



Insight

# A Dive Into Digital Dividends

WILL RINEHART | APRIL 2, 2019

California Governor Gavin Newsom recently called for a data dividend “because we recognize that your data has value, and it belongs to you.” The notion that platforms should pay out a portion of their profits to their users has been gaining steam more broadly, as well. But the idea is misplaced for three primary reasons:

- Advertising revenue does not equal the value of data;
- Even if data is jointly created, joint control isn’t the most efficient outcome; and
- Consumers already benefit from ad-supported platforms to the tune of \$7 trillion a year.

## *Parsing Dividends and Data*

In his State of the State address, [Governor Newsom argued for a new policy](#), saying,

I applaud this legislature for passing the first-in-the-nation digital privacy law last year. But California’s consumers should also be able to share in the wealth that is created from their data. And so I’ve asked my team to develop a proposal for a new data dividend for Californians, because we recognize that your data has value and it belongs to you.

The idea of a digital dividends has been gaining steam. Virtual reality pioneer [Jaron Lanier](#) argued in his 2013 book *Who Owns the Future* that online companies that collect and monetize personal browsing data should compensate their users through micropayments. [Eric Posner and E. Glen Weyl](#) repeated a similar argument in their new book *Radical Markets*, and [Tim Wu of Columbia Law School](#) even advocated for it *The New Yorker* in 2015. In *the New York Times*, it was argued that social media companies should pay users for the data used to train their artificial intelligence systems. [John Thornhill, the Innovation Editor at the Financial Times](#), even suggested that Facebook should launch a permanent fund to finance a universal basic income. On the policy front, Paul Thissen (R-MN), introduced [a bill](#) in the Minnesota State Legislature that would require Internet companies to pay users for their data. Moreover, Jim Steyer, founder and chief executive at Common Sense Media, is currently [drafting legislation for California](#). And the *Economist* [also recently featured](#) the idea. Yet what exactly this idea means is less than clear, and it is worth parsing and comparing the notion of a dividend and the concept of data.

Data ownership seems initially to be analogous to [a common experience](#) of the world. Like collecting apples that have fallen off of a tree, the term “data collection” suggests that Google and Facebook are hoovering up a thing produced by someone else. Yet, if Google didn’t exist, there would be no search data. If Facebook didn’t exist, there wouldn’t be social graph data. Data as a co-created product provides a much richer beginning position to think through the digital dividend.

And while there are parallels, traditional dividends are not a clean analogy for what this idea proposes. Traditional dividends come as a result of assuming the risk of a stock—that is, a discreet ownership stake in a company. Shareholders receive a distribution of the profits through dividends, but the value of their stock could

also drop. Digital dividends, in contrast, aren't related to the cash equity of a company but instead to a more amorphous concept of their value.

Further, there are a range of different kinds of shareholder classes, and [not all classes](#) receive the same dividend payouts. Similarly, there are different classes of data, yet it is not clear how these different classes of data would relate to the proposed data dividend. The first kind of data, which might be called *volunteered data*, is data that is both innate to an individual's profile, such as age and gender, and information they share, such as pictures, videos, news articles, and commentary. *Observed data* comes as a result of user interactions with the volunteered data; it is this class of data that platforms tend to collect in data centers. Last, *inferred data* is the information that comes from analysis of the first two classes, which explains how groups of individuals are interacting with different sets of digital objects.

### *The Value of Data*

Like any other asset, the [value of data](#) lies in its ability to earn revenue, but the relationship between revenue and user data isn't straightforward. [Most valuations of big data](#) simply divide the total market capitalization or revenue of a firm by the total number of users. [In its 2018 annual report](#), Facebook calculated that the average revenue per user was around \$112 in the United States and Canada. Antonio Garcia-Martinez recently used this data point in *Wired* magazine to place [an upper limit](#) to the dividend. And Douglas Melamed [argued in a recent Senate hearing](#) that the upper-bound value should at least be cognizant of [the acquisition cost for advertisements](#)—putting the total user value at around \$16 (although he cautiously noted that this estimate was likely inaccurate). Similarly, when Microsoft bought LinkedIn, for example, [reports suggested](#) that they were buying monthly active users at a rate of \$260.

Yet it is misstep to equate the advertising dollars going to tech platforms with the value of user data. Understanding multi-sided platforms requires understanding the goods traded on the user side and the advertiser side. Advertisers spend money on platforms because [people are there](#), just as advertisers spend money on TV, print, and radio because people watch television, read newspapers, and listen to the radio. On Google, Facebook, Instagram, Twitter, and Reddit, user demand comes as a result of the shared content, which is an experience good. Advertiser demand in turn relies upon total user demand, since they are trying to get their messages to users. For advertisers, the *inference data* explain which groups of people—sorted by age, gender, or location—clicked on a web site, liked a page, shared it, or left the platform.

The demand for users is tightly coupled with demand for advertisers, leading to demand interdependencies, which [were explored by the American Action Forum](#) last year. As noted then,

Demand is tightly integrated between the two side of the platform. Changes in user and advertiser preferences have far outsized effects on the platforms because each side responds to the other. In other words, small changes in price or quality tends to be far more impactful in chasing off both groups from the platforms as compared to one-sided goods.

While data is important to the overall maintenance of the platform, much of this data is valuable only within the platform's relationships.

The bankruptcy proceedings for Caesars Entertainment, a subsidiary of the larger casino company, offer a unique example of this problem. As the assets were being priced in the selloff, the Total Rewards customer loyalty program got valued at nearly \$1 billion, [making it](#) “the most valuable asset in the bitter bankruptcy feud

at Caesars Entertainment Corp.” But the ombudsman’s report acknowledged that it would be a tough sell because of the difficulties in incorporating it into another company’s loyalty program. Although it was Caesars’ most valuable asset and helpful in it generating cash flow for that company, its value to an outside party in generating cashflow was an open question. The data itself, apart from the company’s systems, was not obviously valuable at all.

Some businesses have tried to separate out data from the broader information ecosystem, but they have met with little success. The [pay-to-surf](#) business model was popular in the late 1990s until the dot-com crash swept the companies under. Owen Thomas recalled what [happened in the San Francisco Chronicle](#): “AllAdvantage, a Hayward company that exemplified the approach, had to yank its initial public offering and auction off its assets after blowing through millions of dollars.” Later, both [Handshake](#) and [Datacoup](#) began offering payments for data. But Handshake went under while Datacoup isn’t taking new users. *Wired* editor [Gregory Barber](#) went another route and became his own data entrepreneur. He sold his location data, Apple Health data, and Facebook data, and all he got was a paltry 0.3 cents.

### *Data Innovation*

Why couldn’t Barber sell his data for a large sum? Data is often valued within a relationship, but practically valueless outside of it. There is a term for this phenomenon, [as economist Benjamin Klein explained](#): “Specific assets are assets that have a significantly higher value within a particular transacting relationship than outside the relationship.” Since data is a highly specific asset, granting platforms control should be a more efficient outcome.

How then should ownership of those assets be allocated? A broader legal and economic discussion—with its origin in the merger between Fisher Body, [an automobile parts provider](#) in Detroit, and General Motors in 1926—has sprung up around this question. Before the deal, GM bought car bodies directly from Fisher and then mounted them on frames and sold the completed cars to consumers. In this sense, the car bodies were intermediate goods, in much the same way that data is an important intermediate good. But what if Fisher Body, after signing a long-term contract with GM, decided to ask for more money for their parts? Final production would cease, leading to what is known as the [holdup problem](#).

Much research into contracts, mergers, and the control of assets developed as a result of this scenario, and in 2016, Oliver Hart received the Nobel for Economics as a direct result of this work. As one [review of his work explained](#),

[T]he optimal allocation of property rights—or governance structure—is one that minimizes efficiency losses. Thus, in a situation where party A’s investment is more important than party B’s investment, it is optimal to allocate property rights over the assets to party A, even if this discourages investment by party B.

(In the technical appendix to this paper, the model that Hart and Sanford Grossman helped to pioneer is applied to the platform space.)

Even if data is jointly created, joint control isn’t the most efficient outcome. When one party’s investment in the data does not boost the total value that much, then it is better for the other person to own both assets. In the parlance of economics, the party with higher marginal returns from investment should be given control, which is why platforms, and not users, spend so much time and effort to understand what is happening on the platform. Newsom might want to change this ownership division, but it makes sense from an efficiency standpoint.

Changing it would result in less efficiency.

### *The Opportunity Costs*

Digital dividends aim to distribute the value of data that platforms are capturing to users. But there is an extensive amount of value that the platforms *aren't* capturing. Every hour spent on the site is an hour not spent on other activities. There is an opportunity cost to using the platforms.

Indeed, one common way of valuing free services such as Facebook and Google is to calculate the amount of forgone wages. A conservative estimate from a couple years back suggests that users spend [about 20 hours](#) a month on Facebook. Since the [current average wage is \\$27.71](#), this calculation indicates that people roughly value the site by about \$6600 over the entire year. A study using data from 2016 using similar methods found that American adults consumed 437 billion hours of content on ad-supported media, [worth at least \\$7.1 trillion](#) in terms of foregone wages.

Because users have limited attention and the platforms provide experience goods, an absolute upper limit on the value of people's attention exists. As *The Verge* [reported](#), “[E]ngagement has declined throughout the sector, suggesting that the attention economy has peaked. Consumers simply do not have any more free time to allocate to new attention seeking digital entertainment propositions, which means they have to start prioritising between them.” While Newsom and others want to use the digital dividend as a means of distributing value, this research suggests that consumers already receive tremendous value from their data.

Researchers have also explored the value of platforms in experimental settings. Economist Caleb Fuller [tested users'](#) willingness to pay for Google, and found that most were simply not willing to pay for the service. Still, under generous assumptions, the company could expect somewhere between \$14 and \$15 million per year if it charged a fee. To put that in perspective, the 2017 [total revenue](#) for Google's parent company, Alphabet, was \$136 billion. In a twist on this experiment, [one study found](#) that Facebook users would require more than \$1,000 to deactivate their account for one year. After conducting his own version of these studies, former Chief Regulatory Czar Cass Sunstein [noted](#), “The critical point is that we are now used to getting those goods for free.”

### *Conclusion*

The digital dividend might be a simple solution, but it would likely not help consumers. As noted, many are quick to equate advertising revenue with the value of consumer data, but that move isn't warranted. Even if such a system were implemented, it would be inherently skewed towards the wealthy, since it is this group that advertisers are looking to reach. Consumers already benefit tremendously from ad-supported platforms. Policies meant to rebalance an already unequal relationship where consumers win is likely to harm the ecosystem to their detriment.

---

### *Technical Appendix*

One way to understand this bargain is through the Grossman-Hart-Moore model, which considers a relationship between two risk-neutral parties, a buyer and a seller, or *B* and *S*. For this exercise, let's assume that the buyer of the data, *B*, is the platform, and the seller of the data, *S*, is the user, and again let's just work with the singular transaction. As such, the platform buys data, which is an intermediate good, from the users to create a final output. The value of the final good is  $V(e)$ , which is contingent on  $e$ , a variable for the investment into the

process by the platform. Similarly, the cost of the intermediate good is  $C(i)$ , which is contingent on the investment,  $i$ , in the process conducted by the user.

There are two periods. In the first period, each party undertakes some kind of investment and in the second period, they decide to trade at a specific price,  $p$ . If they don't end up trading, they can turn to others and do so. A key assumption of this model is that the investments in the first time period are not contractible.

The social optimum would involve maximizing the total benefits minus the investment costs:

$$\max_{e,i} \{V(e) - C(i) - e - i\}$$

Optimal investment thus occurs when

$$V'(e) = 1$$

and

$$-C'(i) = 1$$

But in the present set up, each party will only retain half of the gains from trade, such that

$$\max_e \left\{ \frac{V(e)}{2} - e \right\}$$

and

$$\max_i \left\{ \frac{-C(i)}{2} - i \right\}$$

Because the parties will have to bargain over how to split the total surplus, each will get half of the benefits from their investment. See [Aghion and Holden \(2011\)](#) for further details on the Nash bargaining. Thus, each

party will underinvest relative to the first best.

If the parties instead have vertically integrated, the result is slightly different. If, say,  $B$  controls the total gains from the production processes, then  $B$  will invest at their first best level while  $S$  will underinvest. Similarly, if  $S$  were to own total gains, then  $S$  will invest at their first best, while  $B$  will underinvest.

This model yields some interesting insights. It is important to note that, like the rest of the literature in this space, the investment elasticities are key. Since  $S$  or users, have extremely inelastic investment decisions, that is, they don't change that much with the possibility of  $B$  appropriating them, it is the case that  $B$  should own the total gains for the most efficient outcomes.

This makes sense in the case of platforms. The investment that matters the most lies in the inference data of the platform. Users have indeed tried to sell their own "investment," but these transactions don't yield much. Moreover, the relative investments speak to why data ownership efforts are likely to fail. Since the marginal returns for any user  $S$  is much higher when a platform  $B$  controls both, as compared to when users simply "own their data," independent ownership is likely to lead to inefficient gains for all sides.