Learning to Advertise and Differences in Firm Performance

Extended Abstract

Few findings in empirical economics are as robust as the large and persistent differences in productivity across establishments, even within narrowly defined industries (Syverson, 2011). A great deal of effort has gone into explaining these differences, and most of it has focused on the production side of the firm: its human capital, its incentive structures, and its management practices. Yet production is only part of what makes a business succeed. A firm can be superb at making its product and still fail if it cannot sell it, and selling depends on activities—marketing, advertising, sales—that are seldom recorded in conventional data and are notoriously hard to measure. Our aim in this paper is to take a step into this gap, and to ask whether differences in how well firms advertise can help account for the differences in how well they perform.

The question has become more pressing as advertising has moved online. When economists think of advertising, many still picture a spot during a football game or a full page in a newspaper. But roughly three quarters of advertising spending in the United States is now digital, and a digital campaign bears little resemblance to a newspaper page. It is a complex object that the advertiser continually adjusts: whom to target, how much to bid, how to allocate a budget, which creatives to show, and, importantly, what data to collect on the results. It is precisely this complexity that makes advertising online something one can be good or bad at, and something one can learn to do better over time.

To study returns to advertising, we first need a measure of performance—an advertising analogue of the productivity measures used in the broader literature. The natural one is the return on ad spend: the incremental revenue a firm earns for each dollar it spends advertising. Measuring it credibly, however, is difficult, because the revenue that follows an ad may have arrived even if the ad had never been shown. We therefore begin with a large-scale field experiment on Meta's platforms, spanning more than 700,000 campaigns run by over 210,000 advertisers across 25 industries, and comprising 3.94 billion user-ad opportunity pairs. For each campaign, a small, randomly chosen set of targeted users was prevented from seeing the campaign's ads, which allows us to compare the revenue of users who could see the ads with that of otherwise identical users who could not. This yields a clean, causal return on ad spend for each campaign, and it is this measure that anchors everything that follows. On average, we find that a dollar of advertising returned $3.31 in revenue. This benchmark is of interest in its own right, since credible estimates of the returns to digital advertising remain scarce, but for us it is mainly a point of departure: the average conceals enormous variation across advertisers, and it is that variation we set out to understand.

Our first finding is that returns rise markedly with experience. Advertisers who have spent more on advertising in the past, who have run more campaigns, and who have been on the platform longer all earn higher returns. A natural reading of this is that advertisers learn. But it is not the only reading. The same pattern would arise from selection, if advertisers who happen to earn poor returns simply stop advertising, leaving a more effective group behind. To separate learning from selection, we make use of a feature of our setting that distinguishes it from the classic studies of learning by doing (Arrow, 1962; Benkard, 2000), in which a worker or a plant improves by repeating a fixed task. In advertising online, repetition alone changes little, because it is the platform's algorithm, not the advertiser, that keeps optimizing a campaign once its parameters are set. If an advertiser is to learn, it must do so by actively changing what it does. This suggests a sharp test: returns should rise with experience only for advertisers who actually tweak their campaigns, and not for those who leave them untouched. That is exactly what we find. Among campaigns that are updated often, returns climb steadily with age; among those that are rarely or never updated, they

do not. Selection, user familiarity, and algorithmic improvement all predict otherwise, since none of them has any reason to spare the campaigns their owners leave alone. In work nearing completion, we sharpen this test with advertiser fixed effects, comparing each advertiser's own campaigns to one another to rule out selection on advertiser-level attentiveness.

If learning requires acting, then it should matter a great deal what information advertisers have to act on. This is where measurement enters. Advertisers differ in how carefully they track what happens on their websites—how many events they record, and whether they adopt the more robust server-side tools Meta offers. Better measurement can raise returns in two distinct ways: directly, by giving the platform more signal with which to optimize delivery, and indirectly, by making the advertiser's own experimentation informative, since one cannot learn from a change one cannot observe. Consistent with this, advertisers who measure more also tweak more, and show the same rising returns to experience that active tweakers do, while those who measure little show no such pattern. We are now extending this analysis to explore the complementarity between measurement and learning more directly, by asking whether, among advertisers who experiment with equal intensity, those with richer measurement realize greater returns to that experimentation.

Finally, we ask whether these traits (experimentation and an investment in measurement) are related to whether firms survive. We look at survival in two ways. The first is whether an advertiser is still advertising on Meta eight months later. The second, which we have completed recently and which we find more compelling, is whether the web domains a firm advertised are still active roughly four years after the experiment, a measure of survival that has nothing to do with the firm's relationship to Meta. Sorting advertisers by whether they are above or below the median in learning activity and in measurement, we find that survival is strongly ordered: firms high in both are far more likely to be alive, on and off the platform, than firms low in both, with the two intermediate groups in between. The differences are simply too large to be explained by differences in advertising returns alone, given how little the typical advertiser in our sample spends. We read this as evidence that these advertising traits stand in for broader capabilities of the firm—much as Bloom and Van Reenen (2007) find for management practices, though with the advantage that we observe the relevant behavior directly and measure the outcome with an experiment.

The experiment lets us go one step further and ask what these traits really capture. If they were merely badges of a generally good firm, a firm's actual, experimentally measured return on advertising should tell us little more about its survival within any group. But if they reflect a real capacity to notice how one is doing and respond to it, then measured returns should predict survival precisely among firms equipped to observe them, and not among firms that are not. This is what the data show. Among advertisers who learn, or measure, or both, those with higher returns are substantially more likely to survive, on both of our measures. Among advertisers who do neither, returns and survival are unrelated on the platform, and, strikingly, inversely related off it: high returns in this group appear to mark businesses that live and die by the ad channel rather than sound firms that happen to advertise well.

Taken together, our results suggest that knowing how to advertise, and investing in the means to learn how, is a real and measurable part of what separates firms that thrive from firms that do not. It is an aspect of firm performance that the productivity literature has long suspected mattered, but has rarely had the data to examine. We believe our experiment offers an unusually direct look at it, and we hope it encourages further work on the customer-facing side of the firm.