Social media advertiser data sources

Author: Justin Katz
Date: May 2022

0 Overview

This document summarizes data on advertising volumes across online publishers and advertising costs at Facebook and Instagram.

- The first section describes Pathmatics data for ad volumes
- The second and third sections describe RevealBot data for ad prices and a few other preliminary sources
- The fourth section summarizes what I learned from preliminary analysis of the RevealBot pricing data

While it looks long, most of the pages are miscellaneous graphs from the RevealBot analysis.

1 Advertising volumes – Pathmatics

Information from website

Overview:

- Pathmatics collects information on the digital ad market by using established panels and web crawlers.
- The company tracks video and display ads on sites across the web, with special emphasis on Facebook, Instagram, Twitter, and Youtube.

Data fields: for each ad, Pathmatics tracks (among other things):

- The “ad path,” which describes the intermediary platforms used by advertisers and publishers to serve a particular ad
- Advertiser metadata – parent company, brand, category
- Ad metadata – size of ad, text, type, date, device, URL landing page, geography where ad was displayed
- Intermediation information – purchase channel, the provider from whom advertiser purchased impressions (e.g. The Trade Desk) and sales channel, the provider from which a publisher sold impressions (e.g. Rubicon). Distinction between direct buy vs programmatic buys
- Estimated number of impressions
- Estimated dollars spent on individual / group of ads
    - Note of caution: one figure says: “Pathmatics spend estimates are based on our proprietary panel and spend methodology. **They may not accurately estimate**

the spend for any advertiser, and should be used for directional trend analysis."

- Site where the ad was observed (wsj.com, facebook.com, etc)

Standard aggregation tools:

- Individual brands (e.g. Hershey) –
    - Specific ad "creatives" (images, text of ad)
    - Targeted audience by metro and state, estimated spending, estimated impressions
    - Distribution of spending across platforms (e.g. type of ad x site where ad is displayed)
    - Time series of spending on each site
- Activity across industries – e.g. the above aggregated to CPG

Time period: some fields go back to 2012.

Methodology (pretty minimal information on website)

- Collection:
    - Gathering method: Ads from the web using crawlers and panels
    - Channels: Desktop / mobile, video / display, FB, IG, Twitter, Youtube
    - Geographies: US, CA, France, Germany, Italy, UK, Australia, NZ
    - Claims to be representative sample
- Processing:
    - ML to categorize advertisers and brands
    - Predict CPM (cost per mil)

Notes on contacting sales reps

- I've sent a few inquiries through their online form (one request week of 5/2, one week of 5/9) but have not gotten a response.
- Hunt sent an email to their CEO a while back
- If we continue to have trouble contacting directly, we can reach out to the author of this NYT story that uses Pathmatics data to get the contact information of their rep

Outstanding questions

- How representative is the sampling process?
- How are sites besides FB / Instagram / Twitter / Youtube selected?
- Is this still missing large digital ad competitors? Google, TikTok, Snap, Amazon, LinkedIn, etc?
- Which data fields are modeled vs. directly observed? For the fields that are modeled, how are the models disciplined? In particular, how are ad prices and ad spend estimated?

    - Some figures say: “Pathmatics spend estimates are based on our proprietary panel and spend methodology. They may not accurately estimate the spend for any advertiser, and should be used for directional trend analysis.” Are the data accurate for estimating market share and relative spend across platforms?
- Pricing?

2 Advertising prices – RevealBot

Overview of company and data collection

Background:

- RevealBot is a digital ad management platform that helps clients automate and optimize advertising spending across social media platforms
- On its website, the company says it helps manage advertising on Facebook, Instagram, Snapchat, TikTok, and Google

Free ad price data:

- Revealbot posts average weekly ad price data for Facebook and Instagram on its website, broken down by:
    - A variety of cost metrics – cost per impression, click, lead, engagement, install
    - Different ad campaign objectives (by Facebook campaign category)
- According to my correspondence with a sales rep: “Those are numbers from our userbase. We have 15,000 or so clients and automate around $100m of ad spend every month”
- The data go back to Dec 2019
- The website is easy to scrape and I’ve pulled the full available time series as of 5/6/2022

Correspondence:

- I also asked about collaborating / purchasing more granular cuts of the data (by geography, for example). On the afternoon of 5/6/2022, the rep said he would ask the CTO and let me know
- Generally my contact is very responsive

Preliminary data analysis: see section 4

3 Advertising prices – other sources

Data source list:

- Hootsuite
- Optily
- AdEspresso
- RevealBot
- Madgicx

- Adzooma
- TrustAds
- Qwaya

Details

Hootsuite
- Link to ad site
- Link to contact page
- Features: allows management of ad campaigns by prospective advertisers

Optily
- Link to ad site
- Link to contact page
- Tool that allows for integration of sales information and ad campaigns
- Likely has something that tracks ad spend

AdEspresso
- Link to ad site
- Link to contact page
- Run by Hootsuite – analytics engine to manage ad spend across multiple platforms

Revealbot
- Link to data on ad prices

4 Preliminary analysis of RevealBot data

Narrative description of figures

Description of figures:
- Data sources: average weekly ad prices on FB and IG from RevealBot from Dec 2019-May 2022
- Outcome measure: Cost per mille (CPM) – effectively cost per impression
- Categories: average across all ad campaign types, average within specific campaign types. See glossary below for detailed description
- Vertical lines: indicate relevant time periods for the instruments for residual demand under consideration:
    - July 2020 ad boycott
    - Oct-November 2020 election spending. The line corresponds to peak election spending as indicated by Google's political ad spending tracker (which is publically available).

    - Apple release of iOS 14.5, which added App Tracking Transparency. The idea is that the update affected IG advertisers more, since I hypothesize that a greater fraction of use is on Apple devices compared to Facebook.

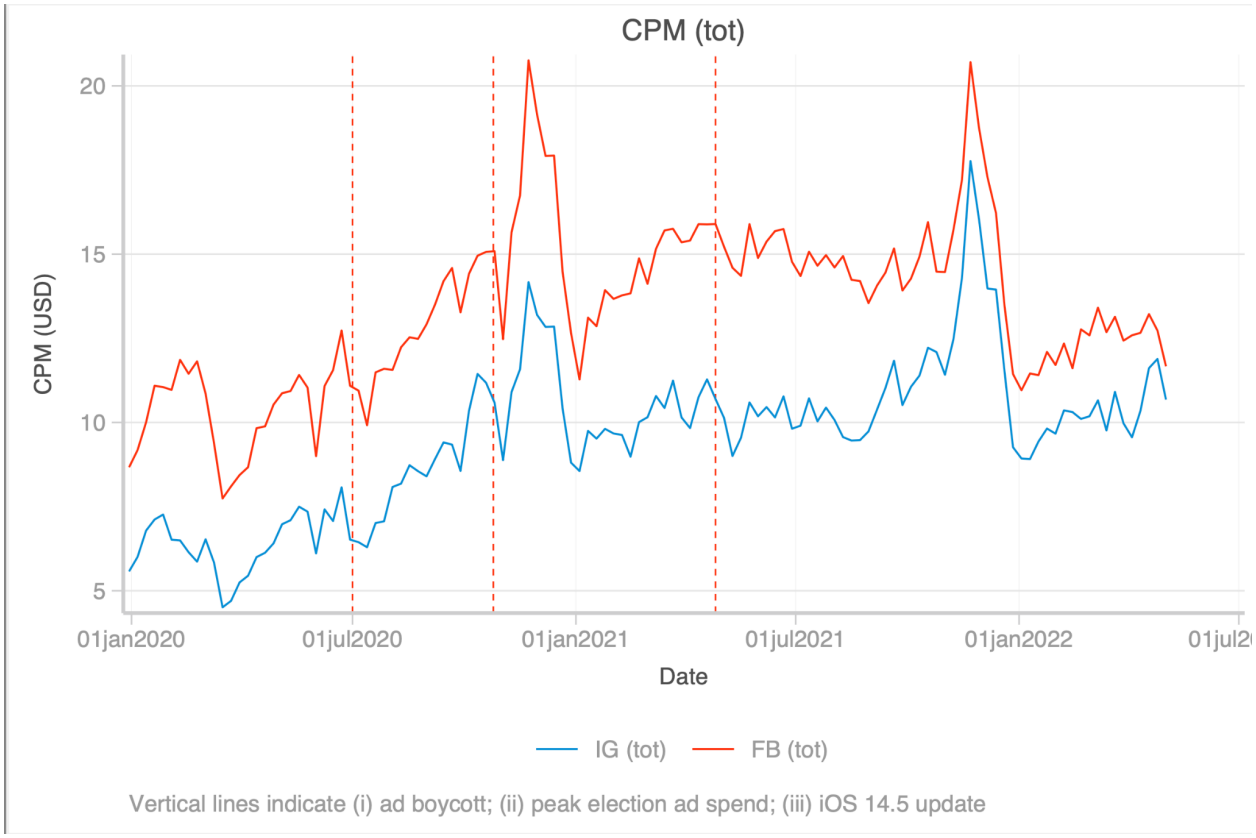
Comments on figures:

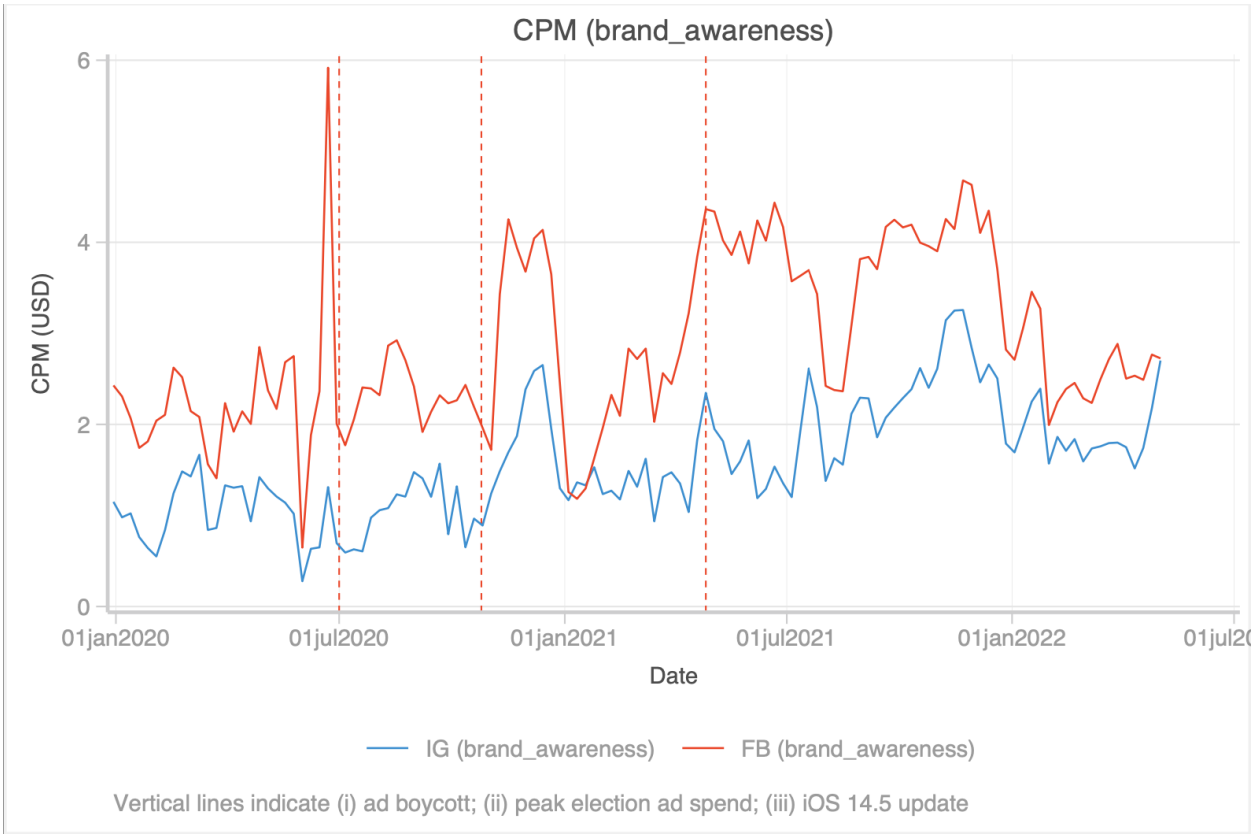
- I think we see spikes in more or less the expected direction for elections. I got the peak election spend from Google's weekly ad spend tracker. Two additional notes:
    - From the Google data, it looks like political ad spend has a slow ramp-up from Aug-Oct, then drops off rapidly (see figure below). So some of the July to election-day increase observable for FB in the cpm_tot figure might be for that reason -- especially since there's no July-Oct gradient in 2021.
    - It looks like there are some confounds for Black Friday / Christmas advertising, so we'd probably have to seasonally adjust.
- The boycott seems to have a more pronounced impact on prices for conversions and traffic.
- Generally, there's a ton of asymmetric price variation on the platform, so I'm optimistic that we can find a credible instrument (even if it requires putting a bit more structure on the demand side).

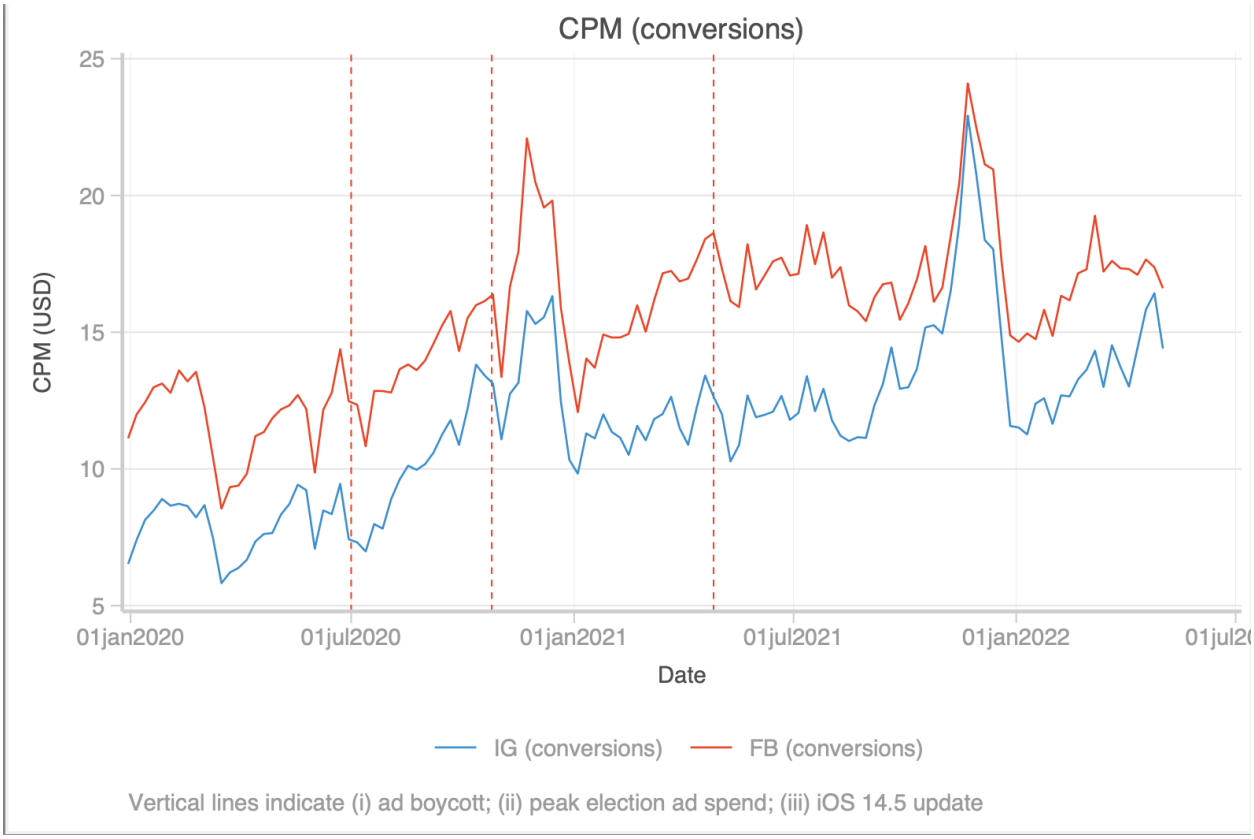
A few other thoughts about instruments / natural experiments:

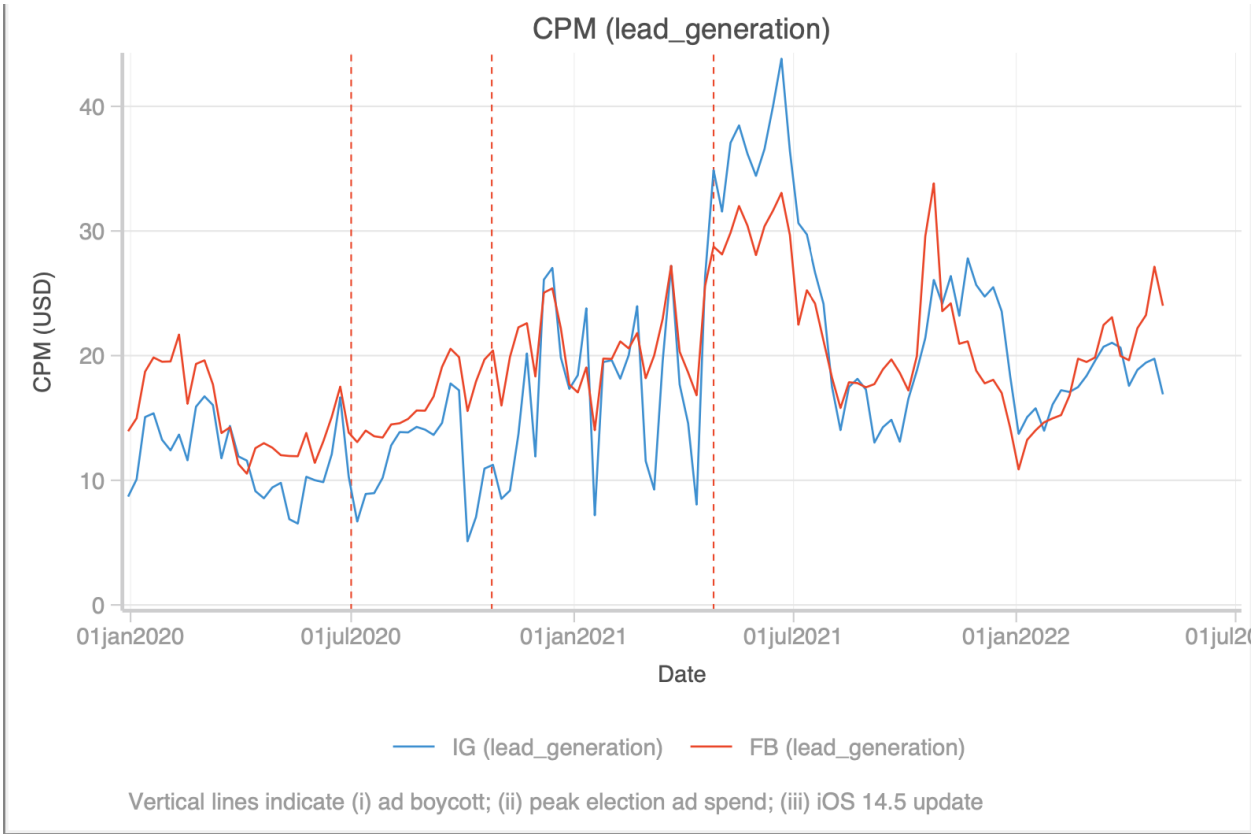
- If we could get costs and quantities for advertising in specific regions / states, then I bet the elections instrument would be way more powerful (most political ads will be targeted in swing states)
- For Apple release of iOS 14.5: there’s some price movement around that time in traffic, video views, lead generation, and brand awareness, although there's a price increase. One possibility is that since the outcome is costs per thousand views, all that's going on is that the update reduces the number of people who see a given ad.
- I wonder if we could use "granular" instruments -- from Pathmatics, we have a time series of ad spend for big firms. Maybe if these firms have huge ad campaigns asymmetric across platforms (e.g. Disney around release of Mandalorian) that would be enough to move prices.

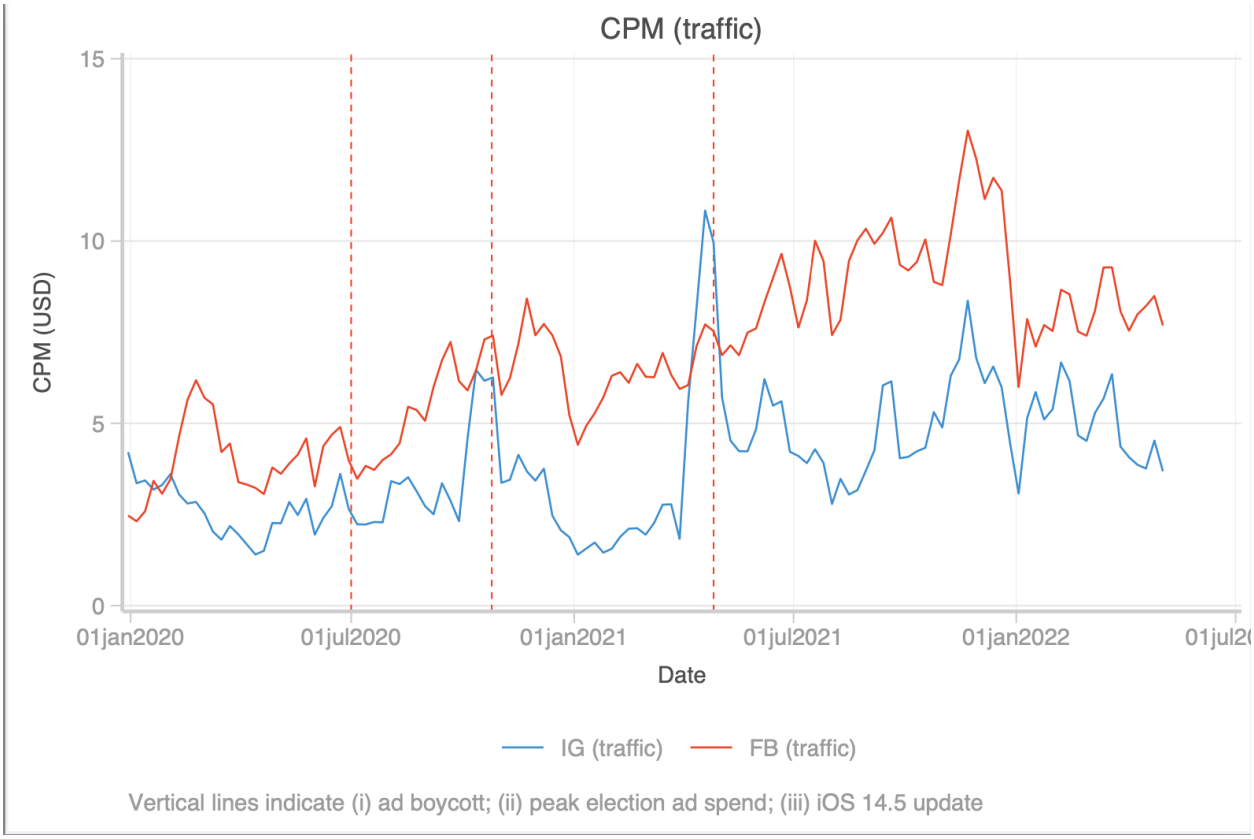
RevealBot CPM figures

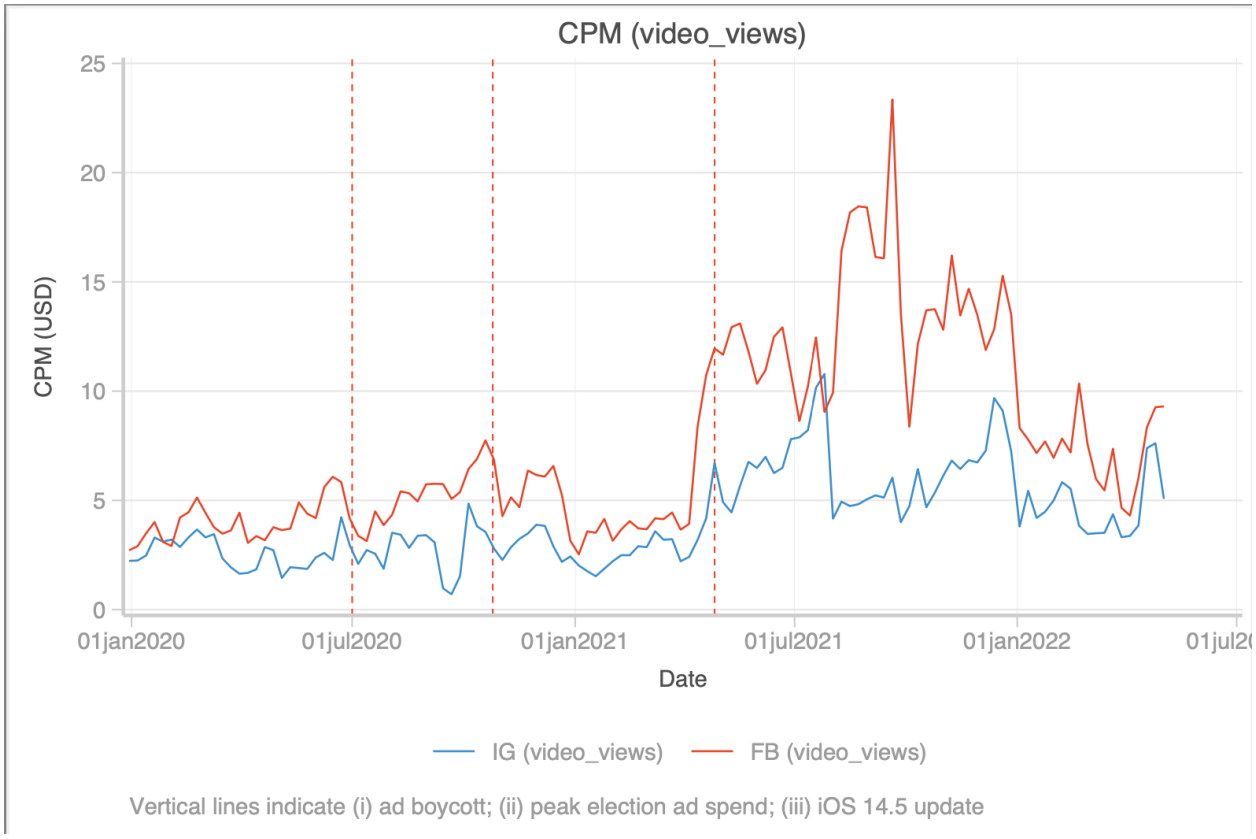

CPM (tot)
CPM (USD)
20
15
10
5
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (tot)
FB (tot)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update

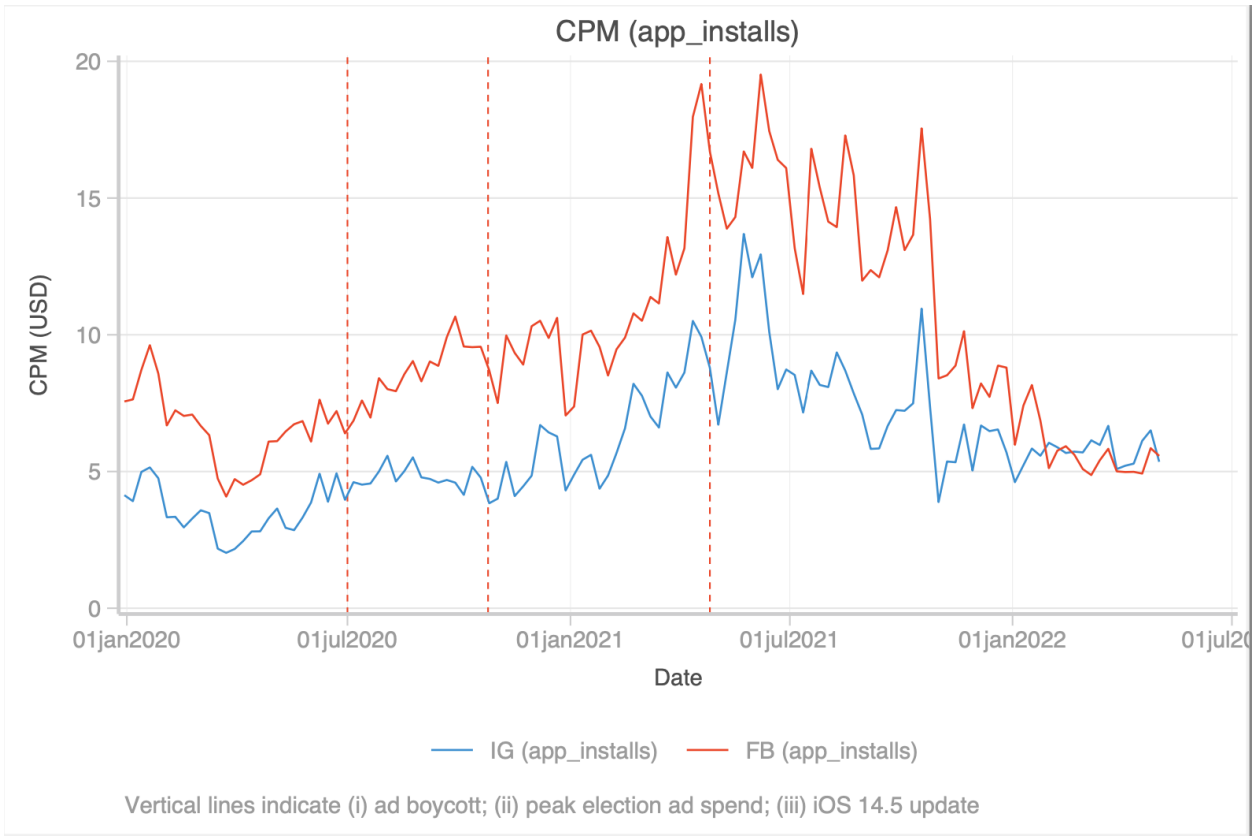

CPM (brand_awareness)
CPM (USD)
6
4
2
0
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (brand_awareness)
FB (brand_awareness)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update


CPM (conversions)
CPM (USD)
25
20
15
10
5
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (conversions)
FB (conversions)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update


CPM (lead_generation)
CPM (USD)
40
30
20
10
0
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (lead_generation)
FB (lead_generation)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update


CPM (traffic)
CPM (USD)
15
10
5
0
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (traffic)
FB (traffic)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update


CPM (video_views)
CPM (USD)
25
20
15
10
5
0
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (video_views)
FB (video_views)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update


CPM (app_installs)
CPM (USD)
20
15
10
5
0
01jan2020
01jul2020
01jan2021
01jul2021
01jan2022
Date
IG (app_installs)
FB (app_installs)
Vertical lines indicate (i) ad boycott; (ii) peak election ad spend; (iii) iOS 14.5 update

Campaign type glossary

The brand awareness objective is for advertisers who want to show ads to people who are more likely to recall them.

The lead generation objective allows businesses to create campaigns to find potential customers for their products or services. Businesses can use these campaigns to gather information from new leads and help them move through the buying process.

The traffic objective is designed to drive people to your website or app. With traffic as your objective, you can create ads that:

Send people to a destination such as a website, app, phone call or Messenger conversation (Website Clicks)

Increase the number of people going to your mobile or desktop app (App Engagement)

The engagement objective is designed to get more people to see and engage with your Facebook post or Page. With engagement as your objective, you can create ads that:

Boost your posts (Post Engagement)
Promote your Page (Page Likes)
Raise attendance at an event on your Page (Event Responses)

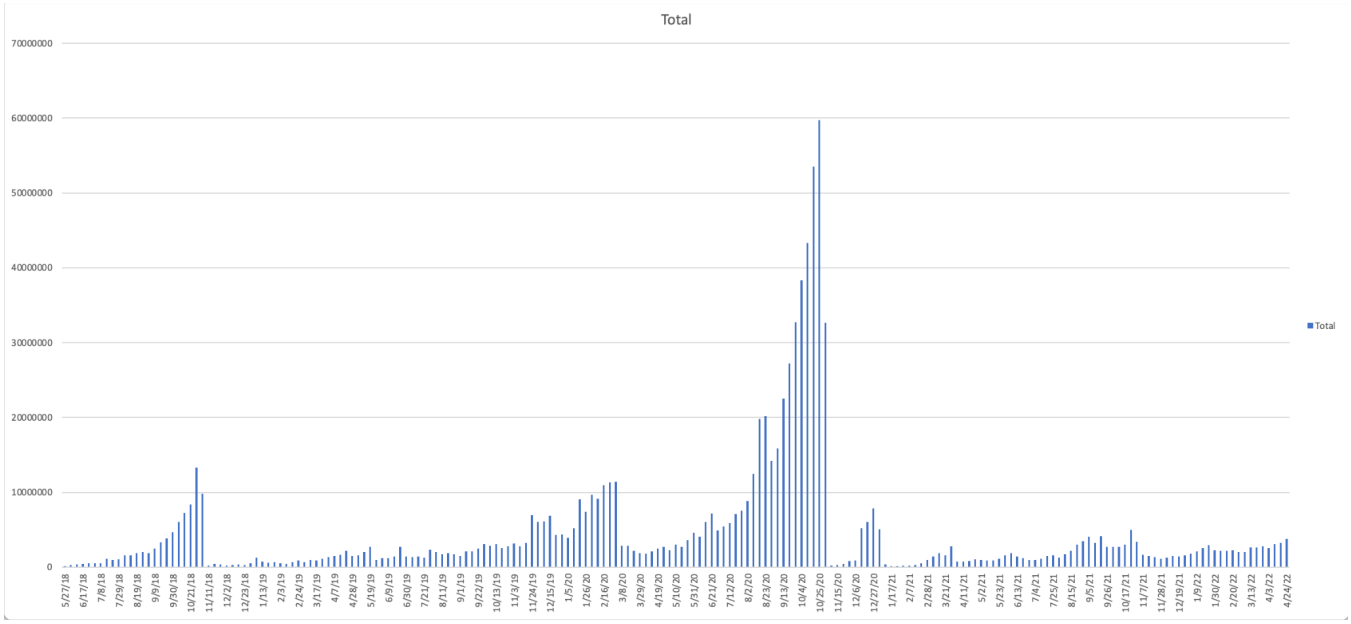
App installs objective: Get more people to install your app, optimize for app events or attract your highest-value customers with value optimization.

The Video Views objective optimizes to get the most plays of your video ad. Show your ads to the people who are most likely to watch for at least 2 continuous seconds, or 15 seconds or longer.

Catalog sales: Show products from your ecommerce store's catalog to generate sales.

The conversions objective is designed to drive valuable actions on your website, in your app or in Messenger from Meta. Actions can include landing page views, purchases, phone calls, messages and more. It requires the Meta Pixel or Conversions API for websites, or App Events for apps.

Google ad spending


Total
70000000
60000000
50000000
40000000
30000000
20000000
10000000
0
Total
5/27/18
6/17/18
7/8/18
7/29/18
8/19/18
9/9/18
9/30/18
10/21/18
11/11/18
12/2/18
12/23/18
1/13/19
2/3/19
2/24/19
3/17/19
4/7/19
4/28/19
5/19/19
6/9/19
6/30/19
7/21/19
8/11/19
9/1/19
9/22/19
10/13/19
11/3/19
11/24/19
12/15/19
1/5/20
1/26/20
2/16/20
3/8/20
3/29/20
4/19/20
5/10/20
5/31/20
6/21/20
7/12/20
8/2/20
8/23/20
9/13/20
10/4/20
10/25/20
11/15/20
12/6/20
12/27/20
1/17/21
2/7/21
2/28/21
3/21/21
4/11/21
5/2/21
5/23/21
6/13/21
7/4/21
7/25/21
8/15/21
9/5/21
9/26/21
10/17/21
11/7/21
11/28/21
12/19/21
1/9/22
1/30/22
2/20/22
3/13/22
4/3/22
4/24/22