

DO CELEBRITY ENDORSEMENTS MATTER? A TWITTER EXPERIMENT PROMOTING VACCINATION IN INDONESIA

VIVI ALATAS*, ARUN G. CHANDRASEKHAR[†], MARKUS MOBIUS[§], BENJAMIN A. OLKEN[‡],
AND CINDY PALADINES**

ABSTRACT. Do celebrity endorsements matter? And if so, how can celebrities communicate effectively? We conduct a nationwide Twitter experiment in Indonesia promoting vaccination. Celebrity messages are 72 percent more likely to be passed on or liked than similar messages without a celebrity’s imprimatur. In total, 66 percent of the celebrity effect comes from authorship, compared to passing on messages. Citing external medical sources decreases retweets by 27 percent. Phone surveys show that those randomly exposed to messaging have fewer incorrect beliefs and report more vaccination among friends and neighbors. The results can inform public health campaigns and celebrity public service more generally.

Date: October 2, 2023.

We thank Ran Abramitsky, Marcella Alsan, Nancy Baym, Emily Breza, Leo Bursztyn, Rebecca Diamond, Dean Eckles, Paul Goldsmith-Pinkham, Ben Golub, Mary Gray, Rema Hanna, Johannes Haushofer, Matt Jackson, Tyler McCormick, Madeline McKelway, Matthew Wai-Poi, Alex Wolitsky, and participants at various seminars for helpful discussions. Data Avicenna, Clotaire Boyer, Aaron Kaye, Nurzanty Khadijah, Devika Lakhote, Eva Lyubich, Sinead Maguire, Lina Marliani, Sebastian Steffen, Vincent Tanutama provided excellent research assistance. We thank Nila Moeloek, then Indonesian Special Envoy for Sustainable Development Goals, Diah Saminarsih, and their team for providing support for this project. This study was approved by IRBs at MIT (Protocol #1406006433) and Stanford (Protocol #31451), and registered in the AEA Social Science Registry (AEARCTR-0000757). Alatas and Paladines were employed by the World Bank during the period this project was launched. Funding for this project came from the Australian Government Department of Foreign Affairs and Trade. The views expressed here are those of the authors only and do not represent those of any of the institutions or individuals acknowledged here.

*Asa Kreativita.

[†]Department of Economics, Stanford University; NBER; J-PAL.

[§]Microsoft Research, New England.

[‡]Department of Economics, MIT; NBER; J-PAL.

**TCW.

1. INTRODUCTION

Social media has allowed celebrities to take an increasing role in social discourse. With millions of followers, celebrities have a channel to spread messages on many issues, including some far removed from their original reason for fame. Their participation in ongoing discussions can make issues prominent and shape the zeitgeist.

Examples abound, from #BlackLivesMatter, for racial justice, to the #IceBucketChallenge, promoting awareness of Lou Gehrig’s disease. Each of these campaigns was initiated by a less-well-known activist, but made prominent in part through celebrity participation. Celebrities are now increasingly being recruited as public health messengers.

Using celebrities effectively, however, depends on what features of the celebrity messaging spur diffusion. Beyond the direct *reach* that celebrities have – they have numerous followers – there are several dimensions that could affect people’s decision to further spread or follow a celebrity’s message. First, how much does celebrity *endorsement*, meaning involvement with the messaging beyond the simple fact that they have numerous followers, matter? Second, how much of an endorsement premium comes from a direct *authorship* effect, as opposed to relaying the message of others? Third, how much does inclusion of *credible sourcing* matter, particularly when celebrities are speaking on a topic removed from their core area of expertise, such as public health?

While celebrity endorsements may have effects for commercial products, whether their endorsement matters in the context of public health remains a matter of public debate, especially since public health issues are often far from their main area of expertise. For instance, in a May 2021 New York Times article (Ives, 2021), epidemiologists and psychologists, based on focus groups, argue that celebrity endorsements may not address COVID-19 vaccine indifference and hesitancy, despite the fact that celebrities were widely recruited to help spread COVID-19 public health messages.¹ Despite widespread interest in using celebrities to promote vaccination, there remains little rigorous empirical evidence on whether this matters, and if so, how to effectively design celebrity outreach campaigns.

Measuring and decomposing the endorsement, authorship, and credible sourcing effects, as well as measuring whether any of these campaigns have offline effects, is challenging, as celebrities’ decisions are typically endogenous, and because people consume such a wide range of information that it is challenging to isolate their impact.

To study these issues, we conducted an experiment through a nationwide immunization campaign on Twitter from 2015-2016 in Indonesia, in collaboration with the Indonesian Government’s Special Ambassador to the United Nations for Millennium Development Goals. Working with the Special Ambassador, we recruited 46 high-profile celebrities and organizations, with a total of over 11 million followers, each of whom gave us access to send up to

¹See <https://www.nytimes.com/2021/05/01/health/vaccinated-celebrities.html>

33 tweets or retweets promoting immunization from their accounts. The content and timing of these tweets was randomly chosen from a bank approved by the Indonesian Ministry of Health, all of which featured a campaign hashtag #AyoImunisasi (“Let’s Immunize”).

To isolate the role of celebrity endorsement and sourcing, we randomly varied (1) Did the celebrity / organization send the tweet directly, or did they retweet a message (drawn randomly from the same tweet library) sent by us from a non-celebrity user’s account?; (2) Did the tweet explicitly cite a public health source? We also randomly varied (3) When did the celebrity tweet? This variation allows us to decompose and measure the relative importance of endorsement, authorship, and credible sourcing effects.

We study the effects of this induced variation in two ways. First, we use online reactions, i.e., likes and retweets, so we can observe the online reactions of every follower to each tweet, and trace out which randomized characteristics of the tweet lead to more approval and diffusion.² Second, we also study the offline effects of exposure to the campaign by conducting phone surveys of Twitter users. By randomly allocating celebrity activity into one of several phases – either before or after the survey – we can examine whether individuals who follow celebrities randomized to tweet before the survey are more likely to have heard of the campaign, updated their beliefs, discussed immunization status with their friends and neighbors, and observed changes in immunization behavior among their friends, relatives, and neighbors.

We chose this setting for several reasons. Twitter is one of the most important mediums of information exchange in the world, with over 1 billion users. Indonesia is quite active on social media; for example, in 2012 its capital, Jakarta, originated the most tweets of any city worldwide. Twitter also has many useful features for our study. Because both the network (i.e., who sees whose tweets) and virtually all information flows over the network (i.e., tweets and retweets) are public, we can precisely map who sees what information, as well as where they saw it, allowing us to observe how much exposure each user had to precise bits of information. By conducting an experiment, randomly varying who tweets what when, we can solve the identification problem of endogenous speaking behavior, as well as disentangle reach vs. endorsement effects. Immunization was chosen as it was a clear public health message, for which celebrities could rely on the Ministry of Health to provide trusted information. And, because of the public nature of Twitter, we can observe people’s responses to information both online (by observing their online “liking” and “retweeting” behavior) and offline (by conducting a phone survey of Twitter users and linking their survey responses to whether they were randomly exposed on Twitter to information prior to the survey).

We begin by using our design, in combination with the structure of how information is passed on Twitter, to distinguish reach from endorsement - i.e. how much additional effect

²Note that on Twitter, a “like” is not pushed to one’s followers, while a “retweet” subsequently passes on the tweet to all of one’s followers.

does the celebrity endorsement *per se* have, above and beyond the fact that the celebrity’s messages are directly seen by many followers. Messages in Twitter are passed on by retweeting a message to one’s followers. Crucially for our design, when a message is retweeted, the follower observes who originally composed the tweet, and who retweeted it directly to the follower, but not any intermediate steps in the path.

To see how this allows us to distinguish reach from endorsement, consider the difference between what happens when 1) we have a celebrity directly compose and tweet a message, compared to 2) when we have a celebrity retweet a message drawn from the same pool of tweets but originated by a normal citizen (whom we henceforth denote as “ordinary Joes and Janes”; these Joes/Janes are also participants in our campaigns). Consider what happens when some celebrity followers (whom we denote F_1) retweet it to their followers, whom we denote F_2 . In the first case, the followers-of-followers (F_2 s) observe that the celebrity authored the message and that F_1 retweeted it. But in the second case, when the celebrity retweeted a Joe/Jane message rather than composed it herself, the followers-of-followers of the celebrity (F_2 s) observe only that the Joe/Jane tweeted and that F_1 then retweeted for F_2 to see. Notice that in this way, F_2 is randomly blinded to the celebrity’s involvement in the latter case, as compared to the former: differences in F_2 ’s behavior therefore correspond to differences due to knowing that the celebrity was involved.³ In the period we study, the ordering of the Twitter feed was strictly chronological, so this design manipulates whether the F_2 s know about the celebrities’ involvement without affecting how prominently the message appeared in the Twitter feed.

We find a substantial endorsement effect. When an individual observes a message through a retweet, and that message was randomized to be originally composed by a celebrity rather than an ordinary individual, there is a 72 percent increase in the number of likes and retweets, compared to similar messages when the celebrity’s involvement was masked. We find similar results even when we restrict attention to those cases where F_1 s are participants in the experiment and we exogenously had them retweet the message, ensuring that whether the F_2 was exposed to the message in the first place was completely exogenous.

The preceding analysis does not distinguish between whether endorsement matters because a downstream individual knows that a celebrity was involved *per se*, or because the celebrity actually authored the message. We therefore next further decompose the endorsement effect to understand the impact of celebrities speaking in their own voice (an ‘authorship’ effect) from the celebrity passing on the message of others. We use the same experimental variation, but look at behavior of the direct followers of celebrities (F_1 s), who see both the celebrities’ directly authored tweets and the celebrities’ retweets. Tweets directly authored by celebrities

³A challenge is that the F_1 decision to retweet may be endogenous. We discuss this issue in detail in Section 3.1.1, and show that the results are largely similar in the subset of cases where F_1 s were also study participants whom we randomly selected and had retweet exogenously, and hence the sample of exposed F_2 s is identical.

are 200 percent more likely to be liked or retweeted, and in fact 280 percent more likely to be retweeted, than comparable messages passed on by comparable celebrities but authored by ordinary users. The majority – 66 - 76 percent – of the celebrity effect therefore comes from celebrities speaking in their own voice.

Celebrities also may choose to bolster their credibility by explicitly citing sources, which may be particularly important in a public health context where celebrities are speaking on subjects far from their own expertise. We find, however, that citing sources actually *reduces* message diffusion. The magnitudes are substantial: randomly attaching a source to a tweet reduces the probability of retweets by 27 percent. One interpretation is the information is less novel if it is sourced; more generally, in Online Appendix B we discuss theoretically how increasing the reliability of information passed has ex ante ambiguous effects on the probability the information is passed.

Taken together, our estimates allow us to decompose the celebrity effect: we estimate that 43 percent of the celebrity effect comes from authorship, 24 percent from endorsement, with the remainder attributable to the intrinsic interest of the message. The results suggest that celebrities can play an important role in the diffusion of public health messages, but to do so, they need to speak in their own voice.

Given that people seem to pay attention to the campaign online – liking and retweeting celebrity messages on immunization – we then turn to whether such a campaign has effects on real-world beliefs, knowledge, and behavior. To study this, we used the timing of the tweets to randomly generate differences in exposure to our campaign. Specifically, we randomized the celebrities into two groups, with the first group assigned to tweet during July and August 2015 (Phase I) and the second group assigned to tweet from November 2015 - February 2016 (Phases II and III). We conducted a phone survey of a subset of followers of our celebrities in between these two groups of tweets. Since we know which of our celebrities each of these followers followed at baseline, this randomization into two phases generates random variation in how many immunization-related tweets from our campaign each individual had potentially been exposed to as of the time of our survey.

The evidence using this variation, while suggestive, indicates that exposure to celebrity endorsements does have measurable effects. We find that people did pay attention: a one standard deviation increase in exposure to the campaign due to our randomization, equivalent to about 15 tweets or retweets showing up on a user’s Twitter feed over a period of about one month, corresponds to a 21 percent increase in the probability that the respondent in the phone survey knows about our hashtag, #AyoImunisasi; a 9 percent increase in the probability they have heard about immunization through Twitter; and a 13 percent increase in the number of times they report having heard about immunization through Twitter. We then show that exposure to the campaign may have increased knowledge about immunization. We asked phone respondents a number of factual questions about immunization (e.g., whether

the vaccine was domestically produced, an important public message for the Government as domestically produced vaccines are known to be halal and hence allowed under Muslim law), all of which were addressed in some of the campaign tweets. A one standard deviation increase in exposure to the campaign corresponds to a 4 percent increase in the probability that the respondent knows that vaccines are domestic (on a base of 58 percent in knowledge in the whole sample), though no increase in the three other dimensions of knowledge we examined.

We then turn to respondents' knowledge of immunization behavior among neighbors, friends, and relatives. Again we find effects of the celebrity pro-immunization campaign: a one standard deviation increase in exposure corresponds to a 5 percent increase in the probability of knowing about one's neighbors' recent immunization behavior. We find no increases in knowledge for friends and relatives. The idea that one would learn about immunization decisions of neighbors is consistent with immunization practices in Indonesia, which take place at *posyandu* meetings, staffed by a midwife, that occur each month in each neighborhood (*dusun* or *RW*) of Indonesia (Olken et al., 2014). Finally, we look at changes in reported immunization decisions of respondents and those of their neighbors, friends, and relatives. We find no effects on one's own immunization decisions, though our statistical power is such that we cannot rule out substantial effects. But casting a wider net, we find that those exposed to the campaign were more likely to report recent immunizations among neighbors, friends, and relatives. In sum, while the estimates in each domain are suggestive, taken together we find consistent evidence that celebrity endorsements actually may affect both offline knowledge about immunization, as well as the knowledge of health status and health-seeking behavior by one's neighbors, friends, and relatives.

Our study builds on several facets of the literature. First, we build on the recent literature on diffusion of information for public policy and the computer science literature on generating online cascades (e.g., Leskovec, Adamic, and Huberman, 2007; Katona, Zubcsek, and Sarvary, 2011; Bakshy, Hofman, Mason, and Watts, 2011; Banerjee, Chandrasekhar, Duflo, and Jackson, 2013; Beaman, BenYishay, Magruder, and Mobarak, 2021; Beaman and Dillon, 2018). While this literature has studied the flow of information over social networks, and how network position affects the flow of information, it has typically been silent on what aspects of the message matter.

Second, we build on the extensive literature that has studied celebrity endorsements, primarily in commercial advertising. This literature has examined the impacts of celebrity endorsements of commercial products on outcomes such as stock prices (e.g., Agrawal and Kamakura, 1995), sales (e.g., Elberse and Verleun, 2012; Garthwaite, 2014), and brand evaluations, and studying various aspects of the celebrity's identity (e.g., gender, attractiveness); see Bergkvist and Zhou (2016) for a comprehensive review. Our study is one of the first to study celebrity effects in the online space through a real-world, large-scale field experiment,

and the first large-scale field experiment we know of to study public health messaging of any type. Indeed, the only study of a similar magnitude we know of is a marketing study by Gong, Zhang, Zhao, and Jiang (2017), who experimentally vary tweets in China on Sina Weibo about TV programs, randomizing whether these tweets were retweeted by influencers. Our study builds on this by *decomposing* the celebrity effect, identifying the value of authorship *per se* as opposed to relaying others’ messages, and identifying the value of credible sources, i.e., health authorities in a public health context.

2. EXPERIMENT

2.1. Setting and Sample. Our study took place in Indonesia in 2015 and 2016 on Twitter, which is one of the most important mediums of information exchange in the world, with over 1 billion users. Indonesia is very active on social media, ranking third worldwide with 130 million Facebook accounts⁴ in 2020 (about half the population), and ranking eighth with 10.6 million Twitter accounts (about 6.4 percent of the population).⁵

The experiment focused on immunization. At the time, Indonesia was trying to improve immunization as part of its drive towards the Millennium Development Goals. A set of 550 tweets was developed in coordination with the Ministry of Health that sought to improve information about immunization. The tweets included information about access (e.g., immunizations are free, available at government clinics, and so on); information about immunization’s importance (e.g., immunizations are crucial to combat child diseases); and information designed to combat common myths about immunization (e.g., vaccines are made domestically in Indonesia, rather than imported). For each tweet, we identified a source (either a specific link or an organization’s Twitter handle). All tweets were approved by the Ministry of Health, and included a common hashtag, #AyoImunisasi (“Let’s Immunize”).

With help from the Indonesian Special Ambassador to the United Nations for Millennium Development Goals, we recruited 37 high-profile Twitter users, whom we denote “celebrities,” with a total of 11 million Twitter followers. These “celebrities” come from many backgrounds, including music stars, TV personalities, actors and actresses, motivational speakers, government officials, and public intellectuals. They have a mean of 262,647 Twitter followers each, with several having more than one million. While these celebrities primarily tweet about things pertaining to their reason for fame, they also comment occasionally on public issues, so tweets about immunization would not necessarily have been unusual.⁶ We also recruited

⁴<https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>

⁵<https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>

⁶For example, three celebrities in our sample (a musician, a TV personality, and a musician’s agent) had recently tweeted about the importance of breakfast, including a link to an article about the health benefits

9 organizations involved in public advocacy and/or health issues in Indonesia with a mean of 132,300 followers each.

In addition, we recruited 1,032 ordinary citizens, primarily Indonesian university students, whom we call “Joes and Janes”. Their role will be to allow us to have everyday individuals compose tweets that are then retweeted by celebrities. Their Twitter profiles are far more typical, with a mean of 511 followers.

Every participant (celebrities and Joes/Janes) consented to signing up with our app that (1) lets us tweet content from their account (13, 23, or 33 times), (2) randomizes the content of the tweets from a large list of 550 immunization tweets approved by the Ministry of Health, and (3) has no scope for editing.⁷ Participants were given two choices: (1) the maximum number of tweets (13, 23, or 33), and (2) a choice of formal or casual Indonesian language (to better approximate their normal writing style).

2.2. Experimental Design. Our experiment is designed to understand which aspects of social media campaigns are important for disseminating a message. Ex ante it may seem obvious, for instance, that sources are better (after all, the information is more credible) and celebrity involvement is better (after all, for a variety of reasons the information may be viewed as more credible). But thinking carefully about the information sharing process demonstrates that, in fact, the effect of each of these design options is actually theoretically ambiguous.⁸

We focus on two main interventions: (1) whether a tweet was tweeted directly by a celebrity, or tweeted by a Joe/Jane and then retweeted by a celebrity; and, (2) for a subset of tweets, whether the tweet included a credible source (i.e., the source link or referring organization’s Twitter handle).⁹ We also randomized the timing of tweets (matching the empirical frequency of local time-of-day of Indonesian tweets), and the content (i.e., which tweet from our pre-prepared bank of approved tweets was tweeted by whom and when).¹⁰

Randomizing whether a tweet was directly tweeted by a celebrity or retweeted – in combination with the particular way Twitter messages are shown – allows us to develop a novel

of children’s breakfast; an athlete tweeted about supporting education for young children; and a musician tweeted in support of Asia Against AIDS.

⁷Celebrities could veto a tweet if they did not want it sent from their account, though this in fact never happened.

⁸Appendix B presents an application of a simple model developed independently by Chandrasekhar, Golub, and Yang (2018) to demonstrate the ambiguity, though certainly other models can be used and this is inessential for the empirical analysis.

⁹A small subset of tweets on topics deemed ‘sensitive’ by the Government always included a source; these are excluded from the analysis of sourcing.

¹⁰Note that in the period we study, a Twitter user saw all tweets and retweets from the users they follow in strict reverse chronological order (i.e., newest tweets appeared first, and so on). Twitter subsequently (in March 2016) applied an algorithm to prioritize the ordering of the tweets, but since in the period we study (July 2015 through February 2016) tweets appeared in strictly chronological order, nothing in our experimental design affects the ordering of tweets in a user’s Twitter feed.

test for isolating celebrity effects *per se*. Specifically, we use the fact that Twitter messages show the identity of only two people: the originator who wrote the tweet, and the person whom you follow who directly passed it to you (call this person F). Other steps along the chain are omitted.

Figure 1 illustrates the design. Consider what happens when some celebrity followers (whom we denote F_1) retweet a message to their followers, whom we denote F_2 . If the celebrity authored the original message (Panel A), the followers-of-followers (F_2 s) observe that the celebrity authored the message and that F_1 retweeted it, as seen in Panel B. But if the celebrity retweeted a message from a Joe/Jane (Panel C), the followers-of-followers of the celebrity (F_2 s) observe only that the Joe/Jane tweeted and that F_1 then retweeted for F_2 to see, as displayed in Panel D. The F_2 does not observe the celebrity involvement at all in the second case.

We therefore can study the retweet behavior of F_2 s – i.e., followers-of-followers-of-celebrities – across these two cases to identify the endorsement effect premium, i.e., the additional effect of their knowing that the celebrity originated the message, holding F_2 's network position fixed.¹¹

A challenge, of course, is that the F_1 decision to retweet may be endogenous. To address this concern, we had some additional study participants exogenously retweet certain celebrity tweets or retweets. We can therefore examine the behavior of the followers (F_2 's) of these exogenously 'Forced Joes/Janes' only, where we know there is no endogeneity issue of the F_1 decision to retweet, and compare this to the effect when we consider the full sample of F_2 's.

We can then further decompose the endorsement effect to understand the impact of celebrities speaking in their own voice (an 'authorship' effect). We use the same experimental variation, but look at behavior of the direct followers of celebrities (F_1 s), who see both the celebrities' directly-authored tweets and the celebrities' retweets.

To identify the impact of sources, we explicitly randomize whether a source was included in the tweet, and examine the behavior of the F_1 s.

Finally, to measure offline effects, celebrities were randomized into two phases, with half tweeting in the first phase (Phase I, July and August 2015) and half in the second phase (November 2015 - February 2016, Phases II and III). In addition, towards the end of the last phase, all tweets / retweets by a celebrity were then retweeted by a randomly selected number of Joes/Janes (Phase III). We conducted a survey between phases (August - October 2015) of a subset of followers of our celebrities and we use this between-celebrity randomization to estimate the impact of the Twitter campaign on offline beliefs and behaviors. A timeline of the experiment is shown in Figure 2.

¹¹Note that on average, a typical F_2 follows only 1.18 F_1 s. Given this, combined with the fact that F_1 behavior is defined over the entire sample, means it is very likely a given F_2 only saw information from the study on the day in question from the F_1 we are interested in, not more generally.

2.3. Data.

2.3.1. *Online data.* We collected detailed data via the Twitter Firehose and API. Before the experiment began, in early 2015, we collected a baseline image of the publicly available Twitter network, including the list of followers of any celebrity participating in our study.

On Twitter, followers can take two primary actions: “likes” and “retweets”. While likes are public, likes are not automatically pushed out as tweets to a user’s followers.

For each the of the 672 total tweets originated by our experiment, we tracked each time the tweet was liked or retweeted by any of the over 7.8 million unique users who followed at least one of the participants in our study for the two weeks following the tweet. When the tweet was retweeted by a celebrity’s follower, we also scraped all of this follower’s followers and their liking and retweeting behavior. We denote those retweets / likes coming from a direct follower of a celebrity as F_1 events, and those retweets / likes coming from a follower of a follower of a celebrity as F_2 events.¹² Table 1 reports descriptive statistics.

2.3.2. *Offline data.* To measure whether online conversations led to offline behavioral changes, we conducted a phone survey on a sample of 2,441 subjects, all of whom followed at least one of our study participants on Twitter. These subjects were recruited primarily via ads placed on the Twitter platform. The phone survey was designed to capture information about immunization (including beliefs in some of the various myths our tweets were intended to counteract), immunization history for children in the family and knowledge of recent immunizations of children of close friends, and questions about immunization and Twitter.

To recruit this sample, we advertised with promoted tweets on Twitter a recruitment to participate in a healthcare survey, targeted to the 7.8 million unique users who followed participants in our study. This process resulted in 2,441 total subjects. All respondents were surveyed by phone during the endline period; we also contacted a subsample of these respondents (approximately 73 percent) by phone for a baseline survey prior to the beginning of Phase I tweets.

Table 1, Panel B reports demographics of our offline survey sample. To gauge the sample selection in our sample, we also present comparable data from the 2014 wave of the SUSENAS, the annual representative Indonesian national household survey. Relative to the nationally representative SUSENAS sample, we see that our demographic is more urban, slightly younger, and has a similar gender composition.

¹²When anyone retweets/likes a tweet in our study, we check at that time all the people that person follows, in order to determine if that person is an F_2 or not, and if so, the shortest path through which they observed the tweet.

Panel C reports baseline statistics for beliefs about vaccinations. We see that there is considerable confusion about the nature and value of vaccines. For instance, only 56 percent of individuals thought that vaccines are domestically made (they are), and only 38.5 percent thought that vaccines are free of cost (they are). This suggests substantial room for improvement on immunization knowledge in our study sample.

2.4. Estimation. We estimate three models. First, to estimate the overall effect of endorsement, we focus on the behavior of followers-of-followers-of-celebrities – i.e., F_2 s – and estimate by Poisson regression, the equation

$$(2.1) \quad E[y_{trcmp} | \mathbf{x}_{trcmp}] = \exp(\alpha \cdot \text{Celeb}_{tcm} + \beta \cdot \log(\text{Followers})_r + \omega_c + \omega_m)$$

where t indexes tweets, r indexes retweeters (i.e., an F_1 who retweeted the tweet t), c indexes celebrities, m indexes the type of message content, and p indexes phase. The variable Celeb_{tcm} is a dummy that takes 1 if the celebrity authored the tweet herself (and hence her identity is visible to F_2), and 0 otherwise (and hence her identity is not visible to F_2). Each observation is a retweet of one of our original tweets, and the dependent variable y_{trcmp} is a count of how many times this retweet was itself either liked or retweeted again by an F_2 . Since y is a count, we estimate a Poisson regression, with robust standard errors clustered at the original tweet (t) level. We control for the log number of followers of F_1 , and for dummies (ω_m) for different types of messages (e.g., dummies for it being about a fact, importance of immunization, etc.). All regressions include celebrity fixed effects (ω_c), which absorb variation in casual/formal style, etc. The coefficient of interest is α , which measures the differential impact of the tweet having been written by the celebrity (as compared to being written by a Joe/Jane) and this being observable to the F_2 deciding whether to retweet. We estimate equation (2.1) for the full sample of F_2 s and for the subset of F_2 's who follow the 'Forced Joe/Janes' (i.e. followers of study participants who exogenously retweeted celebrity tweets and retweets.)

Second, to decompose the celebrity effect further, we restrict attention to direct followers of the celebrity (F_1 individuals), and estimate

$$(2.2) \quad E[y_{tcmp} | \mathbf{x}_{tcmp}] = \exp(\alpha \cdot \text{Celeb}_{tcm} + \omega_c + \omega_m).$$

We now have one observation per tweet, and look at the number of retweets/likes, retweets, or likes by F_1 s who are distance 1 from the celebrity. We continue to include celebrity (ω_c), and message-type (ω_m) fixed effects. We run an analogous regression replacing $Celeb$ with $Source$ to study the impact on F_1 behavior of including public health authority sources.

Finally, to examine offline effects, we turn to our phone survey data. We define $\text{Exposure to Tweets}_i$ as the number of campaign tweets that i is randomized to see through Phase I (normalized to have standard deviation 1). $\text{Potential Exposure}_i$ is the total number of campaign tweets that i could potentially see through the campaign given the celebrities

she follows at baseline (i.e., had all the celebrities she followed been randomized to tweet in Phase I). The experimental design of randomizing celebrities into phases means that, while individuals i may differ in the number of celebrities they follow, Exposure to Tweets $_i$ is random conditional on Potential Exposure $_i$.

We therefore run logistic regressions of the form

$$(2.3) \quad f(y_i) = \alpha + \beta \cdot \text{Exposure to Tweets}_i + \gamma \cdot \text{Potential Exposure}_i + \delta' X_i,$$

where y_i is the outcome for respondent i and $f(\cdot)$ is log-odds, i.e., $\log\left(\frac{P(y_i=1|\mathbf{x}_i)}{1-P(y_i=1|\mathbf{x}_i)}\right)$. X are controls, such as the number of celebrities followed by i , the log of the number of followers of celebrities by i , survey dates, and (in some specifications) demographics and baseline beliefs, selected here and in subsequent regressions by double post-LASSO (Belloni, Chernozhukov, and Hansen, 2014a,b). We report standard errors and p -values clustered at the level of the combination of celebrities followed; further, because of the complex nature of the potential correlation in Exposure to Tweets $_i$ across individuals i induced by partial overlap in which celebrities our survey respondents follow, we present randomization-inference (RI) p -values as well.

3. RESULTS

3.1. Decomposing The Endorsement Effect: Authorship versus Relaying Messages, and Citing Public Health Authorities.

3.1.1. *Measuring the Total Endorsement Effect.* We begin by measuring the size of the endorsement effect, and then decompose it into the value that comes from direct authorship of a message as opposed to a more passive involvement, and examine the value of citing public health officials.

To do so, we begin by examining the behavior of F_2 s, i.e., followers-of-followers. Recall from the discussion above and the illustration in Figure 1 that if the celebrity retweets a message by a Joe/Jane, and this is retweeted by F_1 , F_2 sees the message, sees that it is composed by a Joe/Jane, and knows that F_1 retweeted it. But crucially F_2 does not know that the celebrity had retweeted it: F_2 is likely to be blind to the celebrity's involvement. On the other hand, if the celebrity had written this tweet herself, this would be visible to F_2 .

We analyze this by estimating equation (2.1). Figure 3 presents the results, and Table A.1 in Appendix A reports estimates in table form. We have three main outcome variables: in blue whether an agent either liked or retweeted the tweet, in brown whether an agent retweeted the tweet, and in red whether an agent liked the tweet. The left-hand side graph reports the results on the full sample for each of these dependent variables.

We see large endorsement effects when we look at the entire sample of F_2 's. Having a celebrity compose and tweet the message relative to having a Joe/Jane compose the message and the celebrity retweet it leads to a 72 percent (0.54 log point) increase in the retweet or like rate ($p = 0.001$; note that since this is a Poisson model, the coefficients are interpretable as the change in log number of retweets/likes) by followers-of-followers (F_2 s). The results are qualitatively similar for retweets and likes alone—68 percent (0.52 log points) for retweets and 92 percent (0.66 log points) for likes. We document similar effects of an organization being the originator rather than a Joe/Jane in Table F.1 of Appendix F.¹³

These results imply that, holding the content of the tweet constant (since it is randomized across tweets) and holding the F_2 position in the network constant (since they are all followers-of-followers of the celebrity), having the F_2 be aware of the celebrity's involvement in passing on the message substantially increases the likelihood that the F_2 responds online.

One potential challenge with this regression, however, is that the sample here may be endogenous. That is, when we look at F_2 agents, i.e., those who are at distance two from the celebrity, whether a given agent sees a retweet from his or her F_1 s may be endogenous and respond to our treatment, i.e., which F_1 s choose to retweet the message may be directly affected by the fact that the celebrity composed the message. In equation (2.1), we control for the log number of followers of the F_1 who retweeted the message, and hence the number of F_2 s who could potentially retweet it, so there is no mechanical reason for a bias in equation (2.1). But there may nevertheless be a *compositional* difference in which F_1 s retweet it, which could potentially lead to selection bias of which F_2 s are more likely to see the retweet.

To address this issue, in the last phase of the experiment, we added an additional randomization. We use the subset of Joes/Janes who are also F_1 s, and so direct followers of our celebrities. For some of these Joes/Janes, we randomly had their accounts retweet our celebrities' tweets and retweets in the experiment; that is, we created exogenous F_1 s. For this sample, we can look at how *their* followers – that is, the followers of F_1 Joes/Janes we exogenously forced to retweet a particular tweet – responded as we randomly vary whether the celebrity, an organization, or a Joe/Jane composes the message. We analyze this experiment by estimating equation (2.1) just as we did for the full sample, but here we know that whether an F_2 sees the tweet is exogenous by construction. This is a much smaller sample of F_2 s – it's followers of study participants who were randomly chosen to retweet a message – but has the virtue of being clearly exogenously determined.¹⁴

The right-hand side graph of Figure 3 presents the results for this sample. The point estimates are if anything somewhat larger than in the full sample, and we cannot reject

¹³Recall we only have 9 organizations, which reduces the overall instances of such cases, so we relegate this to an appendix. Also, we condition on non-sensitive tweets for this sample.

¹⁴Another advantage of this exercise is as follows: in the full sample, we also have more observations – i.e. more F_1 retweets – precisely when the celebrity effect may be higher. This is not an issue in the forced Joe/Jane subsample.

equality. Statistical significance is reduced somewhat (p -values of 0.119, 0.111, and 0.107 for likes and retweets, retweets, and likes respectively), but the fact that results are broadly similar to the overall effects of the full sample suggests that the possible endogenous selection of F_1 s in our full sample is not leading to substantial bias.

A second potential confound comes from the fact that a retweet shows how many times the original tweet has been retweeted or liked when the user views it. Since our treatment assignment affects the retweet count, this itself could spur further changes in the likelihood of retweeting. The same randomization of forced Joe/Jane retweets also helps address this issue, because we randomly varied the number of Joes/Janes we forced to retweet a particular tweet. We find that being randomly assigned one, five, ten, or even fifteen extra retweets makes no impact on the number of F_1 or F_2 retweets that the given tweet faces (see Appendix E, Table E.1).

3.1.2. Decomposing Endorsement: Authorship vs. Sending Messages. In the preceding analysis, we isolated whether the F_2 knew the celebrity was involved in the message thread at all. But it groups together celebrity *authorship* with the celebrity *sending* the message, since the F_2 either saw both – a celebrity authored and sent message – or no celebrity involvement altogether.

To decompose this, we can examine the direct followers of the celebrities (i.e., F_1 s). For F_1 s, they know the celebrity is involved either way, but the randomization changes whether the message was directly authored by the celebrity or retweeted.

We estimate equation (2.2), present results in Figure 4, and report estimates in Table A.2 of Appendix A.

We find that authorship matters: tweets authored by celebrities are 200 percent (1.10 log points) more likely to be retweeted/liked than those where the celebrity retweets (in blue, $p < 0.001$), with large effects on both retweets (278 percent increase, 1.329 log points, $p < 0.001$) and likes (123 percent, 0.803 log points, $p < 0.001$).

This fact allows us to decompose the impacts of celebrities. The estimates in Figure 4 imply that 66 percent of the total celebrity effect comes from authorship per se. Combining these estimates with those in the previous section suggests that, on net, 43 percent of the celebrity effect comes from authorship, 24 percent from endorsement, with the remainder attributable to the intrinsic interest of the message.¹⁵

¹⁵To see this, consider the estimates from columns 1 of Tables A.1. and A.2., which are 0.544 and 1.101, respectively. Denote by x the number of retweets/likes a message with no endorsement or authorship gets; this is the inherent interest in the message. A message authored by a celebrity (which combines authorship and endorsement) gets $e^{1.101}x$ retweets/likes. A message endorsed by a celebrity (without authorship) gets $e^{0.518}x$ retweets/likes. Thus, of the total $e^{1.101}x$ retweets, $\frac{x}{e^{1.101}x} = 33\%$ is due to interest in the message, $\frac{e^{0.544}x - x}{e^{1.101}x} = 24\%$ is due to endorsement, and the remainder – 43% is due to authorship. This also implies that $\frac{43\%}{43\% + 24\%} = 66\%$ of the total celebrity effect is due to authorship. For retweets alone, the analogous numbers are 26% due to intrinsic interest, 18% due to endorsement, and 56% due to authorship.

3.1.3. *Citing Public Health Authorities.* Finally, we examine the impact of citing sources. Every tweet in our databank was paired with a source, but we randomized whether this source was included in each tweet. The sources come in several forms. In some cases, the tweet refers to the website or Twitter handle of a trusted authority who has issued that statement. For example, one tweet says “Polio vaccine should be given 4 times at months 1, 2, 3, 4. Are your baby’s polio vaccines complete? @puskomdepkes” where @puskomdepkes is a link to the Twitter handle of the Ministry of Health (known as *DepKes* in Indonesian). In other cases, explicit sources are cited, with a Google shortened link provided.¹⁶

We re-estimate equation (2.2) at the F_1 level, adding a variable for whether the tweet was randomized to include a source.¹⁷ Figure 5 below presents the results, and estimates are reported in Table A.3 of Appendix A.

On average, we find that citing a public health authority reduces the retweet and liking rate by 26.3 percent (-0.306 log points; $p = 0.051$, in blue on Figure 5), with similar effects for likes and retweets. We find that both the retweeting rate and liking rate experience broadly similar declines (retweeting declines by 0.318 log points, $p = 0.048$, in brown; liking declines by 0.277 log points, $p = 0.13$, in red, not significantly different from zero). This result shows that when a message is relayed by citing a public health authority, willingness to pass it on downstream, and perhaps liking the message, declines.

This result may seem surprising, since one might expect that a sourced message may be more reliable. There are, however, several possible explanations for this finding. One possibility is that for an F_1 , passing on a message has both instrumental value (delivering a good message), as well as a signaling value (conveying to followers that the F_1 is able to discern which information is good). In such a model, the fact that the message contains a source reduces the signaling value – because now anyone can figure out that the message was high quality given that it has a source, an F_1 no longer is no longer able to signal her ability to discern quality by passing it on. The effect of sources is therefore theoretically ambiguous in such a model, and depends on whether the signaling or instrumental effects dominate.¹⁸ We illustrate this in Appendix B, building on prior work by the authors (Chandrasekhar et al., 2018). While this is one possible explanation, other mechanisms are certainly possible as well.

¹⁶Note that Twitter automatically produces a short preview of the content if the site linked to has Twitter cards set up. There is one non-Google shortened link used when citing IDAI (Ikatan Dokter Anak Indonesia, the Indonesian Pediatric Society).

¹⁷Note that the number of observations is smaller, because some tweets on topics deemed ‘sensitive’ by the Government always included a source, as noted above. We restrict analysis to tweets for which we randomized whether the source was included.

¹⁸This also suggests a possible test, which is that F_1 ’s whose ability to discern is already established may value the instrumental value more than the signaling value; we regard this as an interesting dimension for future research.

3.2. Does Online Discussion Have Offline Effects?

3.2.1. *Did people hear about the campaign?* We next examine whether an online celebrity endorsement campaign can have measurable offline effects. To investigate this, we estimate equation (2.3). We use the fact that our offline survey was conducted between Phases I and II, so that conditional on the number of our celebrities a user followed, exposure to our campaign as of the time of the phone survey was randomly assigned.

We begin with what can be thought of as akin to a first-stage and which is presented in Figure 6 below (table-form results are in Table A.4. We ask whether respondents were more likely to have heard of our hashtag (#AyoImunisasi) (in blue) or heard about immunization discussions from Twitter if they were randomly more exposed to campaign tweets, conditional on their potential exposure (in brown).

We find that a one-standard deviation increase in exposure to the campaign (15 tweets) corresponds to a 21.2 percent increase in the probability that the respondent had heard of our hashtag relative to a mean of 7.75 percent (0.219 increase in log odds, clustered $p = 0.056$, RI $p = 0.075$).¹⁹ Further, a one-standard deviation increase in exposure corresponds to an 9.2 percent increase in the probability they heard about immunization in general from Twitter relative to a mean of 18.1 percent (0.108 increase in log odds, clustered $p = 0.110$, RI $p = 0.190$).

3.2.2. *Did people then increase their knowledge about immunization facts?* We then ask whether exposure to the campaign led to increased knowledge about immunization, particularly about common vaccine ‘myths’, which were the focus of the campaign.²⁰

Our survey asks questions about several categories of knowledge which were covered by the campaign. We ask whether people know that vaccines are domestically produced, to combat the common rumor in Indonesia that they contain pig products (which would make them unacceptable for Muslims, who represent the bulk of Indonesia’s population; domestic products are known to be halal). We ask whether they believe that natural alternatives (breastfeeding, herbal supplements, alternative supplements) replace the need for immunization. And, we examine knowledge that typical symptoms (mild fevers or swelling) are normal and not a cause for alarm. We also ask about “access” information; in particular, we test whether they know that immunizations are free at government health centers.

Figure 7 below presents the results for each of these categories of information (tabular estimates are shown in Table A.5). We find knowledge effects about domestic production (in blue)—though not on rumors about substitutability, side-effects, nor free access (in brown, red, and green respectively). Seeing 15 campaign tweets in general corresponded to an

¹⁹Note that we report impacts on log-odds in the figures and tables; we translate these into percent increases in the text.

²⁰Myth-dispelling facts comprised 36.7 percent of tweets, and 82.4 percent of all fact-related tweets sent out. Table D.1 in Appendix D shows that tweets concerning myths also diffused more widely than other facts.

increase of 4.3 percent in the probability of correctly answering the domestic question on a base of 57.5 percent (0.101 increase in log odds, clustered $p = 0.122$, RI $p = 0.082$).

3.2.3. *Communication about immunization.* We next ask whether individuals were more likely to know about their neighbors, friends, and relatives' immunization behavior, which would be a byproduct of offline conversations. We ask about knowledge of immunizations since June 2015 to capture the period since the start of the campaign. Immunizations in Indonesia take place at monthly *posyandu* meetings, which occur each month in each neighborhood (usually hamlets, or *dusun*, in rural areas, and neighborhoods known as *rukun warga*, or *RW*, in urban areas; see [Olken, Onishi, and Wong \(2014\)](#)), so if knowledge would increase, one might expect it to be the knowledge about immunization practices of ones' neighbors.²¹

The left-hand side graph of Figure 8 presents the results, which are reported in Panel A of Table A.6 in Appendix A. The blue estimates shows that being exposed to 15 more tweets corresponds to a 4.9 percent increase in the probability of knowing the neighbor's status (0.231 increase in log odds, clustered $p = 0.004$, RI $p = 0.088$) relative to a mean of 0.775. We do not consistently see significant effects on non-neighbor friends. With relatives the point estimates are comparably large, though the estimates are noisier.

3.2.4. *Did exposure lead to changes in reported immunizations?* Finally, we ask whether individuals have more knowledge of actual immunizations among friends, neighbors, and relatives. That is, does the campaign appear to change immunization behavior as reported by our survey respondents?

The right-hand side graph of Figure 8 corresponding to Panel B of Table A.6 presents the results, where the dependent variable is whether the respondent knows of anyone among their friends, neighbors, or family who immunized a child since June 2015.

We condition the sample when we look at knowledge of immunization among neighbors, friends, and family to those who knew whether the vaccination status of the children of members of these individuals, so this effect is above and beyond the effects reported above.²²

The estimate represented in blue shows that when looking at neighbors, an increased exposure by 15 tweets corresponds to a 12.8 percent increase in the number of reported vaccinations (0.194 increase in log odds, clustered $p = 0.071$, RI $p = 0.133$) relative to a mean of 0.356. For friends, in brown, we find an increased exposure of 15 tweets corresponds to a 16.4 percent increase in the number of reported vaccinations (0.246 increase in log odds,

²¹Note that here the sample is restricted to respondents who know friends, relatives, or neighbors with at least one child (ages 0-5) respectively as this is the relevant set.

²²An important caveat is that knowledge of neighbors', friends', and family members' children's vaccination status itself is affected by treatment (as discussed above), so these results should be interpreted with some caution. Interestingly, the fact that we see impacts on reported vaccinations of friends, but knowledge of friends' behavior is unchanged, suggests that perhaps this is less of an issue. More generally, however, this could reflect some combination of changes in immunization and changes in discussions about it.

clustered $p = 0.010$, RI $p = 0.071$) relative to a mean of 0.353. The red dot presents results looking at relatives. An increased exposure by 15 tweets corresponds to a 9.8 percent increase in the number of reported vaccinations among relatives (0.140 increase in log odds, clustered $p = 0.159$, RI $p = 0.06$) relative to a mean of 0.314. Finally, the green dot looks at own behavior, and our estimate is not statistically different from zero. On net, the results in this section are suggestive that an online Twitter campaign can have offline effects on knowledge and, potentially, behavior.

4. CONCLUSION

We study how to design a public health campaign to deploy celebrities effectively by conducting a large-scale, online field experiment. Using our design, we are able to decompose to what extent celebrity authorship *per se* as opposed to passing on messages of others matters, and to show whether citing public health officials amplifies or reduces these effects.

We find the celebrity endorsement matters. Moreover, the vast majority (66-76%) of the increased effect of a celebrity writing a message come from authorship *per se*, as opposed to a celebrity passing on messages from others.

We also find that explicitly referring to public health sources has an adverse effect. This result may seem surprising, since one might expect that a sourced message may be more reliable, but we discuss how this result is possible if including the source reduces the value to the F_1 of retweeting the message as a signal of her ability to discern good from bad messages. At a broader level, these results are consistent with the results on celebrity authorship: messages are most likely to be passed on when they come from the celebrity speaking directly, rather than passing on messages from others.

We also find offline effects: messages are heard, certain myths (about imported vaccines, in particular) become less prevalent, offline communication about vaccination increases, and we see more reported immunization behavior among respondents' neighbors, friends, and relatives.

These findings – particularly the idea that celebrity messaging works best when celebrities speak directly, themselves, rather than merely pass on recommendations from others – are applicable in a wide variety of public health settings, including perhaps in designing information campaigns to encourage ongoing vaccination for COVID-19 or a future pandemic.

REFERENCES

- AGRAWAL, J. AND W. A. KAMAKURA (1995): “The economic worth of celebrity endorsers: An event study analysis,” *Journal of Marketing*, 59, 56–62. 1
- BAKSHY, E., J. M. HOFMAN, W. A. MASON, AND D. J. WATTS (2011): “Everyone’s an influencer: quantifying influence on Twitter,” in *Proceedings of the fourth ACM international conference on Web search and data mining*, ACM, 65–74. 1
- BANERJEE, A., A. G. CHANDRASEKHAR, E. DUFLO, AND M. O. JACKSON (2013): “The diffusion of microfinance,” *Science*, 341, 1236498. 1
- BANERJEE, A. V., E. BREZA, A. G. CHANDRASEKHAR, AND B. GOLUB (2018): “When Less is More: Experimental Evidence on Information Delivery During India’s Demonetization,” Working Paper 24679, National Bureau of Economic Research. B.1
- BEAMAN, L., A. BENYISHAY, J. MAGRUDER, AND A. M. MOBARAK (2021): “Can network theory-based targeting increase technology adoption?” *American Economic Review*, 111, 1918–43. 1
- BEAMAN, L. AND A. DILLON (2018): “Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali,” *Journal of Development Economics*, 133, 147–161. 1
- BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014a): “High-dimensional methods and inference on structural and treatment effects,” *The Journal of Economic Perspectives*, 29–50. 2.4
- (2014b): “Inference on treatment effects after selection among high-dimensional controls,” *The Review of Economic Studies*, 81, 608–650. 2.4
- BERGKVIST, L. AND K. Q. ZHOU (2016): “Celebrity endorsements: a literature review and research agenda,” *International Journal of Advertising*, 35, 642–663. 1
- BURSZTYN, L., G. EGOROV, AND R. JENSEN (2017): “Cool to be smart or smart to be cool? Understanding peer pressure in education,” *The Review of Economic Studies*. B.1
- BURSZTYN, L. AND R. JENSEN (2015): “How does peer pressure affect educational investments?” *Quarterly Journal of Economics*, 130, 1329–1367. B.1
- CHANDRASEKHAR, A. G., B. GOLUB, AND H. YANG (2018): “Signaling, Shame, and Silence in Social Learning,” Tech. rep., National Bureau of Economic Research. 8, 3.1.3, B.1, B.3
- ELBERSE, A. AND J. VERLEUN (2012): “The economic value of celebrity endorsements,” *Journal of Advertising Research*, 52, 149–165. 1
- GARTHWAITE, C. L. (2014): “Demand spillovers, combative advertising, and celebrity endorsements,” *American Economic Journal: Applied Economics*, 6, 76–104. 1
- GONG, S., J. ZHANG, P. ZHAO, AND X. JIANG (2017): “Tweeting as a marketing tool: A field experiment in the TV industry,” *Journal of Marketing Research*, 54, 833–850. 1

- IVES, M. (2021): “Celebrities Are Endorsing Covid Vaccines. Does It Help?” *The New York Times*. 1
- KATONA, Z., P. P. ZUBCSEK, AND M. SARVARY (2011): “Network Effects and Personal Influences: The Diffusion of an Online Social Network,” *Journal of Marketing Research*, 48:3, 425–443. 1
- LESKOVEC, J., L. A. ADAMIC, AND B. A. HUBERMAN (2007): “The dynamics of viral marketing,” *ACM Transactions on the Web (TWEB)*, 1, 5. 1
- OLKEN, B. A., J. ONISHI, AND S. WONG (2014): “Should Aid Reward Performance? Evidence from a field experiment on health and education in Indonesia,” *American Economic Journal: Applied Economics*, 6, 1–34. 1, 3.2.3

FIGURES

FIGURE 1. Identification of the value of endorsement of celebrity involvement.

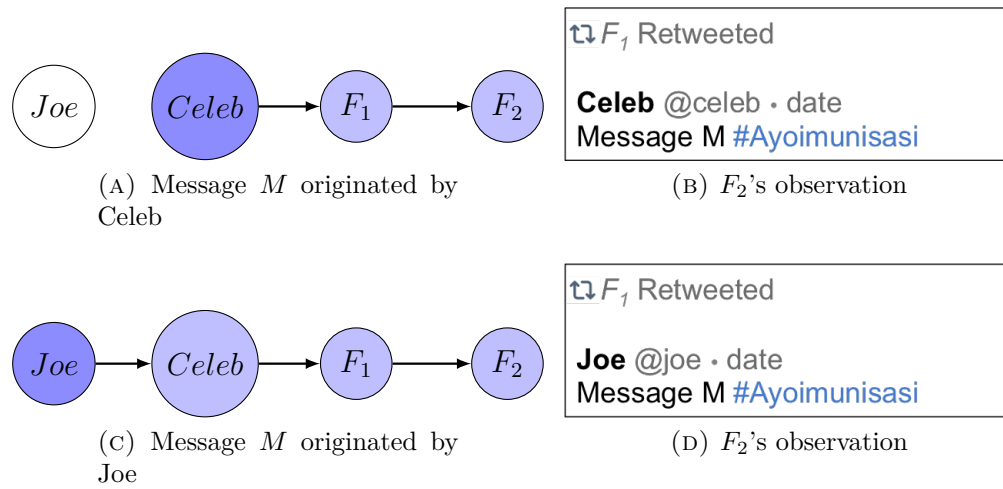


FIGURE 2. Timeline

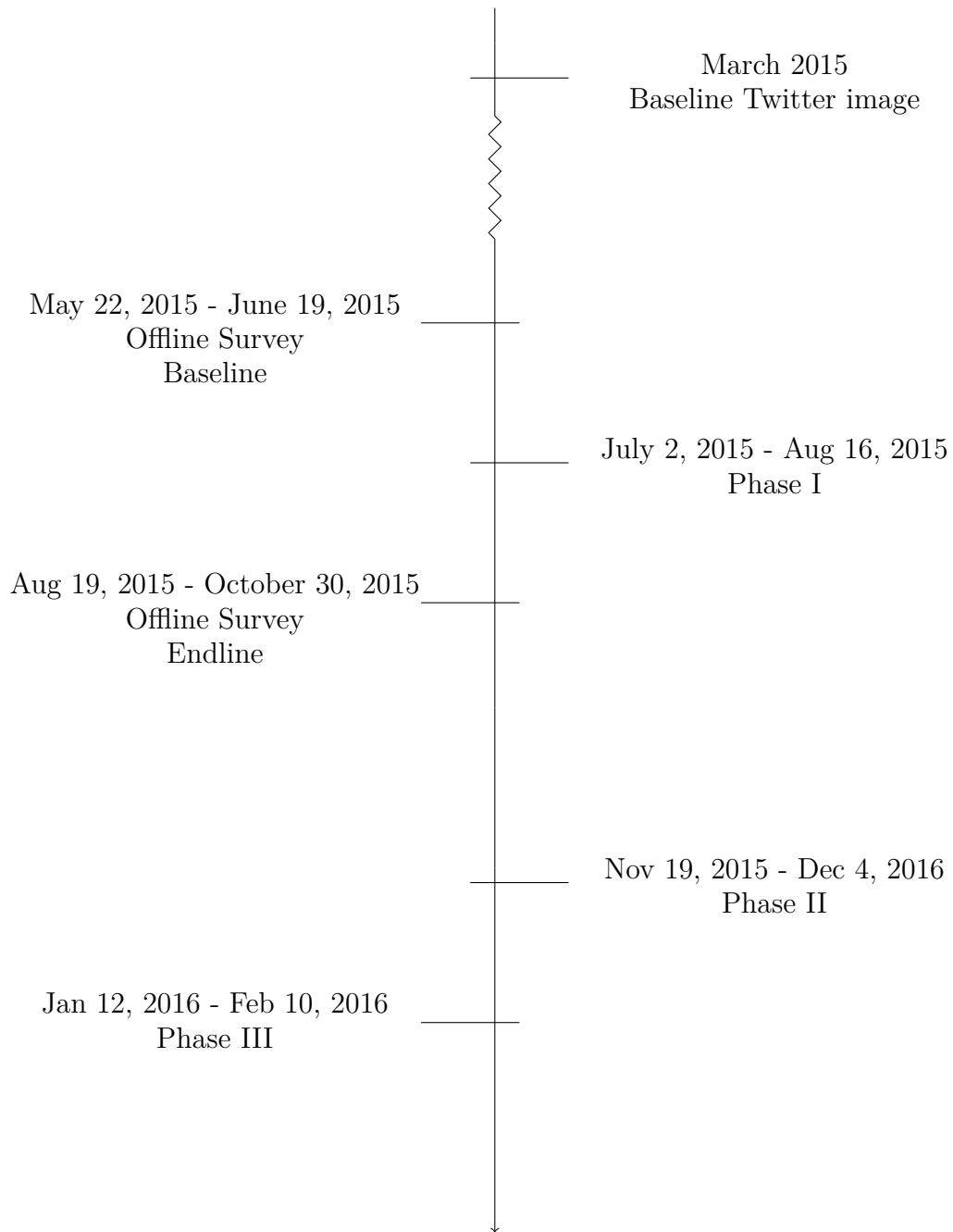
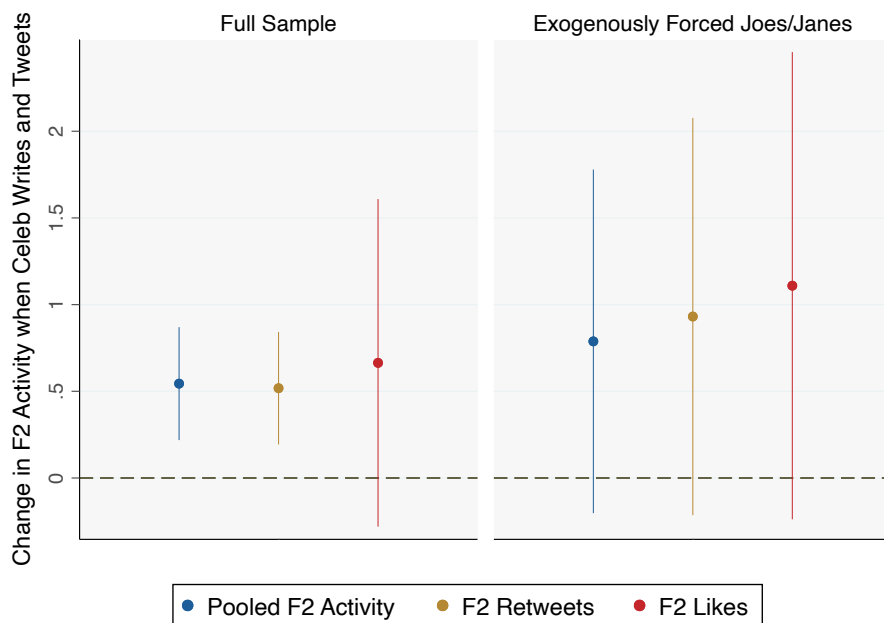
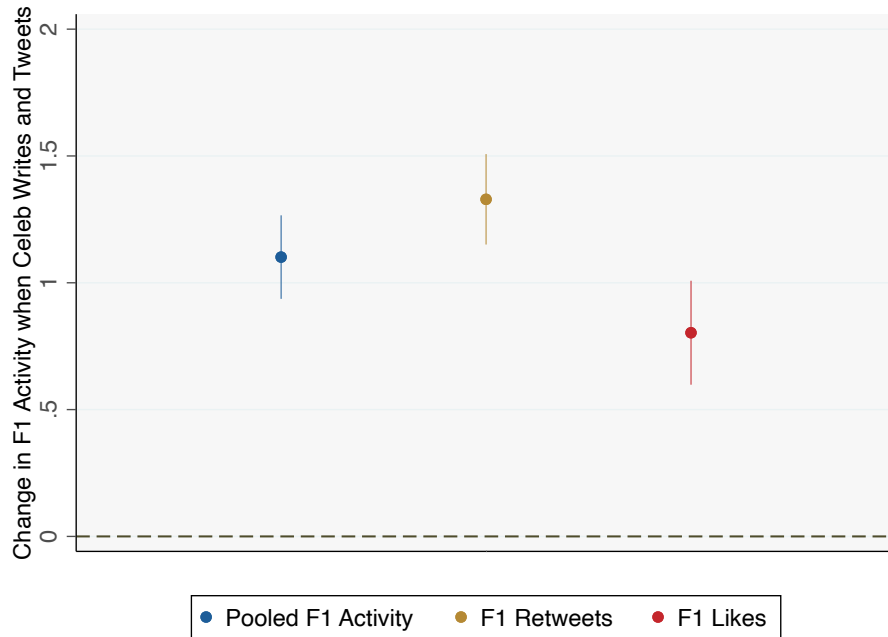


FIGURE 3. Reach vs Endorsement: Value of Celeb Endorsement through Involvement measured by F2 likes/retweets



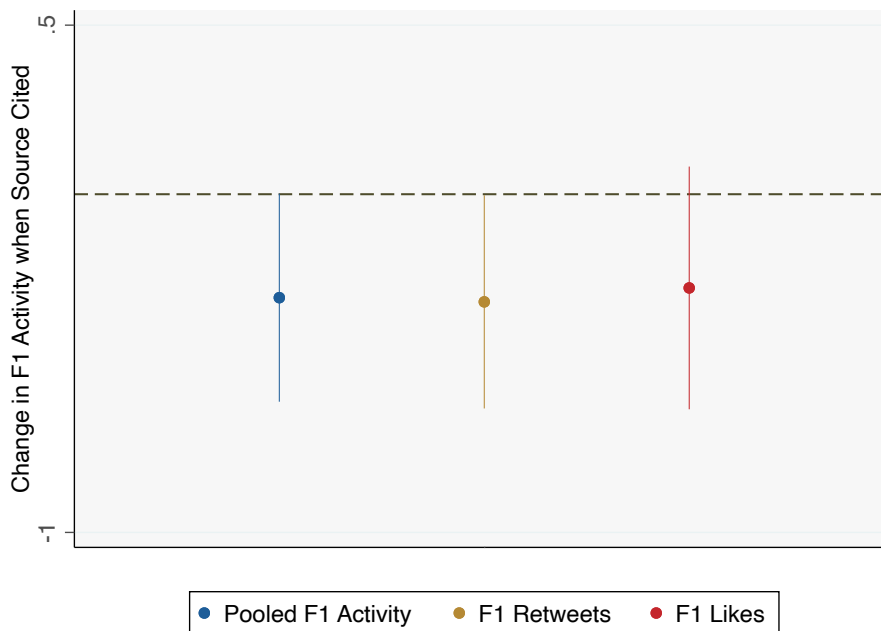
Note: Figure depicts regression estimates with their 95% confidence intervals. Equation model estimated is (2.1). Sample conditions on all tweets originated by Joes/Janes or celebrities. All regressions control for phase, celebrity fixed effects, content fixed effects, and the log number of followers of the F1. These coefficients are also reported in Table A.1.

FIGURE 4. Value of Celeb Endorsement through Composition measured by F1 likes/retweets



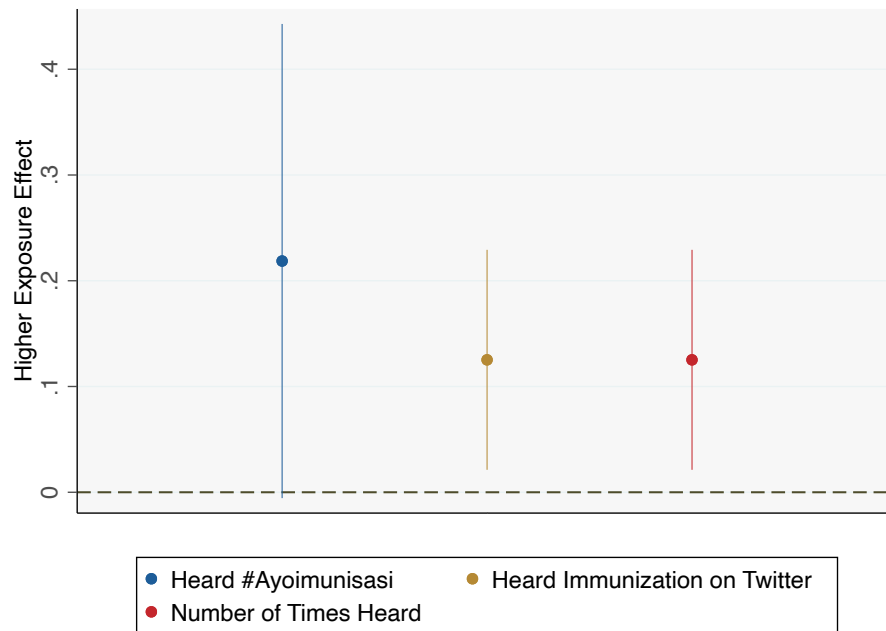
Note: Figure depicts regression estimates with their 95% confidence intervals. Equation model estimated is (2.2). Sample conditions on all tweets originated by Joes/Janes or celebrities. All regressions control for phase, celebrity fixed effects, content fixed effects. Standard errors are clustered at the celebrity/organization level. These coefficients are also reported in Table A.2.

FIGURE 5. Value of Endorsement through Source Citation measured by F1 likes/retweets



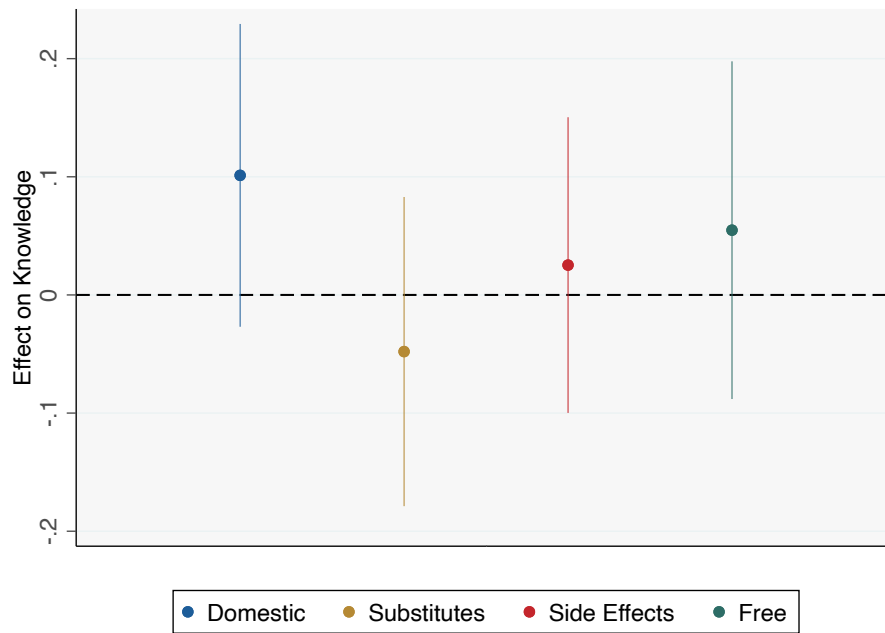
Note: Figure depicts regression estimates with their 95% confidence intervals. Equation model estimated is (2.2). Sample conditions on non-sensitive tweets. All regressions control for phase, celebrity fixed effects, content fixed effects. Standard errors are clustered at the celebrity/organization level. These coefficients are also reported in Table A.3.

FIGURE 6. Did people offline hear about the campaign?



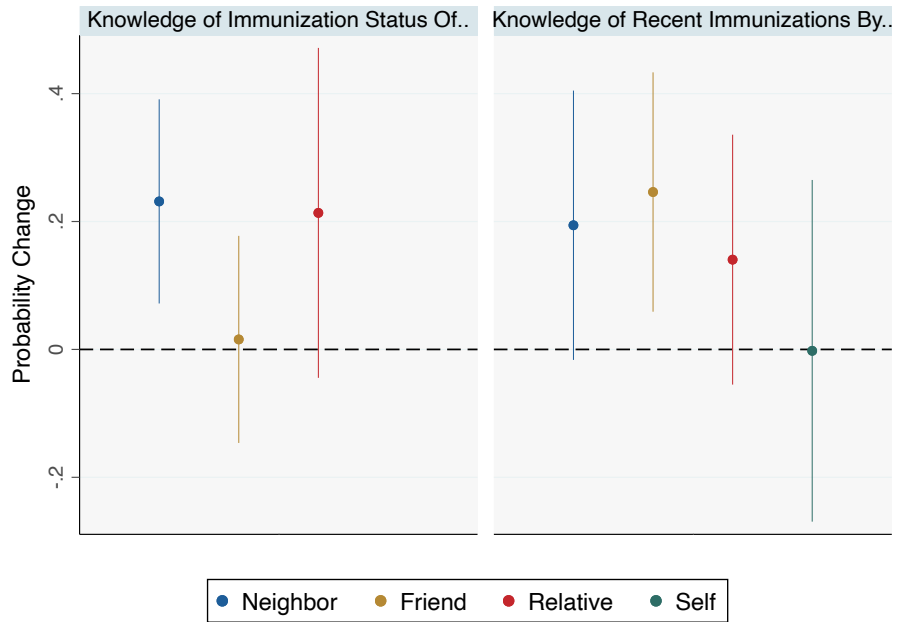
Note: Figure depicts regression estimates with their 95% confidence intervals. Equation model estimated is (2.3). Standard errors are clustered at the combination of celebs followed level. Demographic controls include age, sex, province, dummy for urban area and dummy for having children. One standard deviation of exposure is 14.96 tweets. These coefficients are also reported in Table A.4.

FIGURE 7. Did people offline increase knowledge?



Note: Figure depicts regression estimates with their 95% confidence intervals. Equation model estimated is (2.3). Standard errors are clustered at the combination of celebs followed level. Demographic controls include age, sex, province, dummy for urban area and dummy for having children. One standard deviation of exposure is 14.96 tweets. These coefficients are also reported in Table A.5.

FIGURE 8. Networks and behavior



Note: Figure depicts regression estimates with their 95% confidence intervals. Equation model estimated is (2.3). In both panels (left and right-hand side graphs), standard errors are clustered at the combination of celebs followed level. Demographic controls include age, sex, province, dummy for urban area and dummy for having children. One standard deviation of exposure is 14.96 tweets. In left-hand side Panel, the sample is restricted to respondents who know friends/relatives/neighbors with at least one child (ages 0-5) respectively. In right-hand side Panel, the sample when looking at network members' behaviors (estimates in blue, brown, and red) is restricted to respondents who know the behavior of their network. When looking at own behavior (represented in green), sample is restricted to respondents with children younger than age 2. These coefficients are also reported in Table A.6.

TABLE 1. User summary stats

Panel A: Online user summary

	mean	obs
Followers of celebrities	262648	37
Followers of organizations	145300	9
Followers of Joes/Janes	574	134
Followers of forced Joes/Janes	502	898
Followers of celeb followers	1379	1073

Panel B: Offline user summary

	Sample mean	National Average (SUSENAS) mean
Age	28.416	29.959
Female	0.504	0.499
City (<i>kota</i>)	0.610	0.200

Panel C: Baseline beliefs

	mean	obs
Immunization is important	0.988	886
Immunization is safe	0.944	886
Immunization is beneficial	0.983	886
Breastfeeding can't replace immunization	0.650	886
Supplements can't replace immunization	0.872	886
Herbal supplements can't replace immunization	0.832	886
BCG is a basic vaccine	0.452	622
Hepatitis B is a basic vaccine	0.291	622
DPT is a basic vaccine	0.399	622
HIB is a basic vaccine	0.084	622
Polio is a basic vaccine	0.712	622
Measles is a basic vaccine	0.611	622
Immunization does not cause swelling	0.561	886
Immunization does not cause fever	0.647	886
Vaccines are domestically made	0.561	886
Vaccines are free of cost	0.385	886