Why Us and not Them?
A Theory of Political Fact-Checking

Giovanni Andreottola and Charles Louis-Sidois*

March 3, 2025

Abstract

This paper develops a theory of political fact-checking aimed at understanding fact-checkers' motivations from their observed behavior. We first document two empirical stylized facts: fact-checkers scrutinize one side of the political spectrum more (Republicans) and uncover a higher share of false facts from that side. We then develop a model in which two politicians report facts — true or false — to persuade voters. The fact-checker selects one politician and reveals whether their fact is true. We examine several fact-checker motivations, characterizing the probabilities of each politician being fact-checked and the share of false facts detected. Our results show that the empirical patterns are inconsistent with an impartial fact-checker. Instead, these patterns may be explained by a strong bias in favor of Democrats or a weak bias in favor of Republicans. We also characterize the fact-checking strategy maximizing voters' welfare and discuss the welfare implications of each type of fact-checker.

JEL Classification: D72, L82
Keywords: Fact-checking, Media

1 Introduction

Fact-checkers, that is, journalists who evaluate statements made by politicians, have positioned themselves as arbiters of political debates [Amazeen, 2015]. They claim that their

*Andreottola (corresponding author): Department of Economics, Johannes Kepler University (JKU), Linz. Email: giovanni.andreottola@gmail.com. Louis-Sidois: Department of Economics, Vienna University of Economics and Business (WU). Email: charles.louis-sidois@wu.ac.at. We thank Aleksandr Levkun, Andrea Mattozzi, and Lily Ling Yang for helpful comments and suggestions.

goals are to correct misinformation [Barrera et al., 2020], hold politicians accountable [Mattozzi et al., 2022], and ultimately enhance voter welfare. This portrayal, which assumes that fact-checkers are politically neutral, is contested by parts of the public as well as by some media experts, who accuse them of political bias [Brandtzaeg and Følstad, 2017, Uscinski and Butler, 2013, Palumbo, 2023]. The importance of fact-checkers' motivations in the public debate was once again underscored when Meta suspended fact-checking services in January 2025, citing "too much political bias."[1] Yet, fact-checkers' motivations have received limited attention from scholars, who have rather concentrated on the effects of fact-checking. This paper addresses this gap, proceeding in the spirit of revealed preference theory: while fact-checkers' motivations cannot be directly observed, they can be inferred from their published content. We consider the different motivations commonly discussed in the debate and characterize the equilibrium content published under each. Then, we use empirical data on fact-checking content to assess which theoretical predictions align with it.

In particular, our analysis focuses on two key statistics of the fact-checker's activity that are central to debates about their impartiality: the proportion of articles fact-checking each side of the political spectrum and the share of false facts identified for each side [Marietta et al., 2015, Lim, 2018, Louis-Sidois, 2022]. In Section 2, we analyze the content published by PolitiFact from 2009 to 2013, when it was the primary provider of political fact-checking in the United States, and document two empirical stylized facts: first, one side of the spectrum —Republicans— was checked more frequently, and second, this side exhibited a higher share of false facts. There is an open debate about the interpretation of these imbalances. Fact-checkers, as well as some scholars [Bucciol, 2018, Grinberg et al., 2019, Guess et al., 2020, Mosleh et al., 2024], view them as evidence that Republicans disseminate more false information. Alternatively, critics of fact-checking interpret these imbalances as evidence of bias. This debate is summarized by The Daily Standard's writer Mark Hemingway: "You can believe that Republicans lie more than three times as often as Democrats. Or you can believe that, at a minimum, PolitiFact is engaging in a great deal of selection bias." This leads us to focus on three potential motivations for fact-checkers: first, they might be impartial lie-seekers, motivated solely by detecting false claims; second, they may be lie-seekers with a mild bias, preferring to detect false facts from one side; or third, they may be strongly biased, aiming to detect the false claims of one side, and to the contrary, the true claims of the other.

[1] Liv McMahon, Zoe Kleinman, and Courtney Subramanian, "Facebook and Instagram get rid of fact checkers," *BBC News*, January 7, 2025, https://www.bbc.com/news/articles/cly74mpy8klo. Accessed January 13, 2025.

Our game-theoretic approach yields several key insights into the imbalances observed in fact-checkers' content. First, a neutral lie-seeking fact-checker must lead to equal shares of false facts. This outcome relies on the disciplining effect of fact-checking: if a politician was more likely to spread misinformation, the fact-checker would focus on that politician. Increased scrutiny would, in turn, discipline them, implying that both sides must exhibit equal shares of false facts in equilibrium. Moreover, the empirical imbalances may be explained by both a mild and a strong bias. One possibility is a strong bias in favor of Democrats, which aligns with criticisms leveled against fact-checkers [Brandtzaeg and Følstad, 2017, Palumbo, 2023]. However, we also show that the stylized facts are compatible with a mild bias in favor of Republicans. Thus, our model provides new insights into an important public debate, underscoring that fact-checking cannot be understood without considering the interactions between politicians, voters, and fact-checkers.

We introduce our model of political fact-checking in Section 3. We study a setup where two politicians (both referred to with feminine pronouns) report facts to convince voters (they/them). Each politician may have a true fact or not, and one of them is more likely to lack a fact. If a politician lacks a true fact, she can fabricate and report a false one. Voters do not observe whether the reported facts are true. A fact-checker (he/him) examines the fact reported by one politician and reveals to voters whether it is true. Ultimately, voters decide which politician to elect. Voters prefer to elect a politician with a true fact, but their choice is also influenced by a random popularity shock. Politicians aim to be elected and additionally incur a shame cost if the fact-checker reveals that they reported a false fact. This cost represents additional reputation damages and shame for being caught lying.

We assume that at least one politician has a true fact. If no politician had a true fact in her favor on a given issue, it is likely that the issue would remain out of the political debate. Therefore, we expect all issues relevant for fact-checking to involve at least one politician with a true fact to report.[2]

In Section 4, we consider a neutral lie-seeking fact-checker who derives a fixed payoff for identifying a false fact, no matter which politician it comes from, leading him to maximize the probability of checking a false fact. It aligns with the qualitative literature on fact-checkers' motivations [Graves, 2016, 2017] and is supported by fact-checkers' declarations: in a survey conducted by Singer [2021], "Correct misinformation" was ranked as the most important goal by fact-checkers. The first result is that a neutral lie-seeking fact-checker produces balanced conclusions: the shares of false facts uncovered from both politicians must be equal. Intuitively, the fact-checker must be indifferent between checking either politician, which requires both politicians to have an equal likelihood of report-

[2] We explore alternative fact distributions in Appendix A and summarize them in the conclusion.

ing false facts. Otherwise, the fact-checker would concentrate on one side, discouraging that politician from reporting false facts, and ultimately making her statements unworthy of fact-checking. As a result, the fact-checker reaches balanced conclusions even if politicians have different probabilities of lacking a fact. The politician who is less likely to have a fact is disciplined and 'admits' more often (i.e., refrains from fabricating a false fact) when she lacks one, thereby making the reported facts equally likely to be true. As a result, a neutral lie-seeking fact-checker always finds the same share of false facts for both politicians. This conclusion holds even if the probabilities of politicians lacking facts differ, implying that an unequal distribution of facts is not a sufficient explanation for the imbalances in fact-checking content.

Turning to the probabilities of checking each politician, a neutral lie-seeking fact-checker can lead to balanced verification, meaning that politicians are fact-checked with the same probability. This occurs when the shame cost is above a certain threshold, i.e., when politicians are sufficiently responsive to fact-checking. In this situation, when lacking a true fact, both politicians are indifferent between admitting that they do not have a fact and reporting a false one. If a politician reports a false fact, she is either fact-checked, in which case she loses the election and incurs the shame cost, or the other politician is fact-checked. In the latter case, the other politician's fact is necessarily true, which the fact-checker confirms to voters. In turn, voters also update their beliefs about the fact not checked. Since politicians have the same likelihood of reporting a false fact, the probability that the fact not checked is true does not depend on who reported it. This implies that both politicians face the exact same tradeoff and are made jointly indifferent by a balanced verification. However, when the shame cost is small, both politicians cannot be made jointly indifferent because a balanced verification would not deter them from reporting false facts. In this case, the fact-checker checks more frequently the politician who has the higher probability of lacking a fact.

To sum up, this setup allows us to highlight the interactions between players, but it is not compatible with the empirical stylized facts presented in Section 2, in which the share of false facts uncovered is higher for Republicans.

In Section 5, we introduce a mild form of bias by assuming the fact-checker derives a higher payoff from identifying a false fact from one politician, referred to as the valuable target politician. To maintain the fact-checker's indifference condition, the valuable target politician must report false facts less frequently. This ensures that the expected payoffs from checking each politician remain equal. Consequently, when the fact-checker is mildly biased against a politician, that politician becomes more disciplined and exhibits a lower share of false facts. While intuitive, this prediction may seem surprising, as a lower share of false facts is typically associated with positive coverage.

Hence, politicians have different probabilities of stating false facts, which in turn implies that they should be checked with different probabilities. When the shame cost is high, both politicians are indifferent when lacking a fact. As the valuable target politician is less likely to state false facts, she has a higher probability of being elected when her fact is not checked. This gives her a stronger incentive to report false facts than she would have under a neutral lie-seeking fact-checker, which in turn means she must be fact-checked more often. On the other hand, when the shame cost is low, the fact-checker checks more frequently the politician with a higher probability of lacking a fact, so as to make her indifferent, while the other politician always reports a false fact. If the valuable target politician is less likely to lack a fact, an equilibrium may arise in which such politician is checked less often and at the same time exhibits a lower share of false facts, aligning with the stylized facts.

As a result, we obtain a first potential explanation for the empirical stylized facts, which requires a low shame cost, Democrats being the valuable target politicians, and Republicans being more likely to lack a fact. This mild bias against Democrats might be rationalized by the left-leaning tendencies of fact-checking readers [Shin and Thorson, 2017]: PolitiFact may prioritize correcting misinformation that its audience, likely more exposed to Democratic rhetoric, encounters. This finding shows that the imbalances in fact-checking content are not definitive proof of a pro-Democrat bias, as critics often claim [Brandtzaeg and Følstad, 2017, Palumbo, 2023].

Section 6 examines a stronger form of bias. The fact-checker derives a positive 'verification payoff' for establishing that its preferred politician reported a true fact, whereas it derives a positive payoff from uncovering a false fact reported by the other politician. We find that the equilibrium aligns with the empirical stylized facts under certain conditions, hence providing a second potential explanation. First, the fact-checker must favor Democrats, consistent with critiques often raised against PolitiFact. Second, the verification payoff must be sufficiently high to ensure that Democrats, as the preferred politicians, remain disciplined and report fewer false facts than Republicans. Finally, the shame cost must be low, leading to Republicans being fact-checked more frequently.

Section 7 proposes an analysis of voters' welfare, balancing the two channels through which fact-checking influences voters: first, it directly exposes falsehoods, as documented by Thorson [2016] and Barrera et al. [2020]; second, it disciplines politicians, reducing their likelihood of making false statements [Nyhan and Reifler, 2015, Mattozzi et al., 2022]. We establish that the welfare-optimal strategy has a straightforward form: it involves fact-checking the politician more likely to lack a fact with the minimum scrutiny necessary to prevent false statements. This result implies that a neutral lie-seeking fact-checker cannot be welfare-optimal, as it cannot disproportionately focus on one politician. However, both forms of bias mentioned above can achieve welfare optimality, provided the fact-checker is biased against the politician more likely to lack a fact. Deviations from neutrality can therefore enhance voters' welfare, but they require focusing on a politician who does not report false facts, which is incompatible with the empirical stylized facts.

Our analysis focuses on fact-checker's motivations that are most plausible in the context of the public debate: correcting misinformation, as claimed by fact-checkers themselves, and the forms of bias frequently highlighted by critics. Section 8 concludes with a discussion of additional factors that might influence fact-checkers' content. However, a formal analysis of all potential mechanisms explaining the stylized facts lies beyond the scope of this paper.

Some assumptions should be clarified at this stage. First, we focus on a single fact-checker, which we believe is harmless for our purpose: we aim to identify fact-checker's motivations consistent with the empirical stylized facts, relative to a period when there was only one main provider of political fact-checking. While competition plausibly alters the content, it is unlikely to affect fact-checkers' objectives. Second, we assume that fact-checking only occurs when both politicians report a fact. Indeed, fact-checkers primarily arbitrate debated topics. If one side makes an undisputed claim, it is likely to be trivially true, and therefore unworthy of fact-checking. Third, we assume that the fact-checker chooses whom to fact-check but does not control the outcome of the fact-check. This reflects how fact-checkers operate: they follow strict guidelines enforced by external organizations such as the International Fact-Checking Network. Previous studies have shown that fact-checkers agree when assessing the same fact, suggesting limited control over the conclusion [Amazeen, 2015, Louis-Sidois, 2022]. Finally, we rule out the selection of claims with different likelihoods for each side, as well as the selective disclosure of fact-checking outcomes. Though we cannot test whether such manipulations exist, they would violate the guidelines governing fact-checkers.

A few papers investigate how fact-checkers' present their motivations. In a survey, Singer [2021] finds that fact-checkers identify their primary motivation as correcting misinformation, followed by informing citizens, building public trust, and holding powerful figures accountable. This view is supported by Graves et al. [2016], who show that messages emphasizing the ethical values of fact-checking increases this practice. As a result, qualitative studies on fact-checking typically assume that it aims to correct misinformation [Amazeen, 2013, Graves, 2016, 2017, Bigot, 2019]. However, Uscinski and Butler [2013] and Uscinski [2015] argue that fact-checkers make subjective choices, particularly regarding what content to fact-check. Additionally, readers have questioned fact-checkers' impartiality. Brandtzaeg and Følstad [2017] report frequent accusations of left-wing bias in the comment sections of U.S. fact-checkers and Shin and Thorson [2017] find that fact-checking is predominantly shared by Democrats on Twitter. To our knowledge, the only paper that theoretically analyzes fact-checkers' strategies is Levkun [2022], who studies communication between a sender and a receiver with a strategic fact-checker. While he compares equilibria under different fact-checker preferences, his model features a single politician, whereas we focus on determining which politician is fact-checked. Similarly, the theoretical literature on cheap talk with detectable deception focuses on a unique sender [Dziuda and Salas, 2018, Balbuzanov, 2019, Ederer and Min, 2022, Yang, 2023].

Analyses of fact-checkers' content underscore the importance of investigating their motivations. Marietta et al. [2015] demonstrate that U.S. fact-checkers examine varying numbers of claims from each side of several political debates. Lim [2018] finds limited overlap between claims fact-checked by different U.S. fact-checkers. Additionally, Louis-Sidois [2022] finds that fact-checkers share the leaning of the media outlets with which they are affiliated. Regarding consistency, Louis-Sidois [2022] and Amazeen [2016] report that fact-checkers generally agree when evaluating the same claim, but Lim [2018] and Allen et al. [2021] identify instances of disagreement.

A larger strand of the literature focuses on the effects of fact-checking. It enhances readers' factual knowledge [Chan et al., 2017, Nieminen and Rapeli, 2019, Walter et al., 2020], reduces the spread of misinformation [Henry et al., 2022], and disciplines politicians [Amazeen, 2013, Nyhan and Reifler, 2015, Mattozzi et al., 2022]. It also increases demand for news outlets [Chopra et al., 2022] and contributes to the improvement of journalistic quality [Graves et al., 2016, Bigot, 2019]. However, Thorson [2016], Barrera et al. [2020], and Nyhan et al. [2020] find a limited impact on voters' support for politicians.

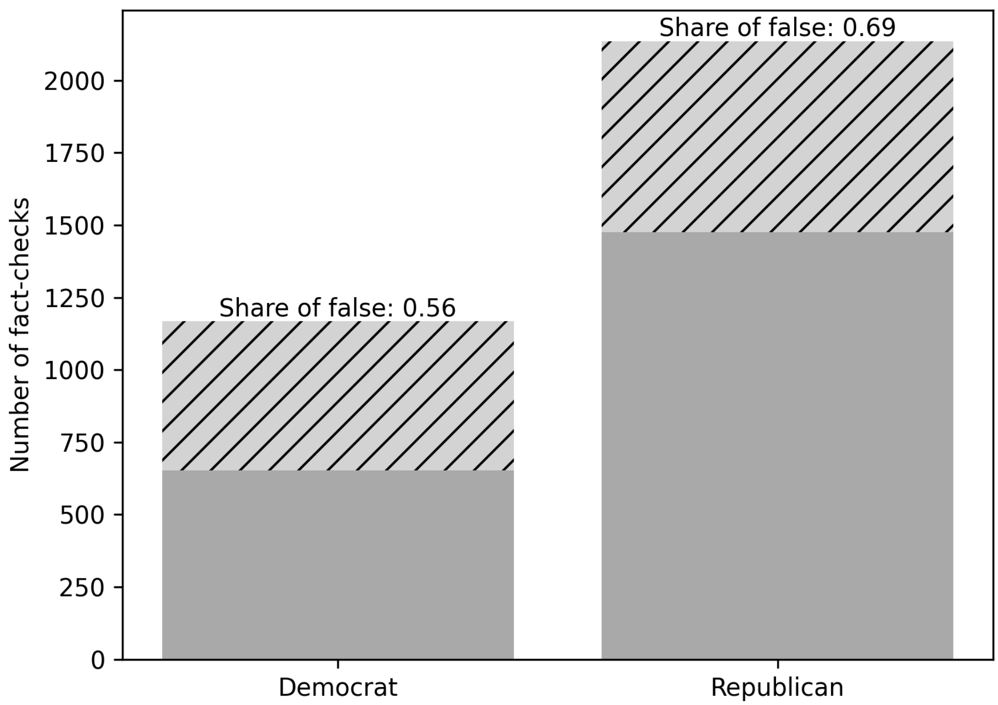
Finally, our paper contributes to the broader literature on the objectives of non-fact-checking media outlets. Profit-motivated media outlets aim to attract readers [Mullainathan and Shleifer, 2005, Gentzkow, 2006, Gentzkow and Shapiro, 2010, Chopra et al., 2023] and satisfy advertisers [Reuter and Zitzewitz, 2006, Di Tella and Franceschelli, 2011, Szeidl and Szucs, 2021, Beattie, 2020, Beattie et al., 2020]. They may also attempt to influence political outcomes, either to promote the interests of their owners [Bagdikian, 2004, DellaVigna and Hermle, 2017, Fize, 2020, Louis-Sidois and Mougin, 2023] or due to their connections with politicians [Besley and Prat, 2006, Ozerturk, 2022]. While these motivations are well-studied, our focus is on the unique aspects of fact-checking.

2 Empirical stylized facts

We examine the content produced by PolitiFact during President Obama's first term, from January 2009 to January 2013. This period is particularly well-suited for documenting the stylized facts, as PolitiFact was the predominant fact-checker in the U.S. at the time, producing three times more articles than other fact-checkers such as FactCheck.org. As a result, the findings from this period are unlikely to be influenced by specialization or competition between different fact-checkers.

We use data from Misra [2022], which identifies the individual who reported the fact assessed in each article published by PolitiFact. We employ OpenAI to determine whether this individual was affiliated with the Democrats or the Republicans. We exclude articles without identifiable political affiliation.

Figure 1: Distribution of fact-checks by PolitiFact, 2009-2013



Note: Number of articles published by PolitiFact on facts reported by Democrats and Republicans from January 2009 until January 2013. Data from Misra [2022]; affiliations assigned by OpenAI. "True": PolitiFact assesses a fact as True or Mostly True. "False": other assessments including pants on fire, false, mostly false, and half-true.

Figure 1 reveals two empirical stylized facts. First, PolitiFact checked a higher number of statements from Republicans (about 2,200) than from Democrats (about 1,200). Second, Republicans exhibited a higher share of false statements (0.69) compared to Democrats (0.56). In other words, the fact-checker scrutinized one side of the political spectrum more frequently and found a greater number of false statements from that side.

We now turn to the model to see which fact-checker's motivations are compatible with these stylized facts.

3 Model

We study a model with two competing politicians, L and R, a fact-checker, and a mass of identical voters. At the beginning of the game, nature decides which politician has a true fact. There are three possible scenarios: with probability $p_l > 0$, only L has a fact (and R lacks a fact); with probability $p_r > 0$, only R has a fact; and with the remaining probability $1 - p_l - p_r > 0$, both L and R have a fact. We assume that at least one politician has a fact, as we believe this is a necessary condition for an issue to be part of the political debate and eligible for fact-checking. The distribution of facts is discussed further in Section 8. Only the politicians observe who has a fact.[3] Politicians then simultaneously decide what to report to the voters. A politician can report a false fact if she does not have a true one. Next, the fact-checker observes the politicians' reports and chooses one report to fact-check. A fact-check reveals to all players whether the fact reported by a politician is true or false. Finally, voters decide which politician to elect.

Regarding payoffs, voters receive 1 if the elected politician has a fact and 0 otherwise. Additionally, an exogenous popularity shock influences the election outcome: voters receive ϵ if politician L is elected. ϵ is drawn immediately before the vote, with $\epsilon \sim U[-1, 1]$. For politicians, the benefit of being elected is normalized to 1, while they incur an additional shame cost of c if the fact-checker reveals they reported a false fact. c is a reduced form for the additional costs associated with being caught fabricating a fact, which includes long-term reputation.

We examine three critical preferences of the fact-checker. In Section 4, the fact-checker maximizes the probability of identifying a false fact. Section 5 introduces a mild bias, where the fact-checker receives different payoffs for detecting false facts from each side. Finally, Section 6 examines a stronger bias, where the fact-checker prefers to check true facts from one side and false facts from the other.

[3] Whether politicians only observe their fact or also observe the other politician's fact does not affect the results.

To sum up, the timing of the game is as follows:

0. Nature chooses who has a fact. With probability p_l (resp. p_r), only the L (resp. R) politician has a fact; with probability $1 - p_l - p_r$ both of them have a fact.
1. Politicians observe the facts and simultaneously report whether they have a fact. A politician lacking a true fact can report a false one.
2. The fact-checker decides which politician to fact-check, which reveals whether the fact is true.
3. The popularity shock ϵ is drawn and voters elect one of the two politicians.

We will characterize the Perfect Bayesian Equilibria (PBE) of this game. To focus on sensible communication strategies, we assume that a politician who has a fact always reports it.[4] As a result, a politician reporting that she lacks a fact is necessarily truthful. For politicians lacking a fact, L (resp. R)'s strategy consists of a probability $f_l \in [0, 1]$ (resp. $f_r \in [0, 1]$) to report a false fact. By construction, fact-checking is irrelevant when one politician reports a fact, as the fact reported is necessarily true. In such cases, we assume that fact-checking does not take place. We characterize the fact-checker's strategy conditional on both politicians reporting a fact. It consists of a probability of checking R, denoted $\sigma_r \in [0, 1]$. L is checked with probability $\sigma_l = 1 - \sigma_r$. To break ties, we assume without loss of generality that L is elected if voters are indifferent between the two politicians.

We briefly analyze a benchmark without fact-checking in Appendix B. In this cheap-talk game, politicians always report having a fact, and voters' beliefs remain fixed at the prior.

4 Neutral lie-seeking fact-checker

We assume that the fact-checker gets a payoff of 1 if he fact-checks a false fact and 0 otherwise. For now, we assume $p_r < p_l$. We discuss $p_r \geq p_l$ at the end of the section. Recalling that politicians report false facts conditional on lacking true ones with probabilities f_l and

[4] As in most signaling games, there may exist equilibria where the meaning of reports is reversed, i.e., a politician reports having no fact when she has a fact and reports having a fact if she does not. Such equilibria are economically counter-intuitive and we do not consider them.

f_r, and focusing on cases where both politicians report a fact,

$$\Pr(L\text{'s fact is false}) = \frac{p_r f_l}{1 - p_l - p_r + p_r f_l + p_l f_r},$$
$$\Pr(R\text{'s fact is false}) = \frac{p_l f_r}{1 - p_l - p_r + p_r f_l + p_l f_r}.$$

Hence, the fact-checker always checks L as long as $\Pr(L\text{'s fact is false}) > \Pr(R\text{'s fact is false})$, and always checks R when the inequality is reversed. For both politicians to be fact-checked with positive probability, it must be that $\Pr(L\text{'s fact is false}) = \Pr(R\text{'s fact is false})$, which holds if

$$p_r f_l = p_l f_r. \tag{1}$$

This condition must be satisfied in equilibrium: to show this by contradiction, suppose there exists an equilibrium in which a politician is always fact-checked. Reporting a false fact implies paying the shame cost c, and voters learn with certainty that the fact is false. Hence, the politician never reports a false fact. Moreover, the other politician is never fact-checked and always reports a fact. However, if that were the case, the fact-checker would deviate and fact-check the politician who might report false facts. This implies that in the presence of a neutral lie-seeking fact-checker, both politicians must be fact-checked with positive probability in equilibrium, and facts must be false with equal probabilities to make the fact-checker indifferent.

Proposition 1 *In any PBE with a neutral lie-seeking fact-checker, conditional on both politicians reporting a fact, each fact is equally likely to be false.*

Proposition 1 implies our first key prediction for content: a neutral lie-seeking fact-checker leads to balanced conclusions, in the sense that the share of false facts is the same for both politicians. Interestingly, (1) also implies that $f_r = (p_r/p_l)f_l < f_l$. Hence, R, who is more likely to lack a fact, is less likely to report a false fact, making politicians equally credible when claiming to have a fact.

So far, we have derived a relationship between f_l and f_r that must be satisfied to make the fact-checker indifferent. To exactly pin down f_l, f_r, as well as the fact-checking probabilities σ_r and $\sigma_l = 1 - \sigma_r$, we now use politicians' incentives. A politician's payoff when admitting she lacks a fact is 0: she loses the election with certainty because the other politician has a fact. Instead, if politician j reports a false fact,[5] she is fact-checked with probability σ_j. In this case, she loses the election and pays the shame cost c. If the

[5] $j \in \{L, R\}$ refers to one politician and $-j$ to the other.

other politician is fact-checked, her fact is proved to be true and voters cross-update their beliefs about the probability of politician j also having a true fact:

$$\Pr(j\text{'s fact is true}| -j\text{'s fact is true}) = \frac{1-p_l-p_r}{1-p_l-p_r+p_{-j}f_j} := \rho_j. \quad (2)$$

Using (1), we obtain $\rho_r = \rho_l := \rho^*$: to make the fact-checker indifferent, reported facts are equally likely to be false. This implies that voters cross-update identically: if R's fact is proved true, the probability that L's fact is true, ρ_l, is equal to ρ_r, the probability that R's fact is true if L's fact is proved true.

As a result, given that $\epsilon \sim U[-1,1]$, a politician who reports a fact but does not get fact-checked wins the election with probability $\rho^*/2$. Hence, politician j is indifferent and can mix if:

$$\sigma_j c = (1-\sigma_j)\frac{\rho^*}{2} \quad (3)$$

With this in mind, let us consider first a candidate equilibrium in which both politicians are indifferent and mix when they do not have a fact. (3) highlights that the payoffs for L and R are the same in each situation if they lack a fact: being fact-checked implies not being elected and incurring the shame cost c, while not being fact-checked implies winning with probability $\rho^*/2$. As a result, the indifference conditions for both politicians can only be satisfied if $\sigma_r = \sigma_l = 1/2$. Hence, in an equilibrium where both politicians mix when lacking a fact, they must be checked with equal probability.

We can now plug $\sigma_j = 1/2$ into (3) to obtain $\rho^* = 2c$. Using the definition of ρ^* in (2), this implies:

$$p_l f_r = p_r f_l = (1-p_l-p_r)\frac{1-2c}{2c}. \quad (4)$$

For now, we assume that $c < 1/2$ so that this expression is positive. We will discuss $c \geq 1/2$ with the other limit case at the end of the section. For both politicians to mix, f_l and f_r must be at most 1. As $p_r < p_l$, $f_l > f_r$ and f_l should not exceed 1, implying that the maximum value for $p_r f_l$ is p_r. Combining this upper bound with (4), we obtain:

$$(1-p_l-p_r)\frac{1-2c}{2c} \leq p_r \Leftrightarrow c \geq \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}. \quad (5)$$

As a result, when the shame cost c is above the threshold defined by (5), there is an equilibrium in which the fact-checker makes both politicians indifferent by checking them with equal probabilities, i.e., we have a balanced verification. Moreover, the politicians mix with interior probabilities pinned down by (4):

$$f_r = \frac{1-p_l-p_r}{p_l}\frac{1-2c}{2c} \qquad \text{and} \qquad f_l = \frac{1-p_l-p_r}{p_r}\frac{1-2c}{2c}. \tag{6}$$

However, if the shame cost c is lower than the threshold in (5), satisfying (4) would require $f_l > 1$. Intuitively, the shame cost is too low to discipline L, the politician more likely to have a fact. We now establish that when this is the case, the equilibrium is such that both R and the fact-checker are indifferent and mix, while L strictly prefers to report a false fact, implying $f_l = 1$.

To make the fact-checker indifferent with $f_l = 1$, R must report a false fact with probability $f_r = p_r/p_l$. This implies that $\rho^* = (1-p_l-p_r)/(1-p_l)$. Moreover, σ_r is pinned down by R's indifference condition given by (3), which yields:

$$\sigma_r = \frac{1-p_l-p_r}{1-p_l-p_r+2c(1-p_l)}. \tag{7}$$

This fact-checking probability implies that R is indifferent while L prefers to report a false fact. The expression in (7) exceeds $1/2$, implying that R is checked more frequently: (7) decreases with c; it approaches 1 as the shame cost c approaches 0 and reaches $1/2$ when c equals the threshold defined in (5). Therefore, when the shame cost is low, the fact-checker is unable to discipline both politicians and focuses on satisfying R's indifference condition, which results in fact-checking R more often than L.

Hence, we have characterized one PBE for each value of $c < 1/2$. To establish that the PBE is unique, we show in the proof appendix that other strategy profiles from politicians cannot be part of a PBE. The results are summarized in the following proposition:

Proposition 2 *Suppose there is a neutral lie-seeking fact-checker, $p_l > p_r$ and $c < \frac{1}{2}$. Then, the game has a unique PBE, such that:*

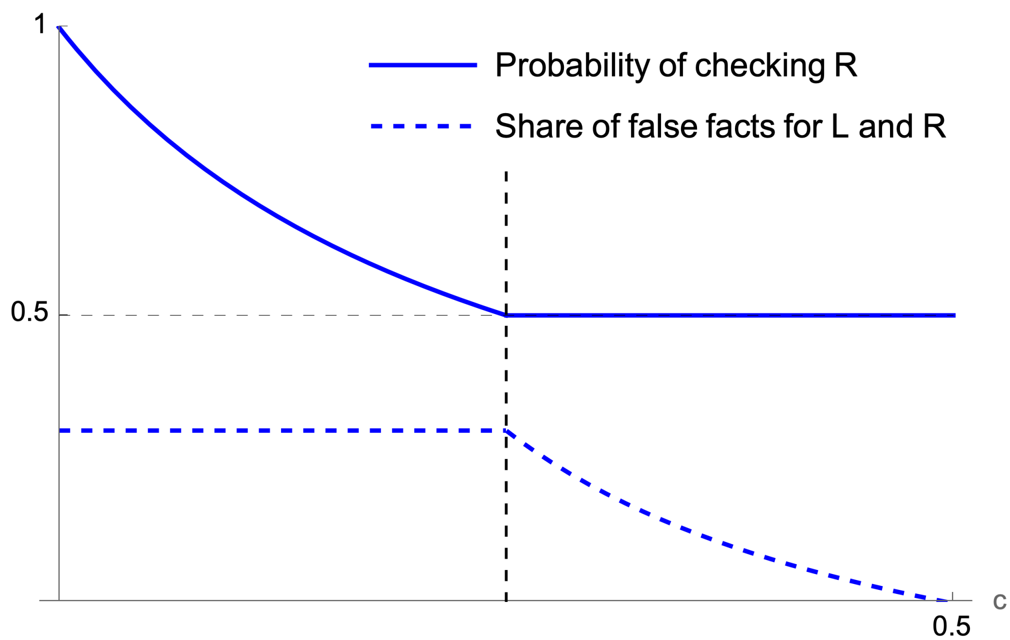
(a) For $c < \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}$, $f_l = 1$ and $f_r = \frac{p_r}{p_l} < 1$. Moreover, $\rho^* = \frac{1-p_l-p_r}{1-p_l}$ and $\sigma_r = \frac{1-p_l-p_r}{1-p_l-p_r+2c(1-p_l)} > \frac{1}{2}$.

(b) For $c \geq \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}$, $f_l = \frac{1-p_l-p_r}{p_r}\frac{1-2c}{2c}$ and $f_r = \frac{1-p_l-p_r}{p_l}\frac{1-2c}{2c}$, with both f_l and f_r smaller than 1. Moreover, $\rho^* = 2c$ and $\sigma_r = \frac{1}{2}$.

We illustrate the implications of Proposition 2 for the fact-checker's content in Figure 2. It shows the share of false facts (dashed curve) and the probability that R is fact-checked (solid curve) as the shame cost, c, varies. The share of false facts is the same for both politicians. When the shame cost is below the threshold defined by (5), represented by the dashed vertical line, the fact-checker is unable to discipline both politicians and prioritizes fact-checking R. The share of false facts remains constant and equal to p_r, as L always reports false facts. When the shame cost exceeds the threshold, politicians are equally likely to be fact-checked and the share of false facts decreases with c.

We assumed that R was more likely to lack a fact to simplify the analysis. Assuming $p_r > p_l$ would yield comparable results, except that $f_r = 1$ in the equilibrium described in Proposition 2.a. As a result, even if politicians differ in their probability of having a fact ex-ante, a neutral lie-seeking fact-checker levels political communication and leads to balanced conclusions, i.e., the same share of false facts, and possibly to balanced verification, i.e., equal probabilities of fact-checking. This setup allows us to build the intuition and present the key mechanisms, but these predictions are not consistent with the empirical stylized facts in Section 2.

Figure 2: Equilibrium content with a neutral lie-seeking fact-checker



Note: Solid: σ_r, probability of checking R as shame cost c varies. Dashed: $p_l f_r = p_r f_l$, share of false facts checked, which is the same for both politicians. Vertical dashed: value of c that distinguishes subcases (a) and (b) in Proposition 2. $p_l = 0.4$ and $p_r = 0.3$.

We now turn to the limit cases. First, when $p_l = p_r$, there is a multiplicity of equilibria in Proposition 2.a: when the shame cost is low, $f_l = f_r = 1$, and there exists a range of σ_r where both politicians strictly prefer reporting false facts. The other results are unaffected.

Second, for a high shame cost $c \geq 1/2$, politicians never report false facts in equilibrium. Both politicians can be checked with equal probabilities, but less balanced fact-checking probabilities can also be consistent with an equilibrium.

5 Mild bias: asymmetric lie-spotting benefits

We now relax the fact-checker's neutrality and assume that the fact-checker's payoff is 1 for checking a false fact from R and v for checking a false fact from L. We focus on $1 < v < p_l/p_r$ for the discussion. The condition $1 < v$ implies that L is a more valuable target, either due to bias against L or because detecting false facts from L is more important, possibly reflecting the composition of the readership. Moreover, we continue to assume that $c < 1/2$ and $p_r < p_l$. The cases $v > p_l/p_r$, $v < 1$, and $p_r \geq p_l$ are considered at the end of the section.

The fact-checker's indifference condition becomes:

$$v p_r f_l = p_l f_r, \tag{8}$$

which implies that, in equilibrium, the probability that L reports a false fact, $p_r f_l$, is lower than the probability that R reports a false fact; otherwise the fact-checker would fact-check L. Crucially, this situation disrupts the balance of conclusions: the more valuable target must be less likely to report false facts in equilibrium. This outcome is intriguing: while the fact-checker may appear biased against L, he disciplines that politician, reducing her share of false facts. In turn, this implies that the cross-updating is not identical, and the indifference conditions for R and L are:

$$\sigma_r c = (1-\sigma_r)\frac{\rho_r}{2} \qquad \text{and} \qquad (1-\sigma_r)c = \sigma_r \frac{\rho_l}{2}. \tag{9}$$

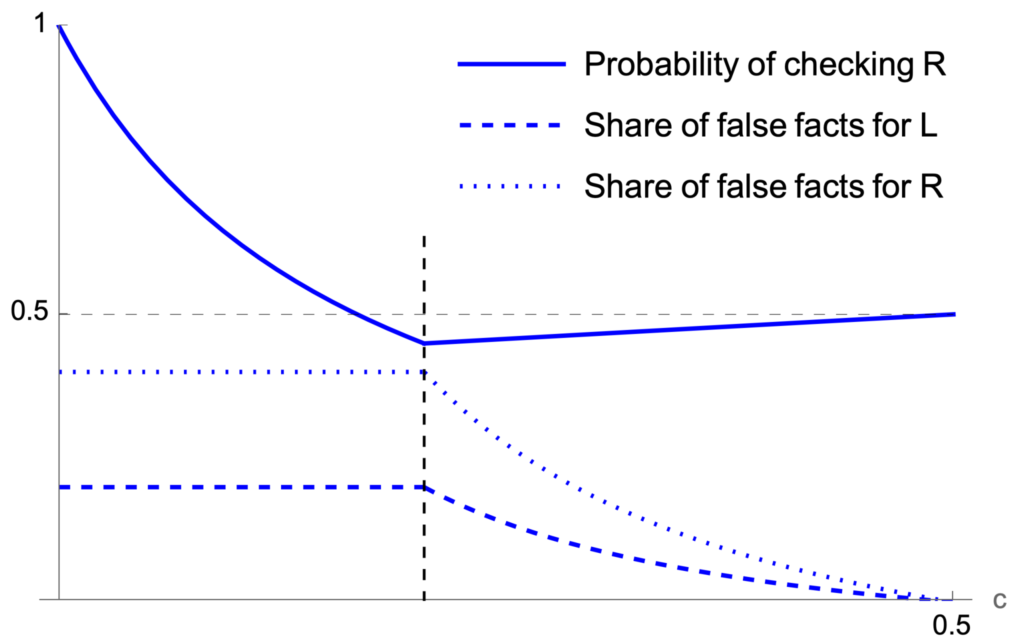
As in Section 4, we use conditions (8) and (9) to determine the mixing probabilities of the fact-checker and politicians. We also obtain two types of PBE: when the shame cost is higher than a threshold, both indifference conditions in (9) are satisfied, and politicians mix. When the shame cost is low, both politicians cannot be disciplined simultaneously. The assumption $v < p_l/p_r$ ensures that, in such a PBE, L reports a false fact with probability 1, while R mixes:

Proposition 3 *Suppose the fact-checker has asymmetric lie-spotting benefits with $1 < v < p_l/p_r$. Moreover, $p_l > p_r$ and $c < \frac{1}{2}$. Defining $\Theta := \sqrt{c^2(v-1)^2 + v} + c(v-1)$, the game has a unique PBE, such that:*

(a) *For* $c < \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}\frac{\Theta}{v}$, $f_l = 1$ *and* $f_r = \frac{vp_r}{p_l} < 1$. *Moreover,* $\rho_l = \frac{1-p_l-p_r}{1-p_l}$, $\rho_r = \frac{1-p_l-p_r}{1-p_l+p_r(v-1)}$ *and* $\sigma_r = \frac{1-p_l-p_r}{1-p_l-p_r+2c(1-p_l+p_r(v-1))}$.

(b) *For* $c \geq \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}\frac{\Theta}{v}$, $f_l = \frac{1-p_l-p_r}{p_r}\frac{\Theta-2cv}{2cv}$ *and* $f_r = \frac{1-p_l-p_r}{p_l}\frac{v-2c\Theta}{2c\Theta}$, *with both* f_l *and* f_r *smaller than* 1. *Moreover,* $\rho_l = 2c\frac{v}{\Theta}$, $\rho_r = 2c\frac{\Theta}{v}$ *and* $\sigma_r = \frac{\Theta}{\Theta+v}$.

Figure 3: Equilibrium content with asymmetric lie-spotting benefits



Note: Solid: σ_r, probability of checking R as shame cost c varies. Dashed: $p_r f_l$, share of false facts checked for L. Dotted: $p_l f_r$, share of false facts for R. Vertical dashed: value of c that distinguishes subcases (a) and (b) in Proposition 3. $p_l = 0.6$, $p_r = 0.2$, and $v = 2$.

The implications for the fact-checker's content are illustrated in Figure 3. When the shame cost exceeds the threshold, as described in Proposition 3.b, the share of false facts is higher for R, yet L is fact-checked more frequently. To understand this, recall that the fact-checker's indifference condition in (8) requires a lower share of false facts for L. This implies that the cross-update is more favorable to L: when the fact-checker reveals that R's fact is true, the probability that L's fact is also true, ρ_l, is higher than ρ_r, the probability that R's fact is true, given that L's fact has been proved true. Thus, if both politicians are fact-checked with equal probabilities, L's expected payoff from reporting a false fact, $\rho_l/4 - c/2$, exceeds that of R due to $\rho_l > \rho_r$. Therefore, to satisfy both indifference conditions in (9), L must be fact-checked more often. As a result, the more valuable target is checked more frequently, which imposes greater discipline and results in a lower share of false facts. Hence, the PBE described in Proposition 3.b does not align with empirical stylized facts, where the politician who is fact-checked more frequently exhibits a higher share of false facts.

However, the PBE described in Proposition 3.a can align with the empirical stylized facts. First, the fact-checker's indifference condition implies that the share of false facts is higher for R. Moreover, following the same logic as in Proposition 2.a, the fact-checker is unable to discipline both politicians and focuses on satisfying R's indifference condition, resulting in more frequent fact-checking of R when the shame cost is low. As a result, R exhibits both a higher likelihood of being fact-checked and a higher share of false facts, which reflects the empirical patterns observed for Republicans.

To summarize, we obtained two key findings in this section. First, a fact-checker with asymmetric lie-spotting benefits leads to a lower share of false facts for the more valuable target. Second, assuming a mild bias in favor of Democrats and a low shame cost, the PBE is compatible with the empirical stylized facts. This bias challenges the claim that fact-checkers favor Democrats [Brandtzaeg and Følstad, 2017, Palumbo, 2023]. However, it might be compatible with the left-leaning tendencies of fact-checking readers [Shin and Thorson, 2017], who are plausibly more exposed to false facts from Democrats. Hence, a fact-checker genuinely committed to combating misinformation might prefer to scrutinize false facts from Democrats to better serve its audience.

Finally, we return to the three cases $v < 1$, $v > p_l/p_r$, and $p_r \geq p_l$. First, if $v < 1$, R is a more valuable target, so she must exhibit a lower share of false facts, which is not compatible with the empirical stylized facts. Second, when $vp_r > p_l$, the fact-checker's preference for L is so high that his indifference condition requires $f_l < f_r$, i.e., it more than compensates for the lower ex-ante probability of a fact for R. Proposition 3.b is unaffected, but the PBE differs when the shame cost is low: as $f_l < f_r$, the fact-checker concentrates on disciplining L, implying that L would also be more likely to be checked in Proposition 3.a. Hence, L exhibits a lower share of false facts and is fact-checked more frequently for all values of c, which is not compatible with the empirical stylized facts. Third, suppose $p_r \geq p_l$. $v > 1$ still implies that L has a lower share of false facts and should be checked more to be indifferent. She is also checked more in Proposition 3.b: the indifference condition of the fact-checker, $vp_r f_l = p_l f_r$, can only be satisfied for $f_l < f_r$ and the fact-checker concentrates on disciplining L. It follows that a mild bias in favor of Democrats can only explain the stylized facts if Republicans are more likely to lack a fact. This hypothesis aligns with a strand of the academic literature [Bucciol, 2018, Grinberg et al., 2019, Guess et al., 2020, Mosleh et al., 2024], although these studies rely on fact-checker content and assume it is unbiased.

6 Strong bias: partisan fact-checker

We now consider a stronger form of bias: the fact-checker derives a positive payoff from confirming the fact of one of the politicians. Specifically, we assume that the fact-checker receives a payoff of 1 for checking a false fact from R, and a certification payoff of b for checking a true fact from L. This setup can be interpreted as a bias against R and in favor of L. We examine a fact-checker with a bias against L at the end of the section. The assumptions $p_r < p_l$ and $c < 1/2$ are unnecessary.

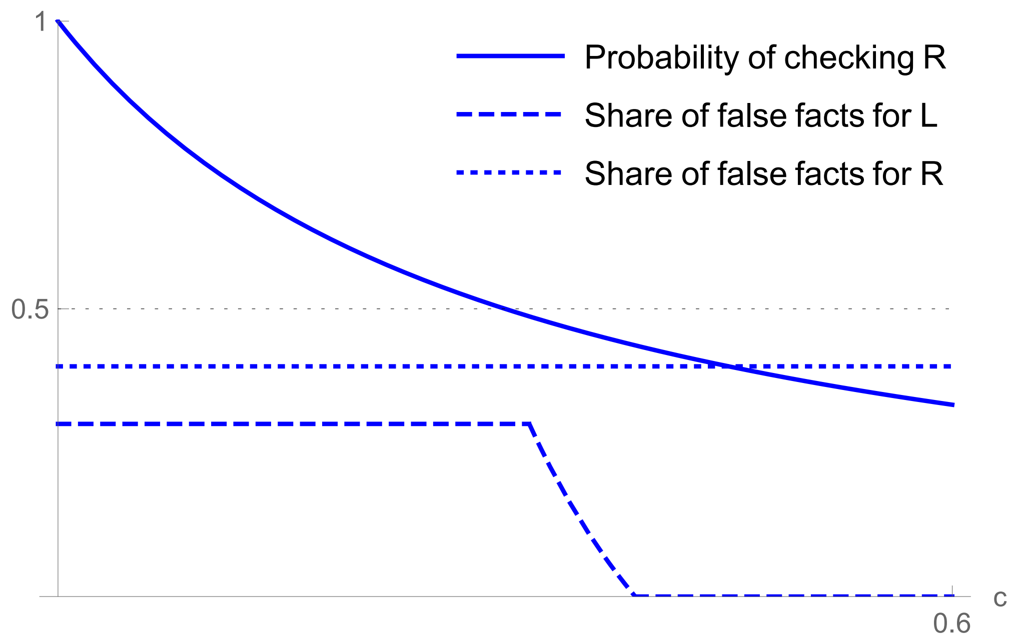
As in the previous sections, checking only R cannot be part of a PBE: R would never report false facts, and the fact-checker would prefer to check L to obtain b when L reports a true fact. However, checking only L can be part of a PBE. In this case, L never reports false facts, while R does so with probability 1. Checking R yields an expected payoff of $p_l/(1-p_r)$ (recall that we focus on cases where both politicians report facts), while checking L yields the certification payoff b. Therefore, if $b \geq p_l/(1-p_r)$, the fact-checker checks only L in equilibrium. Such a PBE is inconsistent with the empirical stylized facts, as only one side would be fact-checked and would exhibit no false facts. When $b < p_l/(1-p_r)$, we have:

Proposition 4 *Suppose the fact-checker is biased in favor of L and derives a certification payoff $b < p_l/(1-p_r)$ when checking a true fact from L. The game has a unique PBE, where $f_r = \frac{b}{1-b} \cdot \frac{1-p_l-p_r}{p_l}$, $\sigma_r = \frac{1-b}{1-b+2c}$, and*

$$f_l = \begin{cases} 1 & \text{if } c < \frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}, \\ \frac{1-p_l-p_r}{p_r}\frac{1-b-4c^2}{4c^2} & \text{if } c \in [\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}, \frac{\sqrt{1-b}}{2}], \\ 0 & \text{if } c > \frac{\sqrt{1-b}}{2}. \end{cases} \tag{10}$$

Figure 3 illustrates the implications for the fact-checker's content. In this setup, the fact-checker's tradeoff no longer depends on the probability that L reports a false fact: from the fact-checker's perspective, this probability reduces the likelihood that L's fact is true, and also decreases the likelihood that R's fact is false. Thus, the fact-checker's indifference condition only depends on f_r, which pins down the equilibrium value of this probability. As the fact-checker's indifference condition does not depend on c, the share of false facts for R is independent of c. In turn, the fact-checker focuses on R's indifference condition. A higher shame cost implies that a lower probability of fact-checking is needed

Figure 4: Equilibrium content with a strongly biased fact-checker



Note: Solid: σ_r, probability of checking R as shame cost c varies. Dashed: $p_r f_l$, share of false facts checked for L. Dotted: $p_l f_r$, share of false facts checked for R. $p_l = 0.1$, $p_r = 0.3$, and $b = 0.4$.

to make R indifferent. Consequently, L becomes more likely to be fact-checked as c rises, leading to a lower share of false facts from L.

Turning to the certification payoff b, the fact-checker becomes less likely to fact-check R as b increases. This outcome is intuitive since b represents the potential payoff from checking L. In turn, a higher b translates into a lower share of false facts from L, who is disciplined, and a higher share of false facts from R.

Most importantly, Figure 3 reveals that, for some parameters, the fact-checker's content aligns with the empirical stylized facts:

Proposition 5 *In the PBE described in Proposition 4 with a fact-checker biased in favor of L, R is checked more than L and exhibits a higher share of false facts if $c < (1-b)/2$ and $b \geq p_r/(1-p_l)$.*[6]

The first condition of Proposition 5 is that the shame cost c is small so that R is checked

[6]The interval $[p_r/(1-p_l), p_l/(1-p_r)]$ is non-empty: $p_l/(1-p_r) < p_r/(1-p_l)$ would imply

$$p_l(1-p_l) - p_r(1-p_r) < 0 \Leftrightarrow (p_l - p_r)(1 - p_l - p_r) < 0,$$

which cannot hold as we assume $p_l > p_r$ and $1 - p_l - p_r > 0$.

more frequently. The second condition is that the certification payoff b should not be too low, ensuring that L is sufficiently fact-checked to exhibit a lower share of false facts. Under these conditions, we obtain a simple explanation for the imbalance in content, which aligns with the claims that fact-checking favors Democrats.

While the explanation for the stylized facts with a mild bias required $p_r < p_l$, the conditions in Proposition 5 are also compatible with L being more (or equally) likely to lack a fact. Indeed, we have $p_r > p_l$ in Figure 3.

We now consider a fact-checker with a strong bias in favor of R: he receives 1 for checking a false fact from L and b for a true fact from R. We assume $b < p_r/(1-p_l)$ to avoid the fact-checker only checking R. The PBE follows a structure similar to Proposition 4. However, there are no parameters for which it is compatible with the stylized facts:

Proposition 6 *With a fact-checker biased in favor of R, R cannot be jointly fact-checked more than L and exhibit a higher share of false facts.*

In equilibrium, the fact-checker focuses on L's indifference condition. Consequently, R is checked more frequently when the shame cost is high, as a small probability of checking L suffices to make her indifferent. However, a higher probability of checking R than L ensures that R is disciplined and cannot exhibit a higher share of false facts. Thus, we conclude that a strong bias in favor of Democrats can align with the stylized facts, whereas a strong bias in favor of Republicans cannot.

7 Welfare analysis

We established that fact-checking plays a key role in disciplining politicians and helping voters elect the politicians best suited for them. However, it remains to be established which fact-checking strategy is actually optimal for voters. In this section, we first characterize the fact-checking strategy that maximizes voters' welfare. Next, we derive the conditions under which the fact-checker exhibiting the preferences studied in the previous sections is welfare optimal. Finally, we discuss how the welfare achieved under the different preferences compares. We focus on $p_r < p_l$ to simplify the discussion.

Proposition 7 *The voter's welfare can be expressed as:*

$$1 + \frac{1 - p_l - p_r}{4}[(1-\sigma_r)\rho_r + \sigma_r \rho_l] \tag{11}$$

The voters' welfare is maximized by a fact-checking strategy such that:

(a) If $c < \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, $\sigma_r = \frac{1}{1+2c}$, such that $f_r = 0$.

(b) If $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, then any $\sigma_r \in \left[1 - \frac{1}{1+2c}, \frac{1}{1+2c}\right]$ maximizes welfare.

Hence, the welfare-optimal strategy has a straightforward form: when the shame cost is small, as in Proposition 7.a, it involves fact-checking the politician more likely to lack a fact to the minimum extent necessary to fully discipline her, resulting in $f_r = 0$, and allocating the remaining capacity to L. This underscores the critical role of the disciplining effect for welfare: it is optimal to fact-check a politician who does not report false facts, even when the other politician does. Such an outcome cannot be achieved by a lie-seeking fact-checker and is also inconsistent with the empirical stylized facts. While a small shame cost requires a precise strategy to maximize welfare, Proposition 7.b reveals that a higher shame cost allows for a range of welfare maximizing fact-checking probabilities: the reason for this is that the expression for welfare (11) depends on $(1-\sigma_r)\rho_r + \sigma_r\rho_l$, and following condition (9), the indifference of both politicians requires $(1-\sigma_r)\rho_r + \sigma_r\rho_l = 2c$. Therefore, changing the distribution of fact-checking while maintaining politicians indifferent has no effect on welfare. It follows that each motivation can achieve welfare optimality:

Proposition 8 *The following holds:*

(a) For a sufficiently large shame cost, specifically $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, a neutral lie-seeking fact-checker and a fact-checker with asymmetric benefits are welfare-optimal.

(b) A strongly biased fact-checker favoring L is welfare-optimal for $c \in \left[\frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}, \frac{\sqrt{1-b}}{2}\right]$ (unless the interval is empty), but for $c > \frac{\sqrt{1-b}}{2}$, a strongly biased fact-checker cannot be welfare-optimal.

Proposition 8.a highlights that a neutral lie-seeking fact-checker and a fact-checker with asymmetric benefits for checking false facts can only maximize welfare for high enough shame cost c: they have no incentive to fact-check a politician that never reports false facts and cannot implement $f_r = 0$. Hence, they cannot be welfare-optimal in the sense of Proposition 7.a. However, when c is large, they fact-check R with a probability converging to $1/2$, which falls within the range of welfare-maximizing fact-checking probabilities specified in Proposition 7.b. Instead, a strongly biased fact-checker favoring L can only be welfare optimal if c is neither too small, nor too large. The intuition behind why c cannot be too small is that a positive probability of R reporting false facts must be compatible with welfare optimality. Moreover, if c is too large, specifically if $c > \frac{\sqrt{1-b}}{2}$, the fact-checker's scrutiny of L, aimed at certifying true statements, leads L to strictly prefer not to report a false fact. This is inefficient since R keeps reporting false facts and should be fact-checked more.

While Proposition 8 focuses on the conditions on the shame cost that make fact-checking welfare-optimal, our next results demonstrate that both forms of bias can approach welfare optimality:

Proposition 9 *The following holds:*

(a) The equilibrium strategy of the fact-checker with asymmetric lie-spotting benefits converges to the welfare-optimal strategy when v, the payoff for checking a false fact from L, approaches zero. Furthermore, welfare is (weakly) decreasing in v.

(b) The equilibrium strategy of a strongly biased fact-checker favoring L converges to the welfare-optimal strategy when b, the certification payoff, approaches zero. Furthermore, welfare is (weakly) decreasing in b.

Proposition 9 establishes that limit cases of bias against the politician more likely to have a fact lead to welfare optimality. Specifically, Proposition 9.a shows that a fact-checker with asymmetric lie-spotting benefits approaches welfare optimality as $v \to 0$, i.e., when R is is a much-preferred target. Similarly, Proposition 9.b establishes that a strongly biased fact-checker favoring L achieves the same result when the certification payoff is very small ($b \to 0$). In both cases, the fact-checker's indifference condition requires f_r to be close to zero, achieving welfare optimality given Proposition 7.a.

Finally, we compare the welfare obtained under the different preferences. There is a simple welfare ranking between a neutral lie-seeking fact-checker and one with asymmetric lie-spotting benefits. By Proposition 9, the voter's welfare decreases with v, the payoff for checking a false fact from L, who is less likely to lack a fact. Since the two fact-checkers are equivalent when $v = 1$, a fact-checker with asymmetric benefits yields higher welfare when $v < 1$ because it better disciplines the politician more likely to lack a false fact. Conversely, a neutral lie-seeking fact-checker yields higher welfare when $v > 1$. However, there is no clear welfare ranking when we compare the strongly biased fact-checker with the others.

8 Discussion and concluding remarks

To sum up, this paper presents a model to examine the motivations of political fact-checkers, offering key insights into the imbalances observed in the content they publish.

Our findings demonstrate that the unequal scrutiny of Republicans and Democrats cannot be solely explained by initial differences in the probabilities of lacking facts, as the disciplining effect of fact-checking should lead to equal shares of false information identified across the political spectrum. Instead, these imbalances can be accounted for by either a strong pro-Democrat bias or a mild pro-Republican bias. These results challenge simple explanations of bias and underscore the need to consider fact-checking as a strategic process involving journalists, politicians, and voters.

We focused on an information structure in which *at least* one politician has a fact, as we believe it is a necessary condition for a topic to be in the public debate. We now discuss alternative information structures. In a setup where *at most* one politician has a fact, there is a trivial and potentially unique equilibrium in which politicians never report false facts, which provides limited insights for our analysis. If a politician deviates and reports a false fact, then at most one reported fact is true. Thus, the fact-check necessarily reveals which politician has the true fact, making such a deviation unprofitable. In Appendix A, we formally analyze the more relevant setup where nature independently determines whether each politician has a fact, focusing on the neutral lie-seeking fact-checker. Hence, both politicians may simultaneously lack a fact. As in our main model, the fact-checker's indifference condition equalizes the probabilities of reporting false facts. However, the politician less likely to have a fact is at least partly disciplined, while the other either always or never reports false facts. This outcome still fails to reconcile the stylized facts.

We conclude with a discussion of some potential alternative explanations. We considered simple forms of bias that do not explicitly include attempts to manipulate the outcome of the election. One interpretation of the strong bias is that the positive certification payoff comes from the increased winning probability of the preferred politician. However, this does not reflect all the effects on the election, as the fact-checker may find a false fact from her preferred politician and reduce her winning probability. An alternative approach would be to assume that the fact-checker maximizes the probability that one side is elected, but our political setup is not suited for investigating such preferences. As both the fact-checker and voters share the same beliefs about the accuracy of each fact, Bayes' law dictates that the expected posterior probability remains equal to the prior probability. Consequently, a fact-check does not alter the expected probability of a fact being correct. Since the popularity shock is uniformly distributed, fact-checking one side over the other does not influence election probabilities on average. One possibility would be to move away from the uniformity assumption of the popularity shock, but it would introduce tractability challenges, and we leave it for future research.

Furthermore, it has been suggested that some politicians may be less affected by being found wrong, either due to their personal traits or the specific reactions of their supporters.[7] However, asymmetric shame costs among politicians, combined with a neutral lie-seeking fact-checker, would also lead to equal shares of false facts, and hence cannot explain the empirical stylized facts. To see this, notice that the shame cost does not enter the fact-checker's indifference condition in (1). Hence, the share of false facts identified by a neutral lie-seeking fact-checker would remain equal in the presence of asymmetric shame costs.

However, we can obtain different shares of false facts if one of the politicians has a negative shame cost. Such a politician always reports false facts when lacking a true one. If we rule out the possibility of reporting a false fact when a true one is available,[8] the structure of the PBE depends on the distribution of facts. First, suppose the politician with a negative shame cost is more likely to have a fact. In this case, there is a unique PBE in which she always reports a false fact when lacking a true one and is consistently checked. The other politician is never checked and, therefore, also always reports false facts. This extreme form of imbalance is incompatible with the stylized facts, as it implies that only one side of the spectrum would be fact-checked. Second, suppose the politician with a negative shame cost is less likely to have a fact. In this scenario, the other politician would be fact-checked if she always reports false facts, leading to an equilibrium in which both the fact-checker and the politician with a positive shame cost are indifferent. As a result, the content would still exhibit a balanced share of false facts under a neutral lie-seeking fact-checker, meaning that a negative shame cost cannot explain the empirical stylized facts.

References

Jennifer Allen, Antonio A Arechar, Gordon Pennycook, and David G Rand. Scaling up fact-checking using the wisdom of crowds. *Science advances*, 7(36):eabf4393, 2021.

Michelle A Amazeen. A critical assessment of fact-checking in 2012. *New America Foundation*, 2013.

Michelle A Amazeen. Revisiting the epistemology of fact-checking. *Critical Review*, 27(1): 1–22, 2015.

[7]This idea is not supported by Mattozzi et al. [2022], who find that politicians from all sides react to fact-checking.

[8]If the politician could report a false fact despite having a true one, she would always do so and be fact-checked every time. The fact-checker's content would be entirely focused on one side of the spectrum, which is inconsistent with the empirical stylized facts.

Michelle A Amazeen. Checking the fact-checkers in 2008: Predicting political ad scrutiny and assessing consistency. *Journal of Political Marketing*, 15(4):433–464, 2016.

Ben H Bagdikian. *The new media monopoly: A completely revised and updated edition with seven new chapters*. Beacon Press, 2004.

Ivan Balbuzanov. Lies and consequences: The effect of lie detection on communication outcomes. *International Journal of Game Theory*, 48(4):1203–1240, 2019.

Oscar Barrera, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya. Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics*, 182:104123, 2020. https://doi.org/10.1016/j.jpubeco.2019.104123.

Graham Beattie. Advertising and media capture: The case of climate change. *Journal of Public Economics*, 188:104219, 2020. https://doi.org/10.1016/j.jpubeco.2020.104219.

Graham Beattie, Ruben Durante, Brian Knight, and Ananya Sen. Advertising spending and media bias: Evidence from news coverage of car safety recalls. *Management Science*, 67(2):698–719, 2020. https://doi.org/10.1287/mnsc.2019.3567.

Timothy Besley and Andrea Prat. Handcuffs for the grabbing hand? Media capture and government accountability. *American Economic Review*, 96(3):720–736, 2006. https://doi.org/10.1257/aer.96.3.720.

Laurent Bigot. *Fact-checking vs fake news: vérifier pour mieux informer*. INA, 2019.

Petter Bae Brandtzaeg and Asbjørn Følstad. Trust and distrust in online fact-checking services. *Communications of the ACM*, 60(9):65–71, 2017.

Alessandro Bucciol. False claims in politics: Evidence from the us. *Research in Economics*, 72(2):196–210, 2018.

Man-pui Sally Chan, Christopher R Jones, Kathleen Hall Jamieson, and Dolores Albarracín. Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological science*, 28(11):1531–1546, 2017.

Felix Chopra, Ingar Haaland, and Christopher Roth. Do people demand fact-checked news? evidence from us democrats. *Journal of Public Economics*, 205:104549, 2022.

Felix Chopra, Ingar Haaland, and Christopher Roth. The demand for news: Accuracy concerns versus belief confirmation motives. *NHH Dept. of Economics Discussion Paper*, (01), 2023.

Stefano DellaVigna and Johannes Hermle. Does conflict of interest lead to biased coverage? Evidence from movie reviews. *Review of Economic Studies*, 84(4):1510–1550, 2017. https://doi.org/10.1093/restud/rdx002.

Rafael Di Tella and Ignacio Franceschelli. Government advertising and media coverage of corruption scandals. *American Economic Journal: Applied Economics*, 3(4):119–151, 2011. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.3.4.119. URL http://pubs.aeaweb.org/doi/10.1257/app.3.4.119. https://doi.org/10.1257/app.3.4.119.

Wioletta Dziuda and Christian Salas. Communication with detectable deceit. *Available at SSRN 3234695*, 2018.

Florian Ederer and Weicheng Min. Bayesian persuasion with lie detection. Technical report, National Bureau of Economic Research, 2022.

Etienne Fize. Aquiring newspapers, acquiring influence? media bias and capture by billionaires. *Unpublished*, 2020.

Jesse M Gentzkow, Matthew and. Media bias and reputation. *Journal of Political Economy*, 114(2):280–316, 2006. https://doi.org/10.1086/499414.

Matthew Gentzkow and Jesse M Shapiro. What drives media slant? Evidence from U.S. daily newspapers. *Econometrica*, 78(1):35–71, 2010. doi: 10.3982/ECTA7195. URL http://web.stanford.edu/{~}gentzkow/research/biasmeas.pdf. https://doi.org/10.3982/ECTA7195.

Lucas Graves. *Deciding what's true*. Columbia University Press, 2016.

Lucas Graves. Anatomy of a fact check: Objective practice and the contested epistemology of fact checking. *Communication, culture & critique*, 10(3):518–537, 2017.

Lucas Graves, Brendan Nyhan, and Jason Reifler. Understanding innovations in journalistic practice: A field experiment examining motivations for fact-checking. *Journal of Communication*, 66(1):102–138, 2016.

Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. Fake news on twitter during the 2016 us presidential election. *Science*, 363(6425):374–378, 2019.

Andrew M Guess, Brendan Nyhan, and Jason Reifler. Exposure to untrustworthy websites in the 2016 us election. *Nature human behaviour*, 4(5):472–480, 2020.

Emeric Henry, Ekaterina Zhuravskaya, and Sergei Guriev. Checking and sharing alt-facts. *American Economic Journal: Economic Policy*, 14(3):55–86, 2022.

Aleksandr Levkun. Communication with strategic fact-checking. 2022.

Chloe Lim. Checking how fact-checkers check. *Research & Politics*, 5(3):2053168018786848, 2018.

Charles Louis-Sidois. Both judge and party? an analysis of the political leaning of fact-checkers. *An Analysis of the Political Leaning of Fact-checkers (June 14, 2022)*, 2022.

Charles Louis-Sidois and Elisa Mougin. Silence the media or the story? theory and evidence of media capture. *European Economic Review*, 158:104533, 2023.

Morgan Marietta, David C Barker, and Todd Bowser. Fact-checking polarized politics: Does the fact-check industry provide consistent guidance on disputed realities? In *The Forum*, volume 13, pages 577–596. De Gruyter, 2015.

Andrea Mattozzi, Samuel Nocito, and Francesco Sobbrio. Fact-checking politicians. 2022.

Rishabh Misra. Politifact fact check dataset, 09 2022.

Mohsen Mosleh, Qi Yang, Tauhid Zaman, Gordon Pennycook, and David G Rand. Differences in misinformation sharing can lead to politically asymmetric sanctions. *Nature*, pages 1–8, 2024.

Sendhil Mullainathan and Andrei Shleifer. The market for news. *American Economic Review*, 95(4):1031–1053, 2005. https://doi.org/10.1257/0002828054825619.

Sakari Nieminen and Lauri Rapeli. Fighting misperceptions and doubting journalists' objectivity: A review of fact-checking literature. *Political Studies Review*, 17(3):296–309, 2019.

Brendan Nyhan and Jason Reifler. The effect of fact-checking on elites: A field experiment on us state legislators. *American Journal of Political Science*, 59(3):628–640, 2015.

Brendan Nyhan, Ethan Porter, Jason Reifler, and Thomas J Wood. Taking fact-checks literally but not seriously? the effects of journalistic fact-checking on factual beliefs and candidate favorability. *Political Behavior*, 42(3):939–960, 2020.

Saltuk Ozerturk. Media access, bias and public opinion. *European Economic Review*, 147: 104161, 2022.

Matt Palumbo. *Fact-Checking the Fact-Checkers: How the Left Hijacked and Weaponized the Fact-Checking Industry*. Liberatio Protocol, 2023.

Jonathan Reuter and Eric Zitzewitz. Do ads influence editors? Advertising and bias in the financial media. *Quarterly Journal of Economics*, 121(1):197–227, 2006. https://doi.org/10.1093/qje/121.1.197.

Jieun Shin and Kjerstin Thorson. Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67(2):233–255, 2017.

Jane B Singer. Border patrol: The rise and role of fact-checkers and their challenge to journalists' normative boundaries. *Journalism*, 22(8):1929–1946, 2021.

Adam Szeidl and Ferenc Szucs. Media capture through favor exchange. *Econometrica*, 89(1):281–310, 2021. doi: https://doi.org/10.3982/ECTA15641. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA15641. https://doi.org/10.3982/ECTA15641.

Emily Thorson. Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, 33(3):460–480, 2016.

Joseph E Uscinski. The epistemology of fact checking (is still naìve): Rejoinder to amazeen. *Critical Review*, 27(2):243–252, 2015.

Joseph E Uscinski and Ryden W Butler. The epistemology of fact checking. *Critical Review*, 25(2):162–180, 2013.

Nathan Walter, Jonathan Cohen, R Lance Holbert, and Yasmin Morag. Fact-checking: A meta-analysis of what works and for whom. *Political Communication*, 37(3):350–375, 2020.

Lily Ling Yang. Information design with costly state verification. 2023.

Proofs

Proposition 1 is proved in the text.

Proposition 2

Proof. We establish in the text that the strategy profile is a PBE. This proof demonstrates that the PBE is unique. The text shows that the fact-checker must be indifferent and mix. Hence, it is sufficient to show that the other strategy profiles of politicians are not compatible with a PBE. We consider in turn these strategy profiles.

- Politician j never reports a false fact, $-j$ reports a false fact with positive probability. $-j$ is checked with probability 1 and prefers to never report a false fact; this cannot be part of a PBE.

- Both politicians never report false facts. This implies $\rho_l = \rho_r = 1$ (facts are always true). Hence, politicians prefer not to report false facts if:

$$\sigma_j c \geq \frac{(1-\sigma_j)}{2}.$$

  However, satisfying this condition for both politicians is only possible if $c \geq \frac{1}{2}$, which we assume does not hold.

- L is indifferent and mixes, R always reports false facts. $f_r = 1$, so $\frac{p_l f_r}{p_r} > 1$ and the indifference condition of the fact-checker requires $f_l > 1$ and cannot be satisfied. Hence, R is always checked and cannot report false facts.

■

Proposition 3

Proof. Following the text argument leading to Proposition 2, we will first characterize the equilibrium in which both fact-checkers are indifferent and mix (Proposition 3.b) and then the equilibrium in which L always reports a false fact (Proposition 3.a). The uniqueness of the PBE results from the argument of the proof of Proposition 2.

- Proposition 3.b. The equilibrium is pinned down by:

$$\begin{aligned} v p_r f_l &= p_l f_r \\ \sigma_r c &= (1-\sigma_r)\frac{\rho_r}{2} \\ (1-\sigma_r) c &= \sigma_r \frac{\rho_l}{2} \end{aligned} \tag{12}$$

Using the latter two conditions:

$$\sigma_r = \frac{\sqrt{\rho_r}}{\sqrt{\rho_r} + \sqrt{\rho_l}} < \frac{1}{2}$$

where the inequality follows from the fact-checker's indifference together with $v > 1$. Hence, R is fact-checked less than L. Moreover, using the definition of ρ_r and ρ_l and the fact-checker's indifference,

$$v\frac{1-\rho_l}{\rho_l} = \frac{1-\rho_r}{\rho_r}$$

and from this:

$$\rho_l = \frac{v\rho_r}{1+(v-1)\rho_r}.$$

Now, from the last two conditions of (12) we can write:

$$\left(\frac{\sigma_r}{1-\sigma_r}\right)^2 = \frac{\rho_r}{\rho_l}$$

which then yields:

$$\left(\frac{\sigma_r}{1-\sigma_r}\right)^2 = \frac{1+(v-1)\rho_r}{v}$$

Using the fact that $\rho_r = \frac{\sigma_r}{1-\sigma_r}2c$, we obtain:

$$v\left(\frac{\sigma_r}{1-\sigma_r}\right)^2 - 2c(v-1)\frac{\sigma_r}{1-\sigma_r} - 1 = 0.$$

This is a quadratic equation in $\frac{\sigma_r}{1-\sigma_r}$. The solution writes:

$$\frac{\sigma_r}{1-\sigma_r} = \frac{\sqrt{c^2(v-1)^2+v} + c(v-1)}{v}.$$

Denoting by $\Theta := \sqrt{c^2(v-1)^2+v} + c(v-1)$, we obtain:

$$\sigma_r = \frac{\Theta}{\Theta + v}.$$

Finally, we use the last two conditions of (12) to retrieve ρ_r and ρ_l:

$$\rho_r = 2c\frac{\Theta}{v} \quad \text{and} \quad \rho_l = 2c\frac{v}{\Theta}.$$

From those, we use (2) to obtain f_l and f_r:

$$f_l = \frac{1-p_l-p_r}{p_r}\frac{1-\rho_l}{\rho_l} = \frac{1-p_l-p_r}{p_r}\frac{v-2c\Theta}{2c\Theta v} = \frac{1-p_l-p_r}{p_r}\frac{\Theta-2cv}{2cv} \tag{13}$$

$$f_r = \frac{1-p_l-p_r}{p_l}\frac{1-\rho_r}{\rho_r} = \frac{1-p_l-p_r}{p_l}\frac{v-2c\Theta}{2c\Theta} \tag{14}$$

- Proposition 3.a.

  As we assume that $vp_r < p_l$, an equilibrium must be such that $f_l = 1$ and $f_r < 1$, as established in the text leading to Proposition 2. Using the expression of f_l characterized by (13) in the interior equilibrium, we find that $f_l = 1$ occurs when:[9]

$$c < \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}\frac{\Theta}{v}.$$

  With $f_l = 1$, the first condition of (12) implies $f_r = \frac{vp_r}{p_l}$. Using (2), $\rho_l = \frac{1-p_l-p_r}{1-p_l}$ and $\rho_r = \frac{1-p_l-p_r}{1-p_l+p_r(v-1)}$. Finally, the second condition of (12) gives $\sigma_r = \frac{1-p_l-p_r}{1-p_l-p_r+2c(1-p_l+p_r(v-1))}$.

■

Proposition 4

Proof. We will first characterize f_r, then σ_r, and f_l. We finally argue that the PBE is unique.

- $f_r = \frac{b}{1-b} \cdot \frac{1-p_l-p_r}{p_l}$.

  As $b < p_l/(1-p_r)$, the fact-checker mixes. His indifference condition is

$$Pr(\text{R's fact is false}) = b \cdot Pr(\text{L's fact is true}),$$

  which gives:

$$\frac{p_l f_r}{1-p_l-p_r+p_r f_l+p_l f_r} = b \cdot \frac{1-p_l-p_r+p_l f_r}{1-p_l-p_r+p_r f_l+p_l f_r} \Rightarrow p_l f_r = b(1-p_l-p_r+p_l f_r).$$

  f_l affects both sides in the same way: an increase in the probability that L reports a false fact decreases the likelihood that L's fact is true, while also reducing the probability that R's fact is false when both report a fact. Hence, the condition does

[9]Notice Θ is a function of c, but $\frac{c}{\Theta}$ is increasing in c, implying that the condition defines a unique threshold value of c.

not depend on f_r and yields the equilibrium condition for f_r:

$$f_r = \frac{b}{1-b} \cdot \frac{1-p_l-p_r}{p_l}. \tag{15}$$

- Proving $\sigma_r = \frac{1-b}{1-b+2c}$.

  (15) increases with b and equals 1 when $b = p_l/(1-p_r)$. Since we assume $b < p_l/(1-p_r)$, we have $f_r < 1$. Thus, R must be indifferent and R's indifference condition determines the equilibrium mixing probability of the fact-checker. The probability that R's fact is true given that L's fact has been proved true is:

  $$\rho_r = \frac{1-p_l-p_r}{1-p_l-p_r+p_l f_r} = 1-b.$$

  Substituting this expression into R's indifference condition, $\sigma_r c = (1-\sigma_r)\rho_r/2$, we obtain:

  $$\sigma_r = \frac{1-b}{1-b+2c}. \tag{16}$$

- Proving $f_l = \begin{cases} 1 & \text{if } c < \frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}, \\ \frac{1-p_l-p_r}{p_r}\frac{1-b-4c^2}{4c^2} & \text{if } c \in \left[\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}, \frac{\sqrt{1-b}}{2}\right], \\ 0 & \text{if } c > \frac{\sqrt{1-b}}{2}. \end{cases}$

  Conditional on lacking a fact, L's expected payoff if she reports a false fact is:

  $$-(1-\sigma_r)c + \sigma_r\frac{\rho_l}{2} = -(1-\sigma_r)c + \sigma_r\frac{1-p_l-p_r}{2(1-p_l-p_r+p_r f_l)},$$

  where we used the expression for ρ_l in (2). Moreover, L's expected payoff is 0 if she admits she lacks a fact. Hence, she would be indifferent if

  $$f_l = \frac{(2c(\sigma_r-1)+\sigma_r)(p_l+p_r-1)}{2cp_r(\sigma_r-1)} = \frac{1-p_l-p_r}{p_r}\frac{1-b-4c^2}{4c^2},$$

  where we substituted σ_r with its equilibrium value. This expression for f_l decreases with c. Moreover, if $c > \frac{\sqrt{1-b}}{2}$, the above expression for f_l would be negative, so $f_l = 0$. If $c < \frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, the expression would exceed 1 and $f_l = 1$.

Finally, to show that the PBE is unique, notice that f_r is uniquely determined by the fact-checker's indifference condition, σ_r is uniquely determined by R's indifference condition, and L has a unique best response to σ_r and f_r. ■

Proposition 5
Proof.

1. The first condition is that the shame cost c is small, such that R is checked more frequently. Using (16), $\sigma_r > 1/2$ requires $c < (1-b)/2$.

2. The second condition guarantees that R has a higher share of false facts than L. First, the case where $f_l = 0$ in (10) is incompatible with $\sigma_r > 1/2$: when R is checked more, L cannot be perfectly disciplined and never report false facts. To see this, recall that $\sigma_r > 1/2$ requires $c < (1-b)/2$. Since $(1-b)/2 < \sqrt{1-b}/2$, $\sigma_r > 1/2$ implies $c < \sqrt{1-b}/2$, which rules out the possibility that $f_l = 0$ in the last case of (10).

   Furthermore, in the intermediate case of (10), where $f_l \in (0,1)$, it is impossible to have both $\sigma_r > 1/2$ and R exhibiting a higher share of false facts. In fact, combining (15) and (10), the condition $p_l f_r > p_r f_l$ holds if

$$\frac{b}{1-b}(1-p_l-p_r) > \frac{1-b-4c^2}{4c^2}(1-p_l-p_r),$$

   which simplifies to $c > (1-b)/2$, and is therefore incompatible with $\sigma_r > 1/2$. Finally, we consider the case where $f_l = 1$. We have $p_l f_r > p_r f_l$ if

$$\frac{b}{1-b}(1-p_l-p_r) > p_r,$$

   which can be rearranged to $b > p_r/(1-p_l)$. Under this condition, we must have

$$\frac{1-b}{2} < \frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}.$$

   The left-hand side represents the highest value of c compatible with $\sigma_r < 1/2$, and the right-hand side is the maximum value of c to have $f_l = 1$. Hence, satisfying both $\sigma_r < 1/2$ and $f_l = 1$ requires $c < (1-b)/2$. In such a case, R is fact-checked more frequently while exhibiting a higher share of false facts.

■

Proposition 6
Proof. Suppose the fact-checker receives 1 for checking a false fact from L and b for a true fact from R. Following the argument of the proof of Proposition 4, there is a unique PBE

where $f_l = \frac{b}{1-b} \cdot \frac{1-p_l-p_r}{p_r}$, $\sigma_l = \frac{1-b}{1-b+2c}$, and

$$f_r = \begin{cases} 1 & \text{if } c < \frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_r}}, \\ \frac{1-p_l-p_r}{p_l}\frac{1-b-4c^2}{4c^2} & \text{if } c \in [\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_r}}, \frac{\sqrt{1-b}}{2}], \\ 0 & \text{if } c > \frac{\sqrt{1-b}}{2}. \end{cases}$$

Using the expression of σ_l, $\sigma_l < 1/2$ (R is checked more) if

$$c > \frac{1-b}{2}. \tag{17}$$

Moreover, $p_l f_r > p_r f_l$ (R is more likely to report a false fact) can only hold for c sufficiently small: $p_r f_l$ does not depend on c, while $p_l f_r$ decreases with c, and for $c > \frac{\sqrt{1-b}}{2}$, $f_r = 0$ implies $p_l f_r = 0 < p_r f_l$. When $c \in [\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_r}}, \frac{\sqrt{1-b}}{2}]$, $p_l f_r > p_r f_l$ if:

$$p_l \cdot \frac{1-p_l-p_r}{p_l}\frac{1-b-4c^2}{4c^2} > p_r \cdot \frac{b}{1-b} \cdot \frac{1-p_l-p_r}{p_r},$$

which simplifies to

$$c < \frac{1-b}{2}. \tag{18}$$

Hence, (17) and (18) cannot simultaneously hold, which shows that R cannot be jointly fact-checked more and exhibit a higher share of false facts. ■

Proposition 7

Proof. We first derive an expression for the expected utility of the voter under a fact-checking distribution (σ_l, σ_r), with $\sigma_l = 1 - \sigma_r$. There are four possible scenarios in which the voter finds herself in at the time of the election. The first two are those in which either one politician admits to not having a fact, or fact-checking uncovers a false fact. In these cases, for all realizations of the popularity shock ϵ the politician with a true fact wins the election, and the voter gets an expected utility of 1. The other two scenarios are those in which fact-checking does not uncover a lie, hence uncertainty remains about whether the other politician has a true fact or not (recall that it is possible that both have a true fact). In these cases, the popularity shock ϵ becomes decisive: denoting by ρ_j, with $j \in \{L, R\}$, the posterior probability that the not fact-checked politician has a fact, the expected utility of the voter in these cases is $1 + \frac{\rho_j^2}{4}$. Choosing the appropriate probabilities for each event, we obtain the following expression:

$$(p_l(1-f_r)+p_lf_r\sigma_r)+p_lf_r(1-\sigma_r)\left(1+\frac{\rho_r^2}{4}\right)+(p_r(1-f_l)+p_rf_l(1-\sigma_r))+p_rf_l\sigma_r\left(1+\frac{\rho_l^2}{4}\right)+$$
$$+(1-p_l-p_r)\left(\sigma_r\left(1+\frac{\rho_l^2}{4}\right)+(1-\sigma_r)\left(1+\frac{\rho_r^2}{4}\right)\right)$$

which can be rearranged to yield:

$$1+\frac{1-p_l-p_r}{4}\left[(1-\sigma_r)\rho_r+\sigma_r\rho_l\right] \tag{19}$$

or, using the fact that $\rho_j=\frac{1-p_l-p_r}{1-p_l-p_r+p_{-j}f_j}$:

$$1+\frac{(1-p_l-p_r)^2}{4}\left[(1-\sigma_r)\frac{1}{1-p_l-p_r+p_lf_r}+\sigma_r\frac{1}{1-p_l-p_r+p_rf_l}\right]$$

which can be rearranged to:

$$1+\frac{(1-p_l-p_r)^2}{4}\left[\frac{1}{1-p_l-p_r+p_lf_r}+\sigma_r\left(\frac{p_lf_r-p_rf_l}{(1-p_l-p_r+p_rf_l)(1-p_l-p_r+p_lf_r)}\right)\right] \tag{20}$$

Having derived an expression for welfare, we can now proceed to analyze the welfare properties of fact-checking. The following series of lemmas answers the question of how welfare responds to the allocation of fact-checking for a given behavior of politicians. Therefore, we will fix f_j and determine the optimal fact-checking strategy.

Lemma 1 *Suppose that $f_r=0$ and $f_l\in(0,1]$. Then, increasing σ_r decreases welfare. In the mirror case of $f_l=0$ and $f_r\in(0,1]$, increasing σ_r increases welfare.*

Proof. In these cases, one politician is fully disciplined by fact-checking, whereas the other one is not. Let us consider $f_r=0$ and $f_l\in(0,1]$ first. Imposing $f_r=0$ in equation (20) and differentiating with respect to σ_r, we obtain:

$$-\frac{1}{1-p_l-p_r}+\frac{1}{1-p_l-p_r+p_rf_l}-\sigma_r\frac{1}{(1-p_l-p_r+p_rf_l)^2}p_r\frac{\partial f_l}{\partial\sigma_r}$$

Which is negative, meaning that welfare is decreasing in σ_r. In the mirror case in which $f_l=0$ and $f_r\in(0,1]$, the same procedure leads us to:

$$\frac{1}{1-p_l-p_r}-\frac{1}{1-p_l-p_r+p_lf_r}-(1-\sigma_r)\frac{1}{(1-p_l-p_r+p_lf_r)^2}p_l\frac{\partial f_r}{\partial\sigma_r}$$

Which is positive, allowing us to conclude that welfare is increasing in σ_r. The intuition is straightforward: if one politician is fully disciplined by fact-checking, all fact-checking devoted to him is wasted, hence it is optimal to allocate as much as possible to the other politician. ■

Lemma 2 *Suppose $f_l = 1$ and $f_r \in (0,1)$. Welfare is increasing in σ_r. In the mirror case of $f_r = 1$ and $f_l \in (0,1)$, welfare is decreasing in σ_r.*

Proof. Unlike the cases described in Lemma 1, the cases covered by this lemma are not straightforward, since there may be a conflict between the discipline and the information effect of fact-checking. Let us consider first the case of $f_l = 1$ and $f_r \in (0,1)$. Using equation (20) and differentiating yields:

$$\frac{1}{1-p_l} - \frac{1}{1-p_l-p_r+p_l f_r} - (1-\sigma_r)\frac{1}{(1-p_l-p_r+p_l f_r)^2} p_l \frac{\partial f_r}{\partial \sigma_r}$$

where the first two terms capture the information effect, whereas the last one is the discipline effect: the latter is positive, but the information effect may be positive or negative. Therefore, we need to use the equilibrium value of f_r given by the indifference of R. This is done more easily using the posteriors ρ_l and ρ_r. Specifically, given $f_l = 1$ we have $\rho_l = \frac{1-p_l-p_r}{1-p_l}$. By the indifference of R, instead, we obtain:

$$\rho_r = 2c\frac{\sigma_r}{1-\sigma_r}$$

Plugging these into expression (11) we finally get:

$$1 + \frac{1-p_l-p_r}{4}\sigma_r\left(2c + \frac{1-p_l-p_r}{1-p_l}\right)$$

which is strictly increasing in σ_r. For the mirror case of $f_r = 1$ and $f_l \in (0,1)$, the procedure is analogous. We use $\rho_l = \frac{1-p_l-p_r}{1-p_r}$ and $\rho_r = 2c\frac{1-\sigma_r}{\sigma_r}$ to get:

$$1 + \frac{1-p_l-p-r}{4}(1-\sigma_r)\left(\frac{1-p_l-p_r}{1-p_r} + 2c\right)$$

Which is strictly decreasing in σ_r. ■

A further case to consider is one where both politicians are not disciplined by fact-checking, hence fact-checking only has an information role.

Lemma 3 *Suppose $f_l = f_r = 1$. In this case, welfare is strictly increasing in σ_r.*

Proof. In this case, fact-checking serves no discipline purpose: both politicians always make up a fact when lacking a valid one, just like in the absence of a fact-checker. Substituting $f_l = f_r = 1$ into equation (20), and recalling that $p_l > p_r$, we can see that welfare is increasing in σ_r. ■

The only remaining case is one in which both politicians fabricate with interior probability, i.e., $f_j \in (0,1)$ for both $j \in \{L, R\}$. In this case, welfare does not depend on σ_r, as the following lemma establishes.

Lemma 4 *Suppose that $f_r \in (0,1)$ and $f_l \in (0,1)$. In this interval, welfare does not depend on σ_r.*

Proof. Since both politicians are mixing when they do not have a fact, it must hold that:

$$\rho_r = 2c\frac{\sigma_r}{1-\sigma_r} \text{ and } \rho_l = 2c\frac{1-\sigma_r}{\sigma_r}$$

Substituting into equation (11) yields:

$$1 + \frac{1-p_l-p_r}{2}c$$

which does not depend on σ_r. ■

Having established how welfare responds for a given behavior of politicians, we now complete the picture by looking at what behavior of politicians is induced by fact-checking. To do so, we now need to connect σ_r to ρ_r, ρ_l, f_l and f_r.

Consider politician R first. From the comparison of the expected utility when fabricating a fake fact versus admitting not to have a valid one, resulting in the indifference condition (3), we obtain that, for $\sigma_r > \frac{\rho_r}{2c+\rho_r}$, politician R does not fabricate, i.e., $f_r = 0$. Substituting into ρ_r, this yields:

$$\sigma_r > \frac{1}{1+2c}$$

At the other opposite, for $\sigma_r < \frac{\rho_r}{2c+\rho_r}$, we have that $f_r = 1$, which yields $\rho_r = \frac{1-p_l-p_r}{1-p_r}$. Substituting, we obtain:

$$\sigma_r < \frac{1}{1+2c\chi_r} \text{ with } \chi_r := \frac{1-p_r}{1-p_l-p_r}$$

Finally, in order for R to choose $f_r \in (0,1)$, we need $\sigma_r = \frac{\rho_r}{2c+\rho_r}$. The same analysis can be done for L. We have $f_l = 0$ for

$$\sigma_l > \frac{1}{1+2c}$$

and $f_l = 1$ for:

$$\sigma_l < \frac{1}{1+2c\chi_l} \text{ with } \chi_l := \frac{1-p_l}{1-p_l-p_r} < \chi_r$$

and obviously $f_l \in (0,1)$ for $\sigma_r = \frac{\rho_l}{2c+\rho_l}$.

Since two thresholds determine the behavior of each politician and there are two politicians, we have four possible scenarios depending on the ranking of these thresholds. Therefore, the next step is to analyze the four possible scenarios and use the results in the previous four lemmas to determine the fact-checking allocation resulting in optimal welfare.

Case 1 : in this scenario, taking place for $c < \frac{1}{2}\frac{1}{\sqrt{\chi_r\chi_l}} = \frac{1}{2}\sqrt{\frac{(1-p_l-p_r)^2}{(1-p_r)(1-p_l)}}$ the ordering of thresholds on σ_r determining the behavior of politicians is the following:

$$\left(1-\frac{1}{1+2c}, 1-\frac{1}{1+2c\chi_l}, \frac{1}{1+2c\chi_r}, \frac{1}{1+2c}\right)$$

This means that at least one politician is fully undisciplined. Specifically, the behavior taking place in each interval is the following:

$$(\{f_r = 1, f_l = 0\}, \{f_r = 1, f_l \in (0,1)\}, \{f_r = 1, f_l = 1\}, \{f_r \in (0,1), f_l = 1\}, \{f_r = 0, f_l = 1\})$$

Case 2 : in this scenario, which takes place for $\frac{1}{2}\sqrt{\frac{(1-p_l-p_r)^2}{(1-p_r)(1-p_l)}} \leq c < \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_r}}$, the ordering of thresholds on σ_r determining the behavior of politicians is the following:

$$\left(1-\frac{1}{1+2c}, \frac{1}{1+2c\chi_r}, 1-\frac{1}{1+2c\chi_l}, \frac{1}{1+2c}\right)$$

This means that it is now possible to discipline both politicians at the same time in the interval $\sigma_r \in [\frac{1}{1+2c\chi_r}, 1-\frac{1}{1+2c\chi_l}]$. Specifically, the behavior taking place in each interval is the following:

$$\Big(\{f_r = 1, f_l = 0\}, \{f_r = 1, f_l \in (0,1)\},$$
$$\{f_r \in (0,1), f_l \in (0,1)\}, \{f_r \in (0,1), f_l = 1\}, \{f_r = 0, f_l = 1\}\Big)$$

Case 3 : in this scenario, which takes place for $\frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_r}} \leq c < \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, the ordering of thresholds on σ_r determining the behavior of politicians is the following:

$$\left(\frac{1}{1+2c\chi_r}, 1-\frac{1}{1+2c}, 1-\frac{1}{1+2c\chi_l}, \frac{1}{1+2c}\right)$$

This means that it is now possible to fully discipline L while at the same time exerting some discipline on R, in the interval $\sigma_r \in [\frac{1}{1+2c\chi_r}, 1 - \frac{1}{1+2c}]$. Specifically, the behavior taking place in each interval is the following:

$$\Big(\{f_r = 1, f_l = 0\}, \{f_r \in (0,1), f_l = 0\},$$
$$\{f_r \in (0,1), f_l \in (0,1)\}, \{f_r \in (0,1), f_l = 1\}, \{f_r = 0, f_l = 1\}\Big)$$

Case 4 : in this scenario, which takes place for $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, the ordering of thresholds on σ_r determining the behavior of politicians is the following:

$$\left(\frac{1}{1+2c\chi_r}, 1 - \frac{1}{1+2c}, \frac{1}{1+2c}, 1 - \frac{1}{1+2c\chi_l}\right)$$

This means that it is now additionally possible to fully discipline R while at the same time exerting some discipline on L, in the interval $\sigma_r \in [\frac{1}{1+2c}, 1 - \frac{1}{1+2c\chi_l}]$. Specifically, the behavior taking place in each interval is the following:

$$\Big(\{f_r = 1, f_l = 0\}, \{f_r \in (0,1), f_l = 0\},$$
$$\{f_r \in (0,1), f_l \in (0,1)\}, \{f_r = 0, f_l \in (0,1)\}, \{f_r = 0, f_l = 1\}\Big)$$

Having described what the four different scenarios are, we can now use Lemma 1 to Lemma 4 to determine the optimal σ_r. To visualize this, we will write 'up', 'down' or '=' depending on whether the preceding lemmas state that welfare increases or decreases in σ_r.

Case 1 : We obtain:

$$(up, down, up, up, down)$$

which implies that maximum welfare can be either at $\sigma_r = 1 - \frac{1}{1+2c}$, or at $\sigma_r = \frac{1}{1+2c}$.

Case 2 : We obtain:

$$(up, down, =, up, down)$$

which implies that maximum welfare can be either at $\sigma_r = 1 - \frac{1}{1+2c}$, or at $\sigma_r = \frac{1}{1+2c}$.

Case 3 : We obtain:

$$(up, up, =, up, down)$$

which implies that maximum welfare is at $\sigma_r = \frac{1}{1+2c}$, that is, the minimum intensity of fact-checking such that $f_r = 0$.

Case 4 : We obtain:

$$(up, up, =, down, down)$$

which implies that maximum welfare is achieved for all $\sigma_r \in \left[1 - \frac{1}{1+2c}, \frac{1}{1+2c}\right]$, that is, any σ_r such that both politicians are indifferent between fabricating and admitting not to have a valid fact.

To finish the proof, for cases 1 and 2 we need to compare the two candidates for maximum welfare, that is, $\sigma_r = 1 - \frac{1}{1+2c}$ and $\sigma_r = \frac{1}{1+2c}$, which implement respectively $(f_r = 1, f_l = 0)$ and $(f_r = 0, f_l = 1)$. Plugging these into expression (20) and recalling that $p_l > p_r$, we obtain that $\sigma_r = \frac{1}{1+2c}$ is optimal, that is, implementing $f_r = 0$ and $f_l = 1$ dominates the mirror case where $f_r = 1$ and $f_l = 0$.

To sum up, we have shown that welfare is maximized by choosing $\sigma_r = \frac{1}{1+2c}$ as long as we are in Case 1 through 3, that is, as long as $c < \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, whereas all $\sigma_r \in \left[1 - \frac{1}{1+2c}, \frac{1}{1+2c}\right]$ are equally optimal when in Case 4, that is whenever $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$. ■

Proposition 8
Proof. The condition $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$ determines whether an equilibrium in which both politicians are indifferent is socially optimal.

In the fully impartial lie-seeker case, the equilibrium is interior as long as $c > \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}$, which is strictly smaller than $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$. Hence, whenever interior is optimal, a lie-seeker is optimal.

In the asymmetric lie-spotting benefits equilibrium, the equilibrium is interior whenever the following condition is satisfied:

$$c \geq \frac{1}{2}\frac{1-p_l-p_r}{1-p_r}\frac{\Theta}{v}$$

where $\Theta \equiv \sqrt{c^2(v-1)^2+v} + c(v-1)$. First, notice that for $v > 1$, $\Theta < v$ and hence the threshold is smaller than the one for the fully impartial lie-seeker. Therefore, for $v > 1$ we have that the fact-checker with asymmetric lie-spotting benefits is optimal for $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, just like in the baseline model.

Second, notice that as $v \to 0$, the threshold for an interior equilibrium converges to $c \geq \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$. This can be seen using de l'Hopital rule on the ratio $\frac{\Theta}{v}$, which delivers a limit for $v \to 0$ of $\frac{1}{2c}$, and solving.

Third, we can easily verify numerically that the threshold described by $c \geq \frac{1}{2}\frac{1-p_l-p_r}{1-p_r}\frac{\Theta}{v}$ is decreasing in v, going from $c = \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$ when $v = 0$ to $c = \frac{1}{2}\frac{1-p_l-p_r}{1-p_l}$.

Therefore, whenever an interior equilibrium is optimal, a lie-spotter with asymmetric benefits delivers it, no matter what v is.

Consider now the second part of the proposition. From Proposition 4, a biased fact-checker implements an interior equilibrium whenever $c \in \left[\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}, \frac{\sqrt{1-b}}{2}\right]$, and the result follows immediately from the fact that $\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}} < \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$. Notice that for $c > \frac{\sqrt{1-b}}{2}$ a biased fact-checker is not optimal since it implements $f_l = 0$ and $f_r \in (0,1)$, which we know from Proposition 7 cannot be optimal. Finally, a biased fact-checker is never optimal if

$$\frac{\sqrt{1-b}}{2} < \frac{1}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$$

■

Proposition 9

Proof. Consider the biased fact-checker first. As $b \to 0$, we can immediately see from Proposition 4 that $\sigma_r \to \frac{1}{1+2c}$, which is socially optimal following Proposition 7.

Af far as the asymmetric lie-spotter is concerned, we can take the limit of the results in Proposition 3. In both subcases a) and b) of the proposition, we have that indeed σ_r converges to $\frac{1}{1+2c}$ as v goes to zero, which we know to be an optimal fact-checking strategy.

We now move to the second part of the proposition, stating that welfare decreases in v under the asymmetric lie-spotter, and it decreases in b under the biased fact-checker. Consider the asymmetric lie-spotter first. In order to derive the result, we need to combine the expression for welfare in Proposition 7 with the values of ρ_l and ρ_r from Proposition 3. Notice that there are two cases to consider, depending on whether shame cost lies below or above the threshold $\frac{1}{2}\frac{1-p_l-p-r}{1-p_l}\frac{\Theta}{v}$. In the former case, simple algebra gives us the following expression:

$$1 + \frac{(1-p_l-p_r)^2}{4}\frac{1}{1-p_l-p_r+2c(1-p_l+p_r(v-1))}\left(\frac{1-p_l-p_r}{1-p_l}+2c\right)$$

which is decreasing in v. Above the threshold, instead, welfare does not depend on v, since:

$$1 + \frac{1-p_l-p_r}{4}\left(\frac{\Theta}{\Theta+v}2c\frac{v}{\Theta} + \frac{v}{\Theta+v}2c\frac{\Theta}{v}\right) = 1 + \frac{1-p_l-p_r}{4}2c$$

Finally, the threshold $\frac{1}{2}\frac{1-p_l-p-r}{1-p_l}\frac{\Theta}{v}$ is decreasing in v, as it can be easily verified from the expression of its derivative:

$$\frac{2c\sqrt{v+c^2\left(v-1\right)^2}+\left(2c^2-1\right)v-2c^2}{2v^2\sqrt{v+c^2\left(v-1\right)^2}}$$

This means that by increasing v, first welfare is strictly decreasing in v, until the threshold $\frac{1}{2}\frac{1-p_l-p-r}{1-p_l}\frac{\Theta}{v}$ is reached, starting from which welfare is constant in v.

The procedure for the biased fact-checker case is analogous. First, notice that the two thresholds in Proposition 4 are decreasing in b, since they are proportional to $\frac{\sqrt{1-b}}{2}$. Below the first threshold $\frac{\sqrt{1-b}}{2}\sqrt{\frac{1-p_l-p_r}{1-p_l}}$, welfare is:

$$1+\frac{(1-p_l-p_r)^2}{4}\frac{1-b}{1-b+2c}\left(\frac{2c}{1-p_l-p_r}+\frac{1}{1-p_l}\right)$$

which is decreasing in b, since $\frac{1-b}{1-b+2c}$ is. Between the two thresholds, welfare is independent of b, since it equals $1+\frac{1-p_l-p_r}{4}2c$, and finally above the second threshold $\frac{\sqrt{1-b}}{2}$, welfare is:

$$1+\frac{(1-p_l-p_r)^2}{4}\frac{1-b}{1-b+2c}\left(\frac{2c}{1-p_l-p_r}+\frac{1}{1-p_l-p_r}\right)$$

which is again decreasing in b.

■

For Online Publication - Additional Materials

A Information Structure

A.1 Independent Facts Information Structure

In the baseline model we assume that whenever a politician lacks a valid fact, the other one must have one. In other words, there is no independent realization determining whether each politician has a fact, but rather an interdependence between politicians' 'types'. Such an interdependence, along with simplifying the analysis, is realistic for most policy issues, but it is nonetheless useful to analyze the case of politicians with independent facts.

In light of this, in this section we extend the model assuming that each politician has a valid fact with probability π_j. The draws are independent, hence we have four possible scenarios: both have a valid fact, with probability $\pi_l\pi_r$, only L (resp. R) has a fact, with probability $\pi_l(1-\pi_r)$ (resp. $\pi_r(1-\pi_l)$) and finally no politician has a fact, with probability $(1-\pi_l)(1-\pi_r)$. Everything else is identical to the baseline model. In particular, the fact-checker is an unbiased lie-spotter, and similarly to the baseline model, we assume that $\pi_l > \pi_r$ (the analysis, however, goes through also under $\pi_l = \pi_r$).

Let us start from the indifference condition of the fact-checker:

$$Prob(L \text{ fabricated}|\text{both report}) = Prob(R \text{ fabricated}|\text{both report}) \tag{21}$$

This can be written as:

$$\begin{aligned}&\frac{(1-\pi_l)f_l(\pi_r+(1-\pi_l)f_r)}{\pi_l(\pi_r+(1-\pi_r)f_r)+(1-\pi_l)f_l(\pi_r+(1-\pi_r)f_r)}\\ =&\frac{(1-\pi_r)f_r(\pi_l+(1-\pi_l)f_l)}{\pi_r(\pi_l+(1-\pi_l)f_l)+(1-\pi_r)f_r(\pi_l+(1-\pi_l)f_l)}\end{aligned} \tag{22}$$

Notice that, as usual, the denominators are the same, since they express the probability that both politicians report a fact. Therefore, we obtain:

$$(1-\pi_l)\pi_r f_l = (1-\pi_r)\pi_l f_r \tag{23}$$

Intuitively, the fact-checker's indifference implies that the joint probability that a candidate fabricates a fact while the other one has a valid one is the same for both candidates. Therefore, as in the baseline model, the candidate who is more likely to lack a valid fact must fabricate less in equilibrium, conditional on not having a fact.

We now introduce the following posterior probabilities:

$$\rho_l^{rF} = Prob(L \text{ has fact}|\text{R fails check}) \tag{24}$$
$$\rho_l^{rP} = Prob(L \text{ has fact}|\text{R passes check}) \tag{25}$$

Clearly, analogous definitions hold for R. The posteriors above write:

$$\rho_l^{rF} = \frac{\pi_l(1-\pi_r)f_r}{\pi_l(1-\pi_r)f_r + (1-\pi_r)(1-\pi_l)f_r f_l} \tag{26}$$
$$\rho_r^{lF} = \frac{\pi_r(1-\pi_l)f_l}{\pi_r(1-\pi_l)f_l + (1-\pi_l)(1-\pi_r)f_l f_r} \tag{27}$$
$$\rho_l^{rP} = \frac{\pi_l\pi_r}{\pi_r\pi_l + \pi_r(1-\pi_l)f_l} \tag{28}$$
$$\rho_r^{lP} = \frac{\pi_l\pi_r}{\pi_r\pi_l + \pi_l(1-\pi_r)f_r} \tag{29}$$

it is immediate to notice that, given the fact-checker's indifference, it must be that:

$$\rho_l^{rF} = \rho_r^{lF} \tag{30}$$
$$\rho_l^{rP} = \rho_r^{lP} \tag{31}$$

Moreover, simplifying the expressions above we also have that:

$$\rho_l^{rF} = \rho_l^{rP} = \frac{\pi_l}{\pi_l + (1-\pi_l)f_l} \tag{32}$$
$$\rho_r^{lF} = \rho_r^{lP} = \frac{\pi_r}{\pi_r + (1-\pi_r)f_r} \tag{33}$$

This is perhaps not very surprising given independence: the posterior probability is simply the posterior probability that the politician has a fact given that she reports having one. In other words, there is no cross-learning given independence.

We are now ready for the final step, concerning the incentives of politicians. Starting with L, and considering the case in which she has no fact, the payoff from admitting to not having a fact is:

$$(1-\pi_r)\left((1-f_r)\frac{1}{2} + f_r\frac{1}{2}\right) + \pi_r * 0 \tag{34}$$

The payoff from fabricating is instead:

$$(1-\pi_r)\left((1-f_r)\left(\frac{1}{2}-c\right)+f_r\left(\sigma_r\frac{1+\rho_l^{rF}}{2}+(1-\sigma_r)\left(\frac{1-\rho_r^{lF}}{2}-c\right)\right)\right)+ \\ +\pi_r\left(\sigma_r\frac{\rho_l^{rP}}{2}+(1-\sigma_r)\left(\frac{1-\rho_r^{lF}}{2}-c\right)\right) \tag{35}$$

Given these preliminaries, we can now pin down the possible equilibria of the game. First of all, notice that in equilibrium, it cannot be the case that both politicians always fabricate. If that were the case, given $\pi_l > \pi_r$, the fact-checker would always check R, and, just like in the baseline model, this would give R the incentive to deviate and admit not having a fact. Let us now consider the polar opposite is concerned, that is, a situation in which politicians never fabricate and $f_l = f_r = 0$.

No Fabrication: in this case, the fact-checker is by construction indifferent, since there are no lies to be detected. Using conditions (34) and (35), we obtain the following inequalities:

$$c > \frac{1}{2}\frac{\pi_r\sigma_r}{1-\pi_r\sigma_r} \qquad \text{for L} \tag{36}$$

$$c > \frac{1}{2}\frac{\pi_l(1-\sigma_r)}{1-\pi_l(1-\sigma_r)} \qquad \text{for R} \tag{37}$$

As a result of that, we have that a no fabrication equilibrium is feasible for c larger than the minimum of the maximum between the two thresholds above, which is achieved (since the two thresholds are respectively increasing and decreasing in σ_r) at the point where:

$$\frac{1}{2}\frac{\pi_r\sigma_r}{1-\pi_r\sigma_r} = \frac{1}{2}\frac{\pi_l(1-\sigma_r)}{1-\pi_l(1-\sigma_r)} \tag{38}$$

yielding:

$$\sigma_r = \frac{\pi_l}{\pi_l+\pi_r} \tag{39}$$

Substituting this value of σ_r in the condition for the incentives of politicians we have that, in order for a no fabrication equilibrium to exist, it must hold that:

$$c > \frac{1}{2}\frac{\pi_l\pi_r}{\pi_l+\pi_r-\pi_l\pi_r} \tag{40}$$

Interior Equilibrium: In order for politicians to be indifferent, (34) and (35) have to be equal, and the same holds for R, with analogous expressions. Exploiting the fact that the posteriors are all the same, and using the notation ρ^* as in the baseline model, we obtain, after some simplifications:

$$\begin{aligned}(1-\pi_r) = (1-\pi_r)(1+f_r\rho^*(2\sigma_r-1)) + \pi_r\left((1-\sigma_r)+\rho^*(2\sigma_r-1)\right) \\ - 2c\left((1-\pi_r)(1-f_r+f_r(1-\sigma_r)) + \pi_r(1-\sigma_r)\right)\end{aligned} \tag{41}$$

This can be further simplified to:

$$2c(1-\sigma_r(\pi_r+(1-\pi_r)f_r)) = \rho^*(2\sigma_r-1)(\pi_r+(1-\pi_r)f_r) + \pi_r(1-\sigma_r) \tag{42}$$

Using the fact that $\rho^* = \frac{\pi_r}{\pi_r+(1-\pi_r)f_r}$, we can finally write:

$$2c(1-\sigma_r(\pi_r+(1-\pi_r)f_r)) = \pi_r\sigma_r \tag{43}$$

or

$$\frac{2c-\pi_r\sigma_r}{2c\sigma_r} = (\pi_r+(1-\pi_r)f_r)\frac{\pi_l}{\pi_l} \tag{44}$$

Using the analogous equation that we can derive for L, that is:

$$\frac{2c-\pi_l(1-\sigma_r)}{2c(1-\sigma_r)} = (\pi_l+(1-\pi_l)f_l)\frac{\pi_r}{\pi_r} \tag{45}$$

and using again the indifference of the fact-checker, which makes the numerators of the right hand sides the same, we obtain:

$$\pi_l\frac{2c-\pi_r\sigma_r}{2c\sigma_r} = \pi_r\frac{2c-\pi_l(1-\sigma_r)}{2c(1-\sigma_r)} \tag{46}$$

and finally:

$$\sigma_r = \frac{\pi_l}{\pi_l+\pi_r} \tag{47}$$

Concerning the politicians' strategies, we obtain:

$$f_r = \frac{2c(\pi_l+\pi_r-\pi_l\pi_r)-\pi_l\pi_r}{2c\pi_l(1-\pi_r)} \tag{48}$$

$$f_l = \frac{2c(\pi_l+\pi_r-\pi_l\pi_r)-\pi_l\pi_r}{2c\pi_r(1-\pi_l)} \tag{49}$$

Finally, the posterior probability ρ^* becomes:

$$\rho^* = \frac{\pi_r}{\pi_r+(1-\pi_r)f_r} = \frac{\pi_r}{\pi_r+(1-\pi_r)\frac{2c(\pi_l+\pi_r-\pi_l\pi_r)-\pi_l\pi_r}{2c\pi_l(1-\pi_r)}} \tag{50}$$

which simplifies to:

$$\rho^* = \frac{2c}{2c-1} \tag{51}$$

Notice that in order for the posterior probability to be positive, we must have $c > \frac{1}{2}$. However, from the expressions for f_r and f_l we also get that $f_l < 1$ as long as:

$$c \leq \frac{\pi_r}{2} < \frac{1}{2} \tag{52}$$

As a result, the interior equilibrium does not exist.

Corner Equilibrium with $f_l = 1$: the last possible equilibrium is one in which $f_l = 1$ and $f_r \in (0,1)$. In this equilibrium, $\rho_l^{rF} = \rho_l^{rP} = \pi_l$ and from the indifference of the fact-checker, $\rho_r^{lF} = \rho_r^{lP} = \pi_l$, too. Using again (34) and (35) and imposing $f_l = 1$ in their analogues for R, we can derive that R is indifferent as long as:

$$\sigma_r = \frac{\pi_l}{\pi_l + 2c} \tag{53}$$

whereas the condition for L to strictly prefer fabricating is simplifies to, analogously to (43) for the interior equilibrium:

$$2c(1 - \sigma_r(\pi_r + (1-\pi_r)f_r)) < \pi_r \sigma_r \tag{54}$$

Using the value of $\sigma_r = \frac{\pi_l}{\pi_l + 2c}$ just derived from the incentives of R, as well as $f_r = \frac{(1-\pi_l)\pi_r}{(1-\pi_r)\pi_r}$, we obtain:

$$2c\left(1 - \sigma_r(\pi_r + (1-\pi_r)\frac{(1-\pi_l)\pi_r}{(1-\pi_r)\pi_l})\right) < \pi_r\sigma_r \tag{55}$$

which can be solved to get:

$$c < \frac{\pi_r}{2} \tag{56}$$

We summarize the results in this section in the following proposition (previous paragraphs serve as proof):

Proposition 10 *In the game with independent facts, the following holds:*

- *An equilibrium such that $f_l = 1$, $f_r = \frac{(1-\pi_l)\pi_r}{(1-\pi_r)\pi_l}$, $\sigma_r = \frac{2c}{2c+\pi_l}$ and $\rho^* = \pi_l$, exists for $c < \frac{\pi_r}{2}$. We call this equilibrium* corner.
- *An equilibrium such that $f_l = f_r = 0$, $\rho^* = 1$ and $\sigma_r = \frac{\pi_l}{\pi_l + \pi_r}$ exists for $c > \frac{1}{2}\frac{\pi_l\pi_r}{\pi_l+\pi_r-\pi_l\pi_r}$. We call this equilibrium* no fabrication.

Notice that since $\frac{\pi_r}{2} > \frac{1}{2}\frac{\pi_l \pi_r}{\pi_l + \pi_r - \pi_l \pi_r}$*, the two above equilibria coexist for* $c \in \left[\frac{1}{2}\frac{\pi_l \pi_r}{\pi_l + \pi_r - \pi_l \pi_r}, \frac{\pi_r}{2}\right]$.

To sum up, in the game with independent facts, politicians are either fully disciplined and never fabricate, or only R is disciplined, whereas L always fabricates when necessary. Unlike in the baseline model, an equilibrium in which both politicians fabricate does not exist.

Full discipline is easier with independent facts than with interdependence (the threshold is $\frac{1}{2}$ in the baseline model and $\frac{1}{2}\frac{\pi_l \pi_r}{\pi_l + \pi_r - \pi_l \pi_r}$ in the game with independent facts).

A.2 'Only One Fact or No Facts' Information Structure

Suppose now that the availability of facts is not independent, but that instead either only one politician has a fact, or none has one. Like in the baseline model, p_j denotes the probability that only politician j has a valid fact. All other ingredients of the model remain unchanged.

In this setup, an equilibrium with full discipline exists for all parameters. The intuition is the following: suppose that politicians are fully disciplined. In equilibrium, at most one politician reports a fact, and therefore the fact-checker never faces a trade-off in terms of whom to check. If a politician deviated and reported a fabricated fact, there would be two possibilities. The first is that he is the only politician to report, in which case the fact-checker would always spot his lie, leading to the shame cost c; the second is that both politicians report following the deviation. However, also in that case there is no upside to fabrication: if the fabricating politician is checked, she ends up losing the election and paying the shame cost, whereas if the other politician is checked, voters conclude that the other report must be false and elect the one whose fact was confirmed to be true.

Note that this game may also have another PBE in which both politicians fabricate with some probability and in which the fact-checker faces a trade-off, but fabrication is never the unique outcome of the game.

B No fact-checking benchmark

In this appendix, we solve the model without the fact-checker. We have a cheap talk game where the unique PBE is as follows:

Lemma 5 *Without fact-checking, politicians always report a false fact conditional on lacking a true one:* $f_l = f_r = 1$*. Moreover, politician* L *is elected with probability* $\frac{1+p_l-p_r}{2}$ *and voters'* *expected welfare is*

$$1 + \frac{(1 - p_l - p_r)^2}{4} - p_l p_r$$

We solve the game backward to obtain Lemma 5. Voters elect L if:

$$\Pr(L \text{ has a fact}) + \epsilon \geq \Pr(R \text{ has a fact}).$$

Given that $\epsilon \sim U[-1, 1]$, L is elected with probability

$$\Pr(L \text{ is elected}) = \frac{1 - \Pr(R \text{ has a fact}) + \Pr(L \text{ has a fact})}{2}.$$

Recall a politician having a fact always reports it. Hence, we focus on a politician lacking a fact. If she admits lacking a fact, voters understand that only the other politician has a fact and elect her. As a result, it is a dominant strategy to report facts, and politicians always do so in equilibrium. Hence, voters' posterior beliefs are equal to their priors: $\Pr(L \text{ has a fact}) = 1 - p_r$ and $\Pr(R \text{ has a fact}) = 1 - p_l$. Hence, L is elected with probability $(1 + p_l - p_r)/2$.

Turning to voters' welfare, R is elected with probability $(1 - p_l + p_r)/2$, in which case voters get 1 with probability $1 - p_l$ because R has a fact and 0 otherwise. When L is elected, voters get 1 with probability $1 - p_r$ and ϵ. As L is elected when $\epsilon \geq p_r - p_l$, the expected value is $(1 + p_r - p_l)/2$. Hence, voters' expected welfare is

$$\frac{1 - p_l + p_r}{2} \cdot (1 - p_l) + \frac{1 + p_l - p_r}{2} \cdot \Big(\frac{1 + p_r - p_l}{2} + 1 - p_r\Big),$$

which simplifies as in Lemma 5.