The Rapid Adoption of Data-Driven Decision-Making

Citation: Brynjolfsson, Erik and McElheran, Kristina. "The Rapid Adoption of Data-Driven Decision-Making." American Economic Review 106, no. 5 (May 2016): 133–139. © 2016 American Economic Association

As Published: http://dx.doi.org/10.1257/aer.p20161016

Publisher: American Economic Association

Persistent URL: http://hdl.handle.net/1721.1/108650

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

American Economic Review: Papers & Proceedings 2016, 106(5): 133–139
http://dx.doi.org/10.1257/aer.p20161016

DIGITIZATION AND INNOVATION ‡

The Rapid Adoption of Data-Driven Decision-Making†

By Erik Brynjolfsson and Kristina McElheran*

Recent years have seen dramatic changes in data storage and processing technologies. New opportunities to collect and leverage data have led many managers to change how they make decisions—relying less on intuition and more on data. As Jim Barksdale, the former CEO of Netscape quipped, "If we have data, let's look at data. If all we have are opinions, let's go with mine."[1] How significant is the emergence of data-driven decision-making (DDD), and who adopts it? In this paper we provide the first systematic empirical study of the diffusion of DDD and the factors influencing its adoption.

We find that the use of DDD in US manufacturing nearly tripled (from 11 percent to 30 percent of plants) between 2005 and 2010. This rapid diffusion is consistent with the higher productivity of DDD adopters identified in a companion paper (Brynjolfsson and McElheran 2016). Yet adoption is uneven. DDD is concentrated in plants with three key advantages: size, high levels of potential complements (particularly information technology and educated workers), and awareness.

Anecdotes abound suggesting that a shift to more data-driven decision-making can improve performance. Practitioner-oriented accounts emphasize that benefits of new data-related technologies are primarily realized through significant changes in management practices (e.g., McAfee and Brynjolfsson 2012). Some econometric evidence also links DDD with superior performance in a modest sample of large public firms (Brynjolfsson, Hitt, and Kim 2011).

This poses a puzzle, however. If DDD is the current "best practice," why don't all firms adopt? Past research points to frictions that may make productivity enhancements slow or costly to implement in certain firms. Awareness of innovative techniques may be difficult to observe and take time to spread (Geroski 2000). Even well-documented advances may depend on costly or subtle complementary adjustments within the firm (Milgrom and Roberts 1990; Blader et al. 2015) or value chain (McElheran 2015), leading to variation in both adoption and performance.

We recently worked with the US Census Bureau to design and field a large scale survey to examine these phenomena in more depth. This paper and our companion study build on the rich literature concerning persistent performance differences in firms[2] to extend our understanding of how firms take advantage of new technologies and which ones are most likely to benefit.

To begin, we shed light on how these new management practices are diffusing among

‡*Discussants:* Kathryn Shaw, Stanford University; Megan Macgarvie, Boston University; Maryann Feldman, University of North Carolina.

* Brynjolfsson: MIT Sloan School of Management, 100 Main Street, Cambridge, MA 02142, and NBER (e-mail: erikb@mit.edu); McElheran: University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6 (e-mail: k.mcelheran@utoronto.ca). Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the US Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

† Go to http://dx.doi.org/10.1257/aer.p20161016 to visit the article page for additional materials and author disclosure statement(s).

[1] http://www.usmedicine.com/editor-in-chief/if-we-have-data-lets-look-at-data-if-all-we-have-are-opinions-lets-go-with-mine/.

[2] See Syverson (2011) for a review. Several prior studies have focused on heterogeneous adoption of information technology (IT) as an explanation (e.g., Brynjolfsson and Hitt 2000 and Dunne et al. 2004), as well as variation in management practices (e.g., Ichniowski, Shaw, and Prennushi 1997; Bloom et al. 2013; Blader et al. 2015), or a combination of the two (e.g., Bresnahan, Brynjolfsson, and Hitt 2002 and Bartel, Ichniowski, and Shaw 2007).

establishments of different types over time. Size—both in terms of employment and belonging to a multi-unit firm—strongly predicts DDD, consistent with economies of scale. We also find evidence that complementarities may be important, as high levels of information technology (IT) and educated workers are correlated with DDD adoption.[3] Finally, the diversity of ways in which plants learn about new management practices strongly predicts adoption, consistent with heterogeneous learning both about particular practices and about what may be necessary for their successful implementation. Yet, the tripling of DDD rates in five years suggests that firms are overcoming barriers to implementation rapidly. This comports with findings in Brynjolfsson and McElheran (2016) that DDD is not only correlated with significantly better performance in a wide range of operational settings, but also exhibits a timing consistent with the causal relationship described in the case literature.

I. Data and Measures

New large-scale data on management practices was released by the US Census Bureau in 2014 (see Bloom et al. 2013). This Management and Organizational Practices Survey (MOPS) was a supplement to the Annual Survey of Manufactures (ASM), which targets roughly 50,000 American manufacturing establishments and provides representative annual coverage of the manufacturing sector.[4] The survey response rate was 78 percent.

While the MOPS was associated with the 2010 ASM, respondents were also asked to report on the state of practices in 2005. Using this quasi-panel structure and linked ASM data from 2005, we explore correlates of changes in DDD over the five-year sample period. In addition, we explore correlations in the full 2010 cross section which contains roughly 34,000 plants. Linking to the 2005 ASM yields a balanced panel of roughly 18,000 observations.[5]

A. *Measures of Data-Driven Decision-Making*

Our qualitative interviews with plant managers emphasize the usefulness of data for managing their operations. To investigate this systematically, the MOPS asked respondents to choose a value on a 5-point scale according to "what best describes the availability of data to support decision making at this establishment," and "what best describes the use of data to support decision making at this establishment." These metrics are highly correlated with each other.

In addition, we leverage information about the number of key performance indicators (KPIs) relating to production, cost, waste, etc. tracked at the establishment. Based on Census' field-testing of the survey instrument as well as our own independent interviews, we interpret the number of KPIs as an indicator of breadth and/or intensity of data gathering at the plant.

Having appropriate targets against which to compare data plays an important role in decision-making (March 1994). Targets contextualize raw data and motivate action in response to the information signal. Based on our qualitative interviews, we interpret the presence of both long-term and short-term targets as indicating a more extensive use of data to monitor and control performance. To meet our definition of DDD, we require that plants report: (i) being in the top two categories for both availability and use of data; (ii) 10 or more KPIs; and (iii) use of both long-term and short term targets.[6]

[3] Brynjolfsson and McElheran (2016) perform formal tests for complementarities with IT investment, worker education, and other aspects of the plant operating environment.

[4] Sampling probability increases with plant size; those with over 1,000 employees are sampled with certainty and accounted for 67 percent of total value shipped in the US manufacturing sector in 2007 (http://www.census.gov/manufacturing/asm/how_the_data_are_collected/index.html).

[5] Exact records counts are suppressed in the interest of disclosure avoidance. We restrict our attention to establishments that have positive value added, positive employment, and positive imputed capital in the ASM. This makes the standard productivity calculations possible and excludes low-quality records that may introduce systematic biases to the estimation. A technical condition for the panel analysis (and to get controls such as age) requires a valid linkage to the Longitudinal Business Database (LBD). In order to keep our sample stable across specifications, we further restrict our analysis to records with complete responses to the data-driven decision-making questions, headquarters status, and a critical mass of the management practices questions (at least five of questions 1, 2, 5, 6, 8, 9, 11, 13, 14, 15, and 16 in the MOPS). See Brynjolfsson and McElheran (2016) for more details.

[6] Relying on this combination of practices to identify DDD is empirically justified by a polychoric principal factor

A rapid adoption of DDD is apparent in our data. Only 11 percent of plants reported reaching our threshold for DDD in 2005, compared to 30 percent in 2010. The tripling of adoption rates occurs in both multi-unit firms and those that have only one plant, but single-unit establishments start and end our sample with less than half the adoption level of their bigger brethren (see Figure 1).

B. *Investment in Information Technology*

Advances in IT have changed what is measurable, analyzable, and communicable within firms. Firms that invest significantly in IT—either concurrently or in prior periods—are likely to have a greater volume of digitized information to draw on. Conversely, firms that shift to being more data-driven are likely to boost their IT to provide better inputs to DDD. To address this, we calculate IT capital stock (hardware and software) for each plant using a perpetual inventory approach (see Brynjolfsson and McElheran 2016 for details).

C. *Structured Management Practices*

Bloom et al. (2013) identify a broad set of management practices correlated with improved performance which they call "structured management." These include some of the same data-oriented metrics that we study, as well as many other practices, including plant-wide dissemination of production targets and performance-based promotion and compensation. To explore the distinction between our more-focused DDD metric and this broader one, we construct a revised structured management index that excludes the data-related measures comprising DDD.[7]

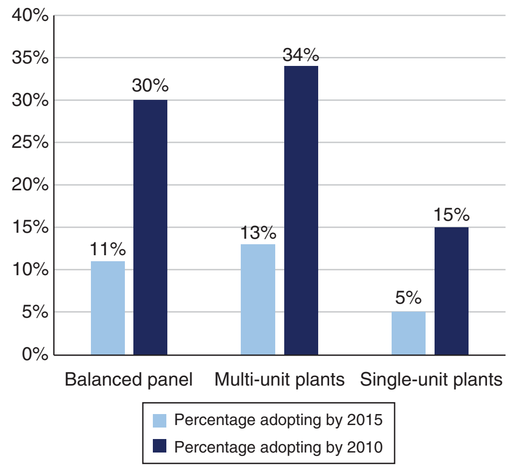

FIGURE 1. ADOPTION OF DATA-DRIVEN DECISION-MAKING IN US MANUFACTURING

D. *Plant Size and Human Capital*

We study two dimensions of size that might affect the likelihood of DDD adoption. The first is total employment at the plant, to account for economies of scale in adoption. The second is an indicator of whether or not the plant belongs to a larger, multi-unit firm. Because effective use of data may depend on higher levels of formal education, we also explore potential complementarities with human capital using the percentage of workers (both managers and non-managers) with bachelor's degrees.

E. *Respondent Characteristics*

Prior work suggests that an important substitute for reliance on objective information might be experience and high levels of tacit information among decision-makers (Porter 1996). We use the reported tenure of the respondent (typically the plant manager) and an indicator of whether or not they are also the CEO to see if this matters for the reported intensity of DDD at the plant.

F. *Learning*

We hypothesize that the many margins of adjustment required for effective DDD may be difficult for firms to discover and implement. To explore whether the patterns in the data

analysis (see Brynjolfsson and McElheran 2016). Applying this technique to these four dimensions yields a single factor with an eigenvalue of 2.28 accounting for 57 percent of the variance in the balanced sample in 2010.

[7] We also eliminate measures that might be particularly confounded with performance, such as the likelihood of reassigning or dismissing underperforming workers quickly and whether performance bonuses were paid in the prior year.

Table 1—Adoption of Data-Driven Decision-Making between 2005 and 2010: Marginal Effects of Probit Regression

Dependent variable:	Indicator of Data-Driven Decision-Making						
	2005 covariates					2010 covariates	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log IT capital stock	0.014***	0.003*	0.003*	0.002	0.007***	0.003	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Structured management	0.064***	0.010					
	(0.021)	(0.021)					
log employment		0.059***	0.040***	0.044***	0.028***	0.057***	0.041***
		(0.004)	(0.004)	(0.005)	(0.006)	(0.004)	(0.004)
Multi-unit status			0.087***	N/A	N/A	N/A	N/A
			(0.010)				
High capital stock (top quartile of four-digit NAICS industry)			0.013	0.014	0.010	0.034***	0.010
			(0.008)	(0.010)	(0.013)	(0.008)	(0.007)
Percent workers with college education			0.058*	0.070*	0.029	0.213***	0.124***
			(0.030)	(0.037)	(0.040)	(0.029)	(0.019)
Respondent reports above-median tenure			−0.052***	−0.052***	−0.048***	−0.015**	−0.011**
			(0.007)	(0.008)	(0.009)	(0.007)	(0.006)
CEO respondent			−0.070***	−0.114***	−0.002	−0.100***	−0.010*
			(0.010)	(0.015)	(0.009)	(0.012)	(0.005)
Number of learning sources			0.017***	0.016***	0.015***	0.029***	0.019***
			(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Sample	Subsample of balanced panel with no adoption of DDD by 2005 ("late" panel)			Multi-unit subsample of "late" panel	Single-unit subsample of "late" panel	Multi-unit subsample of complete 2010 cross section	Single-unit subsample of complete 2010 cross section
Number of establishments	~16,300	~16,300	~16,300	~12,600	~3,700	~22,300	~11,300

Notes: Weighted maximum likelihood probit estimation. Reporting marginal effects calculated at sample means of the covariates. Columns 1–5 have measures from 2005; columns 6 and 7 use 2010 covariates. All columns include controls for age. Robust standard errors are reported in parentheses.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

are consistent with a learning-based diffusion mechanism, we take advantage of a question which asked about the sources from which managers at the plant learn about management practices. Respondents could choose all that apply from: consultants, competitors, suppliers, customers, trade associations or conferences, new employees, and headquarters. We use an index of the number of learning modalities to capture the diversity of pathways by which awareness of DDD may arrive at the firm.

II. Results

To systematically examine the adoption of DDD, we estimate a standard probit model of adoption (David 1969) for two different samples. The first is the sample of plants in our balanced panel that did not clear the threshold for DDD in 2005. The covariates in this analysis come from 2005, allowing us to investigate how early characteristics of the plant may predict later management practices. Columns 1–3 of

Table 1 report the average marginal effects calculated at the means of the covariates; columns 4 and 5 split this by multi-unit status.

The second set of results relies on the complete 2010 cross section with 2010 covariates. Because differences between these two approaches show up mainly when we split the data by multi-unit status, we report only the split: multi-unit plants in columns 6 and single-unit plants in column 7. These findings have the advantage of representing a broader population of plants, particularly the younger, smaller tail of the distribution. Moreover, they reveal that firms were making changes between 2005 and 2010—that are strongly associated with the presence of DDD. We control for industry variation with three-digit NAICS indicators in all specifications in Table 1. Thus, the results should be interpreted as within-industry relationships.

Three key findings emerge. Economies of scale seem to matter—plants with higher employment and those that belong to multi-unit firms are significantly more likely to adopt. Complementary investments may also play an important role; higher IT investment and a greater percentage of educated workers are correlated with DDD, particularly when we look at the high end of the distribution. Thirdly, awareness of these practices and how to implement them has not reached saturation: a greater number of learning modalities has a strong association with DDD adoption.

The first row of Table 1 explores the correlation with IT to test our intuition that this is a technology-led change in practices. As expected, greater investment in IT—both lagged and contemporaneous—is correlated with DDD. The relationship is most consistently significant among single-unit plants, perhaps because this younger population adopted later and may have been more sensitive to advances that made IT more powerful yet less expensive over the time period we observe.

DDD adoption is also correlated with more "structured" management, but the relationship is complex. Using lagged measures in the first column of Table 1, the coefficient is large and significant. However, column 2 suggests that both IT and structured management are correlated with lagged firm size making the marginal relationship difficult to pin down. Moreover, certain structured management practices are sufficiently correlated with both IT use and having educated workers that we exclude it from columns 6 and 7, as well. See Brynjolfsson and McElheran (2016) for a more detailed discussion.

Both total employment and belonging to a larger multi-unit firm have separate and economically large associations with DDD. This relationship holds in both the lagged model (column 3) and in the complete cross section (not reported due to the sample split).[8]

Next, we see that plants with a higher percentage of workers with a college education are more likely to report high levels of data-driven decision-making. The correlation is strongest for multi-unit plants and is even more striking for DDD and education in 2010.

Next we report on features of the plant that may substitute for the use of data and provides some insight into what the alternative to DDD might be. Greater tenure of the respondent is seen in columns 3–7 to be negatively correlated with the adoption of DDD. Similarly, adoption is lower when the CEO fills out the questionnaire, particularly in multi-unit environments. These individuals may simultaneously have high influence on practices at the plant and rely less on data for the types of decisions they make—which may be less routine or well-defined, or perhaps they have less need to rely on the validation of data to imbue their decisions with authority (Porter 1996).

Finally, diffusion of practices will not only be a function of the net benefits of DDD (the assumption underlying the probit model), but also depend on the diffusion of knowledge about practices. Even controlling for many factors at the plant level, firms that learn about new practices from multiple sources are more likely to adopt DDD. These relationships are large enough to be economically important.

Figure 2 shows how each factor contributes to the likelihood of DDD. It begins with the average rate of adoption and layers on the marginal effects from the probit model with lagged covariates (Table 1, column 3; coefficients multiplied by one standard deviation except where indicated) to show the cumulative probabilities for certain types of plants.

[8] While we explore capital intensity as a covariate in Table 1, it is also correlated with size and only has an independent relationship with the likelihood of DDD for the multi-unit plants in column 6.

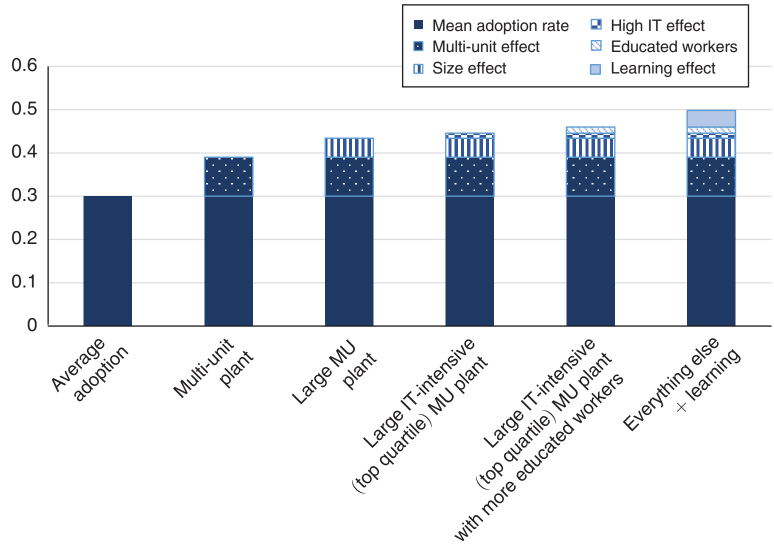

Figure 2. Probability of Adopting DDD by Type of Plant: Marginal Effects of Probit Coefficients

III. Conclusion

Better data creates opportunities to make better decisions. New digital technologies have vastly increased the scale and scope of data available to managers. We find that between 2005 and 2010, the share of manufacturing plants that adopted data-driven decision-making nearly tripled to 30 percent.

Details of DDD adoption patterns reveal that this rapid diffusion is uneven and consistent with three mechanisms that help us to understand the diffusion of management practices, more generally. We find evidence suggesting that economies of scale, complementarities between DDD and both IT and worker education,[9] and firm learning can explain a significant amount of the variation in DDD in recent years.

The rapid diffusion of DDD is consistent with higher productivity from DDD that is found in Brynjolfsson and McElheran (2016). While the effects of DDD are already economically important, there appears to be room for further diffusion of DDD and our model only explains part of the variance. About 70 percent of the plants in our sample had not yet adopted DDD by 2010 and even after controlling for many observable characteristics, there remains significant heterogeneity in the use of DDD. In short, even our very rich window on the phenomenon is still incomplete. A number of potentially salient factors, such as firm culture (e.g., Blader et al. 2015) are beyond easy reach of our data. Our ongoing work aims to uncover other mechanisms that may further explain the adoption and productivity effects of this rapidly diffusing approach to managerial decision making.

[9] More-formal tests for complementarities examine both correlations among practices and their joint correlation with performance (Brynjolfsson and Milgrom 2013). Such tests are conducted in Brynjolfsson and McElheran (2016).

REFERENCES

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw. 2007. “How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills.” *Quarterly Journal of Economics* 122 (4): 1721–58.

Blader, Steven, Claudine Gartenberg, Rebecca Henderson, and Andrea Prat. 2015. “The Real Effects of Relational Contracts.” *American Economic Review* 105 (5): 452–56.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Itay Saporta-Eksten, and John

Van Reenen. 2013. "Management in America." Center for Economic Studies Working Paper 13-01.

Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt. 2002. "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence." *Quarterly Journal of Economics* 117 (1): 339–76.

Brynjolfsson, Erik, and Lorin M. Hitt. 2000. "Beyond Computation: Information Technology, Organizational Transformation and Business Performance." *Journal of Economic Perspectives* 14 (4): 23–48.

Brynjolfsson, Erik, Lorin M. Hitt, and Heekyung Hellen Kim. 2011. "Strength in Numbers: How Does Data-Driven Decision Making Affect Firm Performance?" Unpublished.

Brynjolfsson, Erik, and Kristina McElheran. 2016. "Data in Action: Data-Driven Decision Making in U.S. Manufacturing." Center for Economic Studies Working Paper 16–06.

Brynjolfsson, Erik, and Paul Milgrom. 2013. "Complementarity in Organizations." In *The Handbook of Organizational Economics*, edited by Robert Gibbons and John Roberts, 11–55. Princeton: Princeton University Press.

David, Paul A. 1969. "A Contribution to the Theory of Diffusion." *Memorandum No. 71*, Stanford Center for Research in Economic Growth, Stanford University.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske. 2004. "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment." *Journal of Labor Economics* 22 (2): 397–429.

Geroski, Paul A. 2000. "Models of Technology Diffusion." *Research Policy* 29 (4-5): 603–25.

Ichniowski, Casey, Kathryn Shaw, and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *American Economic Review* 87 (3): 291–313.

March, James G. 1994. *Primer on Decision Making: How Decisions Happen*. New York: Simon and Schuster.

McAfee, Andrew, and Erik Brynjolfsson. 2012. "Big Data: The Management Revolution." *Harvard Business Review* 26 September.

McElheran, Kristina. 2015. "Do Market Leaders Lead in Business Process Innovation? The Case(s) of E-Business Adoption." *Management Science* 61 (6): 1197–1216.

Milgrom, Paul, and John Roberts. 1990. "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economic Review* 80 (3): 511–28.

Porter, Theodore M. 1996. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49 (2): 326–65.

This article has been cited by:

1. Yuanzhu Zhan, Kim Hua Tan, Yina Li, Ying Kei Tse. 2016. Unlocking the power of big data in new product development. *Annals of Operations Research* . [CrossRef]