

Peer Review Information

Journal: Nature Human Behaviour

Manuscript Title: The supply and demand of news during COVID-19 and assessment of questionable sources production

Corresponding author name(s): Pietro Gravino

Reviewer Comments & Decisions:

Decision Letter, initial version:
--

12th July 2021

Dear Dr Gravino,

Thank you once again for your manuscript, entitled "Assessing disinformation through the dynamics of supply and demand in the news ecosystem", and for your patience during the peer review process.

Your Article has now been evaluated by 3 referees. You will see from their comments copied below that, although they find your work of potential interest, they have raised quite substantial concerns. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version if you are willing and able to fully address reviewer and editorial concerns.

We hope you will find the referees' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the referees again in the absence of major revisions. We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

To guide the scope of the revisions, the editors discuss the referee reports in detail within the team, including with the chief editor, with a view to (1) identifying key priorities that should be addressed in revision and (2) overruling referee requests that are deemed beyond the scope of the current study. We hope that you will find the prioritised set of referee points to be useful when revising your study. Please do not hesitate to get in touch if you would like to discuss these issues further.

* In your revision, please carefully address all reviewers' concerns. In particular, Reviewer 1 and Reviewer 2 raise concerns about the characterization of the existing literature and the contribution your study makes. Please follow the suggestion of Reviewer #2 to better discuss the existing literature on news supply and demand, and ensure that the conceptual contribution your study makes is clear. Reviewer 3 raises concerns about the mechanism at work and the possibility of alternative explanations for the results. Please thoroughly address these points, ensuring that the definition of "fake news" is clear and carrying out the additional analyses recommended by Reviewer #3 to compare equivalent types of media (e.g. online vs print), formalize the comparisons of model coefficients as hypothesis testing, and rule out alternative explanations. Reviewer 1 also raises concerns about the validity of Google Trends data as a proxy for news demand, and we ask that you provide additional evidence to support the use of this measure.

* In addition to these concerns raised by the reviewers, from the editorial perspective, we are concerned about the generalizability of your findings and hence how strongly supported your hypothesis is given that your analyses focus on a single topic (COVID-19). We therefore ask that you replicate your main findings in at least one other topic beyond COVID-19.

Finally, your revised manuscript must comply fully with our editorial policies and formatting requirements. Failure to do so will result in your manuscript being returned to you, which will delay its consideration. To assist you in this process, I have attached a checklist that lists all of our requirements. If you have any questions about any of our policies or formatting, please don't hesitate to contact me.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. We understand that the COVID-19 pandemic is causing significant disruptions which may prevent you from carrying out the additional work required for resubmission of your manuscript within this timeframe. If you are unable to submit your revised manuscript within 6 months, please let us know. We will be happy to extend the submission date to enable you to complete your work on the revision.

With your revision, please:

- Include a "Response to the editors and reviewers" document detailing, point-by-point, how you addressed each editor and referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be used by the editors to evaluate your revision and sent back to the reviewers along with the revised manuscript.
- Highlight all changes made to your manuscript or provide us with a version that tracks changes.

Please use the link below to submit your revised manuscript and related files:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

Thank you for the opportunity to review your work. Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Sincerely,

Arunas Radzvilavicius, PhD
Editor
Nature Human Behaviour

Reviewer expertise:

Reviewer #1: misinformation, computational social science

Reviewer #2: communication, disinformation, news ecosystems

Reviewer #3: communication, political science, data science

REVIEWER COMMENTS:

Reviewer #1:

Remarks to the Author:

The authors proposed a framework based on news supply and demand to provide one potential explanation for why people share misinformation. In particular they combined news supply as represented by an Italian news database and demand as captured by Google Trends data for Italy.

While the authors used an interesting methodology to study an important issue in today's society, the paper suffers from fundamental shortcomings in its current form and may not live to the standard of Nature Human Behaviour. The major shortcoming of the paper is the lack of contribution to the theoretical understanding of human behaviour in broader context of consumption and transmit misinformation. It is not clear how findings in terms of macro level analysis of news supply and demand for a particular case advances the current state of the literature.

Additionally, there is much room for improvement in other aspects of the manuscript: (1) the abstract can be improved to clear the contribution of the paper and the method used (2) other reasons why people share misinformation could be discussed (e.g., motivated reasoning, inattention, confusion see Pennycook et al 2021 Nature for related references). (3) Google Trend as a measure could be validated for the purpose of this work. (4) figure 4 could be improved in quality (it is currently hard to read).

Reviewer #2:

Remarks to the Author:

1. Key results.

This paper provides a broader perspective on the news and disinformation ecosystem, by empirically considering both the producers and the consumers of fake news and normal news. This connectivity is an outstanding feature.

2. Validity.

I am not a statistician, so am not best placed to judge this. However, I did not detect any flaws that would prohibit publication.

3. Originality and significance

This paper offers an evidence-based advance in knowledge, in examining an Italian news database (partially annotated for fake news) and news demand (as captured through the Google Trends data for Italy), finding that disinformation is extremely reactive to people's interests and tends to thrive when there is a mismatch between what people are interested in and what news outlets provide. Intuitively, we know this as disinformation researchers, but it is excellent to prove this.

It offers methodological novelty in demonstrating that a suitably defined index can assess the level of disinformation only based on available volumes of news and searches. This is likely to influence the methodological focus of the disinformation field.

The results are of interest to the wider field of information studies, and of immediate interest to the burgeoning field of disinformation and also to journalism studies. It will be of interest to policymakers and governance actors in the realms of disinformation and hybrid threats as well.

4. Data & methodology

The two databases consulted are a good proxy of demand and supply for information. The rank comparison of coronavirus related key words is impactful.

5. Appropriate use of statistics and treatment of uncertainties:

I am not a statistician. However, my basic knowledge of the area allowed me to follow the argument.

6. Custom code

I am not qualified to judge this.

7. Conclusions:

p. 10 - In the discussion, the authors speculate that the more reactive nature of Fake News compared to General News could be a consequence of the different production environments of General News and Fake News, and also speculate that this could be down to different sizes of communities producing these types of news. However, it is unclear why community size would be a good explanation. For instance, what about previous studies showing that Fake News (and false information and affective information) are more viral than other sorts of information on social media, and that Fake News is designed for virality (in its affective tenor that exploits the algorithms of social media)?

8. Suggested improvements

p.2 the definition of Infodemic is not accurate. My understanding of this term is that there is a glut of information, some of it true and some false, so that people do not know what to believe.

p.2 The paragraph characterising the literature on news supply is bizarre. I assume that the authors are referring to the treatment of this field by a sub-discipline or by a particular methodological perspective (I am unsure which ones) – but nowhere do I see any reference to the very many Journalism Studies academic journals, where the topic of news supply is treated thoroughly, from all sorts of angles. The authors must clarify and specify how they are characterising the field of academic scholarship on news supply.

p.2 The paragraph characterising the literature on news demand similarly ignores a wealth of research on news audiences from Journalism Studies, including more recent literature why they may avoid news and why they may be attracted to disinformation. It would seem pertinent to mention this. If the paper is specifically about methodological ways of tracking news demand, then this should be clarified, while also acknowledging the wealth of literature on news demand more generally.

Half way down p.5, there seems to be a missing phrase, as the sentence starts mid-sentence.

9. References: Does this manuscript reference previous literature appropriately? If not, what references should be included or excluded?
see my comments on (8).

10. Clarity and context

Line 1 of the abstract states: 'Social dialogue, the foundation of our democracies, is currently threatened by disinformation and partisanship'. I disagree about the 'partisanship' aspect. Partisanship is fine, and a normal state of affairs, as the political science literature will show. What is not fine is 'polarisation' – especially affective polarisation – which harms social dialogue in many ways. There is a literature on this. I suggest that the authors remove reference to the word 'partisanship' in the abstract.

The rest of the paper is clear.

11. Please indicate any particular part of the manuscript, data, or analyses that you feel is outside the scope of your expertise, or that you were unable to assess fully.
Statistical analysis (I am not a statistician or quantitative expert).

Reviewer #3:
Remarks to the Author:

This paper uses an unprecedented amount of data to investigate the interplay of supply and demand of news, and how that relationship changes based on the trustworthiness of sources. The authors primarily make use of the example of the COVID-19 pandemic as a topic of inquiry. Their primary findings are that (1) Google searches (their proxy for demand) are temporally prior to news content; (2) "Fake News" sources have less inertia and are more impacted by demand; (3) Semantically, "Fake News" is more correlated with searches than general news; (3) The availability of fake news can be predicted temporally by finding when the relationship between Google searches and general news is disturbed (both in terms of autoregression and semantics).

I think this paper is important. As the authors point out, previous studies of news demand are limited by the availability of choices. In the lab researchers define the choice set for participants. In observational tracking studies demand is constrained by what the news itself chooses to print (an excellent point I have never thought about). Their method of using Google trends is a good one. In addition, the analysis in the paper was competently done.

I have one fairly significant question/suggestion, and a few additional minor points.

The thing I struggled with most in the paper was: what is being defined as Fake News here, and how does that change the conclusions of the paper? The manuscript is very short on details. In particular, it's unclear if the fake news consists of individual articles or entire sources. On Pg 6 it states "... we extend our analysis to the news items that were annotated as false or misleading." But on Pg13: "Moreover, the data had a specific annotation for 'fake news' sources. The list of disinformation sources, vetted by the Authority, is provided by fact-checking organizations like bufale.net and butac.it. It had already been employed for other studies on disinformation." In one case the authors talk about fake news items, in the other case, they talk about fake news sources.

The difference matters a great deal. This isn't just a question of "framing", but really one of measurement: the paper claims to be telling us something about disinformation, but is this measure really capturing disinformation? While I trust that these two websites have done some due diligence in identifying sources that, in general, might promote some degree of disinformation, this is no guarantee that all (or even most) of the information presented in these sources is incorrect or fake. So the authors are claiming that the relationships described here are between news demand and specific, verified to be incorrect, information. But what they are really describing is the relationship between news demand and content in particular types of sources. The authors need to be much more transparent about what is happening here. The language of the paper must be changed to indicate that what is being described here is how news demands predicts the amount of coverage in different types of sources, not the production of specifically "fake" news items. After these changes it's not clear to me the degree to which this can be claimed to be a study about "disinformation" explicitly. That the "Fake news" in this study is really "sources that have been deemed to produce some fake news" brings up another potential issue: what is the media of these sources, and how does it compare to the distribution of media in the overall sample? Specifically, I'm not sure what to make of the results if, for example, most of the "fake news" sources are websites but the "General News" category contains a mixture of websites and traditional news sources. The authors hint that this might be the case on page 10 "[General News] are mainly produced by a large and well-established community of professional journalists while the [Fake News] are the outcome of by a scattered multitude of small, unorganized actors."

The finding that "Fake News" sources have less inertia and are more responsive to demand might just because news websites, in general, have less inertia and are more responsive to demand. Can the authors determine the degree to this is true? Can the Fake News sources be compared to a group of sources that are comparable? (Or can controls be put in?) I am worried that people reading this article

will conclude that specific disinformation follows the supply-demand dynamic described here when really it might be that internet sources generally follow this pattern, and the outlets that produce fake news happen to be online.

One small comment I had was for Table 2, which compares the time-series coefficients for General and Fake news. The authors are comparing the magnitude of the coefficients across columns and making claims that coefficients are smaller or larger. This is implicitly a hypothesis test, and the authors should formalize it as such. The difference between these coefficients is a random variable with a distribution that needs to be accounted for. I would suggest performing a bootstrap analysis for the difference between coefficients in rows to determine the degree to which the differences are statistically significant.

Author Rebuttal to Initial comments

Response to the reviewers

Response to the rev. 1

“The authors proposed a framework based on news supply and demand to provide one potential explanation for why people share misinformation. In particular they combined news supply as represented by an Italian news database and demand as captured by Google Trends data for Italy. While the authors used an interesting methodology to study an important issue in today’s society, the paper suffers from fundamental shortcomings in its current form and may not live to the standard of Nature Human Behaviour. The major shortcoming of the paper is the lack of contribution to the theoretical understanding of human behaviour in broader context of consumption and transmit misinformation.”

Reviewer 1 correctly remarks that our study focuses on news supply, i.e., the ensemble of news produced in the country, and news demand, i.e., the request for information by the public, which are both products of human nature. The paper aims to understand their relationships and the difference between information and disinformation production in this system. As a result, we managed to accurately quantify the concentration of disinformation production about COVID-19 in the observed timeframe. Though they represent interesting hints for future works, the distribution, consumption, transmission and sharing on social networks is outside the scope of our paper.

“It is not clear how findings in terms of macro level analysis of news supply and demand for a particular case advances the current state of the literature.”

We politely disagree with Reviewer 1. The significance, importance and originality of our paper were highlighted by Reviewers 2 and 3. In particular, we contend the assessment of disinformation concentration without requiring fake news annotation, thanks to the proposed Combined Index, is a valuable result, whose impact is exemplified on the COVID-19 subject, a case of particular relevance.

We also highlight that, to the best of our knowledge, our results and the methodology we defined in the paper are not present in the current state of the literature.

“Additionally, there is much room for improvement in other aspects of the manuscript: (1) the abstract can be improved to clear the contribution of the paper and the method used”

Thanks to these indications, we have improved the abstract accordingly to clarify the contribution and the methodology used.

“(2) other reasons why people share misinformation could be discussed (e.g., motivated reasoning, inattention, confusion see Pennycook et al 2021 Nature for related references).”

Following the suggestion, we have expanded the literature review about disinformation, including the suggested reference and many others.

“(3) Google Trend as a measure could be validated for the purpose of this work.”

We extended the discussion on state of the art by referencing previous studies on Google Searches that validate the use of Google Trends for this work. In particular, the work of Broder 2002 categorizes the intents underlying the queries in Google Search engine (seeking for information, website access, or a transaction), and Waller 2011 elaborates on that work by defining the queries that, based on the input keywords, are made to search for information.

“(4) figure 4 could be improved in quality (it is currently hard to read).”

Since there is no Fig. 4 in the paper, we understood that this point refers to Supplementary Figure 1 (the fourth figure of the main text). We remade that plot by increasing the font size to improve its readability.

Response to the rev. 2

“Remarks to the Author:”

“1. Key results.

This paper provides a broader perspective on the news and disinformation ecosystem, by empirically considering both the

producers and the consumers of fake news and normal news. This connectivity is an outstanding feature.”

“2. Validity.

I am not a statistician, so am not best placed to judge this. However, I did not detect any flaws that would prohibit publication.”

“3. Originality and significance

This paper offers an evidence-based advance in knowledge, in examining an Italian news database (partially annotated for fake news) and news demand (as captured through the Google Trends data for Italy), finding that disinformation is extremely reactive to people’s interests and tends to thrive when there is a mismatch between what people are interested in and what news outlets provide. Intuitively, we know this as disinformation researchers, but it is excellent to prove this. It offers methodological novelty in demonstrating that a suitably defined index can assess the level of disinformation only based on available volumes of news and searches. This is likely to influence the methodological focus of the disinformation field. The results are of interest to the wider field of information studies, and of immediate interest to the burgeoning field of disinformation and also to journalism studies. It will be of interest to policymakers and governance actors in the realms of disinformation and hybrid threats as well.”

“4. Data & methodology

The two databases consulted are a good proxy of demand and supply for information. The rank comparison of coronavirus related key words is impactful.”

“5. Appropriate use of statistics and treatment of uncertainties:

I am not a statistician. However, my basic knowledge of the area allowed me to follow the argument.”

“6. Custom code

I am not qualified to judge this.”

We thank Reviewer 2 for the points raised so far and deeply appreciate comment number 3. We believe there is nothing more to add from our side.

“7. Conclusions:

p. 10 - In the discussion, the authors speculate that the more reactive nature of Fake News compared to General News could be a consequence of the different production environments of General News and Fake News, and also speculate that this could be down to different sizes of communities producing these types of news. However, it is unclear why community size would be a good explanation. For instance, what about previous studies showing that Fake News (and false information and affective information) are more viral than other sorts of information on social media, and that Fake News is designed for virality (in its affective tenor that exploits the algorithms of social media)? ”

Thanks to this remark, we improved the speculation in the Discussion to explain our hypothesis better. Virality differences between general news and disinformation are the subject of a fascinating open debate in the literature

(see Vosoughi 2018 and Cinelli 2021), which we preferred not to address in this paper. Moreover, this pertains to the diffusion of news (e.g. through social media sharing) rather than news supply or demand, which are our paper’s focus. We believe that the introduction of General News (and Fake News) diffusion in our analysis is an interesting research direction to close the loop between supply and demand. It will be the object of future work.

“8. Suggested improvements

p.2 the definition of Infodemic is not accurate. My understanding of this term is that there is a glut of information, some of it true and some false, so that people do not know what to believe.”

We have edited our definition of Infodemic based on the suggestions and corrected it as such: [...] the term “Infodemic” made a comeback from the SARS epidemic of 2003 [45], to describe the spreading of many pieces of information about the virus SARS-CoV-2, some correct others false, to the extent that overwhelms peoples and hinders their understanding of the phenomenon. *ii*

“p.2 The paragraph characterising the literature on news supply is bizarre. I assume that the authors are referring to the treatment of this field by a sub- discipline or by a particular methodological perspective (I am unsure which ones) – but nowhere do I see any reference to the very many Journalism Studies academic journals, where the topic of news supply is treated thoroughly, from all sorts of angles. The authors must clarify and specify how they are characterising the field of academic scholarship on news supply.”

We are thankful to Reviewer 2 for raising this (and the following) point about a part of the literature we overlooked. We have extended the bibliography (on page 2) to include works from journalism studies that we considered relevant for our work while attempting to convey the extent of this research field:

The news supply can be identified with the overall news production, mainly of officially recognised newspapers. As such, the literature concerning journalism and media information is extensive, and investigated: news linguistics [19, 9]; journalism economics [46]; information coverage [47, 53], often focusing on particular countries or topics [65]; content of news [34] its quality assessment [16] and delivery [17], considering different media sources [33]. Other works investigated the impacts of news and its consumption on, for example, reading behaviour [55], finance [21], and political opinions [28].ⁱⁱ

“p.2 The paragraph characterising the literature on news demand similarly ignores a wealth of research on news audiences from Journalism Studies, including more recent literature why they may avoid news and why they may be attracted to disinformation. It would seem pertinent to mention this. If the paper is specifically about methodological ways of tracking news demand, then this should be clarified, while also acknowledging the wealth of literature on news demand more generally.”

Similarly, as for the previous point, we have extended the bibliography to include journalism studies on news demand and appeal to misinformation. We unbundled the part discussing current methods for tracking news demand. Moreover, we have broadened the discussion about Google Trend’s index, as a proxy for news demand, by referencing past works on the analysis of Google’s searches.

“Half way down p.5, there seems to be a missing phrase, as the sentence starts mid-sentence.”

This has been corrected by replacing the full stop at the end of equation (1) with a comma.

“10. Clarity and context

Line 1 of the abstract states: ‘Social dialogue, the foundation of our democracies, is currently threatened by disinformation and partisanship’. I disagree about the ‘partisanship’ aspect. Partisanship is fine, and a normal state of affairs, as the political science literature will show. What is not fine is ‘polarisation’ – especially affective polarisation – which harms social dialogue in many ways. There is a literature on this. I suggest that the authors remove reference to the word ‘partisanship’ in the abstract.”

Following Reviewer 2’s advice, we have removed the term “partisanship” from that sentence and replaced it with the term “polarisation”.

Response to the rev. 3

“Remarks to the Author:

This paper uses an unprecedented amount of data to investigate the interplay of supply and demand of news, and how that relationship changes based on the trustworthiness of sources. The authors primarily make use of the example of the COVID-19 pandemic as a topic of inquiry. Their primary findings are that (1) Google searches (their proxy for demand) are temporally prior to news content; (2) “Fake News” sources have less inertia and are more impacted by demand; (3) Semantically, “Fake News” is more correlated with searches than general news; (3) The availability of fake news can be predicted temporally by finding when the relationship between Google searches and general news is disturbed (both in terms of autoregression and semantics). I think this paper is important. As the authors point out, previous studies of news demand are limited by the availability of choices. In the lab researchers define the choice set for participants. In observational tracking studies demand is constrained by what the news itself chooses to print (an excellent point I have never thought about). Their method of using Google trends is a good one. In addition, the analysis in the paper was competently done.”

We are deeply thankful to Reviewer 3 for correctly summarizing our work and highlighting its importance and the details motivating our methodology.

“I have one fairly significant question/suggestion, and a few additional minor points. The thing I struggled with most in the paper was: what is

being defined as Fake News here, and how does that change the conclusions of the paper? The manuscript is very short on details. In particular, it's unclear if the fake news consists of individual articles or entire sources. On Pg 6 it states "... we extend our analysis to the news items that were annotated as false or misleading." But on Pg13: "Moreover, the data had a specific annotation for 'fake news' sources. The list of disinformation sources, vetted by the Authority, is provided by fact-checking organizations like bufale.net and butac.it. It had already been employed for other studies on disinformation." In one case the authors talk about fake news items, in the other case, they talk about fake news sources."

Thanks for remarking on the mistake. We corrected the definition based on the suggestions at page 6 and fixed it as such: [...] To this end, we extended our analysis to the news items that were produced by sources annotated as false or misleading

"The difference matters a great deal. This isn't just a question of "framing", but really one of measurement: the paper claims to be telling us something about dis-information, but is this measure really capturing disinformation? While I trust that these two websites have done some due diligence in identifying sources that, in general, might promote some degree of disinformation, this is no guarantee that all (or even most) of the information presented in these sources is incorrect or fake. So the authors are claiming that the relationships described here are between news demand and specific, verified to be incorrect, information. But what they are really describing is the relationship between news demand and content in particular types of sources. The authors need to be much more transparent about what is happening here. The language of the paper must be changed to indicate that what is being described here is how news demands predicts the amount of coverage in different types of sources, not the production of specifically "fake" news items. After these changes it's not clear to me the degree to which this can be claimed to be a study about "disinformation" explicitly."

By highlighting this ambiguity, this comment stimulated us to improve the language of the paper to clarify that what we analyse is not disinformation itself. Our work is focused specifically on production mechanisms, i.e., the disinformation and General News supply and demand dynamics. We modified the language of the paper accordingly to this suggestion.

“That the “Fake news” in this study is really “sources that have been deemed to produce some fake news” brings up another potential issue: what is the media of these sources, and how does it compare to the distribution of media in the overall sample? Specifically, I’m not sure what to make of the results if, for example, most of the “fake news” sources are websites but the “General News” category contains a mixture of websites and traditional news sources. The authors hint that this might be the case on page 10 “[General News] are mainly produced by a large and well-established community of professional journalists while the [Fake News] are the outcome of by a scattered multitude of small, unorganized actors.” The finding that “Fake News” sources have less inertia and are more responsive to demand might just because news websites, in general, have less inertia and are more responsive to demand. Can the authors determine the degree to this is true? Can the Fake News sources be compared to a group of sources that are comparable? (Or can controls be put in?) I am worried that people reading this article will conclude that specific disinformation follows the supply-demand dynamic described here when really it might be that internet sources generally follow this pattern, and the outlets that produce fake news happen to be online.”

The point raised is an interesting one. We assumed that General News is a product of officially recognised journalists belonging to professional associations and, likely, also to newspapers editorial staff, irrespective of the media used to distribute their news production (paper, internet, etc.). We agree that a difference in the diffusion media (online vs paper) is a valid hypothesis that could explain the observed behaviours of news production. We leveraged the annotations about the type of source for all the news outlets to test this hypothesis. We identified the subset of online news suppliers and analysed their behaviour through our improved modelling. The results, reported in Supplemental Tab. S1, show how Fake News suppliers, which mainly are online, display different dynamics from Online News suppliers. The latter are meaningfully different also from General News suppliers, even if the α_1 coefficient (the Inertia) is remarkably compatible. This result contradicts the hypothesis. We integrated the main text to mention this result.

“One small comment I had was for Table 2, which compares the time-series coefficients for General and Fake news. The authors are comparing the magnitude of the coefficients across columns and making claims that coefficients are smaller or larger. This is implicitly a hypothesis test, and the authors should formalize it as such. The difference between these coefficients is a random variable with a distribution that needs to be accounted for. I would suggest performing a

bootstrap analysis for the difference between coefficients in rows to determine the degree to which the differences are statistically significant.”

We agree with the comment raised by Reviewer 3. We have, accordingly, implemented the suggested bootstrap analysis to compare the coefficients reported in Table 2 statistically. The details of the methodology are reported in Materials and Methods. Using 10^5 bootstrap samples (i.e., sampled with replacement), the Fake News supply inertia resulted significantly lower than General News supply inertia. The results of the bootstrap are reported in Supplemental Fig. S6. We also integrated the main text to mention the result.

Decision Letter, first revision:

10th December 2021

Dear Dr Gravino,

RE: "Assessing disinformation through the lens of news supply and demand during the COVID-19 outbreak"

Thank you for submitting your revised manuscript and for all your work on the revision.

Although your manuscript has been revised in response to reviewer comments, it does not fully comply with our editorial policies. Before we can send the manuscript back to our reviewers, we ask that you revise it to ensure that it complies fully with our policies:

* Please report full statistical results supporting your claims regarding differences in coefficient estimates (as in formal hypothesis testing, as requested by one of the reviewers), i.e. report full statistics for your tables in both the main text and the supplement, including exact p values for all reported coefficients alpha and beta (except when the p values are lower than .001, in which case they should report them as lower than .001, rather than the exact value).

* Please merge your response to the editors and to reviewers files into a single file, so that reviewers could also see your responses to editorial requests.

If you are uncertain as to how to address any of the points in the checklist, please don't hesitate to contact me.

Please use the link below to submit your revised manuscript and related files:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

Thank you in advance for attending to these requests and I look forward to receiving your revised manuscript.

Sincerely,

Arunas Radzvilavicius, PhD
Editor
Nature Human Behaviour

Author Rebuttal, first revision:

Response to the editors and reviewers

Response to the editors

“In particular, Reviewer 1 and Reviewer 2 raise concerns about the characterization of the existing literature and the contribution your study makes. Please follow the suggestion of Reviewer 2 to better discuss the existing literature on news supply and demand, and ensure that the conceptual contribution your study makes is clear.”

Following the comments of Reviewers 1 and 2, we have expanded the referenced literature to include for journalism studies about news demand, news supply, and disinformation.

“Reviewer 3 raises concerns about the mechanism at work and the possibility of alternative explanations for the results. Please thoroughly address these points, ensuring that the definition of “fake news” is clear and carrying out the additional analyses recommended by Reviewer 3 to compare equivalent types of media (e.g., online vs print), formalize the comparisons of model coefficients as hypothesis testing, and rule out alternative explanations.”

Following Reviewer 3’s comments, we clarified that the dynamic of news under analysis concerns suppliers, and removed any ambiguity, throughout the paper, about this definition. We have also implemented a statistical hypothesis test, based on bootstrap, to compare coefficients from different models. Finally, we included the comparison of news from online sources with General News and Fake News supply (see Table S1), showing that General News from online sources and Fake News, which are also mainly online, behave differently and therefore the diversity observed between General News and Fake News cannot be explained by a difference in the medium of news distribution.

“Reviewer 1 also raises concerns about the validity of Google Trends data as a proxy for news demand, and we ask that you provide additional evidence to support the use of this measure.”

To back up the use of Google Trends indexes as a proxy for news demand, we have referenced past works on categorising user intentions when searching on Google (Broder 2002; Waller 2011). By retracing their work, the keywords defining the queries in our study can be assumed to be motivated by an informative end, as opposed to navigational (find a website) or transactional (perform some web-mediated activity such as online shopping) purposes, which are unlikely.

“We are concerned about the generalizability of your findings and hence how strongly supported your hypothesis is given that your analyses focus on a single topic (COVID-19). We therefore ask that you replicate your main findings in at least one other topic beyond COVID-19.”

Thanks a lot for raising this important point about the generality of our results. Since we are aware that full generality cannot be claimed at this stage, we modified the text to highlight what can be claimed as general results and which results will need further investigations. A bit of context is in order here to explain the situation better.

The General News dataset we adopted for our work is a unique, multi-modal and incredibly rich resource to analyse the entire landscape of Italian news production from December 2019 to August 2020. This particular time frame was

concurrent with the COVID-19 outbreak and global lockdowns. It isn't easy to assess whether the COVID-19 period only amplified the general dynamics of disinformation or transformed it due to the exceptional character of the situation. In the first case, our results would be fully general, while in the second case, while one cannot rule that a similar dynamics is in place, one would need to investigate a comparable database for a pre-COVID period. What is well-known is that the pandemics triggered a deformation of the public discussion, with an overabundance of data about COVID-19 with a strong shadowing effect on other subjects. National lockdown restrictions impeded many newsworthy events that could be targeted by media attention. Such data unbalance, together with the scarcity of annotations for Fake News sources (see Table S1), enabled us to accurately analyse the relation between General News and Fake News only for the *coronavirus* keyword. In other words, showing that the same results hold for other topics in the dataset is made difficult

due to Fake News annotations scarcity. More importantly, it would also not prove the generalizability of our results because of the particular moment the dataset captures. We believe that the results observed for the *coronavirus* keyword could be nevertheless a powerful asset to improve the quality of news in this crucial domain.

As for the generalizability of our results, we highlighted, throughout the paper, that the analysis on "Dynamics of news supply and demand" (in the Results section) holds for all keywords under analysis and not only for the *coronavirus* keyword.

Moreover, we provide some evidence of generalizability by aggregating together different keywords in a synthetic macro-subject (see the end of the section "Independent detection of Fake News concentration" in Results), to compensate for the reduced size of annotations for Fake News sources. This aggregation leads to promising results, though we agree that this specific point remains open to further investigations.

To sum up, we modified the title, the abstract, and our paper's main text to better tailor the claims and make clear which results are general and which, though intuitively general, will need further studies. We nevertheless hope, through our work, to ignite a new discussion about the analysis of disinformation as a part of the complex dynamics between suppliers and providers of information.

"Include a "Response to the editors and reviewers" document detailing, point-by-point, how you addressed each editor and referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be used by the editors to evaluate your revision and sent back to the reviewers along with the revised manuscript."

"Highlight all changes made to your manuscript or provide us with a version that tracks changes."

Response to the rev. 1

"The authors proposed a framework based on news supply and demand to provide one potential explanation for why people share misinformation. In particular they combined news supply as represented by an Italian news database and demand as captured by Google Trends data for Italy. While the authors used an interesting methodology to study an important issue in today's society, the paper suffers from fundamental shortcomings in its current form and may not live to the standard of Nature Human Behaviour. The major shortcoming of the paper is the lack of contribution to the theoretical understanding

of human behaviour in broader context of consumption and transmit misinformation. ”

Reviewer 1 correctly remarks that our study focuses on news supply, i.e., the ensemble of news produced in the country, and news demand, i.e., the request for information by the public, which are both products of human nature. The paper aims to understand their relationships and the difference between information and disinformation production in this system. As a result, we managed to accurately quantify the concentration of disinformation production about COVID-19 in the observed timeframe. Though they represent interesting hints for future works, the distribution, consumption, transmission and sharing on social networks is outside the scope of our paper.

“It is not clear how findings in terms of macro level analysis of news supply and demand for a particular case advances the current state of the literature.”

We politely disagree with Reviewer 1. The significance, importance and originality of our paper were highlighted by Reviewers 2 and 3. In particular, we contend the assessment of disinformation concentration without requiring fake news annotation, thanks to the proposed Combined Index, is a valuable result, whose impact is exemplified on the COVID-19 subject, a case of particular relevance.

We also highlight that, to the best of our knowledge, our results and the methodology we defined in the paper are not present in the current state of the literature.

“Additionally, there is much room for improvement in other aspects of the manuscript: (1) the abstract can be improved to clear the contribution of the paper and the method used”

Thanks to these indications, we have improved the abstract accordingly to clarify the contribution and the methodology used.

“(2) other reasons why people share misinformation could be discussed (e.g., motivated reasoning, inattention, confusion see Pennycook et al 2021 Nature for related references).”

Following the suggestion, we have expanded the literature review about disinformation, including the suggested reference and many others.

“(3) Google Trend as a measure could be validated for the purpose of this work.”

We extended the discussion on state of the art by referencing previous studies on Google Searches that validate the use of Google Trends for this work. In particular, the work of Broder 2002 categorizes the intents underlying the queries in Google Search engine (seeking for information, website access, or a transaction), and Waller 2011 elaborates on that work by defining the queries that, based on the input keywords, are made to search for information.

“(4) figure 4 could be improved in quality (it is currently hard to read).”

Since there is no Fig. 4 in the paper, we understood that this point refers to Supplementary Figure 1 (the fourth figure of the main text). We remade that plot by increasing the font size to improve its readability.

Response to the rev. 2

“Remarks to the Author:”

“1. Key results.

This paper provides a broader perspective on the news and disinformation ecosystem, by empirically considering both the producers and the consumers of fake news and normal news. This connectivity is an outstanding feature.”

“2. Validity.

I am not a statistician, so am not best placed to judge this. However, I did not detect any flaws that would prohibit publication.”

“3. Originality and significance

This paper offers an evidence-based advance in knowledge, in examining an Italian news database (partially annotated for fake news) and news demand (as captured through the Google Trends data for Italy), finding that disinformation is extremely reactive to people’s interests and tends to thrive when there is a mismatch between what people are interested in and what news outlets provide. Intuitively, we know this as disinformation researchers, but it is excellent to prove this. It offers methodological novelty in demonstrating that a suitably defined index can assess the level of disinformation only based on available volumes of news and searches. This is likely to influence the methodological focus of the disinformation field. The results are of interest to the wider field of information studies, and of immediate interest to the burgeoning field of disinformation and also to journalism studies. It will be of interest to policymakers and governance actors in the realms of disinformation and hybrid threats as well.”

“4. Data & methodology

The two databases consulted are a good proxy of demand and supply for information. The rank comparison of coronavirus related key words is impactful.”

“5. Appropriate use of statistics and treatment of uncertainties:

I am not a statistician. However, my basic knowledge of the area allowed me to follow the argument.”

“6. Custom code

I am not qualified to judge this.”

We thank Reviewer 2 for the points raised so far and deeply appreciate comment number 3. We believe there is nothing more to add from our side.

“7. Conclusions:

p. 10 - In the discussion, the authors speculate that the more reactive nature of Fake News compared to General News could be a consequence of the different production environments of General News and Fake News, and also speculate that this could be down to different sizes of communities producing these types of news. However, it is unclear why community size would be a good explanation. For instance, what about previous studies showing that Fake News (and false information and selective information) are more viral than other sorts of information on social media, and that Fake News is designed for virality (in its selective tenor that exploits the algorithms of social media)? ”

Thanks to this remark, we improved the speculation in the Discussion to explain our hypothesis better. Virality differences between general news and disinformation are the subject of a fascinating open debate in the literature (see Vosoughi 2018 and Cinelli 2021), which we preferred not to address in this paper. Moreover, this pertains to the diffusion of news (e.g. through social media sharing) rather than news supply or demand, which are our paper’s focus. We believe that the introduction of General News (and Fake News) diffusion in our analysis is an interesting research direction to close the loop between supply and demand. It will be the object of future work.

“8. Suggested improvements

p.2 the definition of Infodemic is not accurate. My understanding of this term is that there is a glut of information, some of it true and some false, so that people do not know what to believe.”

We have edited our definition of Infodemic based on the suggestions and corrected it as such: [...] the term “Infodemic” made a comeback from the SARS epidemic of 2003 [45], to describe the spreading of many pieces of information about the virus SARS-CoV-2, some correct others false, to the extent that overwhelms peoples and hinders their understanding of the phenomenon. *ii*

“p.2 The paragraph characterising the literature on news supply is bizarre. I assume that the authors are referring to the treatment of this field by a sub-discipline or by a particular methodological perspective (I am unsure which ones) – but nowhere do I see any reference to the very many Journalism Studies academic journals, where the topic of news supply is treated thoroughly, from all sorts of angles. The authors must clarify and specify how they are characterising the field of academic scholarship on news supply.”

We are thankful to Reviewer 2 for raising this (and the following) point about a part of the literature we overlooked. We have extended the bibliography (on page 2) to include works from journalism studies that we considered relevant for our work while attempting to convey the extent of this research field:

The news supply can be identified with the overall news production, mainly of officially recognised newspapers. As such, the literature concerning journalism and media information is extensive, and investigated: news linguistics [19, 9]; journalism economics [46]; information coverage [47, 53], often focusing on particular countries or topics [65]; content of news [34] its quality assessment [16] and delivery [17], considering different media sources [33]. Other works investigated the impacts of news and its consumption on, for example, reading behaviour [55], finance [21], and political opinions [28]. *ii*

“p.2 The paragraph characterising the literature on news demand similarly ignores a wealth of research on news audiences from Journalism Studies, including more recent literature why they may avoid news and why they may be attracted to disinformation. It would seem pertinent to mention this. If the paper is specifically about methodological ways of tracking news demand, then this should be clarified, while also acknowledging the wealth of literature on news demand more generally.”

Similarly, as for the previous point, we have extended the bibliography to include journalism studies on news demand and appeal to misinformation. We unbundled the part discussing current methods for tracking news demand. Moreover, we have broadened the discussion about Google Trend’s index, as a proxy for news demand, by referencing past works on the analysis of Google’s searches.

“Half way down p.5, there seems to be a missing phrase, as the sentence starts mid-sentence.”

This has been corrected by replacing the full stop at the end of equation (1) with a comma.

“10. Clarity and context

Line 1 of the abstract states: ‘Social dialogue, the foundation of our democracies, is currently threatened by disinformation and partisanship’. I disagree about the ‘partisanship’ aspect. Partisanship is fine, and a normal state of affairs, as the political science literature will show. What is not fine is ‘polarisation’ – especially affective polarisation – which harms social dialogue in many ways. There is a literature on this. I suggest that the authors remove reference to the word ‘partisanship’ in the abstract.”

Following Reviewer 2’s advice, we have removed the term “partisanship” from that sentence and replaced it with the term “polarisation”.

Response to the rev. 3

“Remarks to the Author:

This paper uses an unprecedented amount of data to investigate the interplay of supply and demand of news, and how that relationship changes based on the trustworthiness of sources. The authors primarily make use of the example of the COVID-19 pandemic as a topic of inquiry. Their primary findings are that (1) Google searches (their proxy for demand) are temporally prior to news content; (2) “Fake News” sources have less inertia and are more impacted by demand; (3) Semantically, “Fake News” is more correlated with searches than general news; (3) The availability of fake news can be predicted temporally by finding when the relationship between Google searches and general news is disturbed (both in terms of autoregression and semantics). I think this paper is important. As the authors point out, previous studies of news demand are limited by the availability of choices. In the lab researchers define the choice set for participants. In observational tracking studies demand is constrained by what the news itself chooses to print (an excellent point I have never thought about). Their method of using Google trends is a good one. In addition, the analysis in the paper was competently done.”

We are deeply thankful to Reviewer 3 for correctly summarizing our work and highlighting its importance and the details motivating our methodology.

"I have one fairly significant question/suggestion, and a few additional minor points. The thing I struggled with most in the paper was: what is being defined as Fake News here, and how does that change the conclusions of the paper? The manuscript is very short on details. In particular, it's unclear if the fake news consists of individual articles or entire sources. On Pg 6 it states "... we extend our analysis to the news items that were annotated as false or misleading." But on Pg13: "Moreover, the data had a specific annotation for 'fake news' sources. The list of disinformation sources, vetted by the Authority, is provided by fact-checking organizations like bufale.net and butac.it. It had already been employed for other studies on disinformation." In one case the authors talk about fake news items, in the other case, they talk about fake news sources."

Thanks for remarking on the mistake. We corrected the definition based on the suggestions at page 6 and fixed it as such: [...] To this end, we extended our analysis to the news items that were produced by sources annotated as false or misleading

"The difference matters a great deal. This isn't just a question of "framing", but really one of measurement: the paper claims to be telling us something about dis-information, but is this measure really capturing disinformation? While I trust that these two websites have done some due diligence in identifying sources that, in general, might promote some degree of disinformation, this is no guarantee that all (or even most) of the information presented in these sources is incorrect or fake. So the authors are claiming that the relationships described here are between news demand and specific, verified to be incorrect, information. But what they are really describing is the relationship between news demand and content in particular types of sources. The authors need to be much more transparent about what is happening here. The language of the paper must be changed to indicate that what is being described here is how news demands predicts the amount of coverage in different types of sources, not the production of specifically "fake" news items. After these changes it's not clear to me the degree to which this can be claimed to be a study about "disinformation" explicitly."

By highlighting this ambiguity, this comment stimulated us to improve the language of the paper to clarify that what we analyse is not disinformation itself. Our work is focused

specifically on production mechanisms, i.e., the disinformation and General News supply and demand dynamics. We modified the language of the paper accordingly to this suggestion.

“That the “Fake news” in this study is really “sources that have been deemed to produce some fake news” brings up another potential issue: what is the media of these sources, and how does it compare to the distribution of media in the overall sample? Specifically, I’m not sure what to make of the results if, for example, most of the “fake news” sources are websites but the “General News” category contains a mixture of websites and traditional news sources. The authors hint that this might be the case on page 10 “[General News] are mainly produced by a large and well-established community of professional journalists while the [Fake News] are the outcome of by a scattered multitude of small, unorganized actors.” The finding that “Fake News” sources have less inertia and are more responsive to demand might just because news websites, in general, have less inertia and are more responsive to demand. Can the authors determine the degree to which this is true? Can the Fake News sources be compared to a group of sources that are comparable? (Or can controls be put in?) I am worried that people reading this article will conclude that specific disinformation follows the supply-demand dynamic described here when really it might be that internet sources generally follow this pattern, and the outlets that produce fake news happen to be online.”

The point raised is an interesting one. We assumed that General News is a product of officially recognised journalists belonging to professional associations and, likely, also to newspapers editorial staff, irrespective of the media used to distribute their news production (paper, internet, etc.). We agree that a difference in the distribution media (online vs paper) is a valid hypothesis that could explain the observed behaviours of news production. We leveraged the

annotations about the type of source for all the news outlets to test this hypothesis. We identified the subset of online news suppliers and analysed their behaviour through our improved modelling. The results, reported in Supplemental Tab. S1, show how Fake News suppliers, which mainly are online, display different dynamics from Online News suppliers. The latter are meaningfully different also from General News suppliers, even if the α_1 coefficient (the Inertia) is remarkably compatible. This result contradicts the hypothesis. We integrated the main text to mention this result.

“One small comment I had was for Table 2, which compares the time-series coefficients for General and Fake news. The authors are comparing the magnitude of the coefficients across columns and making claims that coefficients are smaller or larger. This is implicitly a

hypothesis test, and the authors should formalize it as such. The difference between these coefficients is a random variable with a distribution that needs to be accounted for. I would suggest performing a bootstrap analysis for the difference between coefficients in rows to determine the degree to which the differences are statistically significant.”

We agree with the comment raised by Reviewer 3. We have, accordingly, implemented the suggested bootstrap analysis to compare the coefficients reported in Table 2 statistically. The details of the methodology are reported in Materials and Methods. Using 10^5 bootstrap samples (i.e., sampled with replacement), the Fake News supply inertia resulted significantly lower than General News supply inertia. The results of the bootstrap are reported in Supplemental Fig. S6. We also integrated the main text to mention the result.

Decision Letter, second revision:

23rd December 2021

Dear Dr Gravino,

Thank you once again for your manuscript, entitled "Assessing disinformation through the lens of news supply and demand during the COVID-19 outbreak," and for your patience during the peer review process.

Your revised manuscript has now been evaluated by 3 reviewers, whose comments are included at the end of this letter. While two of the original reviewers now recommend publication of your work, Reviewer #3 raises concerns regarding your measurement of disinformation and the reliability of the judgements provided by the fact-checking organizations. We would like to consider your response to these concerns in the form of a revised manuscript before we make a decision on publication.

Your revised manuscript must comply fully with our editorial policies and formatting requirements. Failure to do so will result in your manuscript being returned to you, which will delay its consideration. To assist you in this process, I have attached a checklist that lists all of our requirements. If you have any questions about any of our policies or formatting, please don't hesitate to contact me.

In sum, we invite you to revise your manuscript taking into account the remaining reviewer comments. We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

We hope to receive your revised manuscript within four weeks. We understand that the COVID-19 pandemic is causing significant disruption for many of our authors and reviewers. If you cannot send

your revised manuscript within this time, please let us know - we will be happy to extend the submission date to enable you to complete your work on the revision.

With your revision, please:

- Include a "Response to the editors and reviewers" document detailing, point-by-point, how you addressed each editor and referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be used by the editors to evaluate your revision and sent back to the reviewers along with the revised manuscript.
- Highlight all changes made to your manuscript or provide us with a version that tracks changes.

Please use the link below to submit your revised manuscript and related files:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work. Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further.

Sincerely,

Arunas Radzvilavicius, PhD
Editor
Nature Human Behaviour

Reviewer expertise:

Reviewer #1: misinformation, computational social science

Reviewer #2: communication, disinformation, news ecosystems

Reviewer #3: communication, political science, data science

REVIEWER COMMENTS:

Reviewer #1:

Remarks to the Author:

I would like to thank the authors for reasonably addressing my comments. I do not have further comments on this manuscript.

Reviewer #2:

Remarks to the Author:

The authors have carefully and fully attended to all the critiques and points raised by the reviewers. I am satisfied that this journal paper is ready for publication.

Reviewer #3:

Remarks to the Author:

I appreciate the efforts of the authors to revise their manuscript. Their additions of a test of disinformation sources against online news was well done, and I appreciated their use of a bootstrap for hypothesis testing.

I'm afraid I was a little ambiguous about my larger point in my original review. I highlighted the language around the measurement of news as a particular instance of a much larger problem that I still don't see resolved. My critique was not just about confusing language (which, to some extent, has been fixed), but the fundamental need to communicate to the reader what is being measured here.

Again, this is all the information the reader gets about what is being labeled as disinformation/fake news:

"The list of disinformation sources, vetted by the Authority, is provided by fact-checking organizations like bufale.net and butac.it. It had already been employed for other studies on disinformation [60]."

This is inadequate. How was this information vetted? Why should I trust the "Authority"? The sentence makes it sound like more organizations than these two are involved. Are all these organizations trusted? What proportion of news on these websites is specific disinformation? Are some websites worse than others?.....

The main purpose of this paper is to identify the relationship between news demand and supply of fake news – specific news items that actually contribute to an "infodemic". I still really have no idea what sorts of news items are being represented in the disinformation sources. I've looked up the "other study" that has used this classification but it similarly only says that these sources have been "vetted". I've gone to the original Italian websites, but those don't provide any sort of methodological rigor as to understand how a website ends up on their list (and who knows if I should trust Google's translation of this information).

Put another way: the theory and normative perspective of this paper is about the provision of specific disinformation about the pandemic. But, as it stands, all we know is that there is a relationship between searches and the amount of coronavirus news on a certain type of website. I have no idea if the coronavirus content on those websites is disinformation, or not. It probably is, but it is the authors responsibility to show that they are measuring what they say they are measuring. Assurances of the info being vetted and a citation to one additional paper is not enough to satisfy questions about measurement.

A great deal more information about the content of these websites and the process for their identification as “disinformation” is necessary for me to trust the results of this paper.

Author Rebuttal, second revision:

Response to the editors and reviewers

Response to the editors

“Your revised manuscript has now been evaluated by 3 reviewers, whose comments are included at the end of this letter. While two of the original reviewers now recommend publication of your work, Reviewer 3 raises concerns regarding your measurement of disinformation and the reliability of the judgements provided by the fact-checking organizations. We would like to consider your response to these concerns in the form of a revised manuscript before we make a decision on publication.”

We appreciated the reviewer comments. They allowed us to improve the detail of the explanations about our data and our methodology in this new version of the manuscript. Reviewer 3 concerns are addressed below.

“Your revised manuscript must comply fully with our editorial policies and formatting requirements. Failure to do so will result in your manuscript being returned to you, which will delay its consideration. To assist you in this process, I have attached a checklist that lists all of our requirements. If you have any questions about any of our policies or formatting, please don’t hesitate to contact me. ”

We fine tuned the manuscript to (hopefully) fully comply with the editorial policies and the checklist.

Response to the rev. 3

We wish to thank the reviewer for the comments that helped us to clarify one of the aspects that are often neglected in this kind of studies.

“How was this information vetted? Why should I trust the “Authority”?
[...] Are all these organizations trusted?”

Our language was unclear. In fact, the Authority did not validate the list or the news itself. The Authority carefully explained that it did not verify any of the news (from what we can understand it would be illegal for them to do so), neither it checked the sources in the list. In the new version of the paper we have written a more clear and detailed explanation:

“AGCOM compiled the list of disinformation sources by merging the lists from independent fact-checking organizations such as bufale.net, butac.it, facta.news and pagellapolitica.it. The Authority verified their methodologies and legitimacy through the recognition of the international organisations such as the [International Fact-checkers Organizations](#), the [Duke Reporters’ Lab](#) or the [European Digital Media Observatory](#). The Authority released the list to his scientific partners only after it was used by independent scientific studies [2, 64, 26, 14].”

“Put another way: the theory and normative perspective of this paper is about the provision of specific disinformation about the pandemic. But, as it stands, all we know is that there is a relationship between searches and the amount of coronavirus news on a certain type of website. I have no idea if the coronavirus content on those websites is disinformation, or not.”

Thanks to your previous review’s observations, we clarified that our paper and findings focused on disinformation producers, more than single piece of views marked as disinformation. The results and considerations we draw (less inertia, more semantic precision, etc.) are on the behaviour of the disinformation producers. The “source-based” annotation approach is standard in literature (see [66, 9], and the previous references). In our case, this approach is even more fitting, given our aim to describe disinformation producers behaviour. Claiming that everything they produce is disinformation is not a purpose of our work. A detailed analysis of their production separating the disinformation items from the rest lies beyond the scope of the paper. It would be really interesting, but it would require an annotation of the single news that is not available (and hardly doable, given the size of the data-set).

Decision Letter, third revision:

20th January 2022

Dear Dr Gravino,

Thank you for submitting your revised manuscript, "Assessing disinformation through the lens of news supply and demand during the COVID-19 outbreak". After careful consideration and discussion with my colleagues, I am sorry to have to tell you that we do not feel that the referees' comments have been sufficiently addressed to justify sending this revision back for peer review.

This unusual course of action is taken occasionally to avoid unproductive rounds of review that result in reviewer fatigue and damage the chances of the manuscript obtaining a fair and objective evaluation. Such situations are not in an author's best interest so we try to avoid them when it seems prudent to do so.

While you have added an explanation of how the Authority labels news sources as misinformation, this does not fully address our concerns. In order to consider this manuscript further we would request that you please do your best to fully address the remaining issues. In particular:

1. Provide a much more detailed explanation of how news outlets were classified by organizations. How are individual articles/sources vetted by organizations and what are their protocols? Which criteria are applied? What proportion of new articles coming from these sources has to be false for them to be classified as misinformation? How consistent are these trends over time? This needs to be thorough and complete enough for a broad international audience to understand how fact-checking by these bodies works in Italy.
2. Provide evidence that articles published by sources labeled as disinformation sources are indeed false, as requested by Reviewer #3, or:
3. If providing article-level evidence is unfeasible (point 2), please include a more transparent discussion of the limitations of your approach. Please revise your article throughout to make clear that you look at exposure to unreliable websites, rather than misinformation itself.

Should you be able to adequately respond to these and the reviewers' other concerns, we would be happy to look at a revised manuscript.

We shall hope to receive your revised version as soon as possible. If you anticipate a delay of more than four weeks, however, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Human Behaviour or published elsewhere. Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date may be that of the revised, not the original, version.

Nature Human Behaviour is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit <http://www.springernature.com/orcid>.

If you are not interested in submitting a suitably revised manuscript in the future please let me know immediately so we can close your file. If you have any questions, please contact me.

Please use the link below to submit a suitably revised manuscript and updated response to referees when they are ready.

[REDACTED]

Sincerely,

Arunas Radzvilavicius, PhD
Editor
Nature Human Behaviour

Author Rebuttal, third revision:

Response to the reviewers

“REVIEWER 3: How was this information vetted? Why should I trust the “Authority”?”

Our language was unclear. The Authority neither validated the list of disinformation sources nor each piece of news. Moreover, the Authority carefully explained that it would be illegal for them to do so. AGCOM aggregated, in fact, the list of fake news sources produced by other organisations and verified their vetting procedure and their international recognition. As we specified in the paper:

“MANUSCRIPT (pag. 11): AGCOM compiled the list of disinformation sources by merging the lists from independent fact-checking organizations such as [bustale.net](#), [butac.it](#), [facta.news](#), and [pagellapolitica.it](#). [...] The Authority verified the fact-checking organisations’ methodologies and legitimacy through the recognition of international organisations such as the [International Fact-checkers Organizations](#), the [Duke Reporters’ Lab](#) or the [European Digital Media Observatory](#). The Authority released the list of disinformation sources to its scientific partners only after it was used by independent scientific studies [\[2, 67, 28, 15, 11\]](#). Also, the list provided by the Authority in 2020 was already adopted in one other study [\[17\]](#).”

“REVIEWER 3: Are all these organizations trusted?”

We checked their procedure and found them reliable. However, we are neither professional fact-checkers nor journalists. We assumed they could be trusted because the organisations are internationally recognised (e.g. by [International Fact-checkers Organisations](#)), and comply with the highest standards. The outcome of their activities has been vetted and exploited in several scientific publications [\[2, 67, 28, 15, 11, 17\]](#).

“REVIEWER 3: What proportion of news on these websites is specific disinformation? Are some websites worse than others?”

We checked their procedure and found them reliable. However, we are neither professional fact-checkers nor journalists. We assumed they could be trusted because the organisations are internationally recognised (e.g. by [International Fact-checkers Organisations](#)), and comply with the highest standards. The outcome of their activities has been vetted and exploited in several scientific publications [\[2, 67, 28, 15, 11, 17\]](#).

“REVIEWER 3: What proportion of news on these websites is specific disinformation? Are some websites worse than others?”

This information is not currently available, but we assumed that the variations would be a secondary effect, since we never consider sources singularly but always look at the aggregated production. Future studies will be able to address this kind of issue when the requested data will become available. We discussed these limitations of our study in the manuscript.

“Still, the source-based approach implies some limitations. Sources annotated as fake might not publish just fake news, and news from the annotated sources might be fake by different degrees. In our approach, the sources annotated as disinformation are assumed to be disinforming to the same extent and static for the observed period. In principle, the percentage of fake news might vary from one source to the other and over time. Future studies will address these limitations. For the present work, we assumed that the source annotations, and their extension to the relative news, are reliable enough to represent the disinformation supply, at least at the aggregated level we considered.”

Decision Letter, fourth revision:

23rd February 2022

Dear Dr. Gravino,

Thank you for submitting your revised manuscript, "Assessing disinformation through the lens of news supply and demand during the COVID-19 outbreak", and for all of your work to address reviewer and editor concerns. After careful consideration, we believe that your manuscript is making progress towards addressing these concerns. However, we also believe that a core concern remains insufficiently addressed.

In order to consider this manuscript further we would request that you please do your best to fully address all of the comments of the reviewers. In particular, please ensure that your manuscript makes clear throughout the text that you study exposure to unreliable websites, rather than misinformation. We appreciate that a number of studies in the literature use domain-level metrics to assess disinformation and "fake news". However, our editorial view is that, as highlighted by Reviewer 3, the relationship between these two concepts must be directly demonstrated. Where doing so is not feasible (as in this case), the manuscript must make clear what is measured, rather than using domain-level information as a proxy for misinformation itself. We would ask that you thoroughly revise the text (including section headings that reference "fake news") to address this outstanding point.

We shall hope to receive your revised version as soon as possible. If you anticipate a delay of more than four weeks, however, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Human Behaviour or published elsewhere. Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date may be that of the revised, not the original, version.

Nature Human Behaviour is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit <http://www.springernature.com/orcid>. If you are not interested in submitting a suitably revised manuscript in the future please let me know immediately so we can close your file. If you have any questions, please contact me.

Please use the link below to submit a suitably revised manuscript and updated response to referees when they are ready.

[REDACTED]

Sincerely,
Aisha

Aisha Bradshaw, PhD
Senior Editor
Nature Human Behaviour

Decision Letter, fifth revision:

Our ref: NATHUMBEHAV-210515499E

8th March 2022

Dear Dr. Gravino,

Thank you for submitting your revised manuscript "Dynamics of news supply and demand during the COVID-19 outbreak and assessment of questionable sources production" (NATHUMBEHAV-210515499E). After careful consideration of this latest version of your manuscript, we will be happy in principle to publish it in Nature Human Behaviour, pending minor revisions to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements within two weeks. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Please do not hesitate to contact me if you have any questions.

Sincerely,
Aisha

Aisha Bradshaw, PhD
Senior Editor
Nature Human Behaviour

Final Decision Letter:

Dear Dr Gravino,

We are pleased to inform you that your Article "The supply and demand of news during COVID-19 and assessment of questionable sources production", has now been accepted for publication in Nature Human Behaviour.

Please note that *Nature Human Behaviour* is a Transformative Journal (TJ). Authors whose manuscript was submitted on or after January 1st, 2021, may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. IMPORTANT NOTE:

Articles submitted before January 1st, 2021, are not eligible for Open Access publication. [Find out more about Transformative Journals](#)

Authors may need to take specific actions to achieve [compliance](#) with funder and institutional open access mandates. If your research is supported by a funder that requires immediate open access (e.g. according to [Plan S principles](#)) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route, the journal's standard licensing terms will need to be accepted, including [self-archiving policies](#). Those licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

Once your manuscript is typeset and you have completed the appropriate grant of rights, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately. Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see <http://www.nature.com/nathumbehav/info/gta>). In particular your manuscript must not be published elsewhere and there must be no announcement of the work to any media outlet until the publication date (the day on which it is uploaded onto our web site).

If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

An online order form for reprints of your paper is available at <https://www.nature.com/reprints/author-reprints.html>. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

We welcome the submission of potential cover material (including a short caption of around 40 words) related to your manuscript; suggestions should be sent to Nature Human Behaviour as electronic files (the image should be 300 dpi at 210 x 297 mm in either TIFF or JPEG format). Please note that such pictures should be selected more for their aesthetic appeal than for their scientific content, and that colour images work better than black and white or grayscale images. Please do not try to design a cover with the Nature Human Behaviour logo etc., and please do not submit composites of images related to your work. I am sure you will understand that we cannot

make any promise as to whether any of your suggestions might be selected for the cover of the journal.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

In approximately 10 business days you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required.

You will not receive your proofs until the publishing agreement has been received through our system.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

We look forward to publishing your paper.

With best regards,

Arunas Radzvilavicius, PhD
Editor
Nature Human Behaviour