The Evolution of the COVID-19 Pandemic Through the Lens of Google Searches

Robert Marty (robmarty3@gmail.com)
World Bank

Manuel Ramos Maqueda
mramosmaqueda@worldbank.org

Nausheen Khan
World Bank

Arndt Reichert
University of Hannover

Article

Keywords:

Posted Date: April 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1551521/v1

The Evolution of the COVID-19 Pandemic Through the Lens of Google Searches

Robert Marty* [1], Manuel Ramos Maqueda[1], Nausheen Khan[1] and Arndt Reichert[2]

[1]World Bank
[2]University of Hannover

April 21, 2022

Abstract

Real-time data is essential for policymakers to adapt to a rapidly evolving situation like the COVID-19 pandemic. Relying on Google search interest data across 207 countries and territories, we demonstrate the capacity of publicly-available, real-time data to anticipate COVID-19 cases; evaluate the economic, mental health, and social impacts of containment policies; and identify demand for (mis)information about COVID-19 vaccines. We show that: (1) search interest in COVID-specific symptoms can anticipate rising COVID-19 cases across both high- and low-income settings; (2) countries with more restrictive containment policies experienced larger socio-economic externalities; in addition, lower-income countries experienced less searches for unemployment, but more pronounced mental health externalities; and (3) high vaccination rates are associated with strong demand for information about vaccine appointments and side effects; in some settings, high interest in misinformation search terms is associated with low vaccination rates. Overall, the results demonstrate that real-time search interest data can be a valuable tool for both high- and low-income countries to inform policies across multiple stages of the pandemic.

*Corresponding author: rmarty@worldbank.org

1 Introduction

The COVID-19 pandemic and the associated policy responses have rapidly evolved. Across the world, people have had to adapt in multiple ways. Whether it was adopting preventative measures such as masks or testing; understanding symptoms, risks, and treatment options; adapting lifestyles to social distancing measures; dealing with business closures and unemployment; or learning about vaccines that use new technologies such as mRNA, people throughout the world have been continuously learning how to adapt to the pandemic. In this context, one of the first places people turn to learn about coronavirus symptoms, new policies, and vaccines are online Google searches. These searches generate a vast amount of publicly available information about people's symptoms, fears, concerns, and demand for information. These real-time data can have high value for policymakers who aim to better understand, monitor, and react to citizens' experiences. Not only do Google search patterns provide information when other sources such as standard household surveys become inoperative due to risks of infection, they are also up to date when decisions are actually taken [3].

In this paper, we explore the potential of Google search interest data to inform the policy response to the COVID-19 pandemic across 207 countries and territories throughout the world (henceforth, we use "countries" to refer to both countries and territories). Specifically, we examine the capacity to use this real-time, big data source for three objectives: (1) anticipate rising COVID-19 cases, (2) evaluate the social and economic impact of containment policies, and (3) understand demand for information about vaccines, including misinformation. In our analyses, we examine whether the results vary across countries with different income levels, extending the scope of prior research that focuses on a limited set of high-income countries. We are the first to systematically analyze lower-income contexts with the purpose of bridging the evidence gap across different regions in the world. Like high-income countries, much of the developing world has experienced multiple waves of coronavirus infections and implemented similar policies to curb them. Key differences arguably are substantively lower levels of economic protection through social safety nets, access to vaccines, and the capacity to produce reliable administrative data. Thus, the available evidence may not generalize well to low-income countries.

To make the analysis possible worldwide, we develop a method to identify the most popular language for Google searches in each country or territory, and automatically translate search terms into each country's most popular language using the Google Translate API. We then systematically query the search terms from Google Trends for each country and analyze these data to generate decision-focused evidence that responds to the situation on the ground and can contribute to informing countries' and citizens' responses to the pandemic.

As to the first objective, we find that search interest in COVID-19-specific symptoms such as "loss of

smell" and "loss of taste", as well as more general searches such as "COVID symptoms", strongly correlate with and precede reported coronavirus cases across countries. These results builds off of findings from studies focusing on a limited set of countries [26, 39, 28, 40]. Importantly, our results demonstrate that the estimated correlations are as applicable to lower-income countries as they are for higher-income countries. Leveraging search interest data to forecast illness started with Google Flu Trends (GFT) to track influenza [15], where later work showed the potential of search interest to forecast other illnesses [31, 42, 30, 37, 38, 22]. While GFT showed promise in tracking influenza, it ultimately faced challenges such as significantly overestimating Flu [27]. The challenges faced by GFT emphasize the importance of checking and recalibrating models when forecasting illness, particularly as the nature of illnesses change as with the case with new variants of COVID-19 [25]. To this point, we separately examine 2020 and 2021 data; we observe that the demand for information and utility of Google searches does not decay in advanced stages of the pandemic and as the pandemic evolves. That said, after the surge in cases from the Omicron variant, searches for COVID-19-specific symptoms significantly lost explanatory power; however, searches for more general searches—such as "COVID Symptoms"—remained predictive of COVID-19 cases.

To examine the second research question, we implement a difference-in-differences approach that compares trends in search interest before and after the date containment policies were implemented in both the year containment policies were implemented and in the year before. Our results show that COVID-19 shutdown policies are significantly associated with greater interest for a variety of search terms including unemployment, mental health, and social distancing, and lower interest in search terms for personal relationships and family planning. These results are consistent with studies that focus on the impacts of COVID-19 policies within the United States and Western Europe [20, 7, 36, 4, 6, 11]. By leveraging data across all countries with available data, we examine how impacts of COVID-19 policies vary by country characteristics and the type of policies implemented. We find that impacts on select unemployment and unemployment resources keywords (e.g., "unemployment office") tend to be smaller for low-income countries; however, impacts on mental health keywords tend to be larger for lower-income countries. Furthermore, our results suggest that countries with more restrictive containment policies experienced more pronounced mental health and social consequences (i.e., greater search interest in mental health keywords and less interest in relationship and family planning keywords). Importantly, our results also show that countries with greater economic support experienced less search interest for select mental health indicators, suggesting that economic support may have helped to blunt some negative mental health impacts.

To understand the demand for information about vaccines (third objective), we rely on country-level search interest data as well as subnational data in a case study of the United States. In particular, we examine the extent to which search interest in vaccination appointment information, vaccine side effects,

misinformation about vaccines, and general vaccination terms moves with vaccination rates. We find that searches for vaccine appointments and side effects are highly associated with vaccination rates globally. We find strong correlations between vaccination terms across different income levels; however, correlations tend to be larger—on average—for higher-income countries, which may be explained by greater access to vaccines—and more variation in vaccination rates over time—in wealthier countries. The association between misinformation and vaccinations rates are more mixed across countries, where search interest in select misinformation terms typically only registers in a small number of countries. Overall, these results indicate that search interest in misinformation follows country-specific dynamics. In our case study of the United States, we demonstrate that Google Trends data may be used to understand which popular misinformation terms—such as Ivermectin— are highly correlated with low vaccination rates.

Our findings demonstrate that publicly available, real-time data from Google searches may be used to monitor the spread and consequences of the pandemic across the entire world. These data are immediately available for free across a majority of countries even in cases where access to the internet is not widespread. In some cases, the data may also be used as a contrast to administrative data, such as in instances where countries misreport or altogether deny the presence of COVID-19 cases.

Results

Google search interest correlating with and preceding COVID-19 cases

Figure 1 shows the correlation coefficient between different search terms and COVID-19 cases (Panel A) and whether search interest anticipates future changes in COVID-19 cases across countries (Panel B). Panel 1A shows the distribution of correlations across different search terms using data from 2020 and 2021 separately. Search interest for symptoms that tend to be more unique to COVID-19 tend to have the strongest correlation. These terms include "Loss of Smell" and "Loss of Taste", where the 25th percentile of correlations across countries is above zero and the 75th percentile is about 0.5—with some countries seeing a correlation near one, which indicates that search interest moves nearly perfectly with COVID-19 cases. For search interest in loss of smell and taste, the median correlation in 2021 is slightly lower than in 2020, but still high, indicating that these terms are still useful for capturing COVID-19 cases even later in the pandemic. Search interest for "coronavirus" and "how to treat coronavirus" flip from being negatively correlated with cases in most countries in 2020 to positively correlated in most countries in 2021; we hypothesize that global news was a larger factor in driving search interest in 2020, while 2021 searches were more driven by country-specific news and personal experiences about COVID-19.

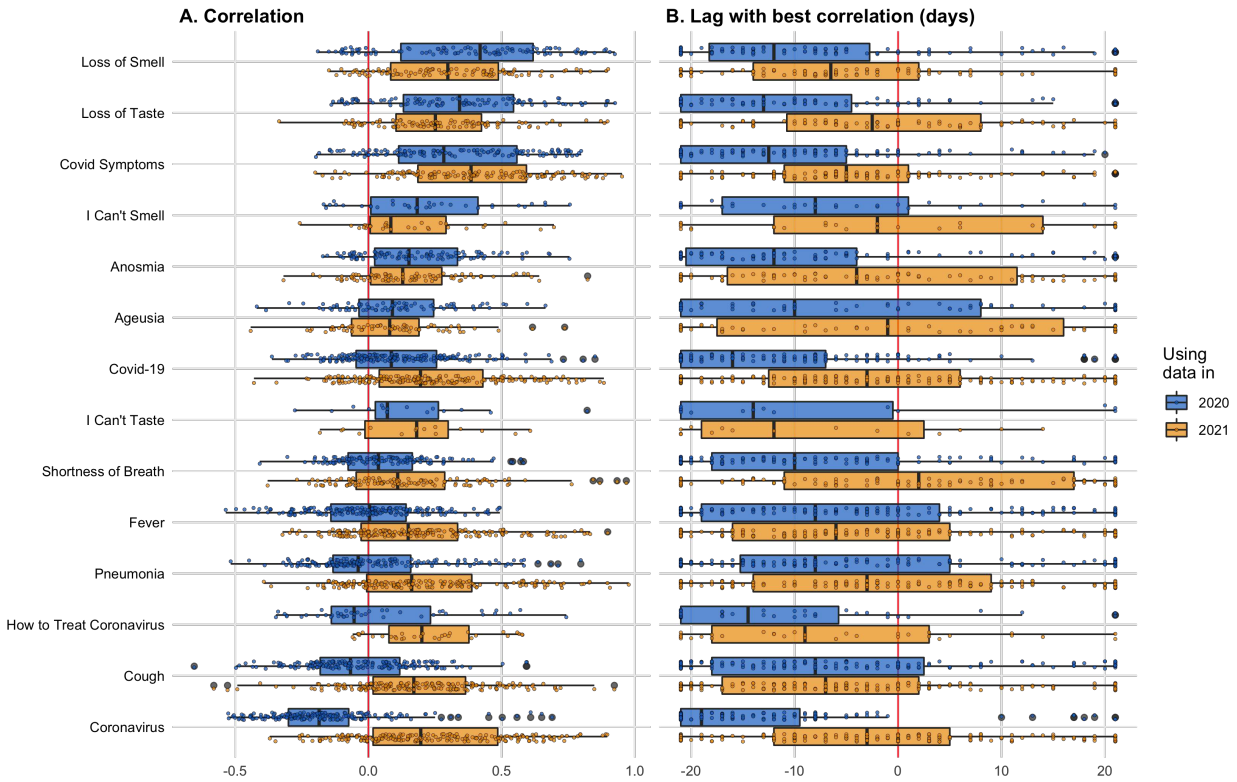

Figure 1: Search interest correlating with and anticipating COVID-19 cases. Panel A shows the correlation between search interest and COVID-19 cases. Panel B shows the lead/lad value of COVID-19 cases that produced the highest correlation with search interest. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.

Searches for more general symptoms—"fever and pneumonia"—tend to see lower median and average correlation values compared to more COVID-19 specific searches; however, the distribution of correlations using "fever" and "pneumonia" are large—where in some countries the correlation is close to one. Appendix S4 shows the distribution of correlations using the highest correlation across different lagged values of COVID-19 cases; here, correlations increase by about 0.1, but the overall patterns stay the same. In addition, appendix S5 shows the spatial distribution of correlations in "Loss of Smell" and a more general symptom ("Fever"), illustrating (1) how the correlation between COVID-19 cases and search interest in "Loss of Smell" is strong throughout different geographic regions and (2) the countries where search interest in "Fever" correlates strongly with cases.

Panel 1B shows the distribution in the lag/lead value of COVID-19 cases that had the best correlation with search interest. Across all search terms, the majority of countries with available data had a negative optimal lag—indicating that trends in search interest precede trends in observed COVID-19 cases. In addition, across search terms, the median lag value was more negative in 2020 than in 2021, i.e., in 2020 search terms anticipated COVID-19 cases by a larger number of days than in 2021. This result may indicate that,

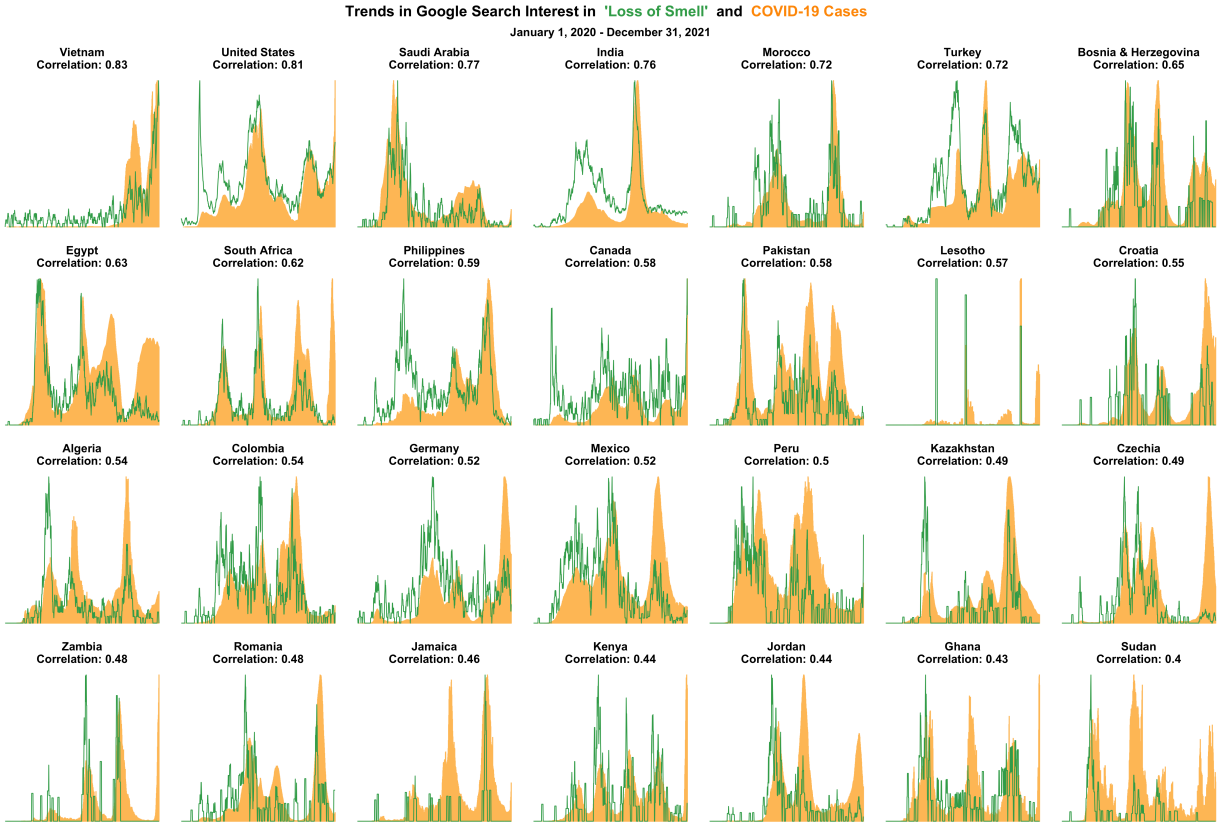

Figure 2: Trends between search interest in "Loss of Smell' and COVID-19 Cases for Select Countries. We show trends for the 28 countries with the highest correlations between search interest in "Loss of Smell" and COVID-19 cases. For both search interest and COVID-19 cases, we plot the 7-day moving average.

as the pandemic has progressed, testing has improved and countries are better able to monitor real-time caseloads. Using data in 2020, the median optimal lags for loss of smell and loss of taste are -12 and -13 days, respectively; using data since 2021, the values are -6.5 and -2.5 days (see appendix S4).

Figure 2 shows trends in search interest in "loss of smell" and COVID-19 cases between January 1, 2020 and December 31, 2021 for the 28 countries with the highest correlations (appendix S6 shows trends for all countries with available data). The figure illustrates how well search interest and cases move together for select countries, even capturing new waves in 2021 due to the Delta variant. The figure includes December 2021, when the Omicron variant caused COVID to surge in many countries (the Omicron variant was first reported on November 24, 2021, in South Africa [32]). The figure shows mixed results as to whether search interest in "loss of smell" captures these most recent waves in cases. For example, loss of smell correlates strongly with cases in South Africa until the surge driven by Omicron, where loss of smell only increases slightly. In contrast, search interest in loss of smell increases with the rapid rise in cases in the United States in December. Appendix S7 shows trends in cases and search interest in "loss of smell", "fever", and "COVID symptoms" for six countries that saw surges in cases in December 2021, where results are similarly mixed;

Table 1: Explaining correlation between search interest in loss of smell and COVID-19 cases

	Dependent variable:									
	Correlation					Best Lag				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total COVID-19 Cases, log	0.04***				0.05**	0.003				−1.08
	(0.01)				(0.02)	(0.59)				(0.92)
Per Pop. Using Internet		−0.0002			0.002		0.08			0.03
		(0.001)			(0.003)		(0.06)			(0.13)
Mobile Cell Sub. per 100			0.0001		−0.0005			0.05		−0.001
			(0.001)		(0.001)			(0.04)		(0.06)
GDP Per Cap, Log				−0.02	−0.07				1.25	1.83
				(0.02)	(0.05)				(0.94)	(2.27)
Constant	−0.21	0.29***	0.25**	0.39**	0.19	−8.53	−12.40***	−13.97***	−19.46**	−10.36
	(0.15)	(0.09)	(0.10)	(0.17)	(0.33)	(7.77)	(4.31)	(4.79)	(8.42)	(16.28)
Observations	109	79	98	105	78	109	79	98	105	78
Adjusted R^2	0.08	-0.01	-0.01	-0.003	0.05	-0.01	0.01	0.01	0.01	-0.01

Models (1)-(5) use the correlation between search interest and cases as the dependent variable, and models (6)-(10) use the lead/lag value that produced the highest correlation as the dependent variable. *p<0.1; **p<0.05; ***p<0.01.

some countries see "loss of smell" or "fever" rising with the recent increase in cases, while others don't. However, while search interest in "loss of smell" became less predictive of COVID-19 with the Omicron variant, search interest in "COVID Symptoms" remained predictive of cases; for example, in South Africa, the United States, and a number of other countries, search interest in "COVID Symptoms" tracks strongly with cases throughout the pandemic, including with the rise in cases in December 2021 (see appendix S7).

Table 1 shows results when regressing country-level indicators on the correlation of COVID-19 cases and search interest of "loss of smell" and the optimal lag of the correlation. Columns (1) and (5) show a positive association between total COVID-19 cases and the strength of the correlation: a 1% increase in COVID-19 cases is associated with a 0.04 to 0.05 increase in the correlation coefficient ($p < 0.05$). However, while there is a significant association, total COVID-19 cases only explain 8% of the variation in the correlation. The other variables we explore—such as Internet usage, mobile subscriptions, and GDP per capita—are not significantly associated with the correlation. Consequently, the results indicate that the correlation coefficients are on average as high in lower-income countries as in higher-income countries. Table 1, columns (6)-(10) shows that no variables are significantly associated with the best lag of the correlation between COVID-19 cases and loss of smell. Appendix S8 shows that results are largely consistent when using search interest in the terms with the second and third highest median correlations: "Loss of Taste" and "COVID Symptoms."

Impact of containment policies on economic, mental health, and social search interest indicators

Figure 3 shows the impact of COVID-19 containment policies on search interest for keywords related to unemployment/economic factors, mental health, relationships and family planning, and social distancing. Panel A shows average trends in search interest in the 90 days before and after the date of containment policies for both 2019 and 2020 for up to 203 countries worldwide with available data. The figure shows sharp changes in search interest across a number of keywords immediately after containment policies were implemented. Panel B shows difference-in-differences results, which shows that containment policies resulted in an increase in unemployment, mental health, and social distancing search terms and a decrease in relationship and family planning search terms (appendix S9 shows that these results are consistent when using an event study approach).

Panel C shows difference-in-differences results that explore the heterogeneity of impacts of containment policies depending on the level of economic support, restrictiveness of policies (proxied by mobility reduction and the stringency index), and GDP per capita. Countries with greater economic support generally saw (1) higher searches for debt and unemployment-related words; (2) lower searches for "anxiety" and "suicide", but significantly higher searches for "boredom", "lonely", "panic", and "social isolation"; (3) lower searches for most relationship and family planning keywords, and (4) higher searchers for social distancing terms. Countries with more restrictive containment policies saw (1) more search interest in a number of unemployment related terms; (2) higher search interest for most mental health-related keywords; (3) lower search interest in "divorce" and "wedding", and (4) mixed results for social distancing terms (a reduction in interest in "social distance" but an increase in "stay at home"). Lastly, lower-income countries (those with lower GDP per capita) saw (1) lower search interest for select unemployment and unemployment resources terms, (2) higher search interest for "anxiety", "anxiety attack", "insomnia", and "suicide", but lower search interest for "boredom" and "panic", (3) higher search interest for all relationship and family planning keywords and (4) lower search interest for "social distance" and "stay at home." Appendix S10 shows that these results are largely consistent when estimating models that rely on data 30, 60 and 120 days before and after the date of the first containment policy.

Overall, the results emphasize a few key trends. First, countries with greater economic support saw higher searches for unemployment keywords and lower searches for mental health keywords. The mechanism for this result could be twofold: on one hand, greater economic support may drive search interest in seeking the economic support (for example, seeking unemployment benefits); alternatively, countries that provided more economic support may have experienced more negative economic impacts which could drive search

interest in unemployment and unemployment resources.

Second, countries with more restrictive containment policies saw (1) higher search interest in mental health keywords, (2) higher search interest in select unemployment keywords and (3) lower search interest in relationship keywords, emphasizing that countries with more restrictive containment policies experienced more pronounced economic, mental health, and social externalities.

Third, greater economic support was only associated with reductions in select mental health keywords, including "anxiety" and "suicide." We may not expect economic support to reduce indicators such as boredom, loneliness, and feelings of social isolation; no matter the level of economic support, people still had to endure lockdowns and social isolation. However, results suggest that economic support may have mitigated the stress of other aspects of the pandemic (for example, anxiety about making ends meet during the pandemic).

Fourth, lower-income countries experiencing greater increases in search interest for a number of mental health terms suggests that there was a larger mental health burden as a result of containment policies in these countries. In terms of the association with unemployment, lower-income countries seeing less search interest in unemployment and unemployment resource terms compared to higher-income countries could indicate that lower-income countries experienced less impacts on unemployment as a result of containment policies; however, it could also suggest that there were less unemployment resources, and thus less resources to search for.

Google search interest to understand vaccine uptake and hesitancy

Figure 4A shows the distribution of within-country correlations between search interest and daily vaccinations across a number of keywords. The figure shows that high vaccination rates are associated with information-seeking about vaccine appointments, vaccine side effects, and general vaccine search terms across most countries. For example, the median within-country correlation between vaccination rates and search interest in "Vaccine Appointment", "Vaccine Reaction", and "Vaccine" are all above 0.25—with correlations in some countries approaching 0.8 and higher. Figure 4B illustrates how new vaccinations move with search interest of select terms. We plot the evolution of search interest for "Vaccine Appointment" and "COVID Vaccine Sick" for the countries with the highest correlation coefficients using these terms. Appendix S11 shows correlations by income level. On average, higher-income countries tend to see higher correlations—which may be explained by higher-income countries having greater access to vaccines (and thus more variation in vaccine rates over time).

Search interest in terms related to misinformation tend to have lower correlations with vaccination rates.

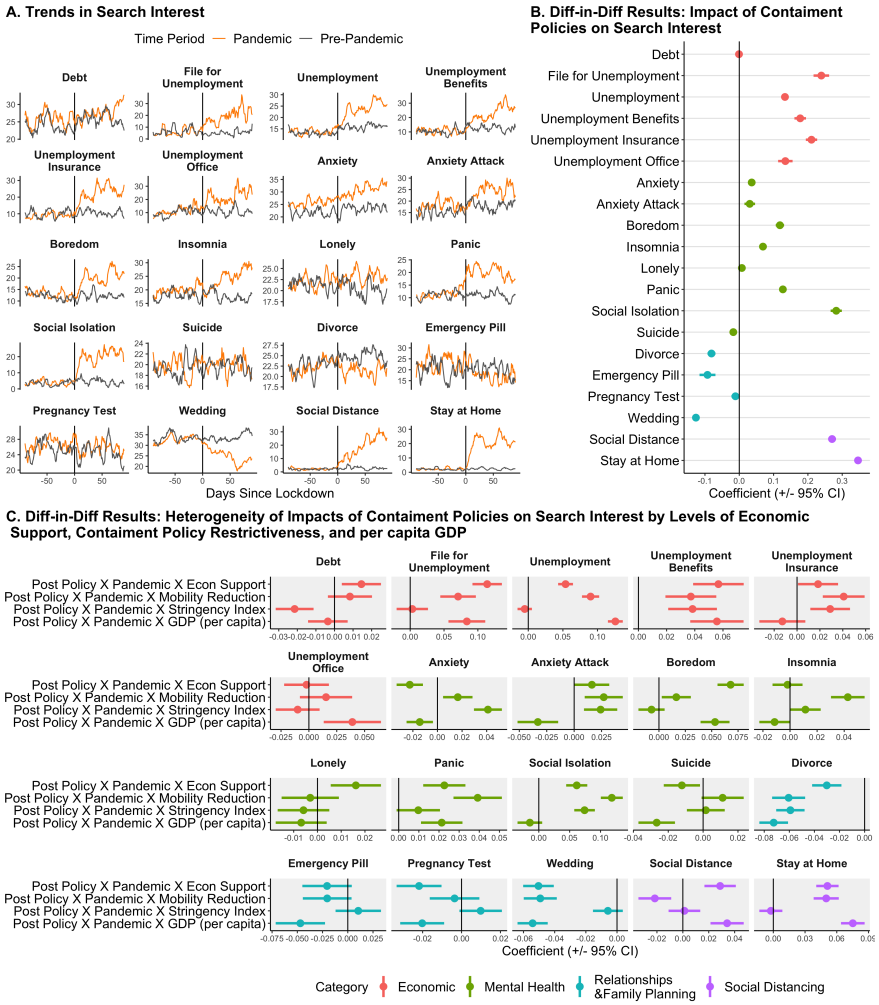

Figure 3: Impact of COVID-19 Policies on Search Interest. Panel A shows average search interest across all countries with available data on Google Trends. Before averaging, search interest values for each country are standardized between 0 and 100. Panel B shows difference-in-differences results showing the overall impact of containment policies across search terms. Only the difference-in-differences coefficient is reported. Panel C shows difference-in-differences results that explore heterogeneity of impacts across containment policy restrictiveness, economic support, and GDP per capita. In panels B and C, point estimates and 95% confidence intervals are shown.

In addition, interest in misinformation keywords is more likely to be missing due to low search interest than other keyword categories; for example, only five countries see high enough search interest in "COVID Vaccine Mercury" for Google to report data and six countries see search interest in "COVID Vaccine Change DNA" (see appendix S1). 163 countries see search interest in "Ivermectin"; however, the correlation is near zero for most countries, with few countries seeing high positive and negative correlations. Overall, the misinformation results suggest that misinformation trends follow more country-specific patterns, as opposed to keywords related to side effects and vaccine appointments, where countries tend to see similar patterns.

While we focus on country-level data, we also conduct a case study of the United States to understand vaccination and search interest at a subnational level. We perform a subnational analysis for three reasons.

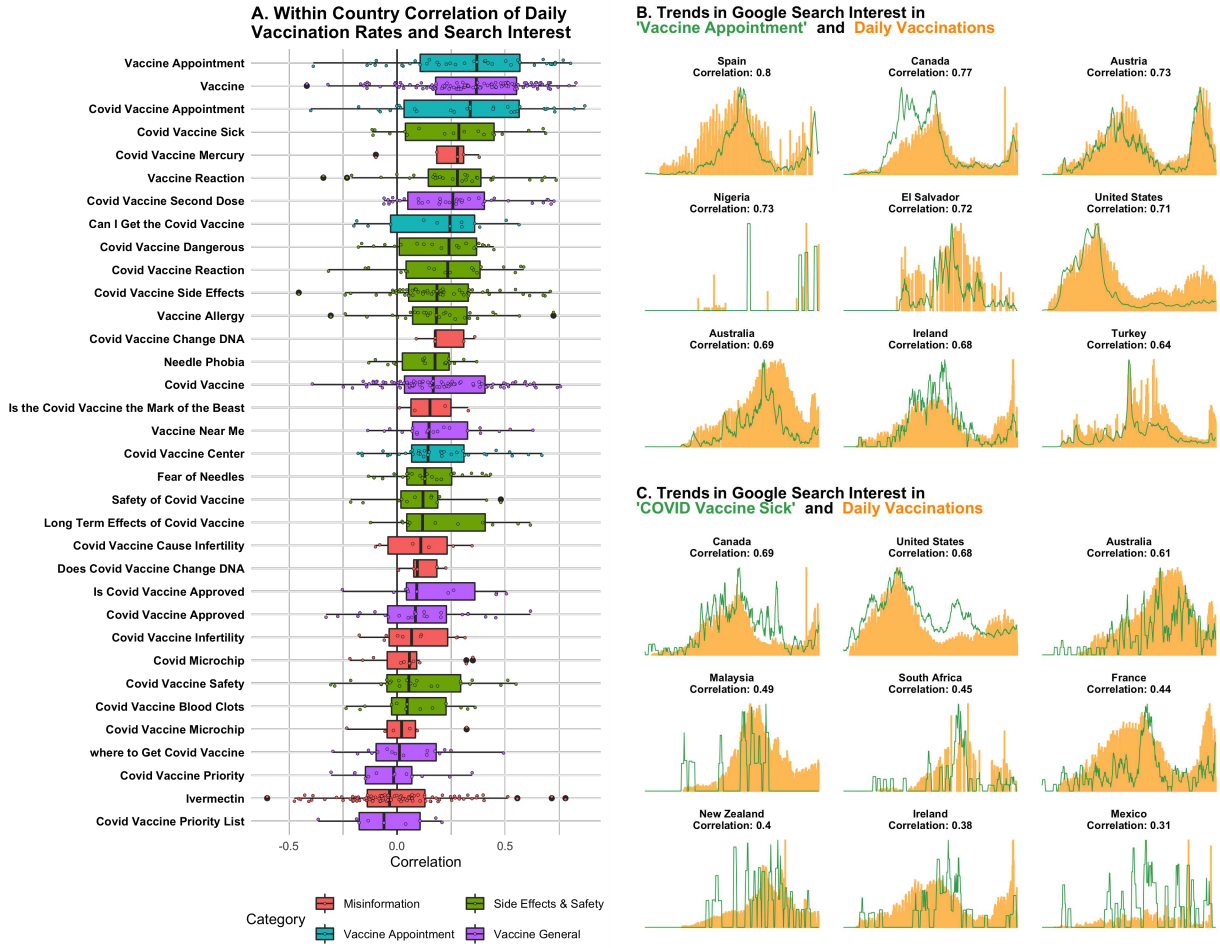

Figure 4: Search interest and COVID-19 vaccination rates. Panel A shows the distribution of the within country correlation of daily vaccination rates and search interest. Panels B and C show how search interest in "Vaccine Appointment" and "COVID Vaccine Sick" move with daily vaccinations respectively. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.

First, we seek to understand whether correlations observed at the country level across time also hold subnationally across geographic regions. Second, comparing search interest across states allows us to understand the association between demand for misinformation and vaccination rates, such as, for instance, which misinformation terms are associated with low vaccination rates, or the characteristics of the states with high search interest for misinformation terms. Third, focusing on a single country like the United States limits the analysis to geographic regions (States) with similar levels of access to vaccines. In the case of the United States there is high vaccine availability yet lower vaccination rates compared to other developed countries (at the time of this writing, the United States is 65% fully vaccinated) [13]. Thus, in our case study low vaccination rates are explained by the decision not to get vaccinated rather than because of the lack of access to vaccines.

Figure 5A shows that demand for information on specific terms across all keyword categories is strongly

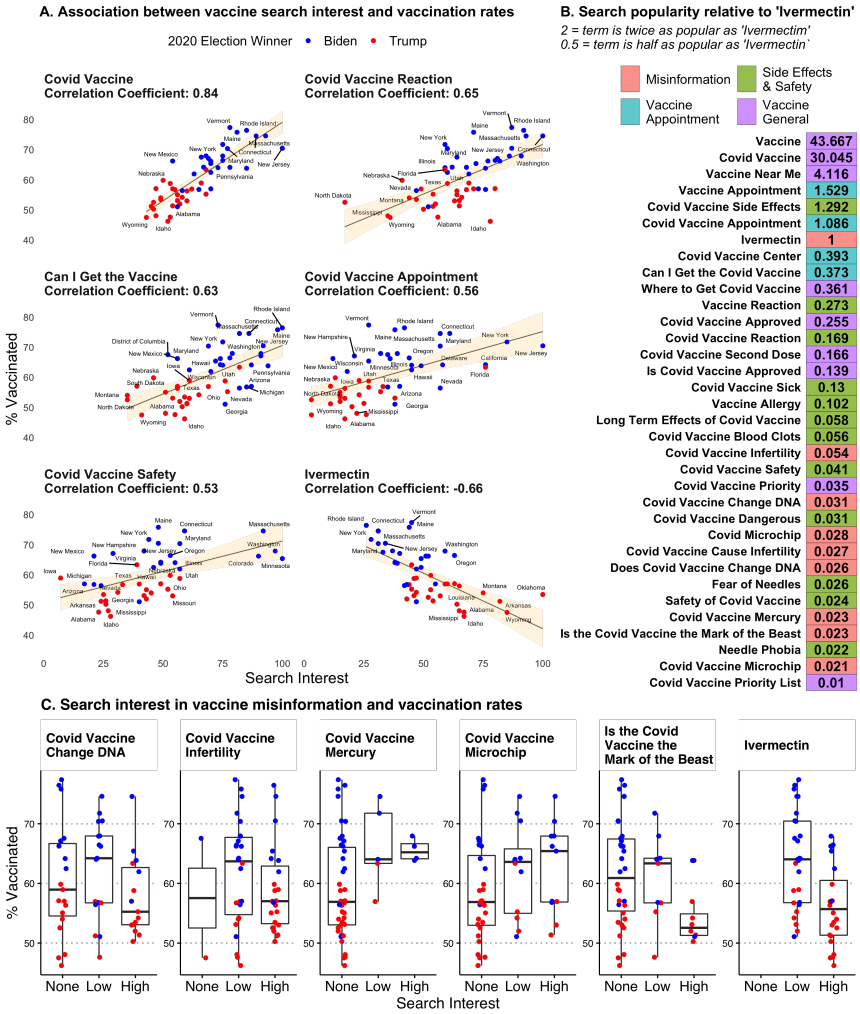

Figure 5: Search interest and COVID-19 vaccination rates in the United States. Search interest captures the popularity of each term across states using the period from December 1, 2020 until December 31, 2021. Panel A shows the association between search interest in select terms and vaccination rates. Panel B shows the search popularity of terms relative to search interest in "Ivermectin." Panel C shows vaccination rates across different levels of search interest in select misinformation terms. The boxplots include: center line, median; box limits, upper and lower quartiles; whiskers, 1.5x interquartile range; points beyond whiskers, outliers.

correlated with vaccination rates in the United States. Search interest in general vaccination terms ("COVID vaccine"), vaccine appointment terms ("Can I get the vaccine" and "covid vaccine appointment"), and vaccine side effect and safety terms ("COVID Vaccine Reaction" and "COVID Vaccine Safety") are strongly correlated with vaccination rates. In contrast, search interest in "Ivermectin" shows a strong, negative association with vaccination rates. The results show that both vaccination rates and search interest are strongly associated with party affiliation: more democratic states are both more vaccinated and see greater search interest in vaccine appointments and side effects, whereas they see lower search interest in "Ivermectin."

Figure 5B shows the popularity of each keyword relative to the popularity of Ivermectin. The most popular search terms are "Vaccine" and "COVID Vaccine". Other keywords such as "Vaccine appointment"

and "COVID Vaccine Side Effects" also experience high volumes of search interest. Even though most misinformation search terms see low search interest, search interest for "Ivermectin" is relatively high compared to all terms and is notably higher compared to other misinformation terms. Search interest for "Vaccine appointment" is only 1.5 times larger than interest in "Ivermectin", whereas "Ivermectin" is over twice as popular as terms such as "where to get the COVID Vaccine" and "Vaccine Reaction" in the United States.

Ivermectin is not only the most popular misinformation term; it is also the misinformation term with the strongest negative association with vaccination rates. Panel C shows that high search interest in misinformation keywords does not necessarily translate into low vaccination rates. For instance, other less popular misinformation terms, such as "COVID Vaccine Change DNA", "COVID Vaccine Infertility", "COVID Vaccine Mercury", and "COVID Vaccine Microchip" are as popular in highly vaccinated states as in states with low vaccination rates. In contrast, "Ivermectin" does see greater search interest in low-vaccinated states, and a strong negative correlation overall. Thus, our results suggest that "Ivermectin" is the most relevant misinformation term in the United States both in terms of popularity as well as in terms of its association with lack of vaccination.

Discussion

In this paper, we demonstrate that Google search data can inform the monitoring and understanding of the COVID-19 pandemic using data across 207 countries. In particular, we show that Google search data can inform COVID-19 policies along three dimensions: (1) capturing COVID-19 caseloads, where trends in search interest across select search terms often precede trends in cases; (2) understanding the social, economic and mental health impact of COVID-19 shutdown policies; and (3) understanding information seeking for COVID-19 vaccines, including demand for vaccine misinformation.

First, we find that search interest in symptoms that are more specific to COVID-19 (particularly, loss of smell and taste) are highly correlated with COVID-19 cases throughout the world. On average, search interest is most highly correlated with COVID-19 cases following a lag of 12 or 13 days in 2020 and a lag of 2 to 6 days in 2021, indicating that searches for COVID-specific symptoms tend to precede trends in COVID-19 cases. This time lag provides valuable time to monitor a possible change in cases and anticipate new waves. This result highlights the potential of Google search data to support the monitoring and prediction of COVID-19 cases throughout the world.

Importantly, we find that the capacity of Google Trends to monitor and anticipate COVID-19 cases is not limited to high-income countries. On average the results are as predictive of new cases in lower-income countries as they are in higher-income countries. Considering the limited resources that developing countries

often have to monitor the spread of the pandemic, the opportunities generated by Google search data are particularly promising for such countries.

Our results also highlight the importance of keeping track of changes in variant-specific symptoms to monitor the spread of the disease, and the contribution of more general search terms. This pattern is visible in the recent surge of the Omicron variant, where search interest in loss of smell and taste is not as associated with COVID-19 cases as with previous variants in select contexts, but where more general searchers such as "COVID symptoms" remained predictive of cases [5]. For example, South Africa was the first country to report the Omicron variant, experiencing a subsequent surge in COVID-19 cases in December 2021. Until then, searches for "loss of smell" strongly tracked COVID-19 cases in the country. However, after the surge in cases from the Omicron variant in December 2021, the country did not experience an associated increase in searches for "loss of smel". In contrast, the United States does see an increase in "loss of smell" in the latest surge, consistent with the fact that the Omicron variant did not become predominant until the last week of December [14]. Furthermore, in South Africa, the United States, and other countries, search interest in "COVID symptoms" increased with the rise in cases in December 2021. These results point to the importance of understanding the specific symptoms of a disease or COVID variant to be able to monitor searches related to such disease, as well as exploring other search terms that may remain predictive of cases despite a change in the symptoms.

Second, we find that internet searches are a valuable tool to understand the impacts of COVID-19 shutdown policies—such as school closures, stay at home requirements, and travel restrictions—across the world. Pooling together all countries with available data and using a difference-in-differences approach, we find that containment policies resulted in an increase in search interest in keywords relating to unemployment, mental health, and social distancing, and a reduction in search interest in keywords related to relationships and family planning.

Additionally, we find notable heterogeneity in the impacts of containment policies across different levels of economic support and mobility restrictions. We find that countries with more restrictive containment policies saw higher search interest in most mental health keywords and select unemployment keywords, and a reduction in search interest in all relationship and family planning keywords. Overall, the results suggest that countries with more restrictive policies experienced more pronounced economic, mental health, and social externalities. We also find suggestive evidence that higher economic support may have helped to blunt select mental health impacts of containment policies, particularly anxiety and suicide.

Third, we show that Google search data can be used to help understand the demand for information (both legitimate and false) about COVID-19 vaccines. Globally, we find that vaccinations are associated with a higher demand for information about (1) vaccine appointments and how to access vaccines and (2)

COVID vaccine side effects. Fewer countries see high search interest in misinformation search terms. The association between search interest in misinformation and vaccination rates is more mixed across countries, indicating that search interest in misinformation follows more country-specific dynamics.

Focusing on the United States, we find that search interest in vaccine appointments and side effects are strongly and positively correlated with both vaccination rates and Democratic party affiliation. States with lower vaccination rates tend to see higher search interest in only select misinformation search terms, particularly "Ivermectin", which is also the most popular misinformation term. In contrast, most other misinformation search terms see no or minimal association with both vaccination rates and party affiliation—suggesting only specific search terms might be relevant at understanding the spread of misinformation and the association with lack of vaccinations. Our results suggest that Google Trends may help identify the most popular and powerful misinformation theories within countries, as is the case with Ivermectin in the United States.

Relying on Google search interests comes with a set of limitations. For example, search interest can be driven by global news events rather than country-specific dynamics. Search interest serves as an imperfect—albeit often strong—proxy for on-the-ground conditions. However, a key advantage is that search interest data is free and immediately accessible in real-time. Moreover, our results emphasize that Google search data can help with monitoring the spread and consequences of the pandemic even in low-resource settings. By providing insights across multiple phases of the pandemic, our results demonstrate the potential of using the widespread expansion of technology to better understand and address pressing, real-time issues that affect citizens throughout the entire world.

Methods

In this section, we first describe Google search interest data, which is used throughout the paper. We then describe the data and methodology used for each of the three sets of analyses: (1) search interest correlating with and preceding cases, (2) impact of containment policies on search interest indicators, and (3) using search interest to understand information-seeking around vaccines.

Google Search Interest Data

Google search interest data was queried using the gTrends R package for the period between September 1, 2018 to December 31, 2021 at the daily level. We query data across 207 countries and for 68 search terms related to COVID-19 symptoms, mental health, possible social and economic consequences of the pandemic, and vaccine-related keywords (appendix S1 provides a list of search terms and the number of countries with

data for each term). Across all analysis, we rely on the seven-day moving average value of search interest, which helps to smooth over spikes in the data.

For each country, we use a translated version of search terms using the language with the highest search activity in Google. We determine which language has the highest search activity by comparing search interest for "fever", "doctor", and "hospital" across the most widely used languages in each country. Translations are done using the Google Translate API; appendix S2 details the specific steps used to determine the language with the highest search activity in each country.

Search interest information from Google reflects the absolute number of searches of a keyword or phrase relative to the total number of searches over a specific geographic area and period of time [19]. Consequently, the search interest indicator reflects the relative popularity of a keyword as opposed to the absolute number of searches. Google scales the search interest value so that the maximum value is 100 for a given query (e.g., search interest for "loss of smell" in the United States is scaled so that maximum value is 100). Google only provides search interest values if there is sufficient search activity; consequently, a search interest value of zero indicates that there could be some search interest for a search term, but there was not sufficient data.

One challenge is that Google only provides daily values of search interest when 270 days or less are queried; given that Google scales values for each query between 0 and 100, the raw search interest values will not be comparable across queries of different date ranges. We create a consistent time series by querying search interest across time periods that overlap, then—relying on search interest in overlapping time periods—scale the data sets to create a consistent time series (see appendix S3 for additional details). We use separate queries for each country and search term; consequently, the search interest data we query is only comparable across time within a specific country and keyword.

Google search interest correlating and predicting COVID-19 cases

The first set of analysis seeks to understand which search terms correlate with COVID-19 cases and whether trends in search interests precede trends in cases. Daily COVID-19 case data comes from the World Health Organization [33].

For this analysis, we focus on keywords relating to the coronavirus generally ("coronavirus", "covid-19", "how to treat coronavirus"), symptoms that are more common to COVID-19 ("loss of taste", "loss of smell", "I can't smell", "I can't taste") and more general symptoms and associated diseases ("fever", "cough", "tired", "pneumonia"). For each country and keyword, we compute the correlation coefficient between search interest and cases to understand which search terms tend to see the highest correlations across countries. When computing the correlation, we use a seven-day moving average of both search interest and cases. The

seven-day moving average helps to smooth over spikes in the data; for example, in instances where COVID-19 cases for the weekend and Monday are all reported on Monday. Second, following [12], [39] and [24], we use a lag correlation analysis to understand whether search interest correlates more strongly with cases when using a lagged value of cases. We compute the correlation between search interest and cases by using lead and lag values of cases up to 21 days; the 21 day lead/lag follows from [24]. Search interest correlating most strongly with lagged cases would suggest that search interest picks up on trends in patterns before they are reflected in official data; search interest correlating most strongly with a lead (positive shift) in cases would suggest that search interest just reacts to on-the-ground conditions.

A potential concern of using Google trends for evaluating the change in COVID-19 cases is that correlations may diminish over time; as the public's understanding of COVID-19 increases over time, there may be less of a need to turn to Google during a rise in cases. To test diminishing correlations, we compute the correlation using data from 2020 and 2021 separately.

We also examine potential factors that may explain where Google search interest provides the best proxy for COVID-19 cases. Here, we regress a set of country-level indicators on the correlation between COVID-19 cases and search interest in select keywords. We use the correlation using data from January 1, 2020 until December 31, 2021. For explanatory variables, we use the total number of COVID-19 cases as of December 31, 2021, metrics that capture online presence (percent of the population that uses internet and mobile cell subscribes per 100, both captured via the World Development Indicators, WDI) and GDP per capita (also captured in the WDIs). We estimate a similar regression using the lead/lag value that produces the highest correlation as the dependent variable. This analysis aims to understand which factors may increase or reduce the time lag between search interest and COVID-19 case reporting.

Impact of COVID-19 containment policies on social, mental health, and economic search interest indicators

Data on containment policies come from the University of Oxford COVID-19 Government Response Tracker dataset [18]. To evaluate the impact of containment policies, we rely on the first date that a country implemented any containment or closure policy, such as school closures, workplace closures, cancellation of public events, restrictions on gatherings, public transportation closures, stay at home requirements, restrictions on internal movements, and international travel controls. We examine the impact of containment policies on search interest for terms related to social distance measures ("social distance" and "stay at home"); mental health ("anxiety", "anxiety attack", "boredom", "insomnia", "lonely", "panic", "social isolation", and "suicide"); relationship and family planning indicators ("divorce", "wedding", "emergency pill", and "pregnancy

test"); and economic, unemployment, and unemployment resource indicators ("debt", "unemployment", "file for unemployment", "unemployment benefits", "unemployment insurance", and "unemployment office."). These search terms have been previously studied for a limited number of countries [4, 8, 29]; however, in this paper we extend the analysis to all countries with available data, which makes it possible to compare the effects across countries with different containment policies.

We explore how impacts of containment policies vary with the (1) restrictiveness of containment policies, (2) the level of government economic support, and (3) income level (GDP per capita). Specifically, we aim to understand whether economic and mental health impacts (proxied by search interest) were larger in countries with more restrictive policies, whether economic support blunted any negative externalities of policies, and whether impacts vary systematically by income level. We use two datasets to measure the restrictiveness of containment policies. First, we use Google COVID-19 Community Mobility Reports, which measure the percent change in mobility relative to pre-pandemic levels to a number of types of locations [16]. We use the average percent change in mobility across retail and recreation locations, groceries and pharmacies, parks, transit stations and workplaces. Second, we use the Stringency Index from the University of Oxford COVID-19 Government Response tracker, a composite measure of the restrictiveness of nine policy measures, including workplace closures, public transport closures, and restrictions on public gathers, among other metrics [18]. To measure the level of government economic support, we use the Economic Support index from the Oxford COVID-19 Government Response tracker, which measures the extent of economic support across metrics such as income support and debt relief.

As our primary model, we implement a difference-in-differences model using data from the year before the beginning of the pandemic. We estimate the following model, using data from 90 days before and after the first containment policy in each country:

$$\begin{aligned} y_{c,d,p} = \beta_0 + \beta_1 \textit{Post Policy Period}_d \times \textit{Pandemic Period}_y + \\ \beta_2 \textit{Post Policy Period}_d \times \textit{Pandemic Period}_y \times \textit{Country Feature}_c + \\ \delta_c + \gamma_d + \epsilon_{c,d,p} \end{aligned} \quad (1)$$

where $y_{c,d,p}$ is the search interest for country c (using the seven-day moving average search interest), d days since the day and month the policy was implemented (irrespective of the year), and where p is the period (either pre-pandemic or pandemic period). *Post Policy Period* is a binary variable that turns on after the day and month the policy was implemented (irrespective of the year) and *Pandemic Period* is a binary

variable that turns on in 2020. *Country Feature* is a country-level variable used to understand how the impact of containment policies varies across countries; we use four features: (1) the Economic Support index, (2) the Stringency index, (3) mobility reduction, and (4) GDP per capita. We run separate regressions for each *Country Feature*. We scale all four features so that they have a mean of zero and standard deviation of one. For economic support, mobility reduction, and stringency index, we use the most extreme value in the 90-day period after the first containment policy was implemented. γ are day of week fixed effects and δ are country fixed effects. In the regression, we also include all individual variables and two-way interactions of $Post\ Policy\ Period_d \times Pandemic\ Period_y \times Country\ Feature_c$. To test the sensitivity of results to the 90-day threshold, we also estimate models using 30, 60 and 120-day thresholds.

To further test the sensitivity of results, we also estimate an event study approach that just relies on the time period from before and after policies were implemented. Here, we estimate the following model:

$$\begin{aligned} y_{i,t} = \beta_0 &+ \sum_{f=-31}^{-2} \beta_f I(date - R_i = f) \\ &+ \sum_{l=0}^{31} \beta_l I(date - R_i = l) \\ &+ \gamma_i + \delta_t + \epsilon_{i,t} \end{aligned}$$

where $y_{i,t}$ is the search interest value for a specific search term for country i on date t t, R_i is the date when the first policy was implemented for country i. We subset the data to 90 days before and after the policy was implemented.

Using Google search interest to understand vaccine information-seeking and vaccine hesitancy

We use data on vaccination uptake to understand how search interest correlates with vaccination rates. We use vaccination data from [23], where we use the percent of people vaccinated on a daily basis. We query search interest for keywords across four categories: (1) general vaccination terms, (2) vaccine side effects and safety, (3) vaccine appointments and (4) misinformation. For misinformation, we use search terms related to a number of false theories, including: (1) misinformation about Ivermectin—an anti-parasitic drug—being an effective cure and preventive drug for COVID despite a lack of evidence, and which has been touted as an effective alternate to the vaccine [34, 21]; (2) false theories about COVID vaccines causing infertility [2, 41], changing one's DNA [9], containing high levels of harmful elements such as mercury [35], or injecting a microchip to be used for tracking [9]; and (3) concerns about the COVID vaccine being the "mark of the

beast" or associated with the devil [10].

Using data across 207 countries with available vaccine and search interest data, we examine whether daily vaccinations correlate with search interest within countries. In computing the correlation, for each country we use data after the first vaccination was given. This analysis allows understanding what information people seek as vaccinations increase. Importantly, the analysis allows assessing the extent to which people demand legitimate (e.g., side effects from the vaccines) versus false (e.g., vaccines changing one's DNA) information about vaccines.

For country-level data we query daily search interest data (where values are comparable across time, within countries); however, for United States state-level data, we query an aggregate measure of search interest for December 1, 2020 to December 31, 2021, where search interest values are directly comparable across states (we use December 1, 2020 as the start date as the first vaccine in the United States was administered in December, 2020 [17]). First, we rely on correlation analysis to observe how search interest in select terms are correlated with vaccination rates. Second, we focus specifically on how misinformation-related search terms relate to vaccination rates. For all misinformation search terms except "Ivermectin", most States see little or no search interest. Consequently, for misinformation terms, instead of implementing a correlation analysis, we compare the distribution in vaccination rates across states with (1) no search interest in the select search term, (2) low search interest (below median search interest among states with positive search interest) and (3) high search interest. In addition, given that political ideology is considered a key factor in understanding vaccine sentiment in the United States [1], we also examine differences across states which voted for Trump or Biden during the 2020 election.

In addition to relating search interest to vaccination rates, we also use Google trends to understand the relative popularity of search terms related to general vaccine interest, vaccine appointments, vaccine side effects, and misinformation. In total, we compare search interest across 34 search terms. Google trends only allows querying up to five search terms at a time to understand their relative popularity, where search interest across different queries with different search terms are not directly comparable. In order to compare all 34 search terms, for each search term we query both the search term and "Ivermectin", then compute a metric of the search popularity of the search term relative to the popularity of "Ivermectin." While the raw search interest values are not comparable, the metric of popularity of the search term relative to "Ivermectin" is comparable across all search terms. We use "Ivermectin" as the reference search term for two reasons: (1) first, we focus the analysis on search interest popularity of search terms relative to misinformation terms and (2) Ivermectin was by far the most popular misinformation search term in Google across the nine misinformation related terms we consider.

Data Availability

The data used in this paper are available at https://github.com/worldbank/covid-gtrends. In addition, the data can be interactively explored in this dashboard: https://datanalytics.worldbank.org/covid_gtrends/

Code Availability

The data used in this paper are available at https://github.com/worldbank/covid-gtrends

References

[1] Which americans are against the jab? *The Economist*.

[2] Jennifer Abbasi. Widespread Misinformation About Infertility Continues to Create COVID-19 Vaccine Hesitancy. *JAMA*, 02 2022.

[3] Clive Bell and Lyn Squire. Providing Policy Makers with Timely Advice: The Timeliness-Rigor Trade-off. *The World Bank Economic Review*, 31(2):553–569, 03 2016.

[4] Lawrence M. Berger, Giulia Ferrari, Marion Leturcq, Lidia Panico, and Anne Solaz. Covid-19 lockdowns and demographically-relevant google trends: A cross-national analysis. *PLOS ONE*, 16(3):1–28, 03 2021.

[5] Lin T. Brandal, Emily MacDonald, Lamprini Veneti, Tine Ravlo, Heidi Lange, Umaer Naseer, Siri Feruglio, Karoline Bragstad, Olav Hungnes, Liz E. Ødeskaug, Frode Hagen, Kristian E. Hanch-Hansen, Andreas Lind, Sara Viksmoen Watle, Arne M. Taxt, Mia Johansen, Line Vold, Preben Aavitsland, Karin Nygård, and Elisabeth H. Madslien. Outbreak caused by the sars-cov-2 omicron variant in norway, november to december 2021. *Eurosurveillance*, 26(50), 2021.

[6] Scott A. Brave, Andrew Butters R., and Michael Fogarty. A Closer Look at the Correlation Between Google Trends and Initial Unemployment Insurance Claims. *Chicago Fed Insights*.

[7] Abel Brodeur, Andrew E. Clark, Sarah Fleche, and Nattavudh Powdthavee. Covid-19, lockdowns and well-being: Evidence from google trends. *Journal of public economics*, 193:104346–104346, Jan 2021. 33281237[pmid].

[8] Abel Brodeur, Andrew E. Clark, Sarah Fleche, and Nattavudh Powdthavee. Covid-19, lockdowns and well-being: Evidence from google trends. *Journal of Public Economics*, 193:104346, 2021.

[9] Mayo Clinic. Covid queries: Altered dna and microchips. `https://newsnetwork.mayoclinic.org/discussion/covid-queries-altered-dna-and-microchips/`. Accessed: 2022-03-07.

[10] Gordon Darroch and Asha C. Gilbert. Some say covid-19 vaccine is the 'mark of the beast.' is there a connection to the bible? *USA Today*.

[11] Sebastian Doerr and Leonardo Gambacorta. Identifying regions at risk with Google Trends: the impact of Covid-19 on US labour markets. *BIS Bulletin*, (8), 4 2020.

[12] Maria Effenberger, Andreas Kronbichler, Jae Il Shin, Gert Mayer, Herbert Tilg, and Paul Perco. Association of the covid-19 pandemic with internet search volumes: A google trends(tm) analysis. *In-*

ternational journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases, 95:192–197, Jun 2020. 32305520[pmid].

[13] Centers for Disease Control and Prevention. COVID-19 vaccinations in the united states. `https://covid.cdc.gov/covid-data-tracker/#vaccinations_vacc-people-onedose-pop-5yr`. Accessed: 2022-03-07.

[14] Centers for Disease Control and Prevention. Variant proportions. `https://covid.cdc.gov/covid-data-tracker/#variant-proportions`. Accessed: 2022-03-07.

[15] Jeremy Ginsberg, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, February 2009.

[16] Google. Covid-19 community mobility reports. `https://www.google.com/covid19/mobility/`. Accessed: 2022-01-15.

[17] Ben Guarino, Ariana Eunjung Cha, Josh Wood, and Griff Witte. 'the weapon that will end the war': First coronavirus vaccine shots given outside trials in u.s. *The Washington Post*.

[18] Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature Human Behaviour*, 5(4):529–538, Apr 2021.

[19] Google Help. FAQ about google trends data. `https://support.google.com/trends/answer/4365533?hl=en`. Accessed: 2022-01-15.

[20] Michael Hoerger, Sarah Alonzi, Laura M. Perry, Hallie M. Voss, Sanjana Easwar, and James I. Gerhart. Impact of the covid-19 pandemic on mental health: Real-time surveillance using google trends. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(6):567–568, 2020.

[21] Pien Huang. How ivermectin became the new focus of the anti-vaccine movement. *NPR*.

[22] Atina Husnayain, Anis Fuad, and Lutfan Lazuardi. Correlation between google trends on dengue fever and national surveillance report in indonesia. *Global Health Action*, 12(1):1552652, 2019. PMID: 31154985.

[23] Our World in Data. Covid-19 vaccine doses administered per 100 people. `https://ourworldindata.org/grapher/covid-vaccination-doses-per-capita`. Accessed: 2022-01-15.

[24] Alberto Jimenez Jimenez, Rosa M. Estevez-Reboredo, Miguel A. Santed, and Victoria Ramos. Covid-19 symptom-related google searches and local covid-19 incidence in spain: Correlational study. *Journal of medical Internet research*, 22(12):e23518–e23518, Dec 2020. 33156803[pmid].

[25] Sasikiran Kandula and Jeffrey Shaman. Reappraising the utility of google flu trends. *PLOS Computational Biology*, 15(8):1–16, 08 2019.

[26] Shyam J. Kurian, Atiq Ur Rehman Bhatti, Mohammed Ali Alvi, Henry H. Ting, Curtis Storlie, Patrick M. Wilson, Nilay D. Shah, Hongfang Liu, and Mohamad Bydon. Correlations between covid-19 cases and google trends data in the united states: A state-by-state analysis. *Mayo Clinic proceedings*, 95(11):2370–2381, Nov 2020. 33164756[pmid].

[27] David Lazer, Ryan Kennedy, Gary King, and Alessandro Vespignani. The parable of google flu: Traps in big data analysis. *Science*, 343(14 March):1203–1205, 2014.

[28] Cuilian Li, Li Jia Chen, Xueyu Chen, Mingzhi Zhang, Chi Pui Pang, and Haoyu Chen. Retrospective analysis of the possibility of predicting the covid-19 outbreak from internet searches and social media data, china, 2020. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin*, 25(10):2000199, Mar 2020. 32183935[pmid].

[29] Tichakunda Mangono, Peter Smittenaar, Yael Caplan, Vincent S Huang, Staci Sutermaster, Hannah Kemp, and Sema K Sgaier. Information-seeking patterns during the covid-19 pandemic across the united states: Longitudinal analysis of google trends data. *J Med Internet Res*, 23(5):e22933, May 2021.

[30] Gabriel J. Milinovich, Ricardo J. Soares Magalhães, and Wenbiao Hu. Role of big data in the early detection of ebola and other emerging infectious diseases. *The Lancet Global Health*, 3(1):e20–e21, Jan 2015.

[31] Sudhakar V. Nuti, Brian Wayda, Isuru Ranasinghe, Sisi Wang, Rachel P. Dreyer, Serene I. Chen, and Karthik Murugiah. The use of google trends in health care research: a systematic review. *PloS one*, 9(10):e109583–e109583, Oct 2014. 25337815[pmid].

[32] World Health Organization. Classification of omicron (b.1.1.529): Sars-cov-2 variant of concern. `https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern`. Accessed: 2022-03-07.

[33] World Health Organization. Who coronavirus (covid-19) dashboard. `https://covid19.who.int/`. Accessed: 2022-01-15.

[34] Maria Popp, Miriam Stegemann, Maria-Inti Metzendorf, Susan Gould, Peter Kranke, Patrick Meybohm, Nicole Skoetz, and Stephanie Weibel. Ivermectin for preventing and treating covid-19. *Cochrane Database of Systematic Reviews*, (7), 2021.

[35] Reuters. Fact check: Covid-19 vaccines do not contain the ingredients listed in these posts. `https://www.reuters.com/article/uk-factcheck-covid-vaccine-ingredients/fact-check-covid-19-vaccines-do-not-contain-the-ingredients-listed-in-these-posts-idUSKBN2AQ2SW`. Accessed: 2022-03-07.

[36] Stefan Stijelja and Brian L. Mishara. COVID-19 and Psychological Distress—Changes in Internet Searches for Mental Health Issues in New York During the Pandemic. *JAMA Internal Medicine*, 180(12):1703–1706, 12 2020.

[37] Yue Teng, Dehua Bi, Guigang Xie, Yuan Jin, Yong Huang, Baihan Lin, Xiaoping An, Dan Feng, and Yigang Tong. Dynamic forecasting of zika epidemics using google trends. *PLOS ONE*, 12(1):1–10, 01 2017.

[38] Nataliya Tkachenko, Sarunkorn Chotvijit, Neha Gupta, Emma Bradley, Charlotte Gilks, Weisi Guo, Henry Crosby, Eliot Shore, Malkiat Thiarai, Rob Procter, and Stephen Jarvis. Google trends can improve surveillance of type 2 diabetes. *Scientific Reports*, 7(1):4993, Jul 2017.

[39] U. Venkatesh and Periyasamy Aravind Gandhi. Prediction of covid-19 outbreaks using google trends in india: A retrospective analysis. *Healthcare informatics research*, 26(3):175–184, Jul 2020. 32819035[pmid].

[40] Abigail Walker, Claire Hopkins, and Pavol Surda. Use of google trends to investigate loss-of-smell–related searches during the covid-19 outbreak. *International Forum of Allergy & Rhinology*, 10(7):839–847, 2020.

[41] Amelia K Wesselink, Elizabeth E Hatch, Kenneth J Rothman, Tanran R Wang, Mary D Willis, Jennifer Yland, Holly M Crowe, Ruth J Geller, Sydney K Willis, Rebecca B Perkins, Annette K Regan, Jessica Levinson, Ellen M Mikkelsen, and Lauren A Wise. A Prospective Cohort Study of COVID-19 Vaccination, SARS-CoV-2 Infection, and Fertility. *American Journal of Epidemiology*, 01 2022. kwac011.

[42] Sean D. Young, Elizabeth A. Torrone, John Urata, and Sevgi O. Aral. Using search engine data as a tool to predict syphilis. *Epidemiology (Cambridge, Mass.)*, 29(4):574–578, Jul 2018. 29864105[pmid].

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- MartySI.pdf