I/S: A JOURNAL OF LAW AND POLICY FOR THE INFORMATION SOCIETY

Big Data, Big Economic Impact?

ANGELA BYERS

I. INTRODUCTION

Big data has become a main priority for companies in the technology circles. Interest is growing in enterprises that want to harness the power of big data and among consumers who want to benefit from big data-driven applications. There are many reasons to believe that the potential impact is both big and broad-based, as there are a consistent set of levers, across industries that enterprises can apply to create value. The impact is even bigger when economic impact is broadened beyond traditional measures of productivity to include consumer surplus. Such is what my co-authors and I found when the McKinsey Global Institute (MGI) studied this topic in depth a few years ago.[1] However, some critics argue today that big data is just a fad[2] as impact has not materialized in macroeconomic metrics. Based on the adoption history of earlier innovations in information technology, and what it took to drive broad-based impact, we see that technology investment is only one ingredient in that success. Maximizing value from big data will require the right talent, policies, and organizational processes and management.

II. HOW BIG DATA CREATES VALUE

When we studied the economic impact of big data three years ago at MGI, we found that this was a broad phenomenon not limited to a

---

[1] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh and Angela Hung Byers, "Big data: The Next Frontier for Innovation, Competition, and Productivity" (McKinsey Global Institute, 2011).

[2] James Glanz, "Is Big Data an Economic Big Dud?" *New York Times*, August 17, 2013.

few industries. We estimated that, in the US, an average company in any sector has at least 100 terabytes (TB)[3] of data, and many have more than 1 petabytes (PB).[3] For comparison, the library of congress has 235 TB of data in 2011.[4]

While some sectors are poised for more gains, companies across sectors can make use of big data to drive value. Based on in-depth case studies across sectors, we found that there were five common themes.   First, creating transparency to big data often exposes variability in performance and results, leading to changed behavior for more economic impact. Scorecards are a common example. Scoring different units at a company, based on key metrics from large bodies of data, can identify underperforming units and agitate the underperformers to bring their performance up to benchmarks. As another example, increasing visibility into healthcare quality and pricing data, will enable patients to make informed choices about healthcare providers.[5] Second, big data enables experimentation, often involving rigorous statistics analyses to identify what option is better. In the online world I work in, A/B testing[6] is a common way to optimize the design of a site or app, from figuring out where to put different modules or how to drive more engagement and a higher click-through-rate, to what designs to drive more completed transactions in the case of e-commerce.   Another example, in healthcare, is the use of comparative effectiveness studies to determine optimal treatment pathways – determining which treatment works best, for whom, and under what circumstances.

A third theme is using big data to segment populations in order to customize actions. This is well known in areas like risk management and marketing – and even there, there continues to be more sophisticated big data techniques like real-time micro-segmentation of customers to target promotions and advertising.   For example,

---

[3] 1 TB = 10^{12} bytes; 1 PB = 10^{15} bytes.

[4] Leslie Johnston, "How many Libraries of Congress does it take?" *The Signal: Digital Preservation* (blog), *Library of Congress,* March 23, 2012, http://blogs.loc.gov/ digitalpreservation/2012/03/how-many-libraries-of-congress-does-it-take/.

[5] Cost of healthcare can vary widely; as an example, an investigative journalist report reveals a 4x difference in the price of an MRI in the Boston area: Martha Bebinger, "How Much For an MRI? $500? $5,000? A Reporter Struggles to Find Out," *Kaiser Health News*, December 9, 2012, http://www.kaiserhealthnews.org/Stories/2012/ December/09/mri-cost-price-comparison-health-insurance.aspx.

[6] Randomized experiment with two variants A and B, where A and B refer to the control and treatment in the experimentation setup.

insurance carriers can use predictive weather data, census data, and claims data to identify drivers who live in neighborhoods that are likely to experience hail damage and to offer those drivers special hail coverage. Application of this technique can be revolutionary elsewhere – e.g., in the public sector, where tailoring actions or service levels to population segments with different needs can actually make more efficient use of limited resources.

A fourth theme is supporting or even replacing human decision making with automated algorithms. Sophisticated analytics can substantially help businesses to optimize decisions across the value chain, from the upfront part – involving the location and investment optimization of fixed assets like plants and oil fields – to manufacturing process optimization, supply chain optimization, to logistics optimization. The era of the Internet of Things, when machines, equipment and other physical assets are embedded with networked sensors and actuators[7], will open up more possibilities for machines to collaborate and to act on information independently.

Fifth, big data enables innovation in business models – often enabling new service offerings. Location and mapping data pretty much created an entire category of location-enabled services from Foursquare to augmented reality type travel companion app. Financial services is another example, where MasterCard has a subsidiary MasterCard Advisors that combines the consumers spending data they have with analytics to do benchmarking for merchants, or produce timely macro-economic indicators.

III. How Big Can Big Data's Impact Be?

All up, the potential economic impact of big data is quite huge. For example:

- In retail, $30-55 billion of annual savings estimated in the US, through applying big data levers to supply chain, operations, merchandising, with hundreds of billions of potential profit shifts.[8]

- In consumer packaged goods, open data is expected to unlock $520 billion to nearly $1.5 trillion in value worldwide, through

---

[7] Actuators generically refers to parts of a system that produce motion

[8] Susan Lund, James Manyika, Scott Nyquist, Lenny Mendonca, and Sreenivas Ramaswamy, "Game Changers: Five Opportunities for U.S. Growth and Renewal," (McKinsey Global Institute, 2013), 75.

improvements in product design, manufacturing, store operations, marketing, sales, and post-sale services, as well as creating consumer surplus through enabling shopper comparisons of prices, quality and attributes.[9]

- In manufacturing, savings in the range of $125-$270 billion for the U.S. may be achieved by leveraging big data across R&D, production and supply chain management.[10]

- In electricity it is possible to achieve $340 – 580 billion in value globally by using open data to optimize generation investment, make generation operation more efficient, optimize investment in transmission and distribution, influence end consumption (e.g., through the use of more energy-efficient products).[11]

- Similar orders of magnitude for potential improvement apply to healthcare and public sector.

Much of the value cited above translates to productivity gains, or GDP per capita growth. Big data enables organizations to raise their efficiency by optimizing labor, equipment and processes – allowing them to reduce inputs for a given output level. But some of the value mentioned above involves increasing output, more and better output, with a given input level, which tends to be harder to quantify.[12] In fact, big data is predicted to be one of the five game changers to boost growth in the US (along with talent, infrastructure, energy/shale, trade).[13]

---

[9] James Manyika, Michael Chui, Diana Farrell, Steve Van Kuiken, Peter Groves, and Elizabeth Almasi Doshi, "Open data: Unlocking innovation and performance with liquid information," (McKinsey Global Institute, 2013), 11.

[10] Ibid., 47.

[11] Ibid., 11.

[12] For example, it is easier to quantify the impact of big data leading to the price reduction of a drug, than the impact of big data leading to the discovery of a new drug, with equivalent cost to an older drug that has higher efficacy and leads to better quality of life for patients.

[13] Manyika and Chui, *Big Data: The Next* Frontier, 1-2.

Another dimension of impact that is hard to quantify is the disruptive potential of big data[14]. As a result of using big data levers, some companies can transform how things are done, in turn challenging companies that maintain business as usual. There are plenty of examples in the online media world that I work in. For example, digital TV/movies/music and the myriad of online distribution models have been challenging incumbent distributors (CDs, cable) for years. As another example, online publishers are mining consumer signals from what they read, where they are, the social signals they send – e.g., what articles they share, what topics are trending on Facebook and Twitter — to serve up personalized, relevant content while not being too repetitive and predictable, thus automating and surpassing what human editors can do. Reader apps (like Zite, Prismatic) started to innovate in that area, and some more incumbent online media sites are following in response. Personalization extends to other areas, like e-commerce, where Amazon, with its famous recommendation engine and scale and operational excellence, is the single major force challenging brick and mortar and many online retailers.

One other dimension of economic impact that is big but not captured in traditional measurements of the economy is consumer surplus. For example, the use of personal location data in enabling real time traffic information to inform navigation will save consumers time and fuel – estimated at $500 billion world-wide[15], far surpassing the economic value estimated to accrue to service providers of location-enabled services. Other surplus is hard to quantify, such as when a local review app makes a consumer's life better by finding her the best coffee shop or hotel option when she travels. As another example, the price transparency consumers can easily have now from pricing comparison engines, and product recommendations based on the mass of consumers in just about every category of retail, enable consumers to make much better purchase decisions. We are likely to be happier with our purchase choices and save money. In health care, many of the big data-driven levers such as comparative effectiveness research, clinical decision support tools, and personalized medicine,

---

[14] There are many reports and books covering similar themes. A few recent publications include "Big Data: a Revolution that will Transform how we Live Work and Think" by Viktor Mayer-Schonberger and Kenneth Cukier, "The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies" by Erik Brynjolfsson and Andrew McAfee, "The 'Big Data' Revolution in Healthcare: Accelerating Value and Innovation" by McKinsey & Company.

[15] Manyika and Chui, *Big Data: The Next Frontier*, 89.

will improve healthcare quality, in addition to reducing healthcare expenditures[16].

III. BUT WHERE IS THE IMPACT?

Big data impact starts with the individual companies, and as it is adopted by competition, it will start to create the cumulative effect that, when measured with rigorous studies, will show up in macroeconomic metrics. By many indications, big data still appear to be in early days. Gartner's "hype cycle" predicts it to be another five to ten years before big data technologies will reach the plateau of productivity.[17] Companies still have some ways to go in having the big data technologies in place. ABI Research projects a 30% growth year over year in Big Data spending world-wide[18].

However, investment in technology is only one barrier to overcome. Another barrier is the talent shortage, from deep analytical talent and supporting engineers, to big data-savvy professionals. A couple of years ago, we predicted a shortfall of over 1.8 million such professionals in the US in 5 years.[19] Since then, many universities have created cross-disciplinary degree in analytics or data science (encompassing courses in statistics, analytics, computer science and math), which over time will help correct the problem.   But organizations will need to think about how to re-train those already in the work force.

In addition, there needs to be ongoing resolution on data policy issues – such as privacy, cybersecurity, and data ownership issues. While consumers do accrue benefits in many cases, as discussed earlier in this article, when private and public sectors make use of big data, they remain wary about the potential implications of having

---

[16] There are large bodies of research studying the potential benefits. A couple examples for reference: Chandra A., Jena A., Skinner J., "The Pragmatist's Guide to Comparative Effectiveness Research," *National Bureau of Economic Research* (2011); Eslami S, de Keizer NF, Abu-Hanna A., "The Impact of Computerized Physician Medication Order Entry in Hospitalized Patients – a Systematic Review," *Internal Journal of Medical Informatics* (2008). Of course, the realization of the potential benefits will be challenging. See Chapter 3a in "Big Data: The Next Frontier for Innovation, Competition, and Productivity" (McKinsey Global Institute, 2011) for an overview of the barriers to overcome in US healthcare industry.

[17] "Gartner's Hype Cycle Special Report for 2013," *Gartner*, August 2013.

[18] "Unlocking the Value of Big Data in Enterprises," *ABI Research* (August 2013).

[19] Manyika and Chui, *Big Data: The Next Frontier*, 1-2.

their own data collected.[20] While in many developed countries, there are agencies charged with setting data policies, they need to keep pace and evolve as the type of data getting collected and analytics applied to them evolve.[21] Private sectors also need to thoughtfully set and clearly communicate expectations of data use with their customers, as the cost of not doing so could result in significant backlash. Cybersecurity remains a major challenge, with many companies realizing protection and prevention are insufficient, and their mitigation plans are nascent and developing.[22]

Last but not least, even with the technology in place, organizational barriers need to be overcome – how to bring data, analytics, frontline tools, and people together to create business value. There may need to be a plan to re-architect the data to turn silo-ed data sets into unambiguous golden-source data. Organizations will need to identify where models will create additional business value, who will need to use them, and how to avoid inconsistencies and unnecessary proliferation of databases as models are scaled up across the enterprise. They will also need to consider intuitive tools that integrate data into day-to-day processes and translate modeling outputs into tangible business actions: for instance, a clear interface for scheduling employees, fine-grained cross-selling suggestions for call-center agents, or a way for marketing managers to make real-time decisions on discounts.

The last point about changing organization and management practices is critical. More than a decade ago, MGI studied the contribution of information technology investment in the US productivity growth in the late nineties.[23] They found that what matters was not the dollars invested in IT (because some sectors invested a lot in IT, but saw declines in productivity), but changing

---

[20] Consumers' concerns about big data are rightfully so, given examples of the past year with stolen credit card data (e.g., Target's case) and the data collection by NSA.

[21] As an example, recent MIT study highlights that anonymized location data, with as few as four locations and times, can be used to uniquely identify an individual.

[22] In a recent survey of IT executives, 34% rated their cyber risk management as nascent, and another 61% rated it as developing. Only 5% rated it as mature or robust. Tucker Bailey, Andrea Del Miglio, Wolf Richter, "The Rising Strategic Risks of Cyberattacks," *McKinsey Quarterly* (May 2014), http://www.mckinsey.com/insights/ business_ technology/ the_rising_strategic_risks_of_cyberattacks.

[23] Angelique Augereau, et al., "US Productivity Growth 1995-2000: Understanding the Contributions of Information Technology Relative to Other Factors," (McKinsey Global Institute, October 2001).

management practices that translated to impact.   Emerging academic research suggests that companies that use data and business analytics to guide decision making are more productive and experience higher returns on equity than competitors that don't.   Brynjolfsson and his colleagues, for example found that the effective use of data and analytics correlated with a 5 to 6 percent improvement in productivity, as well as higher profitability (6%) and market value[24].

IV. CONCLUSION

Big data has the potential to create tens and perhaps hundreds of billions of value in many sectors, as discussed in this article, values that can translate to productivity improvement for economies, determine new winners in the private sector, and bring consumers benefits and conveniences.   However, the barriers to overcome are just as real as the potential of impact.   Technology investment alone will not be sufficient to realize impact.   In addition, shortages of talent, inadequate management practices, and data policy and security issues will need to be overcome so that the pockets of success we are seeing at the individual organization level can be scaled to make a noticeable effect on the macroeconomic scale.

---

[24] Brynjolfsson, Erik, Lorin M. Hitt, and Heekyung Kim, "Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?" (April 22, 2011) available at SSRN: http://ssrn.com/abstract=1819486.