The Value of Data: Consequences for Insight, Innovation & Efficiency in the U.S. Economy

By

John Deighton
Harold M. Brierley Professor of Business Administration
Harvard Business School

And

Peter A. Johnson
Adjunct Professor, Columbia University

October 8, 2013

Legal Notice

Acknowledgements

Substantial contributions to the study were made by Kristine M. Newman and Leora D. Kornfeld.

The authors would like to thank the following research team members for their dedication, skill and professionalism: Heather Ramsey, Aaron Smith, and Douglas Clark.

Table of Contents

Executive Summary

This is a study of the U.S. data-driven marketing economy (DDME,) a complex network of firms that uses data about individuals to make efficient matches between producers and consumers (both individuals and business buyers.)

It finds that, in 2012, producers of goods and services spent about $156 billion, and employed about 676,000 people, to buy marketing services that could not have been performed without individual-level consumer data (ILCD.) Presuming that they did not make the outlays without an expectation of equal or greater benefit, the U.S economy benefitted by at least $156 billion, and likely by much more.

The study found that ILCD was a resource to optimize expenditure on interactive and direct response marketing, both offline and on. It reduced inefficiency in matching producers and customers, or increased effectiveness, or both.

The main work of the DDME was to make marketing more efficient. It did so by allowing producers of goods and services to move spending from mass media to more precisely targeted media and consumers. The value of targeting came from efficient customer selection rather than persuasion. It allowed producers to identify and target those who were most likely to want to buy the producer's offering – the people who did not need costly persuasion – and helped to avoid spending on messaging to people who were not likely to buy.

A second consequence of the DDME was to improve the accountability of marketing investments. Because ILCD let firms measure response to marketing actions, they were able to learn over time which elements of a marketing program produced results and which elements do not, revealing how to eliminate unnecessary communication that customers did not find relevant.

Most of the value created by the DDME was dependent on exchange of data among firms: While some firms used the ILCD of their own customers and did not rent, buy, sell, or exchange data with third parties, they tended to be large and self-sufficient. Exchange and use of ILCD was tightly woven into business practices throughout the DDME, creating a high degree of interdependence among firms and practices. Overall, about 70% of the DDM economy, or $110 billion, and 475,000 jobs, were found to depend directly or indirectly on individual-level data exchanged among firms. If individual-level data could not be transferred among firms, it would cost the U.S. economy significantly more than $110 billion to match buyers and sellers at the same rate and maintain U.S. output at current levels.

Much, perhaps most, of the value from exchange of data came from customer acquisition. Because finding new customers is the business practice most closely tied to demand-generation, exchange of data bears directly on the ability of young U.S. businesses to grow. Small firms rely more on flows of data among firms in the DDME than large firms.

Customer acquisition matters more to them than to large firms with large installed bases of customers and products.

The interviews performed in the course of this report suggest that not all firms acknowledge how much they depend on ILCD. Many participants in the DDME who do not themselves buy or rent data nonetheless derive a large portion of their business from clients – or clients of clients – who do. Many non-participating firms do not necessarily have a clear visibility into their clients involvement in the data marketplace, and overlook or misjudge the extent to which demand for their own services is implicated, if not in the short run, in the medium-run.

The report finds that a substantial portion of all U.S. expenditures on marketing and advertising is affected by the DDME. The sum of measured media advertising expenditures and direct marketing expenditures in the U.S. in 2012 was $298 billion, implying that the DDME was implicated in about half of all marketing communications spending.

The study was concerned with the DDME as it existed in 2012, but reflected on trends going forward. It is anticipated that over the next five years the DDME will grow substantially. Not only will the online sector expand, but offline marketing practices will increasingly rely on individual-level data.

Tables of Findings

Table 1 Total DDME Value Added

Total DDME Value-Added Revenues and Dependence on Data-Exchange	$ Millions	Percent
Total contribution to the data-driven marketing economy	**$156,000**	100%
Value added by services that depend directly on data exchanged or rented among firms	$32,000	21%
Value added by services that indirectly depend on data exchanged or rented among firms	$78,000	50%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$110,000*	71%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$46,000**	29%

Table 2 Total DDME Value Added Employment

Total DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**676,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	139,000	21%
Employment added by services that indirectly depend on data exchanged or rented among firms	336,000	50%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*475,000*	**70%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	201,000	*30%*

Table 3 Summary of DDME Value Added Revenues by Dependence on ILCD Exchanged Among Firms

The Data-Driven Market Economy Value created by firms that rely on individual-level consumer data computed by summing firm revenues net of data that they purchase or media for which they are reimbursed								
DDME Business Segment	Total contribution of the sector to the data-driven marketing economy		Value added by services that depend directly on data exchanged or rented among firms		Value added by services that indirectly depend on data exchanged or rented among firms		Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	
	($ Millions)	% of Total	($ Millions)	% of Total	($ Millions)	% of Total	($ Millions)	% of Total
Agency Holding Companies	$7,000	4%	$1,000	3%	$4,000	5%	$2,000	4%
General Independent Agencies	$6,000	4%	$1,000	3%	$4,000	5%	$1,000	2%
Digital Agencies	$2,000	1%	$0	0%	$1,000	1%	$1,000	2%
Direct/ CRM Agencies	$2,000	1%	$1,000	3%	$1,000	1%	$0	0%
Measurement/Analytics	$3,000	2%	$1,000	3%	$0	0%	$2,000	4%
Digital Audience Assembly	$14,000	9%	$7,000	22%	$4,000	5%	$3,000	7%
Search Audience Assembly	$19,000	12%	$2,000	6%	$2,000	3%	$15,000	33%
Audience Targeting	$4,000	3%	$4,000	13%	$0	0%	$0	0%
Direct/CRM Customer Targeting	$7,000	4%	$3,000	9%	$4,000	5%	$0	0%
Postal Production	$32,000	21%	$1,000	3%	$24,000	31%	$7,000	15%
Email Customer Delivery	$1,000	1%	$1,000	3%	$0	0%	$0	0%
Teleservices	$10,000	6%	$2,000	6%	$6,000	8%	$2,000	4%
Mobile Customer Targeting	$2,000	1%	$0	0%	$0	0%	$2,000	4%
Commerce - eCommerce	$34,000	22%	$4,000	13%	$22,000	28%	$8,000	17%
Commerce - Loyalty	$5,000	3%	$2,000	6%	$1,000	1%	$2,000	4%
Commerce - Fulfillment	$9,000	6%	$0	0%	$4,000	5%	$5,000	11%
Total	**$156,000**	**100%**	**$32,000**	**100%**	**$78,000**	**100%**	**$46,000**	**100%**

Table 4 DDME 2012 Employment: Total, and Dependent on Exchange of Data Among Firms (Direct plus Indirect Dependence.)

DDME Business Segment	**DDME Employment**		**Employment** added by services directly or indirectly dependent on data exchanged or rented among firms	
		% of Total		% of Total
Agency Holding Companies	30,000	4%	21,000	4%
Independent General Agencies	25,000	4%	24,000	5%
Digital Agencies	7,000	1%	6,000	1%
Direct/ CRM Agencies	8,000	1%	11,000	2%
Measurement/Analytics	12,000	2%	6,000	1%
Digital Audience Assembly	62,000	9%	48,000	10%
Search Audience Assembly	82,000	12%	17,000	4%
Audience Targeting	18,000	3%	18,000	4%
Direct/CRM Customer Targeting	31,000	5%	30,000	6%
Postal Production	138,000	20%	110,000	23%
Email Customer Delivery	5,000	1%	4,000	1%
Teleservices	43,000	6%	35,000	7%
Mobile Customer Targeting	11,000	2%	1,000	0%
Commerce - eCommerce	146,000	22%	115,000	24%
Commerce - Loyalty	21,000	3%	15,000	3%
Commerce - Fulfillment	37,000	5%	18,000	4%
TOTAL	**676,000**	**100%**	**475,000**	**100%**

Table 5 Summary of DDME Revenues By Media Ecosystem

The Data-Driven Market Economy Value added by services directly or indirectly dependent on data exchanged or rented among firms				
DDME Business Segment	Online ($ Millions)	Strategic ($Millions)	Offline ($Millions)	Total
Agency Holding Companies	$0	$5,000	$0	$5,000
General Independent Agencies	$3,000	$2,000	$0	$5,000
Digital Agencies	$0	$1,000	$1,000	$2,000
Direct/ CRM Agencies	$0	$3,000	$0	$3,000
Measurement/Analytics	$1,000	$0	$0	$1,000
Digital Audience Assembly	$11,000	$0	$0	$11,000
Search Audience Assembly	$0	$4,000	$0	$4,000
Audience Targeting	$0	$4,000	$0	$4,000
Direct/CRM Customer Targeting	$0	$0	$7,000	$7,000
Postal Production	$0	$25,000	$0	$25,000
Email Customer Delivery	$0	$0	$1,000	$1,000
Teleservices	$0	$0	$8,000	$8,000
Mobile Customer Targeting	$0	$0	$0	$0
Commerce - eCommerce	$0	$0	$26,000	$26,000
Commerce - Loyalty	$0	$0	$3,000	$3,000
Commerce - Fulfillment	$0	$4,000	$0	$4,000
Total	**$15,000**	**$48,000**	**$46,000**	**$109,000**

Table 6 DDME Induced Employment

	DDME Incremental Induced Employment	DDME Incremental Induced Employment (depends directly or indirectly on data exchanged among firms)
Agency Holding Companies	46,240.60	32,590.37
Independent General Agencies	39,232.52	36,847.18
Digital Agencies	10,577.18	8,593.96
Direct/ CRM Agencies	12,198.43	10,368.66
Measurement/Analytics	21,036.84	10,518.42
Digital Audience Assembly	95,894.04	67,125.83
Search Audience Assembly	126,805.22	12,680.52
Audience Targeting	22,994.59	22,994.59
Direct/CRM Customer Targeting	47,072.54	32,950.78
Postal Production	212,512.37	136,556.20
Email Customer Delivery	7,178.39	6,460.55
Teleservices	66,327.33	54,070.04
Mobile Customer Targeting	16,313.37	815.67
Commerce - eCommerce	223,905.63	76,631.70
Commerce - Loyalty	31,992.79	15,996.40
Commerce - Fulfillment	57,519.21	14,379.80
TOTAL	1,037,801.04	539,580.67

1 Introducing the Data-Driven Marketing Economy

1.1 Scope and Objective of Research

This research studies a set of firms that operates as an economic system to collect and apply individual-level data to the problem of finding and cultivating buyers for the products and services of producers. We refer to the system as the Data-Driven Marketing Economy (DDME), and characterize it as an economy because the firms create and exchange value for and with each other by developing and applying data processing technologies.

Commissioned by the Data-Driven Marketing Institute, an initiative of the Direct Marketing Association, the research aimed to answer two questions:

- *How much incremental value does data-driven marketing contribute to the U.S. economy?*
- *How much of this value is accounted for by the flow or transfer of data among firms?*

All firms in the DDME rely on individual-level consumer data (ILCD.) By ILCD we mean an information signal that originates with an individual consumer of a good or service, whether the person is buying to fill personal needs or the person is an institutional agent buying on behalf of a firm. But the signal often cannot be traced to an individual. A distinction that we shall rely on throughout the report is between ILCD that identifies a particular person (for example by their name and street address) and information that cannot identify a person. We follow customary usage and define the first kind of ILCD as personally identifying information (PII) and the second kind as pseudonymized information (PI). An example of PI is information associated with a cookie (file appended to a browser) that identifies the browser but not the person behind the browser. Thus, when technology selects a device as the target of an advertisement or marketing offer, we refer to the data involved as PI.

We use the umbrella term ILCD to refer to both PII and PI because PII defines some but not all channels of interest to this study. While PII media account for a large proportion of the DDME value added, some online advertisers are more interested in particular audiences or market segments than a particular individual (which is why ad units are still traded in thousands of customers (cost per thousand or CPM). In particular, digital channels have addressable audiences that may or may not require integration of truly identifiable data. Finally, however, it should be made clear that the understanding of PI and PII is not fully settled: ILCD may typically involve information that is personally or device identifiable but not always; there are also cases that fall into grey areas, such where marketing may be individually addressable to a person or a device but also not identifiable due to the use of de-identification technologies.

1.2 Overview of Method and Findings

This study investigated directly the revenues generated by over 650 firms for provision of services to producer firms, services that by the criteria above supported one or more ILCD dependent practice. We then studied how these supplier firms differentiated themselves in the marketplace, in order to identify key sub-systems of buyers and sellers of intermediate inputs. Among inputs, we focused first on the cost of media, and second on the cost ILCD, both when it was a tradable good and when it was bundled as part of a service. This analysis allowed us to identify and remove instances of double-counting, and arrive at both a more accurate valuation of the overall size of the DDME, as well as a clearer picture of the degree of interdependence among classes of firms in the exchange of ILCD among owners, users, and servicers in creating the efficiencies provided by the DDME.

By summing the net revenues from the provision of data-driven marketing services for all firms that were investigated as participants in the DDME, and subtracting for costs of inputs among sub-systems, the study found that the data-driven marketing economy represented a minimum of $156 billion in value-added output for the U.S. economy in 2012. This revenue directly supports approximately 676,000 jobs.

Value-added revenues are equivalent to the net amount that producers of goods and services spend on individual-level data services to find customers for their offerings, together with associated employment. As such, we treat this figure as the minimum amount of value added to the U.S. economy by the DDME. In a competitive marketplace, economic theory predicts producers will expand expenditures on inputs up to point the marginal value they receive just equals their marginal cost. Thus the sum of all average surpluses less average cost to producers is likely substantially greater than $156 billion, though assessing the exact extent of this surplus awaits future research.

This report finds that altogether, about 71% of the DDM economy, or $110 billion, and 475,000 jobs, depend directly or indirectly on individual-level data exchanged among firms. The amount of $110 billion represents a lower bound on the inefficiency that would be introduced into the U.S. economy if the DDME had to operate without such data. In other words, if individual-level data could not be transferred among firms, it would cost more than $110 billion more to match buyers and sellers and maintain current levels of economic output.

About $32 billion of the $110 billion is directly contingent on exchanging data among firms. This number is primarily the value to firms involved in the "upstream" transfer of data from data owners towards third-party data users for whom it helps identify prospects, customers, and audiences most likely to find relevance in marketers' offers.

A further portion of revenues, or about $78 billion, is indirectly dependent on servicing individual-level customer data exchanged of among unrelated firms. Firms who earn revenues by servicing data exchanged by unrelated firms tend to be providers either of upstream advisory or creative services, or downstream media production and delivery services. Without exchange of such data elsewhere in the economy, these firms that do

not themselves exchange data would be unable to produce market-making efficiencies to the value of $78 billion.

The DDME represents a substantial portion of total U.S. expenditures on marketing and advertising services. For example, the leading advertising journal *AdAge* indicated that U.S. advertisers spent about $251.2 billion in 2012 including both major "measured" media and marketing services, and the Direct Marketing Association puts the value at about $340 billion.[1] (Estimates of the size of marketing are heavily dependent on definitions of the scope of services and the range of media included. The media research firm Kantar indicated that advertising expenditure on measured (non-targeted) media was about $139.5 billion.[2])

Distribution By Media: Data Is Online and Offline.

Classifying net DDME revenues by kind of data and media, digital display media account for about one quarter of the total size of the DDME. Traditional direct response and interactive media, including postal, email, teleservices and certain mobile-native media account for about half the total size of the DDME by revenue. Strategic firms (primarily agencies and database marketing-services firms) involved across all DDM marketing media account for the remainder.

One of the more surprising and counter-intuitive findings of this report is that a large, even disproportionate amount of the data-driven economy is takes place in the off-line, non-digital economy. This finding becomes more explicable if one analyzes the role of ILCD in two respects: its strategic role in customer, prospecting, or audience *selection*; and its tactical role in marketing communications and media *design, production and distribution*. ILCD is involved in both, but its contribution to the cost and efficiencies varies from medium to medium in both its strategic and tactical applications. For example in the offline world, ILCD first plays a role as a necessary tactical input embedded in a process of production and delivery that is quite expensive, owing to the non-data inputs involved. The study's research revealed this clearly for example when email was compared to direct mail. Email's virtue is that the marginal cost of delivery is virtually zero; unsurprisingly, perhaps, the Direct Marketing Association (DMA) reports that email has the highest return on investment of any direct-to-consumer medium.[3] Yet the study found its net value added at $1 billion, suggesting a paradoxically low level of expenditure by marketers on such an inherently attractive medium. By contrast, direct mail with its reportedly lower return on investment was found in this report to have a net value added of an order of magnitude larger than that of email.

---

[1] The 100 Leading National Advertisers spent an estimated $104.5 billion on U.S. advertising last year, up just 2.8% from 2011, according to Ad Age DataCenter's analysis.
The Top 100 in 2012 accounted for 41.6% of total U.S. measured-media spending. http://adage.com/article/news/big-u-s-advertisers-boost-2012-spending-slim-2-8/242761/

[2] http://kantarmediana.com/intelligence/press/us-advertising-expenditures-increased-3-percent-2012

[3] DMA, "The Power of Direct."

The two roles of ILCD as illustrated in the two cases of email and direct mail help illustrate a general insight into the overall efficiencies produced throughout the DDME. One factor contributing to the ratio is email's extremely low marginal cost, which permits the sending of unwanted irrelevant and unsolicited messages with extraordinarily low rates of response from recipients: the all too ubiquitous spam. In the email channel, ILCD is integrated by responsible marketers in order to improve opt-in rates, message relevance, email open, response, and conversion rates. As such, ILCD becomes a significant component of emails cost for professional marketers in ways that are just not true of free-to-send pure spam. In contrast direct mail campaigns involve costly inputs such as paper, ink, envelopes, content that is increasingly personalized to individual recipients, and in the case of catalogs, high production values like color photography. These inputs make it uneconomical to create and transmit marketing messages except where there is a demonstrable and predictable expectation of relevance to recipients, and a corresponding likelihood of response. This reduces levels of clutter in the channel as a whole, and further helps high-value mailings to stand out from clutter in other media.

Ultimately the value of ILCD is that it changes the natural cost structure of media, making it both targetable and measurable, so that it can be optimized and made efficient and effective, as the case may be: the expensive medium of mail is made efficient by the analysis of individual response rates of prospects and customers, while the inexpensive media of email is made more expensive and yet more effective by the application of data that make opted-in recipients more engaged in its content. ILCD properly understood, involves as much a process of data-driven self-selection and self-limiting among marketers as it does the selection of the prospects and customers they seek to reach.

Data Is Both Fuel And Lubricant

Data is important to the offline economy where the cost of interaction with a customer is much higher than in the digital economy. As noted, marketers still spend more on the postal channel than all digital channels combined. Because it is much more expensive to reach someone by the mail than by email or display advertising, the channel would cease to be viable without data to steer away from people who are not going to buy.

Another way to resolve the paradox of distribution of value-added across the DDME is to ask what do we mean by data? Is data necessarily digital? Or does data, and the value added by data, point to something else? The study investigated many different sectors of the DDME, and found value being added in many ways: it uncovered value created by such different processes as rented lists of prospects, audience segments assembled by online publishers, and personalization of direct mail offers sent to individual addresses. In doing so, the study had to attribute value to the application of data to delivery: via USPS sorting machines, the Facebook servers, the variable data printers that create hundreds of versions of the same edition of a magazine. A key concept to understand the efficiency of the DDME as described in this report is economies of scope, as opposed to economies of scale. Traditional broadcast advertising, or geo-targeted mail can be inexpensive because they offer increasing returns to scale as average per message costs

go down with greater viewership, longer print runs, of a uniform offering. Whereas the sorting functions and individual-destination routing functions built into interactive media and networks like Facebook, Google, USPS offer increasing returns to differentiated, individualized content or product.

And though these costs are often higher on a pure CPM basis, the returns to scope, and the ability to precisely measure cost-effectiveness using by delivery, open, response & conversion rates, ultimately changes the economic equation for many producers in favor of data-driven marketing and media.

Data Is Driving Customer Centricity and Accountability

Formerly, economic activity was producer-centric: engineers invented products, companies bought inputs and hired workers to make them, then – almost as an afterthought – they hired agencies to make ads to persuade people to buy. In the words of Henry Ford, "any customer can have a car painted any color that he wants so long as it is black."[4] Data, and one-to-one marketing, is helping change what is left of that mindset. Data-driven marketing concepts like "customer life time value" and "cost of customer acquisition" have taught a new generation of executives that a firm is less a collection of capital assets than a portfolio of loyal customers whose loyalty is expensive to acquire, difficult to maintain. The marketing function is now about discerning what customers want and need and continually re-engineering the company to provide it. The more firms can develop a three-hundred and sixty degree, multi-channel view of what customers think and feel, the more the customer will truly be in control.

The Value of Data Lies in its being Current

ILCD is time-sensitive. Used to predict individual buying intentions, habits and preferences, it value declines as the time between collection and use increases. It allows the marketer or advertiser to use the underlying medium's inputs more efficiently, e.g., by reducing the cost of envelopes, or in the case of email, using it more effectively – i.e. data hygiene practices improve email's delivery, open rate, etc. In the case of email, data makes the message marginally more expensive, but improves its marginal effectiveness significantly.

Data Flows Enable Small Firms To Overcome Rivals' Economies of Scale

The market for data is rapidly transforming the world of media and marketing both online and off, making them two of the most dynamic sectors of the U.S. economy. One need only think of brands like Google, Apple, Facebook, and Twitter. It is having positive consequences for innovation, and in particular, allowing small businesses to compete effectively with big players, and allowing new market entrants to challenge mature players. And with big companies less likely to take their large base of existing customers

[4] Henry Ford, *My Life and Work*. (1922) Chapter IV, p. 71. http://www.gutenberg.org/dirs/etext05/hnfrd10.txt

for granted, consumer choice and power is enhanced. This consumer value takes many forms, from free content, to direct subsidies in trial offers, to reduced prices and higher levels of service provided to participants in loyalty programs, to coupon-based discounts.

Convergence And Interdependence in the Data-Driven Marketing Economy

Finally, the study finds that data flows have made marketing channels more similar and interdependent than ever before. As a result, many more firms, and a much larger portion of the economy than is widely recognized depend, at least indirectly, on the exchange of data in the marketplace. Evidence of convergence can be seen in the increased pressure for producers to demonstrate measurability and accountability for their marketing outlays. This is causing both new media (for example, social media) as well as old (broadcast and cable TV) to experiment with data-driven procedures to deliver highly individualized segmentations, or demonstrate the impact of ad exposures for offline sales.

But perhaps most surprising is the monetary evidence of interdependence. Our earlier cited figure of $110 billion in value-added revenues attributable to the exchange of data in the marketplace involves two kinds of dependence. Direct dependence, in which a firm provides a service in which data is a component, is less than half of the total market value. The majority of value is captured by firms that do not themselves directly participate in the marketplace exchange of data. Instead, their dependence is indirect, because it arises from their clients' participation in the exchange of data. Thus, some significant portions of firms revenues depends on processes involving crucial business partner to provide demand for their services and employment, but such is the fluidity of the data marketplace, that many of these firms may not be aware of their dependence.

1.3 The Scope of the DDME

Our focus on ILCD does not stand on its own, but takes as its necessary condition the existence of addressable and interactive media and sales channels. Without economically efficient interactive media, it would not be technically possible, or economically efficient, or both, for enterprises to collect and use ILCD as part of marketing.

So while interactive media is necessary to data-driven marketing, it is the application of ILCD that suffices to make marketing within interactive media data-driven. This relationship allows us to define the universe of DDME marketing practices (and their applicable service providers) as all those which enable selective and measurable interactions with customers, prospects or audience members on an individual, one-to-one basis. By "selectively interact," we mean the use of individual attributes to select a set of specified individuals hypothesized to most likely respond to a marketing offer or value an experience, then the use of select media touch points to deliver the offer to, or create the experience for, that exact group of recipients. By "measurably interact," we mean that

the response of the recipient should be linkable to the original record, in order to test and refine the predictive value of the original selection hypothesis.[5]

Our definition of the DDME stands in contrast to, or overlaps with, two other major sectors of the marketing economy.

The first point of comparison is with traditional mass marketing, or what is now more often called "broadcast" advertising. This set of practices is defined primarily by its reliance on non-interactive media (e.g., broadcast TV or radio) to deliver ads to anonymous audiences assembled by independent publishers or broadcast networks. Undifferentiated distribution of these non-interactive media within large geographically defined markets means that advertisers cannot select particular consumers to receive their campaigns. Instead they can only select particular media contexts in which they believe their desired market segments will be more heavily clustered. Similarly, the lack of interactivity means demand must be realized via indirect outlets, typically via retail stores, supermarkets, etc., which means it will be anonymous, and not attributable to individuals who saw (or did not see) the advertising. This means producers cannot measure or attribute success to their advertising at anything except the most general geographic aggregate level, such as for an entire city or state, and can only derive limited insight into what elements of a campaign (media choice, price, product, offer, creative, etc.) worked, or not.

The second comparison is with traditional direct marketing, often called direct response marketing. Long the junior cousin to mass advertising, direct marketing is defined by its use of direct response media, and thus, prima facie, there is considerable overlap with data-driven marketing; in fact, it would be fair to say that data-driven marketing is essentially a subset of direct or direct response marketing. In other words, not everything that is direct response marketing is considered part of the DDME for this study, for within direct response there is a small but not inconsiderable portion of activity that does not use ILCD in one or other of the two steps that for DDM must both be interactive. Examples include direct response TV or radio ads, which are bought by context, not selected for individuals; similarly, we exclude direct response newspaper ads, untargeted resident mail distributed geographically by the post-office, etc., as not being part of the DDME.

1.4 Valuing Data-Driven Marketing

When any set of firms emerges in a competitive marketplace and persuades a considerable portion of buyers to redirect their spending to them, the question must be asked about the nature of the underlying value proposition that drives the benefit they expect to receive from such expenditures. The value proposition of the DDME describes what one might call a shared understanding among vendors and clients about the value of data in optimizing one of the major cost centers for producers in the U.S. economy.

[5] Conversely, if the customer is the one to initiates an interaction, the marketer's data enables the marketer to recognize the individual and respond with the most relevant, value- optimizing offer or experience.

Data has a role in persuasion, but its most immediate application is to help firms to find those who are likely to buy: the people who don't need persuasion. When data flow freely, marketing investments are less likely to be wasted on communicating with people who are not likely to buy.

There can be no efficiency gains in marketing if processes are not measured and tested. Tests need controls; we have to observe not only the people considered to be likely prospects, but also those who are thought to be unlikely prospects, for the scientific method to be applied to marketing investments. The more granular such tests and controls can be conducted, the more efficiently the value-add of each element can be identified, and optimized accordingly.[6]

Thanks to data, new producer firms, new products, and new publishers face lower barriers to entry than they faced in the past. On the one hand, data improves matching of products to customers. On the other hand, the data-driven marketing revolution has eased access to advertising, and advertising-dependent capital. Products and publications that deliver value to users can be launched and grow at a pace unchecked by the need to find a large pool of customers willing to make immediate payment for the value they receive, because entrepreneurs and investors have confidence that customers or audience can be built incrementally, and when they build over time, thy eventually have value to advertisers.

These data-driven activities of segmentation, targeting, and measuring matter to marketers because the cost of interacting with consumers is very substantial, perhaps representing as much as 11% of gross revenue for American businesses today. The national market for advertising and direct marketing was $300 billion in 2012.[7] Expensive as acquiring and analyzing individual-level consumer data can be, the cost of demand-generating interactions with consumers is vastly greater. By acquiring ILCD, and using it for segmentation, targeting and measurement, marketers aim to reduce expenditures on non-valuable interactions as close as possible to zero, while ensuring that the expenditures they do make maximize the number of valuable interactions.

The efficiencies conferred by the components of the data-driven economy are determined in large part by the efficiency of the market mechanism in allocating resources for collecting, exchanging, and enabling data across firm boundaries within the DDME. This smooth functioning partly depends on technology, and partly depends on the legal and regulatory norms allocating rights, responsibilities, prohibitions, and obligations across different classes of market participants.

---

[6] Thus, though long the largest part of the marketing world by expenditure, at least $130 billion annually, the mass marketing ecosystem based on broadcast media is under pressure to provide accountability comparable to that practiced by data-driven marketers.

[7] n.a., DMA Statistical Fact Book 2013. Direct Marketing Association, April 2013.

Technologically speaking at least, the cost of data reproduction and processing is dropping, thanks to advances in digital communications, data storage, and analytic capabilities that mean businesses have more engines than ever before with which to develop and test complex, statistically informed market-optimizing hypotheses about how to optimize the value of their interactions with customers and prospects. The worth of these hypotheses is by its nature reliant on the quantity, quality, and recency of the ILCD they use as fuel. In turn, their value also depends on cost-efficient access to the technologies and communication tools required to execute these strategies and tactics.

Finally, the advent of low cost or open source Web scale analytics software means that data intensive market insights are now more accessible to mid and smaller size enterprise. The benefits of the data revolution are being distributed across the economy (not just to larger enterprises) – and enterprises that better know their own customers add to the value and efficiency of the DDME. The emergence of specialized data providers, supplying specialized, derived or modeled data, means that barriers to entry for specialized niche offerings, for niche producers, can be lower still.

1.5 The Data Driven Marketing Value Chain

Data-driven marketing is one method of marketing, one which seeks to communicate with and influence end-customers using ILCD and interactive media. Its measure is the total expenditure on the applicable marketing services that marketers pay for eligible services to supplier firms within the data-driven marketing economy in order to interact with target customers or audience members via the class of media, sales channels and other media that we define as interactive and individually addressable. Equivalently it is the net revenues received by suppliers for these services from marketers

As seen in the diagram below, therefore, producers stand essentially outside the data-driven economy, placing money into it as expenditure on their major marketing functions: strategy (agencies and research providers); data-driven advertising; prospect and customer relationship marketing; and commerce and fulfillment. Only when they themselves provide a data-driven marketing service to other producers (e.g., when they make their customer data available for licensing) are they treated as inside the DDME.

Similarly, end customers (consumers) stand outside the DDME: their purchase of the goods and services that producers market to them via the interactive sales channels touch points of the DDME are not counted in aggregate, nor attributed to the particular channel in which they are realized. Thus, unlike many analysts who report the amount of online eCommerce sales as though those revenues were "due to" the Internet, we take the problem of media attribution very seriously: seriously enough to focus only on what can be known for reasonably certain, namely what marketers pay to those channels to drive incremental sales.

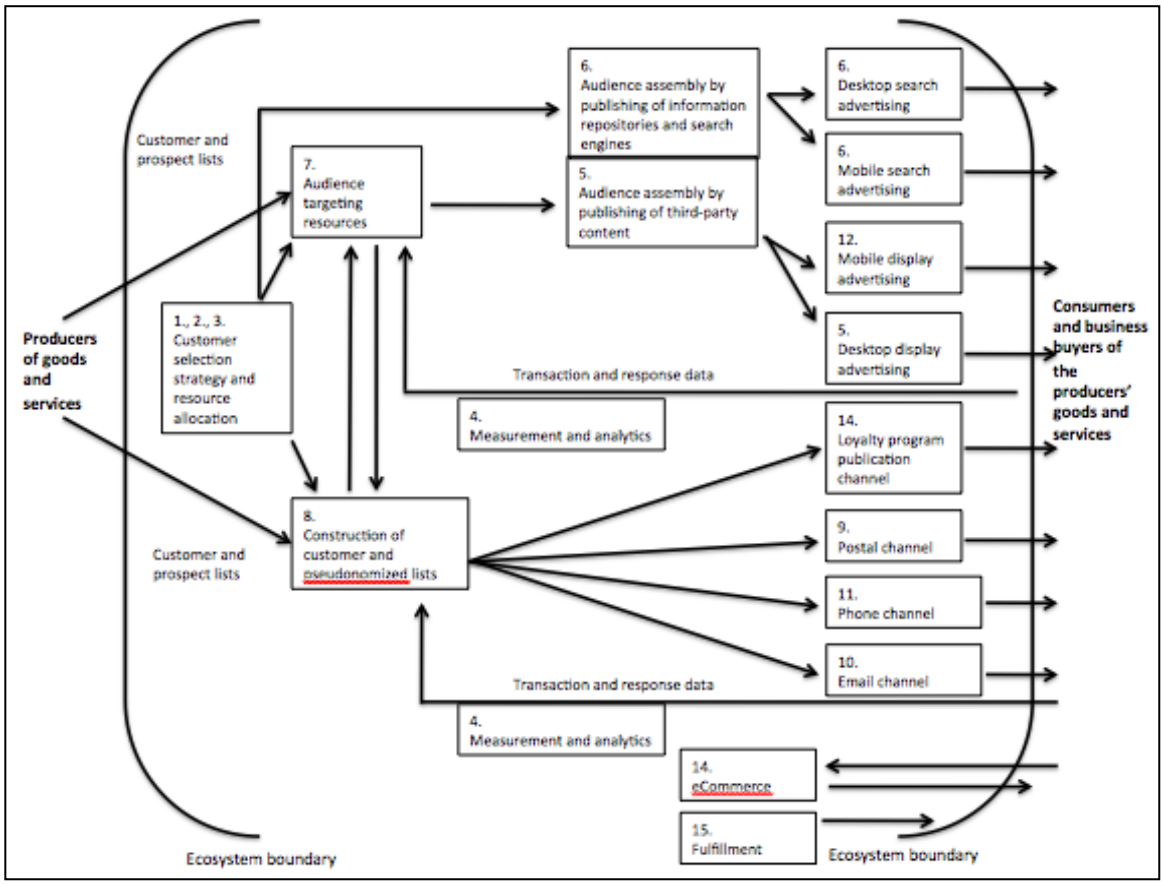

Figure 1: Overview of the Data-Driven Marketing Value Chain

Our definition of value and our approach to measuring is premised on what producers as marketers are willing to pay for services provided by the DDME, because we can safely assume that all other things equal, they would not invest in these services if they did not receive equivalent value in return.

A second advantage follows from the first: arriving at the added value of data-driven marketing requires noticing those juncture points where data crosses firm boundaries, whether as property exchanged with a user in return for monetary or other form of compensation, or as an asset placed in the temporary custody of a business partner for servicing, and then returned. By locating these flows and parsing the exchange in the data marketplace among data owners, data users, and data servicers we can distinguish the exchange of data from the purchases of other goods and services in which data are often intertwined (consulting, media buys, etc.). Thus, we are able to take notice of

where and when data or data processes become the object of market allocation, in what form, and for what purposes.[8]

In other words, observing the and patterns of exchange across the boundaries of the major players in the market allows us to infer which classes of participant in the flow or transfer of data perceive the most value in which classes of data and in which contexts, based on what offer in return. Then, by allocating these value-additions appropriately, we can observe how much the efficient functioning of the market for data and data services contributes added value to the U.S. economy, as distinct from the "underlying value" embedded in marketing, such as media billings, or cost-of goods sold.

1.6 Approach and Findings In Context

This study represents the most thorough, up-to-the moment, and theoretically-informed study yet undertaken of this economically significant activity.

Nonetheless, it must be acknowledged that some aspects of what might be legitimately measured as part of the DDME currently fall outside our current method of valuation, or our definition.

First, value as revenue—whether for data as goods or data as services—does not take into account the incremental benefit (sometimes referred to as incremental campaign lift, delta, marketing return on investment, or ROI) that marketers gain from purchasing DDM services. All other things being equal, one can assume firms in the long run do not spend substantial amounts on inputs unless they expect a return from their last, marginal dollar to match or exceed the value they could achieve by placing that marginal dollar in the market, e.g., at a minimum, simply holding risk-free Treasuries. Thus, were we to measure the value of the data-driven marketing ecosystem using its ROI for marketers, our reported values would in all likelihood exceed the amount reported here, but we cannot hazard a guess as by how much.

Second, the study's insistence on monetary measures revealed in the market for data makes it difficult to measure internal capital investments or overhead expenses on data-driven marketing that marketers conduct in house. Such investments might include marketers' own IT infrastructure or their employment overhead in marketing departments that manage their portfolio of external suppliers. Our resources and time were limited, and the complexities involved in getting at those values were simply too big to be tackled for this project. But we recognize they exist, and that a more complete accounting of the size of the DDME might properly include them.

Moreover, even accepting our "value as net expenditure" measure, there is reason to argue that at least some of the expenditures on marketing – including data-driven

[8] How we handle these issues for classification and measurement purposes is discussed in more detail in chapters two and three.

marketing – are actually contributed (albeit indirectly, and only in small part) by players in the economy other than marketers themselves.

To illustrate: network providers, such as the cable operators or Internet service providers along whose optical fibers the data of the Internet reaches households, principally receive revenues from their subscribers, and marketers and publishers are able to access this network at far less cost than if they had to pay for the whole delivery mechanism themselves. Likewise, many digital publishers and DDME suppliers provide communications services to marketers and consumer alike, but the costs are primarily underwritten (at least in the near term) by venture-capital investments or cross-subsidies from traditional brand media within larger cross-media organizations.

The study's approach to measuring the DDME also produces a conservative result by excluding many marketing practices that some might consider part of the DDME, but this study does not. Its strict definition of the data-driven marketing economy insists on data that is individual and interactive in both inbound and outbound media and channels. What the study characterizes as the DDME does not include, therefore, some practices conventionally measured as part of direct marketing, such as direct response advertising and unpersonalized direct mail.

Similarly, it is important to acknowledge that this study reflects data for one year only. It captures a moment in time – and what may be a fairly unique moment at that. Without a systematic investigation of prior year revenues, this study's capacity to capture the current growth or better yet, anticipate the future size of the DDME a few years hence is limited; but nonetheless, what it did observe strongly suggests that today's DDME is but a harbinger of much larger things to come.

The study's limited purchase on DDME growth rates is particularly significant given that the digital component of the DDME is most obviously a very nascent part of the economy, with large latent potential to further revolutionize the world of advertising. Perhaps an apt comparison might be made to the growth of direct mail as part of the postal ecosystem. Prior to the 1970s, direct mail was a comparatively small and relatively less important part of the overall postal economy. Direct mail's growth, however, soon took off, thanks to the advent of mainframe computing, statistical techniques for analyzing response rates, new accounting metrics such as customer lifetime value, and technologies such as call centers for catalog order placement. Today, over 50% of all mail by volume is now advertising mail, and over 85% of advertising mail is individually targeted. It is entirely possible that data-driven digital advertising may mature in a similar manner in the not-too-distant future.

Just as the study's methodology excludes marketers' surplus (recurring benefits of DDME expenditures) its method of valuation does not factor in the vast (and hard to measure) direct/indirect consumer surplus from the DDME. Such consumer benefits are often said to include things such as improved access to market and pricing information (principally through search); reductions in time spent researching and making purchases increased market competition; and transparency on the price of goods.

Finally, the study is conscious of, but can only measure imperfectly, the numerous spillover benefits -- positive externalities, in the language of economics -- to the U.S. economy from the rise of data-driven marketing. These spillover benefits are not the reason producers undertake such expenditures, and therefore fall outside the scope of our study per se. Nonetheless their wider benefits are real and large, and as such worth mentioning as possible additional measures of value that ideally should be studied further. For example, investments in data-driven marketing continually inspire or support the development of new technology and new ways of storing and analyzing data. Cloud computing is one example and social media monitoring is another.

Data-driven marketing fosters considerable startup entrepreneurship among publishers. Google, founded by two grad students in Susan Wojcicki's garage, is an example of a vast new undertaking inspired by a way for advertisers to capitalize on the purchase intent or pre-purchase research needs of specific in-market individuals. Hadoop database technologies have likewise emerged to serve the needs of the social media publishers, with their vast trove of unstructured (non-numeric) data. Among producers and their agencies, advances in software makes "processed" data an inexpensive and abundant raw material, and the machines that process it decline in cost at steep and steady pace.

Furthermore, it is widely recognized that ad-supported online services have considerable consumer utility (e.g., social networks, map services, voice services) that can be measured either by the value of time spent by consumers on such services (vs. TV or other leisure services); or alternative, from time saved (improved mapping/weather info saves time, travel, gas). And were we to pursue it, it is hardly disputable that compared with 15 years ago, the DDME-supported Internet affords a vastly greater range of free services and content, which consumers spend untold millions of hours consuming.

So, although consumer utility has a long pedigree as an important basis of economic comparison in some precincts of economics, the challenge for researchers is that they involve the consumption of free services, no money changes hands, and so can only be quantified using so-called softer measures than are used in traditional economics and business accounting.

1.7 Induced Economic Impact and Employment Analysis

In recognition of these more outward-facing measures, our second measure of value is employment. Employment is the number of employees in applicable supplier organizations whose job function is directly associated with data-driven–marketing revenues.

Employment is a natural complement to marketer-sourced revenues as a measure of employment, because it allows us to fill in gaps in our revenues where primary data is not available, and because it serves as a check on the validity of data we do have. Differences in the value of revenues in the US, relative to US-based employment in the DDME, can provoke important questions and insights when compared to firms' international revenues, or when online revenues are compared to offline revenues. For

example, we found that the volume of employment appears to exceed the U.S. revenues supporting it. Or to take another comparison, non-mail production employees directly dependent on DDME revenues from the market for data seem comparatively small when compared to traditional off-line employment. Such comparisons point to several insights not revealed by revenues alone.

First, when considering the online economy vs. the offline economy, it is important to recognize the connection between the large number of jobs in digital services that are wholly ad-supported (and DDME dependent); and the number of jobs from startups that depend on venture capital funding. Much of this startup funding is available because of investors' expectation that these firms will be able to engage in high-yield, data-driven advertising in the future. The best recent examples of startups making the pivot to a cash-positive (and data-driven) ad model are Facebook and, more recently, Twitter.

Second, data-driven marketing is a major export industry. The employment analysis allows us to estimate whether the DDME is a net (export) contributor to U.S. economic well-being. As we see below, the firms in the DDME derive a considerable portion of their revenue abroad (sometimes upwards of 15%) while employing nearly all their workers in the US. This strongly confirms what is already widely known, namely that the U.S. leads the world in data science applied to the marketplace. Ideas developed in the U.S. by American statisticians and econometricians, running on US-designed hardware, and coded in algorithms developed and tested in the research offices of U.S. firms, are used to generate revenues throughout the world.

As partial compensation for unavoidable exclusions from our measurement of the DDME's value, the study applied standard economic multipliers to determine an induced impact of the DDME on the wider economy due to the private consumption of DDME employees, outside the DDME itself. Thus in our conclusion we report a set of secondary, induced impacts that represent one possible indication of the extent of yet further value created by the DDME captured beyond the bounds of the system itself.

1.8 Research Overview

To calculate the value of the data-driven marketing's overall contribution to the U.S. economy, we used the definitions above to identify a set of about 650 firms likely to be involved in providing DDME services to marketers. We then performed desk research about our identified firms, using the best available public information in order to determine how much revenue and employment these firms' data-driven marketing services provided. For publicly traded firms, we used their most recent SEC form 10-K filings, identifying the revenue lines of business that met our DDM criteria; for privately held companies, we used sources such as Hoovers, OneSource, and TechBase.

In addition to conducting secondary desk research, we sought to interview a representative, stratified sample of about 10% of our identified companies to gain further insight into their role in the market for ILCD-related services. Their insight, when compared with others in a similar position in the DDME, was used to create general rules of thumb. We formulated rules about which type of firm bought or sold data or data-

employing services, and with which type of firm they conducted this business. We also devised rules about what types of data and services were bought or sold, and to what extent.

The resulting average or typical distributions were generalized to other firms in the interviewees' segments, in order to estimate the cost of goods sold, we would need to calculate the added values of revenue and employment for all such firms, i.e., net of data inputs.

With these derived values, we were thus able generate a value-added picture of the DDME. Such a picture avoids the problem of double-counting, and ensures that each major step in the flow of data in the ecosystem is properly credited with the value it creates, and after subtracting the value of data sourced from elsewhere in the ecosystem. The sum of such value-added revenues and employment across our universe of firms (plus conservative estimates for residual "all other" firms in each segment of the value chain) represents the direct economic impact of the data-driven marketing industry.

1.9 Structure of The Report

Chapters two through five present our findings about the total revenues and employment of each major component of the DDM economy, as well as our estimate of the revenues and employment attributable to the value-added transfer of data among first party owners, third-party users, and third-party servicers. These four chapters cover strategic services (agencies and market research firms); data-driven advertising (Digital display and search); prospect / customer relationship marketing; and commerce and fulfillment. Chapter six provides our conclusion as to the net value-added size of the DDME, and the contribution to this value-added by our major classes of firm. These conclusions include both the business model within the data flow chain as well as the marketing function and type of media.

Finally, supplementing the main body of the report is an appendix containing an extended discussion of the study methodology, and an appendix containing a distribution of value by State.

2 Strategic Services

In this chapter we describe the role, dollar value, and employment created by strategic services in the data-driven marketing economy. We analyze five kinds of firm that give strategic direction to the producers of goods and services as they seek to apply individual-level data to their marketing challenges. First there are large comprehensive agency holding companies. Next there are smaller independent agencies, some providing general agency services and some specializing in digital and direct services. Finally there are those that provide strategic counsel in the form of measurement and analytics. Collectively strategic services generate value of $20 billion to the data-driven economy and employ 82,000 people.

Producers of goods and services, except in the few cases in which economics do not favor it[9], design their strategies to find and retain customers in close consultation with advertising agencies, whether general agencies or those who specialize in particular channels. The strategic decisions that matter most to this report deal with how to allocate resources across alternative communication media. The broadcast advertising media (where purchases do not much rely on individual-level consumer data) and other media such as digital display media and mobile media (where they do rely on such data) are substitutes, so advertising agency strategic counsel is an important input to clients as they choose between broadcast and individual-data-dependent media. Strategic services providers, most of which are advertising agencies, therefore play an important role in determining the value of the data-driven marketing ecosystem. Ultimately, however, their recommendations rest on the two kinds of media's relative efficiency and effectiveness at matching buyers and sellers.

Advertising agencies play a client-facing role in advising marketers and advertisers about what proportion of their strategy should be data-driven. They also play an ecosystem-facing function in advising on the execution of the strategy. In this latter role, they perform for their producer clients what one observer has called "making enterprise sales.[10]" In other words, they recommend to their clients which data service vendors to employ in addition to performing executional services themselves from their stables of subsidiaries. The largest of these services is media buying, including digital display, search, and direct marketing media.

The strategic counsel of agencies is reflected in the way they make their own investment decisions. The split between broadcast capabilities and data-driven capabilities reflects

---

[9] The exceptions tend to occur when producers sell direct to consumers, as credit card issuers and large catalog retailers do. Such firms build substantial internal capabilities and rely less on suppliers. They do buy individual-level data, which we can observe in the revenues of their suppliers, but they use more data services than they buy. We estimate these unobserved amounts from time to time throughout the report.

[10] Wieser, Brian, Pivotal Research Group, 7 June 2013: "The functions that agencies have long-mastered – socializing ideas and making enterprise sales – is much more akin to the services provided by professional services firms such as Accenture, who not coincidentally, have been developing their capabilities in working with marketers over many years. Other companies focused on enterprise technology sales also possess similar competencies, and so the likes of IBM, Oracle and Salesforce.com are each credibly looking for ways to capitalize on the growing importance of information technology [IT] to marketing, and of marketing to IT."

their expectations of the needs of the kinds of clients they serve. Therefore the first step in the valuation of the DDME is to assess how the strategic advisors allocate their own resources.

We begin by reference to the Advertising Age AdAge Datacenter report[11]. It gives United States revenues for the approximately 1,000 advertising agencies of all kinds who report to their survey, from all forms of marketing communications, including advertising, media, digital, marketing services, health care and public relations, as $35.6 billion in 2012. The revenue is earned in two forms: service fees and mark-ups (commissions) on purchased services such as media bought from publishers and research bought from vendors. We need to decide what proportion of this revenue belongs in the data-driven economy.

2.1 Agency Holding Companies

Holding companies are large collections of operating units (some have over 100 entities) spanning a broad range of services including creative agencies, media agencies, direct marketing agencies, market research suppliers, database management and analytics, and others. Holding company groups dominate the provision of strategic services to U.S. producer firms. They earned $22.8 billion of the $35.6 billion of U.S. agency fees and commissions in 2012.

Net Value Added to the Data-Driven Economy by Agency Holding Companies.

The first task was to estimate how much of the $22.8 billion paid to agency holding companies by producers of goods and services was spent on data-driven strategic services, whether strategic or executional. To do so we proceeded as follows. As public companies, all reported the make-up of revenues by lines of business and all reported their U.S revenues from all sources, but none broke them out by operating unit. Therefore we relied on a variety of media reports to estimate the portion of revenues that relied on individual-level consumer data[12] and assumed that the ratio of U.S. data revenues to global data revenues was slightly higher than the ratio of all U.S. revenues to

---

[11] AdAge.com Datacenter Report: "Top Advertising Agencies: Global Rankings, Profiles, and Timelines", www.adage.com/datacenter (via subscription)

[12] Publicis: Like many traditional advertising agencies, Publicis' digital media efforts and mobile ads are now driving innovation and results, especially in the US. The company's digital segment now accounts for more than 35% percent of its total revenue. (Source: Bizmology: http://bizmology.hoovers.com/2013/07/18/digital-mobile-advertising-industry-revenue/)
Publicis Groupe said digital accounted for 30.6% of 2011 worldwide revenue (Source – Ad Age http://adage.com/article/agency-news/digital-star-performer-agency-report-2012/234422/)
WPP's digital revenues were well over U.S. $5 billion in 2012 and represented 33% of total Group revenues of U.S. $16.5 billion. (Source: WPP website, http://www.wpp.com/wpp/investor/financialnews/2013/aug/15/wpp-digital-acquires-minority-stake-in-mutual-mobile-in-the-united-states/)
WPP reported direct, digital and interactive revenue equal to 29.7% of revenue in 2011 (Source: Ad Age http://adage.com/article/agency-news/digital-star-performer-agency-report-2012/234422/)
We found no break-out of Omnicom's digital revenue as a percentage of total. MediaPost reported in 2010: "Digital revenues collectively comprised 18% of Omnicom's overall revenues in 2009. So, a year later, we estimate its proportion at less than rivals', around 20%. (Source: paidcontent http://paidcontent.org/2012/07/31/pc50/16/)

global revenues because data-driven marketing is more developed in the U.S. than globally. By this method we estimated gross revenues in the DDME at $7.5 billion.

From these gross revenues we estimated value added revenues by deducting payments of two kinds; charges customarily passed through to clients which were not passed through (rare in the case of advertising agencies) and data paid for by the agency itself, for its own use, as a general operating expense, whose value plus margin is recouped in providing data-dependent services to multiple clients. Our investigation concluded that these payments were small. We disregarded purchases of custom and syndicated research from audience- and market-research companies because the data were anonymously aggregated and therefore did not qualify as individual-level data. We took account of agency use of their clients' customer data (so-called house file data) and of their clients' proprietary data on prospects, but only to the extent that the clients had acquired the data by trade across the boundaries of their firms. We found that agencies did sometimes buy third-party data to enhance or append to client prospecting and customer retention databases, and to help them to score the databases (to estimate individuals or pseudonymized individuals most likely to respond to the producers' offerings.) And cookie data were used to optimize the design of client website landing page and by agency trading desks for programmatic display ad buying (discussed in Chapter 6.) We estimated agency purchases of individual-level data from other parts of the data-driven economy at $0.6 billion, so we conclude that agency holding companies made a net value-added contribution to the data-driven marketing economy of $7 billion.

How Much Value-Added Revenue Depends on Individual-Level Data?

The second task was to estimate what proportion of these strategic service revenues relied on data that were exchanged or rented across the boundaries of firms in the DDME. As described in Chapter 3, this calculation depends on the media used to earn the revenues. From interviews we conclude that 95% of holding company individual-data-dependent revenues were earned as cross-media or personally identifiable media consulting fees, creative and execution work, or for cross-media commissions. The remaining 5% was earned for services in the pseudonymized digital media: digital display, mobile, and search.

Applying the estimates of how much data is traded among firms in each medium that were identified in Chapter 3, we computed summary inter-firm to be $1 billion. This value is premised on the agencies buying data for their own use in generating the value-added services.

Further, as observed in Chapter 3, data purchases by clients create indirect value for agency holding companies even when they themselves do not exchange or rent data with other firms. Following the principles laid out in Chapter 3, indirect value from data transfers among firms were computed to be $4 billion.

Thus estimates of this sector's value were derived as follows:

Table 7 Agency Holding Companies

Agency Holding Companies DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$7,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$1,000	14%
Value added by services that indirectly depend on data exchanged or rented among firms	$4,000	57%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$5,000*	71%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$2,000**	29%

Table 8 Agency Holding Companies Employment

Agency Holding Companies DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**30,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	5,000	17%
Employment added by services that indirectly depend on data exchanged or rented among firms	16,000	53%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*21,000*	**70%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	9,000	*30%*

2.2 Independent General Agencies

The largest of the general agencies are analyzed in the previous section, but there remain a group of independent general agencies whose combined U.S. revenues were $6.8 billion. Some of these agencies perform work that employs individual-level consumer data, but to a significantly smaller degree than the agency holding companies, which include in their portfolios large specialized direct and digital agencies such as Digitas and Wunderman. Independent general agencies tend not to have large direct and digital divisions, but rather concentrate their services on broadcast communications. Nevertheless the rise of data-driven marketing is changing the proportion of the work

performed that relies directly and indirectly on data not only by agency holding companies but by independent general agencies. Among the direct effects are the inclusion of statistical analysis and modeling into the range of services offered. In addition, as discussed in the section on agency holding companies, there are indirect impacts. While the agencies themselves may not transfer or rent much individual-level data, their clients may do so.

Following the logic of the previous section, we obtain the following estimates of value:

Table 9 Independent General Agencies

General Independent Agencies DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$6,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$1,000	17%
Value added by services that indirectly depend on data exchanged or rented among firms	$4,000	67%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$5,000*	83%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$1,000**	17%

Table 10 Independent General Agencies Employment

General Independent Agencies DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**25,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	6,000	24%
Employment added by services that indirectly depend on data exchanged or rented among firms	18,000	72%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*24,000*	**96%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	2,000	*8%*

2.3 Independent Digital Agencies

The first independent digital agencies were born about the time of the birth of Internet commerce and consumer browsing in the middle 1990s, and, by offering skills to support online advertising and web design, helped offline firms negotiate the transition to online participation. In the years that followed, digital agencies supplied the strategic counsel and to an extent the executional capabilities to drive digital advertising to over one fifth of all U.S. print, television and radio spending[13], and to grow online retailing to 8% of all U.S. retail spending.[14] While the innovation that accounted for the rapid pace of growth of these sectors owed more to developers of search technology (Google) and online commerce technology (Amazon and eBay), and more recently technologies for real time bidding for publisher audiences, digital agencies have played the role of interpreting opportunities for producers to use these technologies to go to market.

As digital advertising and commerce grew, agency holding companies made acquisitions from among the most successful of the pioneers. Nevertheless, partly as a result of a vigorous pace of new agency birth and partly because of determination of some more established digital agencies to remain independent, the 2012 AdAge Datacenter report identified independent digital agencies doing, collectively, $2.0 billion of revenue in the U.S.[15]

We estimate, based on our interviews and secondary analyses, that about half of this revenue is in the form of fees for strategic services, often bundled with creative services, including counsel on allocating resources to new media such as social and mobile platforms, as well as more familiar options like display formats and search. Some part of this revenue could be earned without individual-level consumer data. However the other half, being executional in nature, is much more dependent on such data. It flows from fees earned by supporting clients in media purchases – digital, mobile and search. As these media are more efficient when enabled by individual-level consumer data as Chapters 6 and 7 explain, most of this latter fee revenue forms part of the data-driven economy. Combining the two revenue sources, we conclude that 85% of the revenues of these agencies or $1.70 billion are relevant to this study.

How Much Value-Added Revenue Depends on Exchanged Individual-Level Data?

As described in Chapter 3, the proportion of the $2 billion that was for services in which data was exchanged or rented among firms in the DDME is a function of the media used to deliver the services. Much of the $2 billion is for services performed on client data. We conclude that very little of the DDME value is for services in which agencies buy or rent data themselves, and $1 billion is indirectly reliant on cross-firm transfers.

---

[13] Citation needed
[14] Forrester Research, Understanding Online Shopper Behaviors, U.S. 2011, May 17, 2011
[15] AdAge.com Datacenter Report: "Top Advertising Agencies: Global Rankings, Profiles, and Timelines", www.adage.com/datacenter (via subscription)

Table 11 Independent Digital Agencies

Digital Agencies DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$2,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$-	0%
Value added by services that indirectly depend on data exchanged or rented among firms	$1,000	50%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$1,000*	50%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$1,000**	50%

Table 12 Independent Digital Agencies Employment

Digital Agencies DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**7,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	2,000	29%
Employment added by services that indirectly depend on data exchanged or rented among firms	4,000	57%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*6,000*	**86%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	1,000	*14%*

2.4 Independent Direct Agencies

Digital agencies are, as we have noted, a product of the last 15 years. Direct agencies, by contrast, have existed for well over 100 years. Just as the U.S. created digital market-making media by commercializing the browser in the 1990s, so it created postal market-making media when Montgomery Ward developed the mail order catalog in 1872. The first trade association of direct marketers formed in 1917. In contrast to mass-media-focused general agencies, which advise their clients on how to reach audiences that have

been assembled by third-party publishers, direct agencies advise their clients on how to aggregate markets of specific individuals, a process made economically feasible by a network of individually selectable mailing addresses. As the number of media with individually selectable addresses has expanded, from the postal address to the telephone number, email address, basic mobile phone number as text message destination, and smartphone browser as rich media destination, so the agency function has evolved from predominantly executional to predominantly strategic.

The skills that direct agencies acquired during over a century of working with individual-level data have informed, to a degree, the practice of marketing generally as it has become more data-intensive. Each of the major agency holding companies has acquired once-independent direct agencies, and in some cases the direct marketing executives have become central to the reorientation of general agencies toward digital capability "to better respond, executives say, to the significant shifts in how marketers are seeking to reach consumers.[16]"

Net Value Added by Independent Direct Agencies

Despite this integration, the AdAge Datacenter reports that in 2012 there were independent direct marketing agencies with U.S. revenues of $2.0 billion.[17] Because, as noted, most of their work uses personally identifying information such as a mail address, or information that is easily matched to a personal identity, such as an email address, all of this revenue is considered for inclusion in the DDME. Nevertheless a portion will be deducted. We estimate, based on our interviews and secondary analyses, that direct agencies record as revenue about 5% of revenue that is billed to their clients for media purchases, and that another 5% of revenue is spent on data that appears elsewhere in the DDME as revenue. Therefore we estimate that the contribution of independent direct agencies to the DDME is about 90% of $2.0 billion.

How Much Value-Added Revenue Depends on Individual-Level Data?

Following the procedure in Chapter 3, we estimate the proportion of the $2 billion that was for services in which data was exchanged or rented among firms in the DDME as a function of the media used to deliver the services. We took into account that services designed to find new customers for the client (so-called prospecting services) often make extensive use of data rented from other firms to append to prospect lists to improve their ability to identify people likely to find the client's offering of interest, while services designed to try to retain existing customers use the client's file of existing customers and make less use of rented data from other firms. We found that agency clients with catalogs did 80% of marketing to lists of existing customers with little appended data, while for other direct mail clients as little as 20% was to existing customers and

[16] Elliot, Stuart,"Wunderman, Big WPP Agency, Gets a Major Makeover," New York Times March 8, 2013.
[17] AdAge.com Datacenter Report: "Top Advertising Agencies: Global Rankings, Profiles, and Timelines", www.adage.com/datacenter (via subscription)

appended data played a larger role. Considering the mix of agency clients, we arrived at the following estimates of value for independent direct agencies:

Table 13 Independent Direct/CRM Agencies

Direct/ CRM Agencies DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$2,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$1,000	50%
Value added by services that indirectly depend on data exchanged or rented among firms	$1,000	50%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$2,000*	100%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$-**	0%

Table 14 Independent Direct/CRM Agencies Employment

Direct/ CRM Agencies DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**8,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	6,000	75%
Employment added by services that indirectly depend on data exchanged or rented among firms	2,000	25%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*8,000*	**100%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	0	*0%*

2.5 Measurement and Analytics

The measurement and analytics role is a critical link in the data-driven marketing economy, linking data sources and channels, and reporting actionable information (sometimes in real-time) for the creation of relevant and effective campaigns at the most efficient cost. We estimate its value added to the economy as $3 billion.

The function of these firms is to collect and analyze data on marketing inputs and sales and behavior outputs to increase the effectiveness of marketing practices. While there are other companies in this study that provide some sort of measurement or analytics functions on top of their core capabilities, their value is calculated elsewhere in this study.

Typically these firms collect data from public sources, private 'panels', purchased third party sources, and from marketers' 'owned' data. They then apply a series of services including: database hosting, aggregation and appending of data, data hygiene, merging of offline data with online data, and algorithmic processing, and present the data to a marketer in actionable format. A small list of possible outcomes includes website ranking and usage by demographic information, identification of social media influencers, or purchase behavior by geographic region.

We took as our sample frame the number of measurement & analytics companies active in the U.S. performing the following sub-functions:

- Audience research as currency for media buying
- Sales and campaign performance
- Digital publisher site analytics

The companies that have been considered in-scope use offline and/or online individual-level data for the purposes of collecting, matching or appending data. Some of them share data with marketers; and, in some cases they share it an aggregated level. We include even revenue from aggregated data in our assessment because the exchange supports and, in some cases, entirely enables, the ability to collect data earlier in the data flow.

Net Value Added by Measurement and Analytics Services

Not all measurement & analytics practice areas are included in the data-driven economy. Historically broadcast television researchers combined the two functions of audience measurement and consumer research into a pair of services, which Nielsen, the dominant player, calls in its promotional material the measurement of "what people watch" and "what people buy." As natives to the offline world of broadcast audiences and advertising, their business model involved assembling opt-in panels of consumers. Monitors were used to record what panelists watched, and pantry scanning to record what they had bought. By buying the two sets of panel data together, CPG marketers and retailers in particular could attempt to correlate the former with its impact in the latter, but only at a high, aggregated level, usually no more granular than what was termed a Designated Market Area. These offline audience and market research functions are not included in the DDME economy.

Measurement and analytics firms reliant on individual-level data typically earn revenue in a variety of ways with one or more of the following models: software licensing fees,

fee based models (typically for insights), pay for performance or cost per action, audience cost per thousand or cost per data unit.

When such a firm bills a client for data services in which data are appended, we take care not to count value accounted for elsewhere in the data-driven economy. Based on interviews and analysis we conclude that the cost of these data inputs amount to 10% of revenues.

How Much Value-Added Revenue Depends on Individual-Level Data?

If the cross flow of trade were subtracted, we estimate that the revenues for measurement and analytics firms would be reduced by 50%. They would be unable to access databases with personal information identifiers, including those of the marketers who retained their services. Their services would be constrained to those companies that provide a combination of collection, insights and media buying.

Table 15 Measurement/Analytics

Measurement/ Analytics DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$3,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$1,000	33%
Value added by services that indirectly depend on data exchanged or rented among firms	$-	0%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$1,000*	33%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$2,000**	67%

Table 16 Measurement/Analytics Employment

Measurement/ Analytics DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**12,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	6,000	50%

	Employment added by services that indirectly depend on data exchanged or rented among firms	-	0%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms		*6,000*	**50%**
Remainder:			
	Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	6,000	*50%*

2.6 Future Trends

Agencies. DDME strategic services, whether performed by general agencies or direct or digital agencies, are unlikely to be a substantial source of growth for the DDME. We surmise that the decline in general agency revenues and employment will be offset by growth in digital revenues. Advertising revenue trends tending to be too volatile to be a reliable base for forecasting agency revenues, we gave more attention to employment. The pattern of employment across the entire industry, DDME and broadcast, has shown a persistent decline since 2000. Employment in 2012 was 5% below even 2007 pre-recession levels which were below mid-2000 levels.[18] We anticipate that there will be job migration from general strategic services to digital strategic services, but no significant net recruiting. Our projection is that DDME strategic agency services revenue will grow 5% in total over the next 5 years.

Measurement. According to the DMA Statistical Fact Book, measurement is 9% of an email marketer's budget.[19] Assuming email is a more mature direct marketing medium, we may expect other digital direct marketing media to approach a 9% budget spend as well. Given the growth of the industry, and the complimentary growth of measurement inside a more mature industry, we draw the conclusion that this area will show strong growth from a small base in the next five years.

Overall. We conjecture, with the caveat that our study did not build any empirical base for such a projection, that strategic services revenues will grow to roughly $19 billion over the next 5 years.

[18] Ad Age Datacenter 2012, U.S. Ad Agency Employment Report, December 31, 2012.
[19] n.a., DMA Statistical Fact Book 2013. Direct Marketing Association, April 2013.

3 Audience Assembly and Targeting

In this chapter we describe the role, dollar value, and employment created by Audience Assembly and Targeting services in the data-driven marketing economy. We have identified three types of firms within the DDME; 1) those that focus on digital audience assembly in various channels, 2) those that derive revenues from search audience assembly, 3) and those that focus on audience targeting. The services that make up Audience Assembly and Targeting generate a total value of $37 billion to the data-driven economy and employ 162,000 people.

Audiences vs. Customers

Our report makes a distinction between two apparently quite different ways to use individual-level data in marketing: a $52 billion business to reach customers, and a $37 billion business to reach audiences. In the first instance firms compile lists and a producer of goods or services decides which members of a list it wants to target. In the second instance a publisher assembles an audience and a producer of goods or services decides which members of the audience to target.

Customer targeting reaches a person or pseudonym at a fixed address; a postal address or an email address, for example. Audience targeting reaches a person or pseudonym on the move; in the relatively brief moments when they are visiting a website, watching video, or participating in social media. Customer targeting has been part of consumer experience for as long as there has been postal delivery, and commercially important since the late 1800s, and it has grown to $52 billion of value creation by slow refinement of its methods. It is a culturally familiar marketing practice. As new technologies have created new addresses, for example the telephone, email, and mobile phones, they too became part of the consumer experience, begrudged when intrusive and yet welcomed when relevant. Consumers grasp their costs and benefits.

Audience targeting, or using individual-level data to target an audience member on the move across digital content sites is a very recent practice, beginning in the 1990s and, in its more sophisticated form, no more than a decade old. Yet already it accounts for $37 billion of value. The idea of being targeted on the move is far from culturally familiar, but it uses data of the same kind as is used for customer targeting and for similar purposes, and it is not unrealistic to predict that it may follow the same path to cultural normalization that customer targeting followed. It is the topic of this chapter.

3.1 Assembling and Targeting an Audience

This report is not concerned with the kind of publisher who assembles an audience for broadcast advertising because (with a minor exception that we deal with in section 5.5) broadcast publishers use no data about individuals for marketing. The report is concerned in section 5.2 with the value that digital publishers generate by assembling content whose impact on individuals can be tracked after the individual has been exposed, and in section 5.3 with compilers of indexes that attract individual searchers where again impact on

individuals can be tracked. In both these instances advertising is matched to publisher content, and the only role of individual data is to decide if media purchases are producing results. In section 5.4 we deal with advertising that is selectively exposed to individuals whose profiles are attractive to advertisers.

3.2 Digital Audience Assembly

Digital Audience Assembly refers to the actions of firms who bring together people to read or view the content they publish, attracting audiences that in turn can be sold to advertisers. We estimate that this sector contributes a net $14 billion of value to the data-driven economy, and employs 62,000 people.

Display, Mobile and Social Advertising Publishers

We identified a group of large online publishers consisting of Google, Facebook, Yahoo, Microsoft MSN, AOL and Twitter, whose display advertising revenues are reported to be $7.1 billion[20] and about which, because these firms are all public companies, there exists substantial public information. We enumerated and analyzed another 36 smaller publishers whose revenues from display advertising totaled $5.5 billion[21,] and created a category for all other publishers to account for the $2.5 billion discrepancy from the reported 2012 digital display advertising total of $15 billion[22.] We identified that a small amount of revenue, which we estimated at less than $1 billion, is attributable to promotions to publisher audiences that are not captured in our analysis either in advertising billings nor in advertising agency revenues because it is created within the publisher itself. This kind of activity may be larger, but it is not easy to observe. Consequently our estimate is conservative.

Publisher ad revenues are a component of the data-driven marketing economy for two reasons. Some of the advertising is targeted to individuals or pseudonymized individuals, as we shall discuss in the Audience Targeting section. Second, responses can be linked to particular individuals or pseudonymized addresses.

Publishers benefit from data bought from other members of the DDME, but most is acquired by audience targeting firms, and is accounted for in section 5.4.

To what extent does the value created in this sector require individual-level data to be shared, exchanged or sold across firm boundaries? First, there is some data exchange when display advertising is bought and sold, not in direct negotiation between the sales forces of publishers and buyers in the media affiliates of advertising agencies, but in programmatic buying on platforms of the kind described in section 5.4. This is inevitable because programmatic buying, even when not used for behavioral targeting, operates by collaboration among specialized entities. Second there is data exchange when third party

[20] IAB/PricewaterhouseCoopers Internet Advertising Revenue Report for 2012 and eMarketer March 29, 2013, "Google, Facebook Continue to Lead in Digital Display Earnings."
[21] Drawn from Hoovers, Global OneSource, Crunchbase, and company pages on LinkedIn.
[22] IAB/PricewaterhouseCoopers Internet Advertising Revenue Report, 2012.

measurement firms develop models to estimate the efficiency of ad purchases. That said, the very largest of digital publishers do much less data exchanging than do the smaller publishers, and in the main do no data purchasing or selling at all.

We estimated, for each tier of the digital publishing industry, what proportion of revenues relied on data exchange among firms, and how much was completely independent of such exchange. The conclusions are set out in the table that follows. The display advertising sector is much more sensitive to data exchange than the search advertising sector, because there are far fewer firms in the latter sector and the revenues are more concentrated in firms with enough scale to be highly self-sufficient.

Table 17 Digital Audience Assembly

Digital Audience Assembly DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$14,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$7,000	50%
Value added by services that indirectly depend on data exchanged or rented among firms	$4,000	29%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$11,000*	79%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$3,000**	21%

Table 18 Digital Audience Assembly Employees

Digital Audience Assembly DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**62,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	31,000	50%
Employment added by services that indirectly depend on data exchanged or rented among firms	17,000	27%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*48,000*	**77%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	14,000	*23%*

3.3 Search Audience Assembly

Search engines are viewed in this study as a category of audience assembler or publisher. Instead of publishing content of their own making, they give access to content from a multitude of sources and sell advertising against the results of searches.

We find that this sector contributes $19 billion to the value of the data-driven economy, comprising $17.7 billion on display[23] and $2.0 billion on mobile media. It accounts for 82,000 employees.

The gross expenditure on audience assembly for search requires no reimbursable expenses to be deducted from it, and nor, based on our interviews, do we believe it requires purchase of data credited elsewhere to the data-driven economy. Therefore gross expenditure can be treated as net expenditure.

However we make some allowance for cross-firm data trading. We have seen that display advertising relies substantially on individual-level data traded among firms in the audience targeting sector. Search advertising is much less dependent on this kind of traded or exchanged data. Search results, and the advertising that wins the auction to appear adjacent to search results, can be generated without any such cross-trading at all. However our analysis concludes that we need to make a small allowance for several reasons.

First, search engines are encouraging users to log in before searching. While search results for large search engines can be usefully improved by audience profiles without turning to cross-traded data, that is not true for smaller and specialized search publishers. Second, search companies service Affiliate Marketers, who bid for keywords, and receive money from marketers upon successful sales. In an environment where cross-trade is hindered, an Affiliate Marketer would not be able to act on behalf of a marketer or ecommerce company. Affiliate Marketers spend $4 billion in the U.S. across various advertising channels, of which search is the primary channel. Third, services bundled with third party search analytics form part of our net expenditure. Similarly our net expenditure includes value generated using reconciliation and auditing services between the marketer and the search engine.

On the other hand, in deciding that the allowance should be small we recognize that the set of firms that make up the search audience assembly sector is far smaller than the network that makes up display audience assembly, and the largest firm, with a 75% market share, is quite large enough to be independent of the need to rely on third parties to supply individual-level data.

Consequently we assume that 10% of search audience assembly relies on data that is traded or exchanged among firms in the data-driven economy. Our conclusion is:

---

[23] IAB/PricewaterhouseCoopers Internet Advertising Revenue Report for 2012 and eMarketer_Microsoft_Takes_a_Small_Share_of_US_Search_Ad_Revenues_1009939.pdf

Table 19 Search Audience Assembly

Search Audience Assembly DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$19,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$2,000	11%
Value added by services that indirectly depend on data exchanged or rented among firms	$2,000	11%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$4,000*	21%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$15,000**	79%

Table 20 Search Audience Assembly Employees

Search Audience Assembly DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**82,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	8,000	10%
Employment added by services that indirectly depend on data exchanged or rented among firms	9,000	11%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*17,000*	**21%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	66,000	*80%*

3.4 Audience Targeting

In this section we analyze the audience targeting sector of the data-driven marketing economy, and reach the conclusion that $4 billion of incremental value is generated and 18,000 people are employed. The value comes from eliminating waste in matching sellers to buyers on online publications. This sector is intensely entrepreneurial. We project it to

grow substantially, possibly to double, over the next five years if the proportion of online advertising that relies on behavioral targeting grows in line with recent growth.

The purpose of this section of the report is to develop our projection of the size of the sector. In the process it will explain the birth of customer targeting, how it has evolved to where it is today, and the direction of likely evolution. Although our report values only the U.S. contribution of audience targeting, this sector is the locus of intense technological innovation paying out on a global scale, and some of those who work in research and development are in effect generating export revenues.

Evolution of Audience Targeting

The goal of audience targeting is to put digital advertising in front of only the small proportion of consumers who are very likely to respond to it. This proportion is remarkably small: a coupon distributor, drawing on three-year individual shopper purchase histories, reports that the average supermarket product draws 80% of its sales from only 2.5% of the supermarket's customers[24]. The gain in marketing efficiency from avoiding spillover spending onto uninterested consumers is substantial.

From the birth of the Internet as a marketing environment in the mid-1990s until about 2003, most advertising that appeared on a publisher's website targeted in the manner of broadcast advertising, without regard to data on individuals in the site's audience. As magazine advertising buyers do, the first generation of online advertising media buyers paid attention to the fit between the site's content and the advertiser's target customer. Advertisers of financial services, for example, bought the right to serve their ads on financial news sites. And as in magazine advertising buying, deals were usually struck face-to-face between advertising agency buyers and salespeople representing the larger publishers. The so-called banner advertising of this era did not perform well by comparison with print and television advertising[25] and from 2000 to 2003 U.S. sales of digital display advertising declined from $6.6 billion to $3.2 billion.[26]

New technologies began to transform the medium's attractiveness to advertisers. From the low in 2003 to 2013, display advertising has grown every year to $11.8 billion in 2013, and is forecast to continue to grow, both by taking share from broadcast media and by attracting new firms to advertising.

The technology stimuli took several forms. In 2003 Google's Adsense was launched to automate the process of matching publisher context to advertiser's message, allowing the market to expand to all online publications, including blogs and other properties previously too small to be able to support a face-to-face sales force. Many small publishers and hobbyists, for whom a web presence had been a cost, found that it became

[24] "Catalina In the Digital Age," Harvard Business School Case N9-514-021 August 6, 2013.
[25] Morrissey, Brian. "How the Banner Ad Was Born." Digiday, 12 April 2013. http://digiday.com/agencies/how-the-banner-ad-was-born/
[26] http://www.statista.com.ezp-prod1.hul.harvard.edu/printstat/271500/

a source of profit. Next so-called real time bidding was introduced to allow buyers to bid on online auction for space on sites that were rated as delivering specific contextual cues, bringing efficiency, transparency and scale to the market for digital advertising. The rise of real time bidding was concurrent with a rise in ad networks, arbitragers who bought unsold publisher media at low prices and sold it at higher (but less than full) prices. But until this point, while individual-level audience data were used to judge advertising performance after it had been bought, the data played no part in the buying process.

Individual-level data began to be used to buy advertising in about 2007, when the display market was about $6 billion. It contributed to a doubling of the value of display revenues to publishers in the next 6 years. The trigger was the introduction into real time bidding of a wave of technology termed behavioral targeting. Although behavioral targeting technology by 2012 was used in only 13% of all online display advertising, it was expected to be at 19% one year later.[27]

Within the last year a new form of behavioral targeting, termed retargeting, has been introduced. This is a practice in which past behavior, such as visiting an advertiser's website or showing interest in an ad on one publisher's website, leads a person's browser to be tagged with a cookie so that when the browser visits another publisher's website it can be sent another ad impression.

This brief history makes clear that the pace of innovation during the last 20 years has driven material value creation for small publishers as well as large, and has placed some of the power of advertising, once available only to large firms, in the hands of small firms and entrepreneurs. The pace is not slowing, and new sources of efficiency are likely to be found, but the report will focus on this moment in the trajectory.

Behavioral Data for Audience Targeting

In behavioral targeting, five new kinds of firm play a part in matching producers of goods and services on one hand, to individuals visiting the sites of publishers on the other. They are Demand Side Platforms (DSPs), Supply Side Platforms (SSPs), Data Management Systems (DMSs), Behavioral Data Providers, and Ad Exchanges. (We shall treat them as five discrete kinds of firm, which they often are, but some firms combine functions.) The process of behavioral audience targeting works as follows, typically taking less than half a second to complete.

When a person visits a publisher website (or more precisely when the person's browser calls up a page on the website,) the publisher's web server sends code to the browser instructing it how to assemble and format the page on the person's computer. The code often includes at least one tag inviting advertisers to insert an ad into the page. This tag calls to an advertising server such as Google's Doubleclick or Yahoo's Right Media.

[27] "Share of RTB in Digital Display Ad Spend in the U.S. from 2012 TO 2017" eMarketer Report August 2013 retrieved from Statista September 2013.

Until this point, the process is the same for ads sold with and without individual-level audience data.

When the publisher invites real-time bidding for the right to insert the ad, the ad server calls to an SSP selected by the publisher. (Some ad servers integrate the SSP function.) The SSP calls to one or more Ad Exchanges, which act as auctioneers. Information on the lot to be bid on includes the identity of the publisher and the ad slot dimensions. If the SSP has previously placed a cookie on the person's browser, that information will be included, introducing pseudonymized data into the process for the first time.

The Ad Exchange makes multiple bid requests to DSPs, who represent advertisers who might want to reach the person behind the browser. The amount of the winning bid depends on the value to each advertiser of serving an impression to that browser. The value depends, in turn, on what the advertiser can infer about the person behind the browser and the fit with the advertiser's message.

The pool of information available to make that inference and inform DSP bids is assembled by what we are calling Behavioral Data Providers. They function as aggregators of individual-level data gathered from a variety of sources including browsing histories, and matched to browsers. They use DMP platforms to distribute the data to DSPs. Typically DMPS do not buy data; they analyze and process it, and deploy it to DSPs on software-as-a-service platforms for which they charge license fees.

The identity of the winning advertiser is sent to the SSP and on to the ad server, which serves the ad to the browser of the person where it is assembled into the image of the publisher's website, and generates an invoice for the advertiser to pay the publisher of the website page that attracted the person. Complete assembly of the page and billing occurs within a half second.

Thus individual data, assembled from a number of sources and consequently crossing firm boundaries, determines the value that a publisher receives for the effort of content creation, and determines that the advertiser who most values the prospective customer gets to pay for the opportunity to advertise to that prospect.

If this method maximizes publisher value and minimizes advertiser waste, why is not all online display advertising sold at auction and why does it not all rely on behavioral data? Indeed almost all online search advertising is sold at auction. The answer is that the quality of behavioral data to support valuation of an opportunity to reach a browser is not (or is not yet) as high in display as in search, nor is it as likely to be confirmed by an action such as clicking on the ad, since not all display advertising aims to induce an immediate sale. Its purpose may be to make the prospect aware of a brand, or communicate information for future use. A consumer's search query is better evidence of the prospect's value than information gathered from browsing, and the confirmatory value of a click allows the advertiser to correlate the query to action. Enhanced data, and the ability to track what a pseudonym does after it has been exposed to an ad, even to the point of making a transaction, would increase the efficiency of display ad buying.

A limited amount of data enhanced in this manner is beginning to become available[28]to advertisers. The essential step is to link a pseudonym in a digital medium (where an ad is to be targeted) to an offline pseudonym (where the transaction will occur.) A number of firms offer what is termed a match service to make the link. These firms pay websites on which people register with full personally identifiable information to allow them to create a match between online and offline pseudonyms (for example a cookie on a browser and a frequent shopper transaction record or automobile ownership record.) Industry reports suggest that match rates are not high,[29] and certainly not high enough at present to be the foundation for performance-based ad buying of the kind that has propelled the growth of search advertising,

Valuing Audience Targeting

We identified about 125 firms that played roles in online audience targeting. We classified some as operating across all digital media, but most specialized in one medium or another. The media on which display ads are targeted are display, video, mobile and social.

We sought incremental value, not total value. Because targeting enhances the value of a publisher's inventory, and that value is captured in the revenues of the publisher, it was important not to simply total the gross revenues of these 125 firms. Only the value of behaviorally targeted ad insertions over contextually targeted value is credited to the 125 firms. Relatedly, some targeting firms report in their revenues the sum of all media bought from publishers, even when it is reimbursed by advertisers, while some report only fees. Assumptions were made for each firm's accounting practice and the role of each firm in the value delivery process to get at the increment. Our estimate is that only 45% of the $5.0 billion of reported revenues of these firms are fees net of reimbursable media purchases. Further, based on interviews, we conclude that the firms buy data to a value of $300 million that shows in our report as revenue elsewhere in the data-driven economy, so this amount was deducted, giving a net value added amount of $1.9 billion.

Almost all of this value is created by sophisticated merging of data from many sources. Very little of the work could be performed without cross-firm data exchange among the 125 firms, most quite small (median revenue per firm was $ 4 million, though the largest reported $400 million.) Therefore we attributed all but an incremental amount to direct inter-firm exchange.

The conclusion for this sector is as follows:

[28] https://www.facebook.com/notes/facebook-and-privacy/relevant-ads-that-protect-your-privacy/457827624267125
[29] http://www.adopsinsider.com/online-ad-measurement-tracking/data-management-platforms/data-management-part-iv-syncing-offline-data-to-your-dmp/

Table 21 Audience Targeting

Audience Targeting DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$4,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$4,000	100%
Value added by services that indirectly depend on data exchanged or rented among firms	$-	0%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$4,000*	100%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$-**	0%

Table 22 Audience Targeting Employment

Audience Targeting DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**18,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	17,000	94%
Employment added by services that indirectly depend on data exchanged or rented among firms	1,000	6%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*18,000*	**100%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	0	*0%*

3.5 Broadcast Media

Print publishers possess personally identifiable data on each of their subscribers because they need names and addresses to fulfill subscriptions. Consistent with the principles outlined in Chapter 2 of this report, revenue that they earn by using these data form part of the data-driven marketing economy. However our analysis suggests that these revenues are small today. They come from two main sources.

First, publishers rent their subscriber lists to data brokers. While subscriber lists are important sources of personal data for the rest of the economy, interviews with publishers suggest that, at an average rental price of $70 per thousand names per rental opportunity they do not contribute materially to the value of the data-driven economy.

Second, some publishers produce niche editions of print publications directed at particular individual subscribers who, based on data appended from third-party sources, are valued by particular advertisers. They are able to do so because variable data printing equipment can adapt print products with something of the flexibility that digital products can achieve. Our interviews suggest that the marginal contribution of niche editions to publisher advertising revenues is not material to our study's results given the cost today of variable data printing.

In the broadcast television sector, we found evidence that advertising agency media departments were influenced in the selection of broadcast media programming by individual level consumer data, particularly from social media. Twitter, for example, has access to comments on programming and advertisements from an audience of 40 million people who simultaneously watch television and comment on social media. We did not separately credit value from these data to networks because they are captured in media prices.

Cable and satellite television distributors, like magazine publishers, possess personally identifiable data on each of their subscribers. They have the potential to target advertisements to particular households, and targeting of households can in principle be informed by online and offline behavior just as it is in digital display advertising. Our investigations found evidence of targeting to households defined by demographic profiles, but only small-scale tests of targeting to households defined by individual-level data.

Substantial growth can be expected in the size of the data-driven economy within a two to three year horizon as broadcast media solve the technical impediments to becoming individually addressable media.

3.6 Future Trends

In addition to our expectation that Audience Assembly and Targeting will become culturally familiar as it matures, we also certainly expect that this sector will grow as a practice, and its share of the DDME will grow at an accelerated rate.

While still barely a decade old, companies in the Audience Assembly and Targeting sector command 28% of the total DDME Economy. Even now, the relatively mature technologies of Digital Advertising and Search Advertising are growing at 11%[30] year over year, while nascent marketing technologies like Mobile Advertising and Social Media Advertising have growth trajectories that reach 20% and 30% respectively, year

[30] N.a., DMA Statistical Fact Book 2013. Direct Marketing Association, April 2013.

over year in the next three years[31]. Companies that collect and target audiences in even more nascent markets like wearable technology, implicitly captured in this sector, will accelerate growth even further, such that we expect the DDME Audience Assembly and Targeting revenue, as a whole, to more than double over the next 5 years.

[31] N.a., DMA Statistical Fact Book 2013. Direct Marketing Association, April 2013.

4 Prospect / Customer Relationship Marketing (P/CRM)

As discussed in the previous section, customers are defined as a person or pseudonym at a fixed address. In this chapter we describe the role, dollar value, and employment for the following classifications of firms that specifically focus on targeting customers: 1) list and database service providers, 2) postal media 3) email marketing service providers 4) telephone sales production and delivery, and 5) mobile CRM platform providers and developers. The services that make up Audience Assembly and Targeting generate value of $52 billion to the data-driven economy and employ 228,000 people.

From the domain of digital audience assembly and targeting, with Pseudonymous Information, we turn to the world of classic prospect and customer media. This is where the collecting, analysis and targeting of personally identifiable information as the foundation of producers' go-to-market strategy occurs. If audience assembly and targeting is nascent, dynamic, and growing, it is doing so from a comparatively small base. Customer relationship marketing, on the other hand, is mature and stable, resulting in significant net value-added revenues and employment.

As noted in earlier chapters, the classic media of PII are those where a name can be associated with the address appropriate for that media. For example, a name residing at a street address, an email *handle@domain*, a person reachable on a landline phone number, or a mobile number for texting. (We exclude mobile phone numbers used for voice transmissions because legal restrictions apply.) Emergent within this category are producers' proprietary mobile apps, since once the apps are downloaded, they can become the basis of direct one to one interactions with the customer, and opt-in SMS (text) messaging.

Within this broad functional category, we can distinguish two sub-functions. First, we describe a set of third-party firms specialized in the market for PII-based data as goods and services, that we call List and Database Services Providers. Our second set of data service providers specialize in producing and delivering marketing communications for specific prospect and customer-relationship media. Because these media represent such a large proportion of the total DDME, we treat the firms supporting each medium separately, as follows: 1) direct mail and catalog production and delivery providers (the postal medium); 2) email marketing service providers; 3) teleservices bureaus; and 4) mobile marketing platforms and apps.

4.1 List and Database Service Providers

The list and database marketing services segment comprises two distinct subsystems of players instrumental to making the market for PII-based ILCD: 1) List brokers and managers; and 2) database marketing input providers, often called data brokers. Note that this is a potentially problematic term that should be used with caution and will therefore be qualified later in this chapter.

List Managers and Brokers. This industry has a long pedigree as an ancillary service to direct mail and catalog marketing. It emerged when owners of ILCD realized that the source of a list of names contained more predictive power than the names and addresses themselves. After all, anyone could get a list of names and addresses from a telephone book; but a list of names and addresses of subscribers to a golf magazine has predictive value for producers of golf-equipment, and vice versa. Further, marketers realized that subscriber data from magazine publishers and customer data (from direct-order retailers) had the added advantage of being behavioral and transactional. In other words, it could reveal what people actually buy, rather than what they said they would buy. Such data was therefore found to be a better predictor of future purchase propensity. The list market expanded to include the other PII-based media once it was realized they functioned the same way. Specialized firms for list markets for specific media exist, but are relatively rare, since most users of the list market (producers) seek a multi-channel approach to reaching customers and prospects.

The list marketplace comprises three parties: list brokers, list managers, and list owners. List brokers are third-party data servicers whose role is to select and license lists of names and addresses or occasionally ILCD data appends on behalf of users (normally producers). Brokers receive a standard 20% commission, paid by the list user, out of the total amount the user pays for the rental. List managers promote list owners' lists to interested users or their brokers, and negotiate fees and terms of use on the list owner's behalf. They receive an industry standard 10% commission, paid by the list owner (i.e. their clients), but calculated on the total rental price paid by the list user. Thus, the remaining 70% of the total price of the licensed list is received by the list owner as rental income.

Value-added revenues to the DDME from list brokers and managers. Our sample frame consists of several hundred list brokers and managers on the Direct Marketing Association membership rolls, together with several hundred others contained in its prospect file. We assume that others exist outside the Direct Marketing Association's frame of reference. Our desk research into a sample of the Direct Marketing Association members suggested that the great majority of brokers and managers were fairly small, often having net revenues of below $10 million annually.

To calculate total revenue for all brokers and managers, we factored in an estimate of the total demand for lists (volume rented times average price) weighted across the P /CRM media, then set the cap for broker and manager revenues at 30% of this value (since 70% of the fees paid by list users goes to the list owner.)

We can estimate demand for lists most clearly in the market for postal and email names, thanks in part to annual price indices provided by Worlddata to listpriceindex.com.[32] Depending on the type and source of the list, postal names average about $100 - $110

[32] "2013 Trends Revealed in the Worldata Summer List Price Index." ListPriceIndex.com, August 2013. http://www.listpriceindex.com/currentlpi.htm

cost per thousand, and the least expensive lists appears to be $70 - $75.[33] Email lists vary more widely in their price, especially for business-to-business email lists, where lists of opted-in prospects for very high-ticket purchases (such as IT infrastructure) can run to many hundreds of dollars per thousand names. Rates for teleservices lists are less transparent, but we assume that pricing falls within the same range as postal and email lists.

Since list rentals are for one-time interaction, rental volume can be estimated based on the 2012 delivery volumes from the United States Postal Service's (USPS) for mail pieces attributed to list rentals.[34] This estimate requires examining the proportion of direct mail that is individually addressable, then determining which portion of it is for customer acquisition. Within this amount, the volume of list-rental customer acquisition mailings must be disaggregated from mailings based on owners' opted-in prospect databases, as well as mailings that rely on the syndicated databases of data-driven marketing input providers (see below). Information on the proportion of direct mail that is individually addressable can be gleaned from the USPS *Revenue, Piece, and Weight Report* for 2012. Research firms such as Mintel collect statistics on the distribution of direct mail through various types of acquisition and retention (house file) campaigns.[35]

Cost of Data Goods Sold / Determining Added Value. Several features of the list market make it difficult to measure added value. First, list users (fee-paying lessees) are often also list owners (fee-receiver lessors). Calculating the net of what a producer or agency rents from what they earn risks having the net sum to zero – a result that if taken at face value would paradoxically suggest there was no list market at all. Further, sellers (in our terms, data owners) can be compensated other than by money, i.e. they can pool or swap lists with other list owners, or can barter their lists for goods in kind, often advertising. In our calculations we assume conservatively that such non-monetary transactions are rare and include only a small increment for this value.

Though neither list brokers nor list managers are owners or users of data, their business model is one hundred percent indirectly dependent on servicing the exchange of data between owners and users. List brokers and managers are generally media- and data-agnostic across P/CRM channels.

Database Marketing Service Providers (DMSPs)

---

[33] These are base values for the names and addresses. However, users often seek to increase list response rates by filtering for specific attributes: age range, household income, gender, education, ethnicity, or special interests (for example, golf enthusiast.)

[34] To protect the intellectual property of the original list owner, rental agreements typically authorize a one- time use only; only if the prospect responds to the user (in other words, establishes a direct relationship with the new user by making a purchase, an inquiry, etc.) does the user acquire outright ownership of the name. To ensure the list is used only one time, list managers impose very strict terms and conditions on list users. For example, the copy of the rented list may not be merged or installed with the user's database but can only be serviced as part of campaign production by trusted third-party data servicers, such as database marketing service providers or printers who possess the list in escrow, then destroy or return it to the owner after its use. Rented lists are also normally pre-seeded with fictional names so that the campaign material can be inspected for compliance with the rental agreement.

[35] See the DMA 2013 Statistical Fact Book, 2013.

Database marketing service suppliers differ from list brokers and managers in that they are (third-party) owners of PII that they market upstream to users such as other DMSPs, producers, retailers, publishers and agencies. As their business involves owning data, not just servicing it, we avoid using the term data broker. However, DMSPs can also earn a significant portion of their revenue as third-party data servicers. Sometimes called data aggregators or data compilers, their business model involves a data-owner facing role, and a data-user facing role. In the first role, DMSPs begin as data users who license ILCD (mostly PII) from a vast array of first party owners. Then, they combine (through compilation or aggregation) the data into unified databases, matching inputs by name and contact address across multiple media. It is this data processing that makes them third-party data owners.

In their user-facing role, DMSPs license access to their proprietary databases on a syndicated basis (i.e., grant access to multiple subscribers at one time) to agencies, producers, and publishers. The fees often permit unlimited use of their data for extended license periods. Some of the data goods they provide are demographically segmented lists that can be used exclusively for acquisition campaigns. Data-appending is the other type of data-as-marketable-good that DMSPs provide. Appending extend the number of fields per record on users' pre-existing house file or proprietary prospect databases. DMSPs can also append data in the form of address in different media to make users databases more multichannel.

Apart from offering data goods to users, DMSPs also provide contact data management services, such as helping marketers identify and integrate customer or prospect or audience records across media. Such services often involve performing an analysis to identify matching contact or attribute data in their proprietary databases, and then appending matched data to a user's database. In this case, the DMSP derives income for both providing the data as a good and for performing a service on it, though the price of the data goods may be bundled into the fee for the service. DMSPs can also provide advisory or analytic services in which they are not involved as data owners, since their expertise in database management can be used to build and host clients' proprietary (owner) databases. They can also perform analytic processes necessary for effective segmentation, modeling, and scoring for such data-owners.

Contribution to the Net Value Added of the DDME. Pricing for DMSPs is usually an annual or monthly license fee for unlimited or scope-specified access to the syndicated database for the duration. A flat fee is charged for a general installation of pure prospect records (new names with contact info and/or multiple attribute fields of demographic, interest, or purchase propensity data, or appended data) to user-owned databases. There is then an additional per-unit charge for certain additional or specified variables. Our interviews led us to believe that once adjustments are made for differences in scale and

scope of activity, prices for licensing and installation of data goods are broadly consistent with those charged in the list rental market.[36]

Cost of Data Goods Sold. A considerable portion of the data that DMSPs provide to users must be frequently licensed and refreshed from original owners. DMSPs may compensate some of the first-party-data owners from whom they obtain ILDC using a revenue share model. Though arrangements vary depending on the data transferred to the end data user, as do the terms of license between DMSPs and original owner of their data goods, DMSPs generally build in the cost of royalties payable back to the original data owner, plus overhead and markup, within the base licensing fees charged. Occasionally DMSPs may charge royalties directly to the data user, though we assume this happens only for a handful of their largest-volume subscribers.

Distribution by Dependence on the transfer of data across firm boundaries. Almost all of the net value added by DMSPs depends on the marketed transfer of data-goods across firm boundaries. We classify all of DMSPs' licensing and appending of data inputs as market-dependent on the demand side. On the supply side, the vast bulk of their segmentation, analytics, modeling and scoring services involves using data they themselves licensed, or involves servicing data licensed by clients. Thus, their servicing revenues are also entirely dependent on the marketplace for data. Only the revenues derived from servicing owned data for retention campaigns—e.g., postal production, database design or hosting, or email deployment—can be seen as not dependent on the data market.

Table 23 List Services and Database Marketing Input Providers

List Services and Database Marketing Input Providers DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$7,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$3,000	43%
Value added by services that indirectly depend on data exchanged or rented among firms	$4,000	57%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$7,000*	100%
Remainder:		

[36] Prepackaged segmentations (e.g., "urban sophisticates,", "empty nesters") for acquisition campaigns are similar to list rentals except that profiling attributes or dimensions are more extensive, are more demographically categorical and less transactional, and are pre-selected by the DMSP. We estimate the inclusion of dimension increases the licensing fee by the equivalent of about $5 CPM per dimension for pre-assembled segmentations involving at least ten major dimensions, such as age bracket, gender, ethnicity, etc. This would likely translate into per mailing cost of a targeted campaign of about $150 - $160 CPM, or fifteen to sixteen cents per individual name.

	Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$-**	0%

Table 24 List Services and Database Marketing Input Providers Employment

List Services and Database Marketing Input Providers DDME Value-Added Employment and Dependence on Data-Exchange		Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues		**31,000**	**100%**
	Employment added by services that depend directly on data exchanged or rented among firms	12,000	39%
	Employment added by services that indirectly depend on data exchanged or rented among firms	18,000	58%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms		*30,000*	**97%**
Remainder:			
	Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	1,000	*3%*

4.2 Postal Media: Direct Mail & Catalog Producers and Deliverers

In the individually addressed, direct response advertising mail and catalog production and delivery subsystem, we find a quasi-monopoly, comprised of the USPS and a very large number of firms involved in direct mail production processes. These latter firms are chiefly printers, letter shops, and computer services providers.

This subsystem proves to be the largest component of the data-driven marketing economy by a wide margin. This may seem counterintuitive for those who automatically equate data with digits, not with paper. But this can be explained by considering that individual level communications requires first and foremost a delivery network capable of sorting and routing millions of individual messages accurately at a cost-efficient scale. In the world of digital marketing, sorting and routing is provided by the servers, routers, and cable connections of the Internet backbone; however, marketers (producers) do not pay to use this backbone.

The equivalent of the Internet sorting and delivery backbone in the offline world is the United States Postal Service. Instead of routing zeroes and ones through cables, the USPS sorts and carries ILCD-addressed mail using a network of plants, equipment, vehicles and personnel. However, unlike the Internet, firms using postal as their go-to-market strategy must pay for this service directly on a CPM basis. This CPM is vastly more expensive than other media: $260 or thereabouts, depending of the specific type of delivery service selected. The production costs are equally as expensive: paper, ink,

envelopes, binding of catalogs, full-color photography. By comparison, the cost of assembling digits on a screen and pressing the send key can be a low-budget activity.

Given this price differential, the question really becomes, why does postal remain such a large part of the DDME? In other words, why haven't producers diverted all of their marketing dollars into digital media, which is a more streamlined process, with cheaper CPMs?[37] The answer requires looking deeper into the accounting practices that shape the economics of media for marketing purposes.

Essentially, P/CRM marketers are able to look beyond the cost of producing and delivering the media, and look at what their expenditure buys them in terms of sales. This means measuring campaign response rates, conversion rates, average order sizes, repeat purchases, etc. In postal, these measurable metrics are much higher than they are in digital, often by orders of magnitude. Because they are all highly predictable, and because mail can be purchased in precise quantities to target precise sets of likely buyers, producers spending substantial sums on a specific amount of postal media can be fairly confident to still turn a profit.

Net Value Added to DDME By Postal Delivery and Production

Since all firms involved in the postal production and delivery process charge their customers on a CPM basis, we feel we can overlook differences that might otherwise be important, and treat them as one for purposes of calculating their total contribution to the DDME.

We begin by determining USPS value-added revenues for sorting and delivering individually addressable direct mail.[38] The USPS' 2012 *Revenue, Pieces, and Weight Report* provides a detailed breakout of all USPS revenues by product format, volume, and price of service performed. Direct mail revenues principally fall within two USPS delivery rate categories: First Class Mail (FCM) and Standard Mail A and B (SMA and SMB, with B the rate for catalogs.) We assume no more than 10% of FCM by volume may be regarded as advertising mail, and none of such mail needs to be excluded as non-individually addressed, or broadcast ad mail. (USPS requires all FCM to be individually addressed.)

Standard Mail is where most advertising mail is found. Within Standard Mail there are numerous sub-classifications, based in part of degree of addressability, which is one of our principal criteria for DDM. We include all revenues for standard mail and catalog rate classifications where individual addressability of the ad mail is required (this

[37] The answer to this question is more complex, since it involves a more detailed accounting of digital media's metrics for response, conversion, average order size, etc., than we can give here.

[38] The USPS earns a small amount of revenues through a profit sharing arrangement in the form of a marketing program to consumers associated with its change of address service. We assume this amount to be less than $10 million, and disregard it. Separately, the USPS also receives fees for doing similar data processing work for subscriber periodicals and package delivery. These are discussed in other sections of this report.

excludes all sub-classifications with revenue attributable to "saturation" mail, which does not incorporate ILCD in its addressing.) Revenues for these individually addressed classes show USPS net value-added contribution to the DDME in the amount of $17.053 billion.

The next large contribution to the value added by postal production within the DDME comes from the U.S. printing industry. To better understand this contribution, we identified a range of firms at least partly involved in the printing of direct mail and catalog pieces. We then took their total U.S. revenues as reported by public sources at $22.2 billion, or about 28.6% of the entire U.S. printing industry (by value). The value of all other firms in the printing sector is estimated at $55 billion.[39] However, revenues from printing direct mail and catalog accounts for only a portion of printer revenues. Many lines of revenue, such as printing of business mail for bill presentment and payment, flyers, corporate reports, books, and services to consumers fall outside the scope of the DDME, and must be excluded.[40] Thus, once adjustments are made, we estimate that the value of the direct mail and catalog portion of total printer revenues amounts to 15.5%, which translates into about $12.03 billion.

Finally, computer service bureaus and letter shops account for the remainder of the net value-added in the postal production and delivery process, at $3.13 billion and $1.09 billion, respectively. They provide the crucial connective tissue between printing and insertion into the delivery process.

Letter shop services focus on post-printing production tasks that include high speed matching, folding and inserting of mail into envelopes, labeling of envelopes (given that the mail piece is typically personalized by name, and the envelope address must match), binding processes (for catalogs and other large mail formats), card affixing, gluing, stamp affixing, and pre-sorting and delivery into the postal service sorting stream.

Computer services bureaus specialize in performing services that are closer to the delivery process than the production process. Their original focus was on address hygiene services to improve postal audience targeting. Hygiene involves address consolidation, correction and standardization, and appending to meet USPS requirements. Processes include address validation, correction, and/or standardization, adding of address elements such as Zip+4, Carrier Route Codes, Line-of-Travel data, delivery point bar-coding, identifying of undeliverable and non-forwardable addresses. Computer service bureaus may also perform suppression processes to comply with 'Do Not Mail' lists or other best practice obligations.

Deductions for Data Costs. On the delivery side, the USPS incurs enormous input costs including employee wages, benefits, and pensions, vehicle and fuel purchases, and capital

[39] According to an IBISWorld report of August, 2013, the total gross revenues of the U.S. printing industry were $77.7 billion.

[40] We also exclude revenues from printing of magazines from postal production. The contribution of certain classes of magazines to the DDME is discussed earlier, in chapters three and five.

investments in offices, plant and equipment, etc. However, we were not able to locate within the USPS any significant pass-through purchases of media, nor of ILCD goods or processing within the DDME data marketplace, so we do not subtract any monies for cost of data goods or data services sold. Printers, letter shops and computer services bureaus are intensive servicers of data, but only computer services bureaus appear to make any significant data purchases that become revenues elsewhere in the DDME.

Distribution by Dependence on Marketing of Data Across Firm Boundaries.

All three sets of firms involved in postal production can be regarded principally as data servicers, since their primary role is to execute list-based campaigns planned and paid for by data owners and third party users. Computer services bureaus are a partial exception, because they use data from the DDME data market to provide their hygiene and related services.

As for the remaining postal production and delivery segments, we estimate that 75% of their revenues are indirectly dependent on the DDME data market. For example, in the printing industry, monies are realized from services that are extremely data-intensive. Direct mail is increasingly personalized throughout the content of the mail piece – i.e., variable data digital printing now allows literally dozens, even hundreds of individual fields within a single communication to be customized to specific recipients. Such personalization requires a significant investment in ILCD, but virtually none of the data involved is directly licensed by or from the producers themselves. These revenues are those derived from servicing owners' acquisition mailings, and house file mailings which use appended data or processed data. These campaign types account for essentially 65% of the total mail volume serviced by postal producers and delivery, or about 60% by added value.[41]

Table 25 Postal Production and Delivery

Postal Production DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$32,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$1,000	3%
Value added by services that indirectly depend on data exchanged or rented among firms	$24,000	75%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$25,000*	78%

[41] The difference is accounted for by the proportion of prospect mail between direct mail and catalog (higher in the former) and the difference in production costs between them (higher in the latter.)

Remainder:			
	Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$7,000**	22%

Table 26 Postal Production and Delivery - Employees

Postal Production DDME Value-Added Employment and Dependence on Data-Exchange		Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues		**138,000**	**100%**
	Employment added by services that depend directly on data exchanged or rented among firms	5,000	4%
	Employment added by services that indirectly depend on data exchanged or rented among firms	106,000	77%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms		*110,000*	**80%**
Remainder:			
	Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	28,000	*20%*

4.3 Email Production and Delivery

Email marketing service providers (ESPs) provide creative and deployment services. Producers use email marketing heavily for opt-in acquisition and retention marketing, especially in business-to-business verticals. Nonetheless, the low cost of production and delivery mean producers' total outlays on email services remain relatively low compared with other CRM channels.[42]

Net Value Added to the DDME.

We estimate that email marketing expenditure represents about $2.0 B in gross DDME revenues in the US.[43] Of this total, we examined firms providing $900M U.S. DDME revenue. Some provider firms specialize in email marketing delivery, while other firms offer email delivery platforms as well as other services; many of these overlap with direct

[42] Firms who own and provide the network (the numerous Internet Service Providers, cable network providers, mobile phone operators, and all their supporting infrastructure) do not receive compensation for delivery of email – email senders essentially 'free ride' on a backbone built and paid for by others. Second, we estimate that only about 60% of messages sent by firms are deployed through outsourced email service providers. The other 40% is deployed by producers' internal marketing departments, where the cost remains invisible. This is most likely to be true of high-volume emailers because the economics of 'insourcing' bulk email delivery.

[43] Scevak, Niki with VanBoskirk, Shar. *Forrester Research Email Marketing Forecast*, 2011 to 2016 (US). Forrester Research, 24 March 2011.
http://www.forrester.com/Forrester+Research+Email+Marketing+Forecast+2011+To+2016+US/fulltext/-/E-RES59101

mail. These companies tend to have similar sizes, and for them, email appears to account for less than 50% of firm revenue. Nonetheless, as explained in chapter three, we include all eligible DDME revenues. However, eligible revenues do not include providing transactional or support messaging, such as notifications or alerts involved in confirming a payment, a balance update, a notification of delivery, responses to service or technical inquiries. These we believe are relatively small in volume, except in the financial services sector; therefore we exclude 10% of the gross value as outside the scope of the DDME, leaving about $1.7 billion in gross revenues.

We estimate that data purchased by ESPs is roughly 35% of total eligible revenues. Data purchase costs are fairly high in order to improve the deliverability of email. Without relevant subject lines, or current and up-to-date email addresses, senders of email risk having their messages blocked by recipients' Internet Service Providers (ISPs) or email service as spam. They also risk having them or ignored and left unopened, even when the email reaches the inbox. In addition, a certain amount of expenditure is incurred to append email addresses to houseflies that are otherwise postal address only. Thus the final value added by email marketing is approximately $1 billion.

Distribution by Dependence on DDME Data Exchange.

We classify half of email service provider revenue as directly dependent on the use of data transferred among firms and data servicers. Of the remaining amount, 80% can be attributed to supporting campaigns that use data acquired by others (usually producers) to augment service providers' own proprietary prospect and house file databases.

Table 27 Email Production and Delivery

Email Production and Delivery DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$1,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$1,000	100%
Value added by services that indirectly depend on data exchanged or rented among firms	$-	0%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$1,000*	100%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$-**	0%

Table 28 Email Production and Delivery - Employment

Email Production and Delivery DDME Value-Added Employment and Dependence on Data-Exchange		Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues		**5,000**	**100%**
	Employment added by services that depend directly on data exchanged or rented among firms	2,000	40%
	Employment added by services that indirectly depend on data exchanged or rented among firms	2,000	40%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms		*4,000*	**80%**
Remainder:			
	Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	0	*0%*

4.4 Telephone Sales Production and Delivery

Here we measure revenues earned by supporting clients' outbound telemarketing campaigns that target consumer landlines. We also include marketers' expenditure on inbound call center activity, which involves upselling or cross-selling efforts.[44]

Net Value Added to the DDM Economy.

Teleservices are very expensive on a cost per thousand basis, and service bureaus are either paid by hour or paid by completed action, such as number of conversions, orders placed, total sales volume achieved, etc. We examined a half dozen teleservices bureaus said to have about $9.4 billion in gross revenues. IBISWorld's U.S. Teleservices report indicates that total industry revenues in 2012 were likely about $18 billion, so we make an allowance of $8.6 billion for all other firms.

To arrive at our net DDM revenues for teleservices, we make a number of deductions for ineligible services. We exclude call center activity performed abroad for non-US clients. We also exclude U.S. work that is not connected to one of our eligible marketing functions. Thus we exclude work such as survey research, appointment scheduling, physician answering services, or work for political campaigns. More significantly we exclude fees received for geographically targeted campaigns, and any inbound call-center

[44] Calculations involving the value added to marketers by contact centers are similar in principle to those outlined for ecommerce retailers in the next chapter. After subtracting the cost of goods sold, one must make a determination about how much of the difference relative to the transactions represents the value contributed by data driven marketing – i.e., the use of ILCD to segment and target the consumer while he is placing his order, so that 'incremental' revenues are generated (upsells or cross-sells). One significant difference is that call centers do not purchase media in order to drive traffic – they are bought because other media need an outbound channel to drive traffic to, for order-taking.

activity that is not involved in upselling or cross-selling. We do so because either the cost of such inbound order-taking should properly be attributed to the outbound medium responsible for triggering the original call, or because the call represents organic demand that cannot be attributed to any producer marketing program. We also exclude telemarketing using random-digit dialing. This yields about $12.3 billion in Net DDME revenues.

Deductions for cost of data. When using the paid-by-action model, service bureaus have a strong incentive to enhance the probable success of those actions by purchasing more data. Data is used to categorize respondents to outbound calls by likelihood to purchase, using factors like household income indexed against the value of item offered, in order to minimize operational expenditure on unlikely prospects. Other data uses include recognizing and routing incoming calls from high value customers towards more experienced sales agents, providing sales agents with personalized information about the customer to improve the sales climate, and providing real-time information about new offers, wait times for delivery, etc. Data is also used to display different call scripts to operators (sales agents) or to refine offer configurations. We deduct about $900 million for data goods and services sourced from within the DDME marketplace, for a net value added to the DDME from teleservices of $10 billion.

Attribution of Value Added by Dependence on Data Market

For all inbound call center activity, we assume that teleservices bureaus service data owned by others (the client whom the caller is trying to reach being the owner.) For outbound activity the teleservices bureaus are primarily servicing producer provided data. Most outbound calls are not cold calls made to unknown prospects; few producers do pure prospecting via phone. Most are opted-in prospects who have registered on proprietary websites, ecommerce sites, etc. However, our interviews indicate that a significant portion of campaigns that are serviced by telemarketing bureaus use data appended by the list owner.

Table 29 Teleservices

Teleservices DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$10,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$2,000	20%
Value added by services that indirectly depend on data exchanged or rented among firms	$6,000	60%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$8,000*	80%
Remainder:		

	Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$2,000**	20%

Table 30 Teleservices Employment

Teleservices DDME Value-Added Employment and Dependence on Data-Exchange		Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues		**43,000**	**100%**
	Employment added by services that depend directly on data exchanged or rented among firms	10,000	23%
	Employment added by services that indirectly depend on data exchanged or rented among firms	25,000	58%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms		*35,000*	**81%**
Remainder:			
	Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	8,000	*19%*

4.5 Mobile CRM Platform Providers and Developers

Mobile CRM platform providers and developers embrace two emergent media that exhibit contact properties very similar in function to established CRM media such as postal or email. These firms provide platforms (software as a service, sometimes called SaaS), or develop custom direct one-to-one communications using opt-in SMS (text) messaging to cell phones.

Value-Added Contribution to the DDME.

We identified at least 36 companies that provided mobile SMS and app-based CRM services. Together, they were assessed at an annual revenue valuation of $3 billion in U.S. gross revenues. Among the firms we investigated directly, many offered services (usually related to mobile advertising) that fell outside this particular service area, or required us to make adjustments for pass-through revenues. The resulting revenues per firm were quite small: none of them accounted for more than $20 million in net revenues.

After adjusting for discrepancies in service offering and accounting, we had to take account of revenues accounted for by 'all other' firms in this sector. There are many of these, mostly small start-ups and developers that eluded our initial inventory of firms. To estimate the remaining value attributable to all other firms and developers, we compared reports from the Interactive Advertising Bureau (IAB) and the Mobile Marketing Association. From these, we accepted that total expenditure on SMS messaging

platforms was about $846 million, and that expenditure on mobile apps was $1.73 billion. Together they account for $2.5 billion in net revenues.

Deductions for Cost of Data Goods or Services. It is theoretically possible for SMS and even apps to support a market for list rentals or appending, for example, but this market does not yet exist, largely because the providers of mobile networks and app operating systems do not allow it. There are no significant revenues we need to deduct for data that is bought elsewhere in the marketplace. Thus, we estimate there to be a modest indirect dependence on data exchanged among firms of less than 5% of total revenues.

Table 31 Mobile CRM

Mobile Customer Targeting DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$2,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$-	0%
Value added by services that indirectly depend on data exchanged or rented among firms	$-	0%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$-*	0%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$2,000**	100%

Table 32 Mobile CRM Employment

Mobile Customer Targeting DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**11,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	-	0%
Employment added by services that indirectly depend on data exchanged or rented among firms	1,000	9%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*1,000*	**9%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	10,000	*91%*

4.6 Future Trends

Prospect/customer relationship marketing is the sector with the highest DDME value at a combined $52 billion; however, this value is derived from disparate channels at various levels of maturity, for example, the sector that services postal production is valued at 16 times higher than the sector that services mobile customer targeting. Over time, we believe that this sector will remain relatively flat as mature services decline or become more efficient, and newer services grow to take their place.

List Managers and Brokers. The firms that provide list management and broker service will most certainly experience growth in the future, however, we expect consolidation and merger and acquisition activity in this space to the point that some of these firms may shift outside the pure list management and broker definition.

Postal Production. We expect postal production to remain relevant in this sector due to its consistent return on investment, with continued little to no growth[45]. Marketers will work further toward efficiency in this channel through the use of targeting mechanisms and we expect that the distribution of revenues with the sector moves slightly in favor of companies that service data rather than those that service production.

Email Production and Delivery. The email sector will remain a relatively small portion of prospect/ customer relationship marketing with growth of only 6% between 2011 and 2012.[46]

Telephone Sales Production and Delivery. As expected, telephone marketing has experienced an 18% decline since 2006[47]. More efficient methods of servicing customers in digital channels (text message, email, corporate websites) will continue to replace the value generated by telephone marketing in the DDME.

Mobile CRM Platform Providers and Developers. The mobile CRM sector has remained small due to technology barriers in accessing and targeting customers. SMS in the U.S. was tightly controlled as a marketing channel, and will likely stabilize or decline. However, as new technologies in the mobile space open up access to customers, we expect that this channel's DDME value will grow significantly.

Overall. We expect that the existing $52 billion of value in the prospect/ customer relationship marketing sector will grow modestly, while the distribution of value within the sector will change; more mature channels will stabilize and newer channels will experience growth.

---

[45] 2% or less per DMA Fact book
[46] DMA Fact book
[47] DMA Fact book

5 Data-Driven Commerce & Fulfillment

As discussed in earlier sections of the report, our last significant marketing function involves marketer-supported efforts to complete and fulfill transactions with end customers that were initiated by outbound marketing campaigns. For over a decade, data-driven marketing has come to play an increasingly significant role in reducing latent frictions involved in handling inbound order inquiries, order placement, and payment (the commerce sub-function) and delivering the product to the customer.

Fulfillment is part of the data-driven marketing economy, and the companion to data-driven commerce, because with remote retailing, there is no face-to-face transfer of the physical goods when payment is made.

5.1 Remote Retailers: Call Centers, and Mail Order[48]

The prelude to data as a facilitator of prospect conversion and customer retention in the sales process began when the offline world of direct response mail-order houses realized that prospects were more likely to respond if they included a postage-paid direct reply card or envelope, thus transferring the cost of inbound communications from the customer to their own outbound communications efforts. Given that the cost of a reply card or envelope could raise the cost of a direct mail piece by 20 percent, it became all the more imperative that direct marketers use data to identify individuals most likely to respond.

Catalogers then extended this dynamic to the call center, where the introduction of the toll-free 800-number removed the disincentive caused by high long-distance tariffs in earlier decades. In addition to the expenditures involved in paying for customers' line charges, marketers' outbound programs now had to budget for a variety of additional costs. These costs included the wages of sales associates who were needed to answer questions, and, to entice customers to buy. This cost estimate required careful parsing of response rates, conversion rates, average order sizes, and other factors, lest there be too many, or too few, associates for the volume of traffic.

Thus, it is because the costs of inbound commerce in these channels are normally budgeted into the cost of the outbound campaign media, that we estimate their value-added revenues in the previous chapter discussion of postal and teleservices marketing.

[48] As discussed elsewhere in this report, our interest in this section is to understand the value contributed by marketing services providers, not the value of sales achieved from consumers as a result of using such services. Thus we treat retail channel providers whether online or off, virtual or brick and mortar, as service providers like any other.

5.2 Ecommerce

Remote retailing began to evolve as a new business model more than a decade ago, with the introduction of eCommerce sites.

Online retailers form part of the data-driven marketing economy because they create value by performing services for manufacturers in four categories:

- First they attract customers. In this category they advertise on search engines, buy display advertising, and buy media to retarget visitors who browsed their websites but did not buy.
- Second, they perform show-rooming and product-inquiry services: they assemble displays on websites by photographing objects, writing descriptive prose, sometimes assembling and displaying user reviews of merchandise, and sometimes offering recommendations ("people who bought this buy that.")
- Third, they transact: they offer shopping carts and checkout, and initiate payments processing.
- Finally, they ship to customers, either by drop-shipment (sending an order to the manufacturer or wholesaler for shipment to the customer) or by shipping from their own inventory.

These eCommerce sites combine into a single process many previously separate features of remote retailing. In one site, they combined the store window function of the outbound print catalog, the product research and inquiry function provided by the sales associate of the inbound call center, and the order-processing of the mail order house. With the lower costs of media and marketing afforded by the Internet, many eCommerce sites discovered they could sell digital goods or offline services (goods such as music, video, ebooks, and services such as airline travel, hotel accommodations, and restaurant reservations), using automated computer processes. Some firms found they could apply data at scale to tailor and manage these internal processes, along with managing the costs of fulfillment for physical goods. The fulfillment process is discussed in greater detail later in this chapter.

After a decade or so of development, the eCommerce world has evolved into two distinct business types: large scale diversified sites (such as Amazon.com and iTunes), and small firms supplying niches (such as BassPro for fishing related equipment, mousepadsonline.com for mouse pads, hotpicksusa.com for guitar picks). The analysis that follows makes estimates based on these two types of eCommerce business.

Net Value Added to the DDME

From external research, we take the total value of online retailer revenues in the U.S. at consumer prices in 2012 to have been $225 billion.[49]

We estimate the value of the above-listed retailer services supplied to producers to range between 15% and 20% of revenues, inferred from the commissions that online retailers charge manufacturers[50] plus an allowance for free shipping (borne by the retailer), and estimate that the value is distributed across the four categories as follows:

- Customer acquisition 5%
- Merchandising services 2% to 7%
- Transaction 3%
- Shipping 5%

Subtraction of Cost of Data Exchanged Between Firms.

For our small niche retailers (all but the very largest of online retailers), deductions primarily involve costs of customer acquisition in the form of targetable online media bought elsewhere in the DDME, so it is captured elsewhere in this study and must be deducted. When, for example, a retailer retargets a shopper, it buys the inputs it needs to perform the service from publishers and display ad exchanges, and the payments are reflected in the publisher and ad exchange and network sections of the study.

The exception to this rule is the very largest of online retailers, whose scale is such that they operate their own ad exchanges and networks. We estimate that four online retailers, Amazon, Staples, Apple iTunes, and Wal-Mart, have reached this scale, and their services, internal to each, are not otherwise recorded in this study. Their online sales total $87 billion. We attribute 20% of the U.S. sales of these four retailers to the data-driven economy, or $17 billion.

For the remaining $138 billion of retail revenue, we estimate retailer services contribute 12% of sales to the data-driven economy, or $17 billion.

Allocation of Value By Dependence on Data Exchanged Among Firms

The value created by our four large-scale retailers is dependent on data exchanged across firm boundaries in much the same way that is discussed in the section of the study on Audience Targeting and Communications (see chapter 7). That is, their advertising revenues primarily depend on use of a manufacturer's 1st party cookies. To a lesser extent the value they provide is enhanced by appending third-party-compiled data to

[49] Internet Retailer & U.S. Dept. of Commerce http://www.internetretailer.com/trends/sales/
[50] From retailer websites. eBay charges its sellers a 10% commission on retail receipts to a limit of $250 per item, Etsy charges 6.5%, and Amazon charges a referral fee that varies by product but averages about 15%, plus a small listing fee.

cookies, to categorize website visitors into high- and low-potential prospective customers. Thus the 5% of revenue from the top four retailers is dependent on the data marketplace of approximately $4 billion. The value created by the remaining retailer services is not particularly dependent on data exchanged across firm boundaries. However, similar to fulfillment services below, up to 65% of the value realized by eCommerce service providers depends on the proportion of data exchange in outbound campaigns by third-parties that drive demand to the site, or up to $22 billion.

Table 33 eCommerce

eCommerce DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$34,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$4,000	12%
Value added by services that indirectly depend on data exchanged or rented among firms	$22,000	65%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$26,000*	76%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$8,000**	24%

Table 34 eCommerce Employment

eCommerce DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**146,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	18,000	12%
Employment added by services that indirectly depend on data exchanged or rented among firms	97,000	66%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*115,000*	**79%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	31,000	*21%*

5.3 Brick and Mortar Retailing: Individually Addressable Touch points

A decade or more ago, the world of brick-and-mortar retailing would not have been included in a study of individual-level marketing practices. Simply put, virtually all steps in the retailing funnel were anonymous. That would have included everything from foot traffic through in-store browsing and research inquiries with salesclerks, through final transactions at the cash registers of supermarkets and other retail outlets. Retailers did not know which individuals walking by their store came inside to browse. Among browsers, they couldn't distinguish passive browsers from active seekers, nor could they log the identities of shoppers who became purchasers of specific products.

Even had it occurred to retailers to become data owners, at the time there was no incentive to do so, since the marketer would not have been able to use the data to manage advertisement purchases in the anonymous broadcast media driving retail sales.

Automatic digital processes have since changed retailing's anonymous landscape into one much more amenable to economically efficient data ownership and analysis. The first and most important step in retail's use of individually addressable touch points came with the loyalty card. This tool initially transformed the hotel guest or airline passengers check-in, but went on to transform the retail point of sale.[51] Here we focus on loyalty programs associated with the point of sale as the most developed.

With loyalty programs, the chief challenge for this study was to define what exactly should be considered DDME revenues. To find a value for outsourced loyalty programs consistent with our marketer-expenditure approach, we estimated the portion of offline retail, travel, and accommodation sales that might be associated with an outsourced loyalty management program.

In doing so, we discovered that many marketers and retailers were investing heavily in the capital side of the program (investments included servers, POS terminals, and software, for example) without necessarily investing in the peopleware and data segmentation capability to leverage the tools to their maximum advantage.

Loyalty programs rely on the exchange of individual-level customer data between marketers and loyalty management companies to drive $48 billion of perceived value to U.S. consumers.

Loyalty programs primarily include card-based retail programs, travel and hospitality point and mileage systems, and financial services loyalty programs[52]. All of these form part of the data-driven marketing economy through their creation of value in the following categories:

---

[51] Other steps in retailing are now increasingly addressable: the presentation of loyalty cards helps personalize experiences, and retailers talk of using closed-circuit monitors to individualize the in-store shopping experience. Reportedly, casinos now provide loyalty-based debit cards to loyal customers so their "one-armed bandits" will customize their payoffs to the customer's preference for small frequent wins vs. big infrequent wins.

[52] Hlavinka, Kelly and Sullivan, Jim. *The Billion Member March: The 2011 COLLOQUY Loyalty Census: Growth and Trends in Loyalty Program Membership and Activity.* LoyaltyOne and Colloquy, , April 2011. http://www.colloquy.com/files/2011-COLLOQUY-Census-Talk-White-Paper.pdf

- They manage a system of loyalty points that have a value to the brand, to consumers, and to other marketers
- They create an exchange for the points between consumers and marketers
- They deliver insight against existing store and brand databases
- They deliver offline and online targeted offers

Scope of Loyalty Programs

To find a value for outsourced loyalty management programs consistent with our marketer-expenditure approach, we took two approaches:

- We estimated the portion of offline retail, travel, and accommodation sales that might be associated with an outsourced loyalty management program.
- We estimated that 10% of the perceived value received by customers was attributable to revenue generated by outsourced loyalty management programs.

All revenues for loyalty management program are considered a part of the ecosystem. We conclude that the value of loyalty management program companies in the U.S. DDME at USD 5 Billion.

Table 35 Loyalty DDME Revenues

Commerce- Loyalty DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$5,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$2,000	40%
Value added by services that indirectly depend on data exchanged or rented among firms	$1,000	20%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$3,000*	60%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$2,000**	40%

Table 36 Loyalty DDME Revenues Employment

Commerce- Loyalty DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**21,000**	**100%**

	Employment added by services that depend directly on data exchanged or rented among firms	10,000	48%
	Employment added by services that indirectly depend on data exchanged or rented among firms	4,000	19%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms		*15,000*	**71%**
Remainder:			
	Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	6,000	*29%*

Third party loyalty program management firms rely on the whole on their clients' owned customer and transactional data, and rarely, if ever, purchase additional data. If data is appended, it is often demographic data that can be sourced for free. Therefore we account for cost of data inputs at 0%.

We assume that approximately 50% of the revenues for a loyalty management program company are generated from information collected and held at the ILCD level of detail, while the other 50% is collected and held at an aggregated level.

Fifty-two percent of best-in-class marketers use loyalty data for marketing campaigns[53]. Additionally, we discovered that many marketers and retailers were investing heavily in the capital side of the program (investments included servers, POS terminals, and software, for example) without necessarily investing in the peopleware and data segmentation capability to leverage the tools to their maximum advantage.

5.4 Fulfillment: Parcel Delivery

The last step in the value chain that relates to our study is the delivery of offline goods into the hands of purchasers. Our focus here are data-driven shippers: firms such as FedEx and UPS, who, in addition to the USPS parcel delivery, provide ILCD-based sorting, routing, and carrying services to provide cost-effective delivery of small packages to millions of individual recipients at widely dispersed locations every day.

[53] 2013 DMA Statistical Fact Book, Source: Neolane, Inc. and DMA, "Big Data," 2012.

Table 37 Fulfillment

Commerce- Fulfillment DDME Value-Added Revenues and Dependence on Data-Exchange	Millions	Percent
Total contribution of the sector to the data-driven marketing economy	$9,000	100%
Value added by services that depend directly on data exchanged or rented among firms	$-	0%
Value added by services that indirectly depend on data exchanged or rented among firms	$4,000	44%
Subtotal: Combined value added by services directly or indirectly dependent on data exchanged or rented among firms	*$4,000*	44%
Remainder:		
Value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	**$5,000**	56%

Table 38 Fulfillment Employment

Commerce- Fulfillment DDME Value-Added Employment and Dependence on Data-Exchange	Employees	Percent
Total Employment Attributable to DDME Value-Added Revenues	**37,000**	**100%**
Employment added by services that depend directly on data exchanged or rented among firms	-	0%
Employment added by services that indirectly depend on data exchanged or rented among firms	18,000	49%
Subtotal: Combined Employment added by services directly or indirectly dependent on data exchanged or rented among firms	*18,000*	**49%**
Remainder:		
Employment added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms	19,000	*51%*

To determine how much of these shippers' gross revenues belong to the DDME requires two considerations. How much of their delivery volume is attributable to orders resulting from DDME go-to-market efforts by producers? Then, of this amount, how much of the cost of such shipments is borne by the marketer, as opposed to being paid for by the customer?

On the first point, we began by identifying shippers' U.S. revenue stream that most closely approximated small-parcel delivery to individual consumers. We then took a very conservative figure of 10% of this value as the minimum amount of revenues attributable to orders generated by DDME processes, though in fact this percentage is likely significantly higher. We then cut this figure in half, to account for those shipments

where the cost is charged to the recipient. While this exact proportion is difficult to determine, we believe that free shipping is increasingly the norm in remote sales, thanks in part to the popularity of services like Amazon Prime.

Deductions for Data Inputs or Services. Our investigations lead us to believe that logistics firms purchase or acquire little to no ILCD as part of the fulfillment process. If they do, it is only to update delivery addresses, for example, and is quite small; therefore, we make no adjustments to total value added.

Attribution of Revenues by Dependence on Data Exchanged in the Ecosystem. As servicers of targeting-data for individual addresses, the shippers providing fulfillment services are not directly dependent on data for their own use. Their chief dependence, if any, can only be indirect. In other words, their dependence is only on data exchanged by other players as part of an earlier marketing function. This dependence reflects by consumers in response to prospecting campaigns in any media, together with the portion of shipments attributable to consumer purchases made in response to house file campaigns enhanced with ILCD. Thus, we estimate that about one half of fulfillment's revenues can be classified as indirectly dependent on the exchange of data among firms in the DDME.

5.5 Future Trends

The DDME value for Data-Driven Commerce and Fulfillment grows both from the increase in digital transactions, and through the ability of marketers to identify consumers from one end of the value chain to the other, thus adding traditional transactions into the fold of Data-Driven Commerce.

eCommerce. According to eMarketer, eCommerce in the retail space is currently growing 14% year over year [54] while purchases on the mobile phone, still only 3% of total ecommerce sales, will jump 158% over the next four years[55]. Companies that service these transactions will grow alongside hoping to create security and efficiency and deliver measurable results to marketers.

Loyalty. Loyalty programs that strive to "close the loop" in to identifying a consumer at the offline point of purchase may grow; however, the companies in this sector are narrowly defined and in the future this sector may take on a broader definition that includes Data Management Systems that manage transactional information at the Point of Sale. Without a redefinition of the category, or extensive mergers and acquisitions in this space, we would expect Loyalty to grow relatively slowly in the DDME.

Fulfillment. Lastly, we expect Fulfillment's growth rate to be lifted by the strong growth of ecommerce for retail.

---

[54] eMarketer, April 2013
[55] Forrester Research, Jan 2013

Overall. Data-Driven Commerce and Fulfillment currently accounts for $48 billion in the DDME. We expect the growth of these companies to track closely with eCommerce sales at about 15% year over year, with an eye toward the opportunity for additional growth where there are advancements in eCommerce technologies and processes as well as the opportunity to move more revenue into the DDME when digital goods are adopted over physical goods.

6 Conclusion: Consequences for Efficiency, Insight, and Innovation

Our study finds that producers of goods and services pay $156 billion a year for marketing services that could not be provided without individual-level consumer data. We infer that the U.S. economy benefits by more, and probably substantially more, than this outlay, else the outlay would not be made.

Some of the benefits are direct. They benefit the U.S. economy by allowing markets to operate more efficiently. Data lifts a dead weight from the U.S. economy in every industry when data-driven marketing reduces the cost of matching products and services to the consumers and business buyers who want them. Participants in the DDME gain these efficiencies by shifting marketing dollars from mass media to interactive, data-driven media. At the same time, they rely on data to continually optimize their mix among interactive media. Our study also found evidence that the footprint of the DDME itself was growing, as formerly broadcast media adopt data-enabling technologies and practices – a trend that will likely promote "data-driven adaptation" by current mass marketers.

Other benefits are indirect. They are positive externalities or spillovers from DDME activity to the U.S. economy. While the outlay of $156 billion is made in expectation of efficiency gains, we do not assume that it is made in expectation of these indirect benefits because the beneficiaries of the indirect benefits are, in many cases, not those who make the investments. The benefits are, nonetheless, part of the value of data. We discuss them in this chapter under the headings of insight and innovation.

Our study finds that 676,000 people are employed in jobs that depend on individual-level data. We find that a further 1,038,000 people owe their employment to the DDME jobs. This second category of jobs includes employment in supplier industries beyond the DDME, as well as in sectors where DDME workers spend their paychecks, and people and in local, state and federal government. While these 1,713,000 jobs do not exist solely because of the ability to use individual-level consumer data, data's role in supporting this level of employment appears substantial. If DDME marketers, publishers and networks did not use ILCD, the economy would be less efficient and could require either more people to do the work the DDME currently performs, or fewer, depending on whether the inefficiencies could be compensated for by labor. However, it is a reasonable inference that the costs of less efficient assembly and targeting of audiences, customers and prospects would be felt elsewhere in the U.S. economy, as would job losses. These derivative implications do not form part of our study. All that can be said with confidence is that the DDME requires a wide array of well-compensated, highly specialized, and cutting-edge technical and analytic skills, and reductions in the DDME's demand for such skills would affect the employment climate for up to 675,000 people.

Efficiency

Marketing is a notoriously inefficient business function and makes up a large component of what consumers and businesses pay for goods and services. A survey of the chief marketing officers of 410 U.S. firms reported that 11.3% of firm revenues were spent on marketing activities.[56]

The cost of marketing bears directly on consumer welfare. Consider for instance the consumer products sector, where the same survey of chief marketing officers found that marketing absorbed 12.9% of manufacturer revenues.[57] A manufacturer's recommended retail price is typically about twice its price to the distributor. If so, marketing costs inflate the price that consumers pay for food and household products by $25 in every $100 spent because marketing costs are marked up in retail margins. That is not to imply that these marketing costs are wasteful, because matching consumers to products they may want but not know about is a value-creating task. But if go-to-market strategies could be made more efficient, at least some of this value would be captured by business and consumer buyers through lower prices, as increasingly efficient marketers fight to increase market share.[58]

Though data can drive both efficiency and effectiveness, our study revealed that the former was the more observable. In sector after sector, we found a consensus that the value of targeting data comes from efficient customer selection rather than persuasion. In other words, it helps producers identify and target those who are likely to buy – the people who do not need costly persuasion – and helps avoid spending on messaging to people who are not likely to buy. The measurability of individual-level data is attractive to marketers because it shows which elements of a marketing program produce results and which elements do not, revealing how to eliminate ineffective communication.

Insight

Beyond analytic insights, which lead directly to refinement of marketing methods as discussed extensively in this report, there are insights from data about individuals that reveal entirely new business opportunities or socially important results.

Aggregations of Individual Actions

Google Flu Trends represents one instance in which ILCD-supported collective intelligence provides social benefit. Launched in 2008, Google Flu Trends was created to analyze individual searches related to flu, pseudonymized in the form of Internet protocol addresses, and plot the searchers' locations. Once parameterized, the data was used to develop a model to predict flu outbreaks across specific localities of the U.S.[59] The model

---

[56] cmosurvey.org, *The CMO Survey*. Results by Firm and Industry Characteristics, Topic 3: Marketing Spending. August 2013, page 99.
[57] Ibid.
[58] There can be no precise measure of the efficiency of marketing because there is no obvious perfect market matching price to use as a criterion; any gain in efficiency, however, benefits consumers and business buyers.
[59] It computed a time series for weekly queries entered from 2003 to 2008 for each state in the U.S. and normalized each by dividing it by the number of all Google search queries in that state. A linear model computed the log-odds of

outperformed the U.S. Centers for Disease Control (CDC) through to 2012, and some months thereafter, as it was being updated.[60]

When finally refined, a collective intelligence model such as Google Flu Trends offers the prospect of deploying individual-level data to generate a new kind of aggregate insights not previously possible. This kind of collective intelligence produces socially or commercially useful insights by relying on pseudonymous data to improve the model's signal-to-noise ratio for both inputs and outputs in ways that anonymous data on its own cannot.

As an example of how useful this is, consider the implications of Google Flu Trends for local health officials and doctors who want to combat disease outbreaks. Officials can reasonably assume that one of the first things people do when they begin to feel ill will be to google their symptoms. Researchers see a spike in searches. But to interpret this spike, they must understand whether the total number of searches results from a small group of people performing multiple searchers, or from a large group of individuals each performing a low number of searches. Anonymous aggregate data can't distinguish between the two situations, but aggregate data built from PI (device-identified data) can. Furthermore, if a spike in searches reveals an outbreak, anonymous aggregated data cannot pinpoint precise locations and tends to reveal an overall trend. By contrast, PI-based aggregate data that preserves the location value of IP addresses in searches. This could allow practitioners to identify which communities will most likely see crowded emergency rooms and a run on flu vaccine.

Similar innovations in inferences from collective intelligence have been developed from the intersection of ILCD and social media sharing. At least since Gladwell's *Tipping Point*, marketers have been keen to understand what makes a movie, book, fashion, or tune become popular in the aggregate – or in other words, go viral. As a small example of what might be possible in the way of developing new sources of insight, a 2012 article reported findings from research conducted via MusicLab into the influence of others' online behavior on music consumption. Using ILCD to control for music consumers' exposure to songs that others in their network were downloading, researchers found that consuming music proceeded in two decisions. The first was whether to sample a new song for a few seconds, and the second, whether to download it. ILCD-based aggregate allowed researchers to connect the dots between two individual decisions and see that social interactions influenced the decision to sample, but not the decision to download. (Unsurprisingly, what influenced the decision to download was whether the individual liked what they heard.) [61].

---

visits to a physician for treatment of influenza-like illness, using influenza-related search queries for each week for each state as the predictor of visits.

[60] In the 2013 season the model continued to predict regional outbreaks sooner than the CDC, but recalibration was required because it overstated the level of flu prevalence.

[61] C Krumme, M Cebrian, G Pickard, S Pentland "Quantifying Social Influence in an Online Cultural Market," - PloS one, 2012 - dx.plos.org

For marketers such pseudonymous aggregate insights could one day become routine sources of forward-looking strategies with which to anticipate and possibly even influence emergent market trends, possibly in real time. If true, the long-term implication of ILCD's growing role in market insight is that an increasing share of its value will be realized in the form of effectiveness, not just efficiency.

Such data-driven expansions in the power of social insight are already being actively pursued by governments and individual public agencies. For example, the United Nations' Global Pulse initiative draws data from pseudonymized social media, blogs, and mobile phone records and anonymized financial transactions, to detect spikes in major social issues such as unemployment or food prices.[62] The Alliance for Useful Evidence, a publicly funded network in Britain that champions the use of evidence in social policy and practice, proposes that social media data, whether PII or not, be used extensively at all levels of government. It quotes Jamie Bartlett, director, Centre for the Analysis of Social Media at the British policy institute Demos, "In 10-15 years' time it's entirely possible that every government department will have a social media analysis unit.[63]"

Various kinds of individual-level data are already in widespread use, although the practice is not widely recognized. The instance that is at once most familiar and most surprising is the unique challenge presented by voting. On the one hand, voting is meant to be private, to preserve the individual's freedom to choose without fear or favor; in aggregate, the results produce majority rule. On the other hand, democracy works on the principle of "one citizen, one vote." A fair vote should exclude non-citizens, and prevent citizens from voting more than once. This requires information about specific individuals How do successful democracies reconcile this seeming contradiction between the need for an anonymous aggregate outcome, and individual-level fairness? Societies use voter registration, a form of PII, or the practice of dipping a voter's finger into ink, which for election officers and poll-monitors is a form of PI, used to minimize repeat-voting fraud

While voting is a mature public policy application of individual-level data, political organizations in the United States have pioneered the application of ILCD to electing policy-makers to office, and mobilizing the pressures of public opinion on them once elected. Without ILCD, options for campaign strategies are constrained by what is affordable and effective in mass media. The high cost of television and radio advertising has driven up the cost of campaigning even as the lowest common denominator aggregate focus of the medium creates virtually irresistible incentives to indulge in negative, ad-hominem smear tactics against opponents, usually to the neglect of all but the most generically phrased policy issues. But for the last several decades, mailing lists, telephone numbers, email, and social media, have managed to restore a degree of specificity that can make tailored, issue-by-issue appeals to voters. Especially in swing states or districts where a plurality of swing voters can tip a race in either direction, it is unlikely that such voters will go for long without hearing from both sides on issues that

---

[62] Jason Leavey, "How can social media data be used to improve services?" Guardian Professional, October 3, 2013
[63] Jason Leavey, "Social Media and Public Policy: What is the Evidence? www.alliance4usefulevidence.org, September 2013.

matter to them.[64] Which is not to say such individually oriented, interest based messaging will necessarily be any more inspiring, truthful, or positive, but it won't be from sweeping individual interests under the rug of the so-called "median" single-dimensional voter discovered when mass media campaigning entered its heyday.[65]

Innovation

Data inspires new technology design and new ways of storing and analyzing data. It fosters startup entrepreneurship. Data is an inexpensive and abundant raw material, and the machines that process it are cheap to buy. The bridge between an idea and its implementation at scale is considerably shorter in an information economy than in an industrial economy.

Data-Driven Marketing Lowers Barriers to Entry

Innovation in the age of data faces lower barriers to entry than it faced in the industrial age, and that is often because data gathered by innovators has secondary value to other marketers and advertisers. With the confidence in revenue obtainable from advertisers interested in their audiences, entrepreneurs and investors can launch innovative products and services at low prices, in free and premium-priced (so-called freemium) alternative formats, or in advertising-supported versions at no cost.[66] Indeed psychologists have found that when products are free because they rely on revenues from other firms who can profit from the traffic, consumers attach unusual value to the free price.[67]

Data-Driven Marketing is a Significant Export Industry

The U.S. leads the world in data science applied to the marketplace. Ideas developed in the U.S. by American statisticians and econometricians, running on US-designed hardware, and coded in algorithms developed and tested in the research offices of U.S. firms, are used throughout the world. The interviews conducted in the course of this study found that when a firm earned revenue beyond the U.S. from services that relied on ILCD, a majority, and in most cases all, of its research and development employees were located in the U.S.

Online Audience Targeting has Revived Content Publishing

---

[64] The growing role of data-driven political campaigning can be observed through the change in topics covered by *Campaigns and Elections,* probably the leading journal for U.S. election campaign managers and strategists. Their reporting of the data-driven Obama presidential victory of 2012 can be found at http://www.campaignsandelections.com/magazine/us-edition/362452/the-nuts-and-bolts-of-obamaand39s-datadriven-campaign.thtml

[65] Downs, Anthony (1957). "An Economic Theory of Political Action in a Democracy". Journal of Political Economy 65: 135–150.

[66] Anderson, Free. Hyperion Press 2009.

[67] Shampanier, Mazar, Ariely "Zero as a Special Price: The True Value of Free Products." Marketing Science November/December 2007 vol. 26 no. 6 742-757.

Internet publishers compete with offline broadcast media for advertising revenues since consumers have demonstrated a clear preference for free content in the online world as they do in the offline world. Until fairly recently, however, digital publications were disadvantaged in advertising markets because they could offer advertisers neither the broad reach and frequency of exposure that brand marketers wanted from untargeted media, nor insight into their audiences that would allow them to compete in relevance with direct response channels such as email or search. By turning to the data marketplace, more publishers are applying data-driven audience targeting to Pseudonymous Information (PI) about their audiences to create specialized segmentations that are more attractive to advertisers because they can drive scalable relevance. This development is allowing many more publishers than portals and search engines to thrive on the popularity of their online content offerings.

New business models for industries based on digital goods and services

Revenues in such offline publishing industries as newspapers, books, photography, and music have experienced declines, in some cases in excess of 50%, in the decade or so since broadband Internet access became widely adopted. Much of the decline came from the ability to create and share large files online, and in tandem the development of peer to peer (P2P) networks. These networks were, for several years, a phenomenon of the Internet black market (technically illegal but widely used).

Into this void stepped a new generation of Internet entrepreneurs who saw how data could create a data-driven business model for content. These businesses recognized that universal uploading, friction-free sharing, on-demand content access, and large amounts of cloud storage eliminated the need to purchase raw material like vinyl or invest in expensive production and distributions facilities. Instead, businesses selling digital goods could leverage infinitesimally tiny marginal production costs, virtually limitless digital shelf space and cheap bandwidth to reframe how users evaluated, priced and bought content. By unbundling content from albums, books, or TV seasons, etc., it could be price and purchased a la carte. The same technology economics made if feasible to retain an enormous "back list" of niche offerings, and diversify into a vastly wider range of emergent or niche offerings, and quickly update or refresh content or offerings in a way that was simply not feasible in the old content economy that brought the underlying information to market by encoding it in physical objects produced in factories, shipped by boat train or truck, and stacked on displays.[68]

The impact of this new data-driven model on the music industry alone has been substantial. After a decade of steep revenue declines the IFPI (International Federation of the Phonographic Industry) reported that in 2012 for the first time aggregate global recorded music revenues increased. They reached $16.5 billion, and one component of the total, digital revenues, grew for the second consecutive year from the combined sources of downloads, subscriptions, and advertising-supported services.[69] Of course, the

[68] This phenomenon was most famously brought to the world's attention by Christopher Anderson, The Long Tail.
[69] Digital Music Report 2013, IFPI (International Federation of the Phonographic Industry),

online music industry is now a main driver of the overall growth of the DDME, and a harbinger of how many other categories of goods and services are likely to become increasingly data-driven in their go-to-market strategies.

In our research, we came across echoes of the entrepreneurial transformation of vinyl-driven music into data-driven music in many other content-centric industries, as new start-ups, many of which are venture-funded, exploit their understanding of the efficiencies of data to re-engineer the cost side of their business models. Examples are Dropbox[70], Pandora, Imgur[71], Pinterest[72], and Instagram, with tens to hundreds of millions of users, and growth rates that defy pre-digital economic experience. And in their drive to drive physical costs out of the market and so lower prices to consumers, what has become clear is that keeping costs to consumers as close to zero as possible, requires data capture as integral to these firms strategies for developing sustainable revenue streams sourced from data-driven marketers or advertisers. The data may be first party only, between firm and user, or may be augmented by demographic information, online behavior, or social network data, but whatever form it takes, in a marketplace still shaped by the Napster utopia of free content, firms cannot forego data-driven marketing altogether.

Digital entrepreneurship from time to time introduces new data practices into marketing.

A previous generation of entrepreneurs relied on mailing lists to recruit customers by means of direct marketing. Some subsequent innovative uses of ILCD include distribution across established user communities (as Zynga did with Facebook's community), cultivation of so-called power users within its own community (as Twitter has done) and mutually beneficially matched exchange of labor, Duolingo does when it marries free language learning to crowd-sourced text translation across a range of languages.[73] Previously the business model of such a site might have been by subscription or free for basic offerings, with premium features for those willing to pay for them, or a large enough audience would have been necessary for advertising to cover the costs of running the site. Just as creative use of PII was important to the generations of entrepreneurs that created dominant firms of the 20th century such as Sears Roebuck and niche firms such as NordicTrack, so PI is a resource for these digital entrepreneurs.

Final Reflections

---

http://www.ifpi.org/content/section_resources/dmr2013.html, accessed October 6, 2013

[70] Donna Tam, "Dropbox is like Microsoft in the 90s, says startup's CEO", http://news.cnet.com/8301-1001_3-57602059-92/dropbox-is-like-microsoft-in-the-90s-says-startups-ceo/, September 9, 2013, accessed October 6, 2013

[71] Alice Truong, "Imgur, Reddit's de facto image host, hits 100 million monthly users", http://www.fastcompany.com/3018525/fast-feed/imgur-reddits-de-facto-image-host-hits-100-million-monthly-users, September 26, 2013, accessed October 6, 2013

[72] Rachel King, "SAP, Netbase: Pinterest is fasting growing social site ever", http://www.zdnet.com/sap-netbase-pinterest-is-fastest-growing-social-site-ever-7000002473/, August 13, 2012, accessed October 6, 2013

[73] Parmy Olson, "Duonlingo takes language learning to the next level, by crowdsourcing new languages", http://www.forbes.com/sites/parmyolson/2013/09/25/duolingo-takes-online-teaching-to-next-level-by-crowd-sourcing-new-languages/, September 25, 2013, accessed October 6, 2013

It may well be that our report has detected only the first lapping of a data-driven future on the shore of decision-making across all aspects of social life. The results we report here have focused almost exclusively on value or employment added by the application of ILCD to managerial decision-making and outsourced expenditures by firms. Yet the real growth these figures betoken may be more qualitative and attitudinal, shaping the habits, practices and expectations of consumers themselves. By this we mean to raise the possibility that data may efface the very distinction between marketers (assumed to data-driven, or eventually so) and consumers (not data-driven.) As technological innovation drives lower costs, it does so for everyone.

The innovations uncovered in our researches help us imagine that the day may not be far off when the phrase "data-driven" becomes routinely applied to individual decision-making in all facets of one's private life at home, school, courtship, leisure - anywhere a resource must be optimized. Data is already being sought by ordinary individuals to use their time, money, and attention better. The purposes and opportunities of data, big, small, individual, aggregate or otherwise, as potential opportunities for the individuals currently known as consumers are still in embryo. But they may be no more embryonic than personal computing was forty years ago. Then, computing was personified by the alien, impersonal, institutional mainframe HAL of the movie *2001: A Space Odyssey*, But no sooner had HAL terrorized Dave, than a now famous band of college drop-outs in Silicon Valley began tinkering with motherboards and displays in ways not even Arthur C Clarke could have foreseen.

Generations of computer chips later, every American fifteen year old has in their pocket a smart device with which to text friends across the globe, customize a fantasy world, or stream a library of Sci-Fi flicks in less time than it takes to watch previews in a movie theatre. In short, while we cannot demonstrate it, we can imagine the ubiquity of individual level data-driven decision-making may soon be regarded not as a surprise, but an everyday expectation, even a right, by a new generation of equal and empowered market participants.

Appendices

Appendix 1: Methodology

A1.1 Assessing the Size of the Data-Driven Marketing Economy

In this chapter we describe how the firms that we believe constitute the data-driven marketing economy (DDME) – among them the 645 firms that we investigated individually – can be considered the basis for the valuation of the DDME.

It is our belief that while these firms are competitive with one another, they may be regarded as a distinct, cooperating economic system that creates and applies consumer data to the task of operating an efficient market for producers' offerings.

A1.2 Defining the Universe of Firms

What makes a firm part of the DDME is the intersection of three things: support for at least one function of producers' go-to-market strategy; a specific reliance on interactive media as the basis of that strategy; and finally, the use of individual-level customer data (ILCD), whether as pseudonymous information (PI), or personally identifiable information (PII).[74]

Marketing Function Areas

Our first level of system classified the suppliers in the DDME into six areas of support for producers' go-to-market functions. All of marketing's oft-cited four Ps (product, price, place, and promotion) can involve interactive media and ILCD. However, for practical reasons we exclude product and price, and choose to focus on promotion and place. New product development and pricing determinations are often internal to the producer, and so invisible to us. This leaves us with only outsourced expenditure on outbound marketing communications and expenditure on inbound order taking or sales transacting and fulfillment (both online and offline) as activities we can measure.

- *Strategy:* With few exceptions, producers typically draw on agencies for guidance and counsel for their go-to-market strategy. The agencies can be general agencies that advise on the proportion of spending on data-driven methods, and specialist agencies that advise on only the data-driven portion of spending. The line between strategy and execution is not sharp. Some execution firms give

[74] See note in the Methodology Appendix regarding the "inclusion test" we developed.

strategy counsel, for example, and strategy agencies sometimes participate in the selection of execution vendors, and often acquire execution firms.

- *Measurement and Analytics:* A system that relies on ILCD can improve over time because the producer or its agent can measure different things. They may examine targeted individuals' response to a specific offer, or compare those results to the responses of recipients who do not meet targeting criteria. It is also possible to focus on the targeting criteria to measure efficiency. The firms grouped in the Measurement and Analytics subsystem compete to offer better ways to judge the performance of the other marketing functions. While the other functions in some cases offer measurement and analytics, objectivity dictates that there will always be specialized measurement and analytics firms that are independent of the other functions and thus able to offer dispassionate evaluations.

- *Publishing – Audience and Subscriber Assembly*: There is a long history of firms that assembled audiences to whom producers could advertise, such as publishers and television and radio networks. To be included in the DDME, these firms must assemble audience members in such a way that they are identified and addressable at the individual level. The online editions of products of traditional publishers such as Gannett and News Corp are included here, as are publishers whose products exist only online, such as AOL and Buzzfeed. We treat social media such as Facebook and Pinterest in this category, as well as video publishers such as Hulu and audio publishers such as Pandora and Spotify.

- *Audience Targeting and Communications:* The work of assembling audiences existed long before consumers could be individually differentiated, but the work of targeting individuals in the audience is new. It arose as soon as there was technology to let producers select particular members of a publisher's audience, or address different members with different messages. We group under this heading firms that make it possible for advertisers to bid in real time for individual visitors to publisher websites, termed real-time bidding for online display advertising. Also included are firms that drive multimedia campaigns to publisher sites to target particular visitors. While most of these firms are fairly new, quite small, and generally unknown outside their niches, they represent a highly dynamic and entrepreneurial sector. These firms all operate in online sectors of the economy.

- *Customer Targeting and Communications:* While some marketing functions rely on attracting people to become audiences, others don't. The firms grouped under this heading support the use of direct mail and email, and they locate people at addresses that are either personally identifiable and/or are proxies, such as hashed email addresses, mobile phone fingerprints, and check-ins. This group contains some of the largest data compilers, as well as some very large and sophisticated printers, whose ability to print variants of a format make individual customer targeting possible. This group also contains some small firms such as niche list brokers. Most firms in this category operate in the offline sectors of the economy.

- *Commerce and Fulfillment (Online and Offline):* Producers rely on third-party functions to complete transactions. These third parties are typically retailers but may also be travel agents, auctions, payment processors, and loyalty management services. If the path to purchase is influenced by ILCD but the transaction occurs in the relative anonymity of a retail store, the commerce function does not form part of the DDME. If it occurs in a retail store with a loyalty program, the offline retail function makes use of ILCD. If it occurs online, the retailer likely relies on ILCD to perform the function, whether internally (as Amazon does with its own display advertising network), or by buying consumer and audience targeting functions from specialized firms.

Interactive Media and Touch points

The marketing function areas listed above are universal, or nearly so. However, the economic rationale for a firm to collect, store, analyze or transfer ILCD is not – at least not yet. The use of ILCD depends on the availability and use of a specific set of media that we define as interactive. It is this media dependence that makes data-driven marketing a subset of all marketing.

Media interactivity has three features, all of which must be present for us to consider a media part of the DDM media ecosystem: addressability, identifiability, and measurability.

Interactivity, while not entirely new, has only recently become a significant aspect of the media landscape. Before the age of mass computing, the only individual-level addressable media were the postal service and telephone, and they played a comparatively small role in marketing. In that era, many marketing functions had no choice but to employ media that were broadcast in nature.

A1.3 Researching Firms: Eligible DDME Revenues

The revenues eligible for inclusion as part of the DDME comprise specific marketing practices that meet our criteria outlined above. Since many firms do not simply sell one type of good, or provide one type of service, it was necessary to develop a robust set of service offering categories that would allow us to determine what did or did not fall within the scope of our research for any of the over 600 firms we investigated directly. Parsing firm revenues into appropriate categories not only helped us ensure that the revenue that we report as DDME was as accurate as possible, but it also helped us understand the dynamics of the exchanges of services among firms and segments. This allowed us to arrive at a clearer understanding of how data is used to create efficiencies and competitive advantages.

For each firm we investigated, we began with the company's highest revenues, typically global revenues from all sources. We then reduced this figure through a series of steps that excluded revenues that did not fall within the DDME:

- *Base: Global Revenue ($Millions).* As many large U.S. firms report revenues from abroad as part of their total firm revenues in their public documents, we take this amount as our initial base. By comparing the ratio of non-US revenues within total revenues to the share of U.S. employment within total employment, we can ascertain the extent to which firms' U.S. employees are generating export income for the U.S. economy.

- *US Revenue ($Millions).* This amount is the base for ascertaining the share of firm revenues to be apportioned to the DDM economy as gross.

- *Gross DDME Revenues.* In this step we exclude revenues from non-eligible products and services as well as revenues from non-eligible sources, particularly from consumers. The balance represents the amount of reported firm revenues attributable to DDM services provided to marketers before adjusting for revenues received from or transferred to other segments in the DDME value chain.

- *Net U.S. DDME Revenue.* Some firms involved with DDME claim pass-through billings for media buying as part of their top-line revenues, even though these amounts do not remain in their own pocket. This step ensures that differences in accounting practices are accounted for so that we achieve a true apples-to-apples basis for comparing firms' revenues, using only what properly accrues to them.

- The final result of the steps above is net DDME revenue, which represents the amount that, when summed, is our preliminary measure of the size of the data-driven marketing economy.

The value in the last step above becomes our point of departure for calculating our ultimate goal, namely, the total value-added of the DDME. This required us to analyze net DDME revenues into practices where data was either directly or indirectly involved, in order to identify and correct for over-counting. This analytic step was essentially one of classification involving determining the dependence of the DDME on data transfers among different types of players.

The DDME practices we settled upon for this analysis were as follows:

We began with a set of practices that were strategic in nature:

- *Income from Direct Transfer of, or Access to, Data Goods or Bundled Data Services.* In general, this category included revenues directly attributable to the licensing or rental, exchange, sharing or transferring of ILCD among third party firms for their own use. We defined as a direct sale any revenues priced on a CPM basis where the denominator of the sale was based on PII or PI, such as number of names rented, or number of individual demographic attributes appended.

- *Agency Services.* This category reflects revenues from fees for services such as brand positioning and target market identification, along with a range of more ILC-oriented activities, such as segmenting and analyzing marketing databases.

The next set of practices to which net DDME revenues could be apportioned was in digital advertising and publishing. These were primarily dependent on PI:

- *Display Advertising Sales: (Desktop, Mobile, Video, and Social.)* This primarily involves fees to online publishers or other types of business such as ad networks that buy publisher inventory. To avoid double counting, we deducted revenues received from upstream organizations that went to pay for media cost of goods sold, i.e., revenue that wound up in the hands of publishers for their inventory. These amounts are therefore considered publisher revenue. For social, this category includes revenues received from all ad units and ad targeting options available within social publishers, such as promoted user-generated content about a brand, or sponsored Facebook Likes.

- *Search Ad Production and Delivery (desktop or mobile).* Although seemingly a form of display advertising, these ads are purchased very differently, namely, on the basis of an auction for keyword search terms.

The next set of practices would classify net DDME revenues from marketing services involving the production of media for prospect and customer relationship marketing, particularly the classic direct response media:

- *Mobile Services Platform Provision and Development.* This includes tools and services such as marketer-to-consumer SMS messaging, mobile apps, mobile smart barcodes, and payment services. Here we record revenues received for the development and maintenance of branded apps for marketers or publishers. In the latter case, this amount is treated as a cost for the publisher and is deducted from any mobile app display revenue they receive.

- *Email Services Deployment and Delivery.* This refers to customer relationship management services fees, fees for deployment maintenance of email campaigns, and revenues for ad units targeted to audiences.

- *Direct Mail & Catalog Production and Delivery.* This category comprises volume-based revenues received by producers of individually addressed, direct response advertising mail or catalogs. Services include printing, letter shop services (mail package assembly, bundling, insertion into the mail stream, franking) and postage charges for delivery (including those from the U.S. Postal Service), and are apportioned to their appropriate segment. In other words, any composite revenues that are bundled together by an upstream provider are deducted and assign separately by crediting them to the appropriate segment (e.g., postage charges and envelopes). We exclude revenues associated with producing

broadcast, geotargeted "resident" or "occupant" mail, which is not eligible for inclusion anywhere within the DDME.

- *Telephone Marketing (including inbound call-center services).* This category represents volume, hourly rate, or performance-based fees received for managing or conducting outbound telemarketing campaigns to landline phone numbers. Inbound call-center revenues are recorded here, rather than below, as part of commerce and fulfillment, because these costs, such as telephone line charges and wages for sales associates needed to answer live calls, are normally predicted by the marketer based on expected response rates to outbound programs, and budgeted as part of that program.

The next set of eligible practices represents revenues accrued by retailers and similar sales channels for services provided to marketers when a consumer is in the process of purchasing a product or service, or when the order is filled:

- *Loyalty Program Services.* Revenues accruing to sales channels from marketers for the use of ILCD in a loyalty program or for being able to target customers at the point of sale by recognizing them from their loyalty card. These revenues accrue to brick-and-mortar retailers, airlines, hotels, and others that maintain or outsource a loyalty card program. The challenge faced here is to estimate the amount a retailer receives from marketers, as opposed to amounts received from consumers; we also have to adjust for costs incurred as fees paid to outsourced loyalty services vendors.
- *Commerce Services.* The amount that we record here represents net revenues for sales commissions, from marketplaces or revenues attributable to order-taking, upselling, credit-card processing, etc. For ecommerce retailers who merchandise, we seek the incremental value provided by the services' data capabilities, net of cost of goods sold.
- *Performance, SEO, and Affiliate Marketing Services*. These services are often performed for ecommerce retailers.
- *Fulfillment Services.* This includes the revenues for pulling, handling, warehousing, and delivering any parcels shipped to end customers in response to orders generated by an eligible DDM practice. The revenues exclude any shipping or handling charges borne by the customer themselves.

Together, the values in each of these categories sum to the base of net DDME revenues from which they were derived.

A1.4 Aggregating Revenues and Employment for All Firms

Our final task was to aggregate the net revenues of each firm that we had just allocated by practice area, while making adjustments for revenues that were double-counted within the DDM economy.

This third step involved identifying each firm's business model and arranging them into subsystems that would maximize internal similarity and external dissimilarity. To do so,

we asked: to whom does the firm sell? From whom do they buy? How did they receive compensation for what they produced? Business model categorization was critical in understanding the value a firm added, since deductions for cost-of-DDME practices depended in part on knowing where the firm's segment was located in the value stack. By using general principles derived from both our desk and interview research, we estimated the percentage of each business model's revenues that were due to the purchase of data-practice services from elsewhere in the system, and applied these percentages to all firms in that grouping, along with a conservative estimate of the number of "all other" firms in the category.[75]

Ultimately, once these subtractions were made for the cost of goods sold, the sum of all adjusted revenues for all individually investigated firms, plus the sum of all the "all other" firm estimates, represented the DDME's net value added.

Employment Calculations

For employment, the analytic steps are essentially similar to those above for revenues, using similar criteria for inclusion and allocation. The crucial issue we paid attention to was the portion of net DDME employment that was U.S. vs. global. Also, we assessed employees per million dollars of net revenue from similarly placed public firms, in order to use employment estimates of privately held firms as a quality check on our assumptions regarding these firms' revenues.

A1.5 Calculating Dependence on the Exchange of ILCD

Our second question addresses the value-added contribution of marketable data and data services within the data-driven marketing economy. Answering it required us to describe the key players in the market for ILCD and ILCD services, and how these ILCD services integrate into a process whose added value can be summed and apportioned according to market forces.

In this chapter we outline the logic by which we applied on a firm-by-firm, segment-by-segment basis to arrive at what we could regard as reasonably accurate answers to our two key research questions.

6.6 The Players: First and Third Parties; Data Owners; Users and Servicers

[75] Secondarily, even where we had good reports on firm revenues, knowing how they share function and media, along with knowing their business model was often useful to estimate the portion of their revenue or employment that were to be included in the DDME, or attributed to a particular DDM practice or interactive media. For example, some types of publishers report company revenues that include direct-to-consumer revenues, such as eg magazine publishers' print-copy newsstand sales, and offline or online subscriptions (the digital pay wall, for those segments that have them);). All of these must be excluded to arrive at a firm's net DDME revenues, in order to leave only revenues sourced directly from marketers (advertising, list rentals, etc.).

To meet the challenge raised by Mandel and others while still using market methods to value ILCD, our approach implicitly treats ILCD as a tangible good that can be marketed or distributed according to the basic principles that apply to any form of intellectual property protected by copyright law. At the same time, we also treat it as the basis for providing intangible services – literally, various kinds of data processing performed for a fee, or in the expectation of some compensation.[76]

To understand the function and efficiencies achieved by the market for individual-level consumer data, we use the concepts of data subjects, data owners, first and third parties to supplement the idea of a value chain of marketing services expenditures.

Our investigations suggested that the flow of ILCD begins with the collection of customer information by an organization (such as a publisher, marketer, or public entity) with which they have a direct relationship. In the industry, these firms are normally referred to as first parties, and the industry regards a firm's customer records, their ILCD, as their intellectual property. Subject to laws and customs there is latitude as to how much customer data to collect, store, analyze or transfer. Often, however, marketers were themselves the end-user of their own data, as they leverage it to create and nurture ongoing marketing relationships with their existing customers and proprietary prospects.

It is on the transfer to third parties that we focused to identify the exchange market for data, whether as a good, or as the basis for a service. First parties can transfer ILCD to two kinds of third parties, whom the participants in the data exchange recognize as having different rights with respect to the transferred data. For convenience, we call these two types of third party *servicers* and *users*:

- A third-party *servicer* is distinct from the first party: a marketing supplier to whom a first party has outsourced the execution of one of its proprietary business processes. To perform the delegated activity, the servicer takes temporary possession of the ILCD, or gains temporary access to it, but solely to carry out the first party's stipulated business purpose. As servicer, the third party receives compensation for the service performed on or with the data, but the servicer receives no intellectual property in the data itself. An example of a third-party servicer is a printer who receives a copy of a marketer's customer file in order to print personalized direct mail. The printer, who is required to maintain the file securely and confidentially, would normally return or destroy all copies after the printing was complete, and under no circumstances could the printer use the list to market its own company, or make the list available to other third parties.

[76] Data can be measured as a marketable good when copies can be obtained from the holder according to terms that it sets and compensation it receives in exchange. Data as a good can take the form of consumer content, such as a copy of a ringtone downloaded onto a mobile phone for a fee. Likewise, it can be a business asset, such as a copy of a golf magazine's mailing list that a golf resort might rent. As the basis for a service, such as data processing, data can also be the basis for enabling content delivery. For instance, consider the servers owned by Internet service providers that route the packets of data that construct a website to the browser of a viewer who requests them. Or, consider the example of a marketing agency that uses a statistical model to derive propensity scores about the purchase habits of a catalog company's customers, and then appends those scores to the catalog's database.

- A third-party *user* is a firm who licenses a copy of, or access to, some or all of a first party's ILCD for its own business benefit in exchange for compensation.[77] An example of a third-party user is a catalog company that has rented a magazine publisher's list of subscribers, or a data compiler that obtains public records of fishing licenses (and thereby acquires title to the data) so that "fishing" can be appended as a demographic or lifestyle attribute in its syndicated prospect database. Access to this prospect database with its numerous compiled attributes can then be licensed onwards to other third-party users, again for a fee.

How does this classification of firms into first and third parties, owners and licensers, servicers, and users help us arrive at a sharper evaluation of the efficiency and functioning of the DDME? Essentially, these concepts help allocate the value-added of data created in the marketplace based on fundamental market dynamics. In other words, it this allows us to give credit where credit is due, based on the market value (monetary or otherwise) as perceived by different classes of market participants.

Once we recognize that data has market value in more than just its raw, tangible-goods form, our research can reveal, among other insights, how much of the value-added of ILCD is hidden in the profit margin of firms that transform it into high value processing services. For these firms, unprocessed data is an important and often indispensable input. Some high-margin services include, for example, appending, analyzing, scoring and modeling data to accurately and efficiently select and reach individual customers, and to measure their response. These tasks are most often performed by direct and customer relationship management (CRM) agencies and are often directly dependent on the servicer's ability to purchase or access data from third parties.

This approach also throws light on the role of third-party servicers firms, whose value-added contribution to the DDME is largely generated indirectly by the market for data among other parties in the DDME (i.e., first parties and third-party users). This is particularly the case for those activities where a marketer as third-party data user has purchased data or data services on their own initiative, and has incorporated the data so it is recognized as part of their own proprietary data, such as when they outsource the data to a servicer. One concrete example of this include when an agency is asked to analyze or score a marketer's database that includes names the marketer has rented. Similarly, a printer may be retained to produce a new direct mail campaign using a client's file that contains both the client's own list as well as a list licensed from a syndicated prospect database.

In summary, our goal is to arrive at the follow values, for both revenues and employment:

[77] In most instances, it is appropriate to speak of "licensing" or "rental " of data, rather than selling data, since the first party normally retains full possession and copyright of the data after its sharing with a third- party.

Table 39: Sample Summary and Breakout of DDME Valuations

	Total Net (value added) contribution to the total value of the data-driven marketing economy	**Of which:** 1. Value added by services that depend directly on data exchanged or rented among firms	2. Value added by services that indirectly depend on data exchanged or rented among firms	3. Remaining value added by services that do not depend on data exchanged or rented among firms because it is generated and captured within single firms
$ Revenues				
# Employees				

A1.8 Types of Dependence of DDME Revenues on the Exchange of Data

Answering our second set of questions required us to assess how much of a segment's typical distribution DDME revenue and employment was specifically attributable to services that directly or indirectly relied upon any given set of data flows.

Dependence is classified into two kinds.

- Direct dependencies are those where at least some proportion of the net value-add arose from the sale of data-goods or data services.

- Indirect dependencies are revenues arising from third-party services provided to other segments up- or downstream in the value chain who themselves had used data from third-parties.

All other practices or services we regarded as neither directly nor indirectly dependent on the exchange of data across firm boundaries.

Revenues from directly dependent DDME practices include:

- *Prospect List Rentals*. This may be in the form of rental income to a list owner, or fees paid to a manager who markets the list on the owner's behalf (this is paid by the list owner). It also includes the fees paid to the list broker, who rents the list on behalf of a client; licensing revenues to syndicated prospect databases containing aggregated / compiled lists; and the estimated value of revenue received from participation in a co-operative database. This latter estimate is necessary, because no monetary transaction takes place.

- *Data Appending*. This refers to any service in which a marketer or advertiser acquires new fields of data that update or are added to existing ILCD customer or audience records that do not entirely originate with that marketer or advertiser. This may include demographic or interest data, information classifying the individual by purchase propensity (scores from scoring models, if the model relied on third-party data); location data; and data for database hygiene such as NCOA. This also includes appending addresses to existing databases if the appended data is obtained from a third party. Merger of records across separate channels entirely owned by marketer would be counted as part of first-party database services.

- *Merging of Online / Offline Identities.* Services that provide this involve collecting information about customers, prospects, and audience members in one media and merging with a database, or using it to communicate with the same individual in another media. An example of this is the practice of identifying online site visitors by cross-referencing anonymous or hashed email addresses with a direct mail database.

- *Behavioral Advertising / Retargeting* Behavioral advertising refers to placing third-party cookies (or using similar device identifiers), e.g., in the browser of a site visitor by a third party, so that visitors can tracked and / or advertised to when they visit other sites, as done by ad networks or ecommerce sites. When a first-party (publisher or ecommerce) cookie is used to track the consumer to a new publisher site, the first-party cookie becomes a third-party cookie, and the practice is retargeting.

- *Affiliate Marketing Revenues*. Revenues earned by marketers for providing a marketing opportunity to another marketer. This would include, for example, incorporating offers within an ecommerce site and sharing ILCD data in the process.

For each affected supplier firm, we identified the amount of revenue and employment directly attributable to exchanged services, and then anonymously summarized it by business segment. The sum of such revenues and employment from across all affected ecosystem segments represents the value added within the U.S. economy directly attributable to the compensated exchange of data among firms.

The disaggregation of net DDM practice revenues then enables us to evaluate the revenues of the remaining practices, whether as a cost of the input, or as a driver of demand for services. This serves as the basis for our analysis of revenues that are attributable to practices where data is purchased as a good and then transformed into a bundled or processed service.

The categories of indirectly dependent services therefore include any revenues for performing services on a client's data where at least some of client's data has been acquired from third parties for the client's own use.

- Analytics, Scoring, Segmentation Servicing
- Acquisition Campaign / Media Production or Delivery
- Data-Enhanced House file Campaign Production or Delivery
- Sales Channel Services

All other revenues that are not directly or indirectly dependent on revenues exchanged across firm boundaries remain part of the ecosystem, but their precise monetary value cannot be determined without a price mechanism or its equivalent in exchange.

Appendix 2: Summary of Employment by State

Here we provide two break outs of U.S. DDME economic activity by all fifty states, plus the District of Columbia.

In the first breakout, we assume that DDME revenues are derived from marketers uniformly throughout the country, and therefore should be apportioned to the individual states by their share of national GDP.

Table 40 DDME Revenues By State Share of GDP

	State	% GDP	DDME Value Added Revenue	DDME Revenue Dependent on Data Exchange Among Firms
1	California	13.34	$21,000	$15,000
2	Texas	8.92	$14,000	$10,000
3	New York	7.68	$12,000	$8,000
4	Florida	5.2	$8,000	$6,000
5	Illinois	4.44	$7,000	$5,000
6	Pennsylvania	3.97	$6,000	$4,000
7	New Jersey	3.42	$5,000	$4,000
8	Ohio	3.33	$5,000	$4,000
9	Virginia	2.95	$5,000	$3,000
10	North Carolina	2.81	$4,000	$3,000
11	Georgia	2.79	$4,000	$3,000
12	Massachusetts	2.6	$4,000	$3,000
13	Michigan	2.57	$4,000	$3,000
14	Washington	2.42	$4,000	$3,000
15	Maryland	2.07	$3,000	$2,000
16	Indiana	1.84	$3,000	$2,000
17	Minnesota	1.84	$3,000	$2,000
18	Arizona	1.8	$3,000	$2,000
19	Colorado	1.79	$3,000	$2,000
20	Wisconsin	1.73	$3,000	$2,000
21	Tennessee	1.72	$3,000	$2,000
22	Missouri	1.7	$3,000	$2,000
23	Connecticut	1.61	$2,000	$2,000
24	Louisiana	1.47	$2,000	$2,000
25	Alabama	1.2	$2,000	$1,000

26	Oregon	1.16	$2,000	$1,000
27	South Carolina	1.13	$2,000	$1,000
28	Kentucky	1.11	$2,000	$1,000
29	Oklahoma	1.11	$2,000	$1,000
30	Iowa	1.01	$2,000	$1,000
31	Kansas	0.89	$1,000	$1,000
32	Nevada	0.88	$1,000	$1,000
33	Utah	0.81	$1,000	$1,000
34	Arkansas	0.73	$1,000	$1,000
35	District of Columbia	0.72	$1,000	$1,000
36	Mississippi	0.68	$1,000	$1,000
37	Nebraska	0.62	$1,000	$1,000
38	New Mexico	0.52	$1,000	$1,000
39	Hawaii	0.47	$1,000	$1,000
40	West Virginia	0.46	$1,000	$1,000
41	Delaware	0.43	$1,000	$0
42	New Hampshire	0.42	$1,000	$0
43	Idaho	0.38	$1,000	$0
44	Maine	0.37	$1,000	$0
45	Rhode Island	0.34	$1,000	$0
46	Alaska	0.31	$0	$0
47	South Dakota	0.27	$0	$0
48	Wyoming	0.26	$0	$0
49	Montana	0.26	$0	$0
50	North Dakota	0.23	$0	$0
51	Vermont	0.18	$0	$0
	Total United States	**100**	**$156,000**	**$110,000**

Table 41 DDME Employment by State Share of GDP

	State	% GDP	DDME Employment	DDME Employment Dependent on Exchange of Data Among Firms
1	California	13.34	89,000	63,000
2	Texas	8.92	60,000	42,000
3	New York	7.68	51,000	36,000
4	Florida	5.2	35,000	25,000
5	Illinois	4.44	30,000	21,000
6	Pennsylvania	3.97	27,000	19,000
7	New Jersey	3.42	23,000	16,000
8	Ohio	3.33	22,000	16,000
9	Virginia	2.95	20,000	14,000
10	North Carolina	2.81	19,000	13,000

11	Georgia	2.79	19,000	13,000
12	Massachusetts	2.6	17,000	12,000
13	Michigan	2.57	17,000	12,000
14	Washington	2.42	16,000	11,000
15	Maryland	2.07	14,000	10,000
16	Indiana	1.84	12,000	9,000
17	Minnesota	1.84	12,000	9,000
18	Arizona	1.8	12,000	9,000
19	Colorado	1.79	12,000	8,000
20	Wisconsin	1.73	12,000	8,000
21	Tennessee	1.72	12,000	8,000
22	Missouri	1.7	11,000	8,000
23	Connecticut	1.61	11,000	8,000
24	Louisiana	1.47	10,000	7,000
25	Alabama	1.2	8,000	6,000
26	Oregon	1.16	8,000	5,000
27	South Carolina	1.13	8,000	5,000
28	Kentucky	1.11	7,000	5,000
29	Oklahoma	1.11	7,000	5,000
30	Iowa	1.01	7,000	5,000
31	Kansas	0.89	6,000	4,000
32	Nevada	0.88	6,000	4,000
33	Utah	0.81	5,000	4,000
34	Arkansas	0.73	5,000	3,000
35	District of Columbia	0.72	5,000	3,000
36	Mississippi	0.68	5,000	3,000
37	Nebraska	0.62	4,000	3,000
38	New Mexico	0.52	3,000	2,000
39	Hawaii	0.47	3,000	2,000
40	West Virginia	0.46	3,000	2,000
41	Delaware	0.43	3,000	2,000
42	New Hampshire	0.42	3,000	2,000
43	Idaho	0.38	3,000	2,000
44	Maine	0.37	2,000	2,000
45	Rhode Island	0.34	2,000	2,000
46	Alaska	0.31	2,000	1,000
47	South Dakota	0.27	2,000	1,000
48	Wyoming	0.26	2,000	1,000
49	Montana	0.26	2,000	1,000
50	North Dakota	0.23	2,000	1,000
51	Vermont	0.18	1,000	1,000
	Total United States	**100**	**676,000**	**478,000**

In a second distribution of revenues and employees by state, one which uses the firm's headquarters as the basis of assigning economic activity we discover a strong distribution of revenues and employment in the states of California, the District of Columbia, New York, Washington and Texas, where 88% of the total DDME employees are attributed and over 66% of the DDME revenues are attributed[78].

Twenty-four states have over 1000 employees in the DDME economy and 18 states have over $1 Billion in revenues associated with DDME.

Table 42 DDME Revenue Activity by State of Firm HQ

	State	DDME Value Added Revenue ($ Millions)	% of U.S. DDME Revenue	DDME Revenue Dependent on Data Exchange Among Firms ($ Millions)	% of DDME Revenue Dependent on Exchange of Data Among Firms
1	California	$49,000	31.5%	$25,000	22.6%
2	District of Columbia	$26,000	16.4%	$18,000	16.4%
3	New York	$12,000	7.9%	$11,000	10.2%
4	Washington	$10,000	6.4%	$5,000	4.7%
5	Texas	$5,000	3.5%	$4,000	3.6%
6	Illinois	$3,000	1.8%	$2,000	2.2%
7	Massachusetts	$3,000	1.7%	$3,000	2.3%
8	Ohio	$2,000	1.3%	$2,000	1.8%
9	South Carolina	$2,000	1.1%	$1,000	1.3%
10	Wisconsin	$1,000	0.9%	$1,000	1.2%
11	North Carolina	$1,000	0.8%	$1,000	1.0%
12	Georgia	$1,000	0.8%	$1,000	1.3%
13	Pennsylvania	$1,000	0.7%	$1,000	0.9%
14	Arizona	$1,000	0.6%	$0	0.4%
15	Florida	$1,000	0.6%	$1,000	0.6%
16	Minnesota	$1,000	0.5%	$1,000	0.6%
17	Colorado	$1,000	0.5%	$1,000	0.6%
18	New Jersey	$1,000	0.5%	$1,000	0.6%
19	Tennessee	$0	0.3%	$0	0.5%
20	Virginal	$0	0.3%	$0	0.3%
21	Maryland	$0	0.3%	$1,000	0.5%

[78] In the Summary by State table, employees and revenues are associated with the state in which a company's U.S. headquarters resides, or with the primary U.S. office. Additionally, summarized rows without a state are allocated in the "Other" row, however, the distribution of these employees and revenues can be reasonably expected to match the distribution in other states.

22	Nebraska	$0	0.3%	$0	0.4%
23	Michigan	$0	0.2%	$0	0.3%
24	Connecticut	$0	0.2%	$0	0.1%
25	Oregon	$0	0.2%	$0	0.2%
26	Indiana	$0	0.1%	$0	0.2%
27	Iowa	$0	0.0%	$0	0.0%
28	Oklahoma	$0	0.0%	$0	0.0%
29	Kentucky	$0	0.0%	$0	0.0%
30	Missouri	$0	0.0%	$0	0.0%
31	Vermont	$0	0.0%	$0	0.0%
32	Nevada	$0	0.0%	$0	0.0%
33	New Hampshire	$0	0.0%	$0	0.0%
34	Rhode Island	$0	0.0%	$0	0.0%
35	Alabama	$0	0.0%	$0	0.0%
36	Kansas	$0	0.0%	$0	0.0%
37	Mississippi	$0	0.0%	$0	0.0%
38	Utah	$0	0.0%	$0	0.0%
39	Unclassified	$32,000	20.71%	$28,000	25.05%
	Total United States	**$156,000**	**100.0%**	**$110,000**	**100.0%**

Table 43 DDME Employment by State of Firm HQ

	State	DDME Employment	% of U.S. DDME Employment	DDME Employment Dependent on Exchange of Data Among Firms	% of U.S. DDME Employment Dependent on Exchange of Data Among Firms
1	District of Columbia	466,000	68.9%	241,000	50.3%
2	California	61,000	9.0%	51,000	10.7%
3	New York	32,000	4.8%	48,000	10.1%
4	Washington	23,000	3.5%	17,000	3.6%
5	Texas	11,000	1.6%	13,000	2.8%
6	Illinois	7,000	1.1%	9,000	1.8%
7	North Carolina	5,000	0.8%	7,000	1.4%
8	Massachusetts	5,000	0.7%	8,000	1.7%
9	Ohio	3,000	0.5%	5,000	1.0%
10	Wisconsin	3,000	0.4%	4,000	0.8%
11	Pennsylvania	3,000	0.4%	4,000	0.8%
12	New Jersey	3,000	0.4%	4,000	0.9%
13	Michigan	2,000	0.3%	1,000	0.3%
14	South Carolina	2,000	0.3%	3,000	0.6%
15	Colorado	2,000	0.3%	1,000	0.3%
16	Georgia	2,000	0.3%	3,000	0.7%
17	Florida	1,000	0.2%	2,000	0.4%
18	Nebraska	1,000	0.2%	2,000	0.3%

19	Arizona	1,000	0.2%	1,000	0.1%
20	Minnesota	1,000	0.1%	1,000	0.3%
21	Maryland	1,000	0.1%	2,000	0.3%
22	Virginal	1,000	0.1%	1,000	0.2%
23	Tennessee	1,000	0.1%	1,000	0.2%
24	Indiana	1,000	0.1%	1,000	0.2%
25	Connecticut	-	0.1%	-	0.1%
26	Oregon	-	0.1%	-	0.2%
27	Utah	-	0.1%	-	0.0%
28	Iowa	-	0.0%	-	0.0%
29	Alabama	-	0.0%	-	0.0%
30	Missouri	-	0.0%	-	0.0%
31	Oklahoma	-	0.0%	-	0.0%
32	Kentucky	-	0.0%	-	0.0%
33	New Hampshire	-	0.0%	-	0.0%
34	Vermont	-	0.0%	-	0.0%
35	Nevada	-	0.0%	-	0.0%
36	Rhode Island	-	0.0%	-	0.0%
37	Kansas	-	0.0%	-	0.0%
38	Mississippi	-	0.0%	-	0.0%
39	Unclassified	37,000	5.5%	47,000	9.8%
	Total United States	**676,000**	**100.00%**	**478,000**	**100.00%**

About The Authors

John Deighton is the Harold M. Brierley Professor of Business Administration at Harvard Business School. He is an authority on consumer behavior and marketing, with a focus on digital and direct marketing. He initiated and has led the HBS Executive Education program in digital marketing as well as the elective MBA course, Digital Marketing Strategy.

His current research deals with digital and direct marketing. He has studied the ecosystems of the Internet economy, the interplay of social media and conventional television ratings systems, the propagation of viral videos online and offline, and database marketing.

His research on marketing management and consumer behavior has been published in a variety of journals including the Journal of Consumer Research, the Journal of Marketing Research, the Journal of Marketing, Organizational Behavior and Human Decision Processes, and the Harvard Business Review. His research has also received a number of commendations, including the American Marketing Association's Best Article Award for an article in the Journal of Marketing and an honorable mention from the Journal of interactive Marketing. He received the European Case Clearing House Award in Marketing (2012), the Edward N. Mayer, Jr. Award for Education Leadership (2011), the Direct Marketing Education Foundation Robert B. Clarke Outstanding Educator Award (2002), the University of Chicago's Hillel J. Einhorn Excellence in Teaching Award (1995). He has been a visiting scholar at the University of Tokyo, Duke University's Fuqua School of Business, and the Judge School of Business at Cambridge University.

He is the immediate past editor of the Journal of Consumer Research, a leading outlet for scholarly research on consumer behavior, and was the founding co-editor of the Journal of Interactive Marketing, which reports academic research on marketing and the Internet. He is the Executive Director of the Marketing Science Institute, a board member of the Direct Marketing Education Foundation, and a Director of the Berkman Center for Internet and Society at Harvard University. He has been with HBS since 1994 and received the Greenhill Award for outstanding service to the school.

Prior to joining HBS, he was on the faculties of the University of Chicago and the Tuck School of Business (Dartmouth College). He has a Ph.D. in Marketing from the Wharton School, University of Pennsylvania and an MBA from the University of Cape Town. He also has a B.Sc. in Chemical Engineering from the University of Natal. His applied research includes consulting with a number of U.S and international corporations.

Peter A. Johnson is President and CEO of m-Lightenment and Adjunct Professor at Columbia University. Prior to launching m-Lightenment, Peter was Vice President of Market Intelligence and Strategy for the Mobile Marketing Association, where he led its world-wide market research initiatives. Before that, he was Vice President of Research at the Direct Marketing Association. After receiving his PhD from Cornell University, Peter became a full time assistant professor at Columbia University in New York City

from 1991 to 2001, and subsequently remained affiliated as adjunct professor. He has also been a visiting professor at McGill University and Bryn Mawr College, and has guest lectured at the business schools of Columbia, Villanova, Pace, Baruch, and New York Universities, and the Union Institute.

For the last dozen years, Dr. Johnson's professional activities and research have focused on the economic impact and effectiveness of emergent marketing media, especially those used in direct, interactive, social, and mobile marketing; and how their effectiveness is affected by consumer privacy behavior, marketer best practices, and US and international privacy regulations.

A frequent invited panelist and key-note speaker at policy and business conferences, he has testified to the US Congress on legislation affecting Internet marketing, and co-authored a brief to the US Supreme Court on the value of consumer data to the marketing industry. He has served on the Board of Directors of the US Media Ratings Council and is a former member of the Board of Directors of the Marketing Accountability Standards Board. He has held numerous academic awards and research positions, including at the Universities of Freiburg and Heidelberg, Germany; the Institute for Human Sciences in Vienna, Austria; the US Federal Reserve; and the US Library of Congress.

About the Direct Marketing Institute

The Direct Marketing Association is the world's largest trade association dedicated to advancing and protecting responsible data-driven marketing. Founded in 1917, DMA represents thousands of companies and nonprofit organizations that use and support data-driven marketing practices and techniques.

About the Data-Driven Marketing Institute

The Data-Driven Marketing Institute, an initiative of the Direct Marketing Association, conducts independent, academic research that answers vital questions about how data is changing our way of life. DDMI also provides the platform for academic research to inform policy debates surround the use of data to power data-driven marketing and innovation. In short, DDMI demonstrates that being pro-data is being pro-consumer.