	(0.021)	-0.011	(0.021)	-0.008	(0.020)	0.063	-0.008	-0.055
Topic baseline opinion		0.222	(1.646)	1.677	(1.668)	-1.835	(1.632)	-0.039	0.094	-0.055
N of observations:		522		500		520				

Notes: The table presents the means and standard errors for each covariate specified, and the standardized difference between treatment groups for the "Juneteenth" news issue to assess balance. Treatment branches are marked in column headers, with "Republican" ("Democrat") indicating being exposed to news on the issue lead by Republican-leaning (Democrat-leaning) images, and "Neutral" indicating non-partisan leading images.

A.5 Experiment: Heterogeneity by baseline opinion, topic salience and prior knowledge

A.2.0.1 Images and baseline opinion

Do images exert a different effect on readers who previously expressed intermediate or extreme opinions? To respond, I explore the heterogeneity of treatment effects across terciles of respondents' baseline opinion on each issue. Overall, the results suggest that respondents who at baseline held more moderate opinions may be less affected by leading images, compared to respondents who belong to the first/third terciles. There are however no universal patterns arising across issues, and no other notable interactions between baseline opinion and the magnitude of treatment effects.

The following paragraphs discuss the heterogeneity of treatment effect separately for each news issue. Appendix Table A.2.6 reports the estimates from an OLS regression of the respondents' updated opinion on treatments interacted with terciles of baseline opinion distribution.

Police funding. For this issue, respondents who ex ante chose the lowest police budget update their response by further lowering the budget. Within this group, those who were exposed to Dem-leaning images chose an even lower budget than both those exposed to Rep-leaning images ($p = 0.067$) and those exposed to neutral images ($p = 0.015$). Respondents in the intermediate tercile of baseline opinion, who expressed the mildest variations to the Police budget (in either direction), do not exhibit statistically different reactions to treatments. Finally, respondents who ex-ante were choosing the highest Police funding reduce the budget significantly more if exposed to news with Dem-leaning images as opposed to Rep-leaning ones (p-value = 0.036), and even more so if exposed to neutral images as opposed to Rep-leaning ones (p-value = 0.006).

Covid measures. For this issue, Rep-leaning and Dem-leaning images exert statistically different effects only among respondents who ex ante express the lowest judgement on the adequacy of anti-covid measures implemented in March 2020. Among those, people exposed to Republican-leaning images have a significantly more positive opinion on the Government's measures (p-value = 0.034) than individuals exposed to Democrat-leaning images. In the same group of respondents, there is no significant difference between those exposed to neutral images and the others. No differences across treatment branches exist in the middle and higher terciles of baseline opinion.

Iran deal. Respondents who ex ante express the lowest belief in the success of a US-Iran nuclear deal decrease their judgement on the likelihood of success significantly more if exposed to Republican-leaning images than if exposed to either neutral images (p-value = 0.007) or Dem-leaning images (p-value = 0.086). In the intermediate tercile of baseline opinion there is a difference between Rep-leaning and Dem-leaning images (p-value = 0.043), and no other difference across treatment branches. Once again, no differences across treatment branches exist in the higher tercile of baseline opinion.

Inflation. Respondents who ex ante express the highest belief in the regress of inflation by June 2022 exhibit the largest upward opinion update if exposed to neutral images as opposed to Rep-leaning ones (p-value = 0.018). Otherwise, there are no other significant differences across treatment branches ineither of the terciles of baseline opinion.

A.2.0.2 Images and issue salience

Does the effect of images depend on how relevant is the news issue to the individual respondent? I investigate the relationship between issue salience and treatments by interacting the treatment indicators with the distribution terciles of perceived issue salience. Overall, respondents in the lowest and highest tercile of the perceived issue salience appear mildly more susceptible to the effect of leading images, relative to the respondents in the intermediate tercile. However, the evidence is inconclusive as to whether issue salience is a strong predictor of respondent' sensibility to the images leading the news.

Appendix Table A.2.8 reports the estimates from an OLS regression of the respondents' updated opinion on these interaction terms. The following paragraphs discuss the heterogeneity of treatment effect separately for each news issue.

Police funding. Respondents who are in the lowest tercile for perceived relevance of the police funding issue update their response by lowering the budget comparatively more if exposed to Dem-leaning images than to neutral images, albeit the difference is only modestly significant (p-value = 0.099). Respondents who perceive the issue as more relevant to them (i.e. those in the highest tercile of perceived issue salience) decrease the desired police budget by comparatively more if exposed to Dem-leaning images as opposed to Rep-leaning images (p-value = 0.013). No other significant differences exists across treatment branches in either groups. Similarly, individuals in the intermediate salience tercile do not exhibit significantly different opinion updates across any of the treatment branches.

Covid measures. For this issue, none of the terciles of issue salience display significant differences in the effects across treatment branches.

Iran deal. Rep-leaning images and neutral images have significantly different effects both in the first and in the third salience tercile, with Republican-leaning images producing a relatively lower perceived likelihood of success of a US-Iran nuclear deal (p-values = 0.057 in the lowest salience group, and 0.070 in the highest salience group). No significant effect exist between these two treatment branches in the intermediate tercile; in this group, instead, the effect of Republican-leaning and Democrat-leaning images is significantly different, with the latter eliciting a higher perceived likelihood of success of the deal (p-value = 0.044).

Inflation. For this issue, respondents in the highest salience tercile display a significantly different response to Republican-leaning and neutral images. In fact, the latter induce a relatively higher perceived likelihood of inflation to return to pre-pandemic levels by June 2022 (p-value = 0.076). No other statistically significant differences exist across treatment branches in any of the salience terciles.

A.2.0.3 Images and opinion development

Does the effect of images depend on news readers' stage of opinion development? Does it depend on the knowledge about the issue? To answer these questions I explore whether the effect of images varies between respondents whose prior knowledge and opinion are more vs. less consolidated. Those are directly measured with a question before treatment takes place. While no neat patterns arise, image variation seems to affect highly knowledgeable respondents more often than others (3 news issues displaying significant differences across branches, vs. 1 news issue for least knowledgeable respondents). The evidence on whether prior issue knowledge is a determinant factor is however inconclusive.

Appendix Table A.2.7 reports the estimates from an OLS regression of the respondents' updated opinion interacted with high and low levels of prior knowledge on the news issue. The following paragraphs discuss the heterogeneity of treatment effect separately for each news issue.

Police funding. Respondents who consider themselves not very knowledgeable about the issue do not update their response differently across the treatment branches. Vice versa, respondents who consider themselves highly knowledgeable about the issue update their response by lowering the desired Police budget comparatively more if exposed to Dem-leaning images than to Rep-leaning images (p-value = 0.006), and if exposed to neutral images than to Rep-leaning ones (p-value = 0.073). The effect of any image type never differs between the least and the most knowledgeable respondents.

Covid measures. For this issue, neither the most knowledgeable nor the least knowledgeable respondents' update their response differently across the treatment branches. Moreover, the effect of any image type never differs between the least and the most knowledgeable respondents.

Iran deal. Respondents who consider themselves not very knowledgeable about the issue update their response by increasing the perceived likelihood of success of a US-Iran deal relatively more if exposed to neutral images than to Rep-leaning images (p-value = 0.049). Vice versa, respondents who consider themselves highly aware about the US-Iran deal update their response by increasing the perceived likelihood of success comparatively more if exposed to Dem-leaning images than to Rep-leaning images (p-value = 0.053). No other significant differences exist among treatment coefficients within either knowledge groups. Moreover, the effect of an image never differs between the least and the most knowledgeable respondents.

Inflation. Respondents who consider themselves not very knowledgeable about the issue do not update their response differently across the treatment branches. Vice versa, respondents who consider themselves knowledgeable about the issue update their response by increasing the perceived likelihood of inflation to return to pre-pandemic levels by June 2022 comparatively more if exposed to neutral images than to Rep-leaning images (p-value = 0.054). No other significant differences exist among treatment coefficients within either knowledge groups. Moreover, the effect Democrat-leaning images is mildly different between the least and the most knowledgeable respondents, with a slightly smaller positive effect in the latter group (p= 0.092).

TABLE (A.2.6)
Heterogeneity analysis by tercile of baseline opinion on the issue

	(1) Police funds	(2) Covid measures	(3) Iran deal	(4) Inflation	(5) Juneteenth
Dependent variable: Opinion difference					
Lowest baseline opinion x Dem-leaning images (D)	12.554	2.138	0.105	-2.888	-1.965
	(3.362)	(4.662)	(2.870)	(3.720)	(3.529)
Lowest baseline opinion x neutral images (N)	7.375	0.409	1.544	-3.352	-0.000
	(3.375)	(4.888)	(2.842)	(3.706)	(3.380)
Lowest baseline opinion x Rep-leaning images (R)	7.854	-3.715	-3.153	-3.577	-0.888
	(3.729)	(4.517)	(3.000)	(3.800)	(3.506)
Medium baseline opinion x Dem-leaning images (D)	6.624	-2.498	1.108	0.125	1.232
	(2.437)	(2.622)	(1.885)	(2.395)	(1.813)
Medium baseline opinion x neutral images (N)	7.431	-4.320	-1.083	1.703	0.799
	(2.521)	(2.692)	(1.903)	(2.499)	(1.850)
Medium baseline opinion x Rep-leaning images (R)	4.628	-4.247	-2.021	-0.662	1.207
	(2.562)	(2.616)	(1.946)	(2.499)	(1.713)
Highest baseline opinion x Dem-leaning images (D)	4.582	0.288	-0.794	3.552	0.860
	(2.184)	(2.367)	(1.613)	(2.369)	(0.924)
Highest baseline opinion x neutral images (N)	6.462	-1.969	1.004	5.445	1.481
	(2.366)	(2.250)	(1.519)	(2.293)	(0.952)
Constant	-4.536	-23.977	15.283	20.849	17.378
	(17.908)	(9.630)	(9.085)	(7.958)	(6.435)
Observations	1436	1491	1510	1505	1414
Treatment-independent controls	Y	Y	Y	Y	Y

Notes: The Table presents the OLS estimates for the effect of the Democrat-leaning (D), neutral (N) and Republican-leaning (R) news-leading images interacted with the terciles of respondents' first opinion on the news issue, i.e. that expressed before the treatment exposure. The dependent variable is respondents' opinion after exposure to the news (column headers indicate the relevant news issue). Treatment-independent controls are the same as in the main specification. Heteroskedasticity-robust standard errors are in parentheses.

TABLE (A.2.7)
Heterogeneity analysis by level of self-reported knowledge of the issue

	(1) Police funds	(2) Covid measures	(3) Iran deal	(4) Inflation	(5) Juneteenth
Dependent variable: Opinion difference					
Lowest knowledge x Dem-leaning images (D)	3.573	5.511	1.508	3.636	0.036
	(1.842)	(3.123)	(1.400)	(1.692)	(1.193)
Lowest knowledge x neutral images (N)	3.437	4.474	2.707	2.949	0.844
	(2.118)	(3.369)	(1.403)	(1.727)	(1.148)
Lowest knowledge x Rep-leaning images (R)	1.214	3.114	0.216	1.018	-0.172
	(2.145)	(3.133)	(1.350)	(1.830)	(1.153)
Highest knowledge x Dem-leaning images (D)	5.080	3.030	2.787	0.813	-0.795
	(1.855)	(2.090)	(1.439)	(1.687)	(1.157)
Highest knowledge x neutral images (N)	3.240	0.108	2.239	3.036	-0.279
	(1.803)	(2.122)	(1.472)	(1.574)	(1.363)
Constant	-0.409	-31.033	13.792	13.671	16.434
	(17.627)	(8.207)	(8.704)	(6.948)	(4.977)
Observations	1436	1491	1510	1505	1414
Treatment-independent controls	Y	Y	Y	Y	Y

Notes: The Table presents the OLS estimates for the effect of the Democrat-leaning (D), neutral (N) and Republican-leaning (R) news-leading images interacted with indicators for two levels of (self-reported) knowledge on the issue prior to the news exposure. The dependent variable is respondents' opinion after treatment exposure (column headers indicate the relevant news issue). All dependent variables are adjusted so that the highest value corresponds to the Democrats' ideological position, hence positive coefficients indicate a pro-Democratic opinion shift. Treatment-independent controls are the same as in the main specification. Heteroskedasticity-robust standard errors are in parentheses.

TABLE (A.2.8)
Heterogeneity analysis by level of subjective salience of the issue

	(1) Police funds	(2) Covid measures	(3) Iran deal	(4) Inflation	(5) Juneteenth
Dependent variable: Opinion difference					
Lowest salience x Dem-leaning images (D)	3.707	6.523	-1.574	2.808	-7.221
	(2.449)	(2.437)	(1.772)	(2.303)	(2.066)
Lowest salience x neutral images (N)	4.446	5.313	0.314	2.984	-4.130
	(2.898)	(2.551)	(1.691)	(2.299)	(1.926)
Lowest salience x Rep-leaning images (R)	0.286	3.319	-2.531	0.674	-5.581
	(2.807)	(2.588)	(1.579)	(2.318)	(2.146)
Medium salience x Dem-leaning images (D)	1.781	3.876	3.167	4.352	-2.480
	(2.547)	(2.451)	(1.602)	(2.275)	(1.473)
Medium salience x neutral images (N)	0.784	0.711	0.812	3.862	-1.556
	(2.685)	(2.406)	(1.677)	(2.285)	(1.695)
Medium salience x Rep-leaning images (R)	0.384	0.682	-0.301	2.668	-4.360
	(2.697)	(2.470)	(1.658)	(2.381)	(1.467)
Highest salience x Dem-leaning images (D)	6.260	1.334	1.491	1.248	-0.104
	(2.522)	(2.557)	(1.728)	(2.403)	(1.583)
Highest salience x neutral images (N)	3.655	0.187	3.238	4.142	-2.251
	(2.394)	(2.592)	(1.783)	(2.334)	(1.427)
Constant	-0.179	-30.384	16.298	14.877	20.526
	(17.686)	(8.476)	(8.858)	(6.963)	(5.184)
Observations	1436	1491	1510	1505	1414
Treatment-independent controls	Y	Y	Y	Y	Y

Notes: The Table presents the OLS estimates for the effect of the Democrat-leaning (D), neutral (N) and Republican-leaning (R) news-leading images interacted with indicators for the level of subjective salience assigned by respondents to the news issue (salience is measured before the treatment exposure). The dependent variable is respondents' opinion after exposure to the news (column headers indicate the relevant news issue). All dependent variables are adjusted so that the highest value corresponds to the Democrats' ideological position, hence positive coefficients indicate a pro-Democratic opinion shift. Treatment-independent controls are the same as in the main specification. Heteroskedasticity-robust standard errors are in parentheses.