Big data diffusion in procurement: do organizational characteristics matter?

Andrea S. Patrucco, PhD

Florida International University
College of Business
Department of Marketing and Logistics

Daniel Trabucchi, PhD
Valentina Galante, MSc
Laura Salvadori, MSc

Politecnico di Milano
School of Management
Department of Management, Economics and Industrial Engineering



POLITECNICO
MILANO 1863
SCHOOL OF MANAGEMENT

FIU | Business
FLORIDA INTERNATIONAL UNIVERSITY

Highlights

- Big data use in procurement is an area that has enormous potential, yet to be exploited
- We collect data from 222 procurement and supply chain professionals worldwide about procurement big data use in their organizations
- Findings show that:
  - Big data in procurement are used to support strategic activities (spend analysis, demand planning, and forecasting, performance measurement)
  - Primarily internal sources of information are used to support big data analytics (ERP, historical data)
  - Big data use in procurement contributes to improving performance
  - A higher absorptive capacity of procurement organizations favors more extensive adoption and use of big data

1. Background of the research

In the current era, known as the information age or big data revolution, more data is generated than the amount that is managed, indeed, "*it is believed that the total global data volume doubles in eighteen months*"[1]. While in the past, the recording, storing, and managing of data was complicated today, innovations in information technology and data digitalization allow us to manage and utilize a vast amount of data more efficiently and effectively. In current highly competitive and dynamic markets, big data exploitation would enable companies to obtain a significant competitive advantage, enhancing their decision-making skills and ability to adapt to the environment quickly[2].

The possible applications of big data in the company environment are many and not yet thoroughly explored. The potential exploitation and usage of big data in companies include multiple areas such as customer relationship management, new product development, operations management, and supply chain optimization, and decision-making support. While the application of big data analytics in activities such as sales, marketing, and distribution has been growing over the years – and so has been the scholars' attention – diffusion has been slower in supply chains and supply chain management, particularly in the purchasing and supply management area[3].

Thanks to the increasing relevance of outsourced activities and the strategic role of the supply network, procurement represents a key contributor to market success in almost every supply chain. The purchase of goods and services can represent 40-80% of the revenues in manufacturing companies, thus making this area full of saving opportunities and margin increase in case of optimized decisions[4]. This is why informed strategic decision-making is fundamental, which can be enabled only by complete, structured, and updated data. Despite this, procurement seems to be less mature among the different supply chain processes for what concerns big data applications, even though it is expected to be the most growing area in the following years[5].

For these reasons, the goal of this research project is to better investigate this subject by studying if and how much big data might support procurement managers and contribute to improving procurement performance, and if (and what) procurement organization characteristics can contribute to a more or less intensive adoption. We can formalize the objectives of the study through the following questions:

How do companies use big data to support procurement decisions and activities?

Are there procurement organizational aspects that influence a more or less intense adoption of big data in the procurement area?

2. Survey instrument characteristics

To answer these questions, in 2020, the research team designed a survey with the objective to assess, on an international cross-industry basis, the level and type of use of big data in the procurement area.

---

[1] Kaur, H., & Singh, S. P. (2018). Heuristic modeling for sustainable procurement and logistics in a supply chain using big data. *Computers & Operations Research*, *98*, 301-321.

[2] Sanders, N. R. (2016). How to use big data to drive your supply chain. *California Management Review*, *58*(3), 26-48.

[3] Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, *176*, 98-110.

[4] CAPS Research (2017). The metrics of Supply Management – Cross-Industry 2017 Report. Available at: https://www.capsresearch.org/media/2282/2017-08-31_the_metrics_of_supply_management_-cross-industry-_report_2017.pdf

[5] Tiwari, S., Wee, H. M., & Daryanto, Y. (2018). Big data analytics in supply chain management between 2010 and 2016: Insights to industries. *Computers & Industrial Engineering*, *115*, 319-330.

The survey, which was intended to be answered by procurement professionals (or supply chain professionals aware of the characteristics of their procurement departments), included questions grouped into five sections:

A. **General information** – this section included questions related to the characteristics of the respondent (e.g., role in the company) and their organization (industry, employees, turnover, purchase-to-sales ratio)
B. **Characteristics of the procurement organization** – this section included questions related to the level of centralization of procurement and how procurement professionals use to manage information and knowledge
C. **The role of big data in the procurement process** – this section included questions related to roles defined to collect and analyze data in procurement; the extent of use, the purpose, and the type of sources of big data to support procurement decisions and activities; the contribution of the use of big data to the improvement of procurement performance.

The survey was initially sent out to supply chain professionals with whom the research team had already contacts in place from previous research projects and distributed through other channels (e.g., Linkedin).

2.1 Sample characteristics

We collected 296 responses with completion > 60%; however, 74 needed to be eliminated due to the low reliability of answers and/or missing values on critical questions.
We ended up with 222 usable responses.

The main descriptive characteristics of the companies included in the sample are reported in *Table 1*.

Table 1. Sample characteristics

Descriptive		Freq.	%
Size (employees)			
	> 1,000	84	*37.8%*
	501-1,000	57	*25.7%*
	251-500	36	*16.2%*
	50-250	25	*11.3%*
	< 50	84	*37.8%*
Size ($ million)			
	> 1,000	55	*24.8%*
	501-1,000	26	*11.7%*
	251-500	19	*8.6%*
	51-250	60	*27.0%*
	11-50	41	*18.5%*
	< 10	21	*9.5%*
Industry			
	Manufacturing	154	*69.4%*
	Services	43	*19.4%*
	Projects	23	*10.4%*

Purchase-to-sales ratio		
> 70%	21	*9.5%*
41-70%	82	*36.9%*
31-40%	45	*20.3%*
21-30%	46	*20.7%*
< 20%	28	*12.6%*
Type of procurement organization		
Fully centralized	125	*56.3%*
Hybrid	80	*36.0%*
Fully decentralized	17	*7.7%*
Role for big data management in procurement		
Dedicated role in the procurement department	67	*30.2%*
Category managers/buyers in their day-to-day activities	136	*61.3%*
No one	19	*8.6%*
Total	**222**	***100***

The sample includes mostly medium-big companies, characterized by a high (a single department manages, i.e., all, or most of the purchases for the whole company) or hybrid (i.e., procurement is done by one central department but also by divisions/plants for some specific purchases) level of centralization. The sample primarily includes companies in the manufacturing industries.
In some cases (30.2%), these companies have a dedicated role in the procurement department to systematically collect and analyze procurement data; in most cases (61.3%), these activities are part of the category managers and/or buyers' routine. Surprisingly, only in very few cases (8.6%), the respondents declare that their company does not establish formal responsibilities for structured procurement data collection and analysis.
In this regard, we tried to understand better the sample distribution for what concerns the level of centralization and the responsibility for procurement data (*Table 2*).

Table 2. Level of procurement centralization and responsibility for data collection and analysis

		Responsibility for procurement data collection and analysis			
		Dedicated role	*Category managers/Buyers (as part of their activities)*	*No dedicated role*	
	Centralized	*17.1%*	*35.1%*	*4.1%*	***125 (56.3%)***
	Hybrid	*11.3%*	*20.7%*	*4.1%*	***80 (36%)***

Level of procurement centralization	*Decentralized*	*1.8%*	*5.4%*	*0.5%*	***17 (7.7%)***
		67 (30.2%)	***136 (61.3%)***	***19 (8.6%)***	**222**

As we can see, very few companies do not have a systematic approach for procurement data collection and analysis (8.6% in the sample). Only 1 is a decentralized case (i.e., where the division/plant levels manage all, or almost all, purchases) – an absence of responsibility occurs mostly in centralized/hybrid procurement organizations.

3. Descriptive statistics for the main questionnaire items

3.1 Procurement organization characteristics: absorptive capacity of procurement departments

As explained in the previous section, the questionnaire captured the dimensions of the purchasing organization in two ways: the level of centralization and the approach to managing knowledge and information.

The latter, particularly, is helpful to measure the concept of "absorptive capacity", defined as the procurement organization's ability to "*identify, assimilate, transform, and use external knowledge, research and practice. In other words, absorptive capacity is the measure of the rate at which an organization can learn and use scientific, technological, or other knowledge that exists outside of the organisation itself. It is a measure of an organisations ability to learn.*"[6]

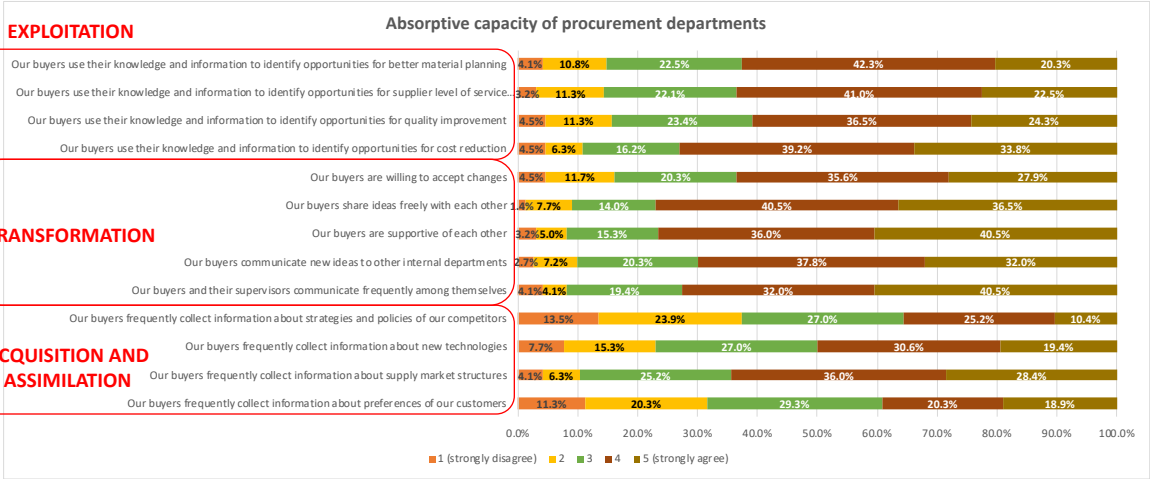
In line with the model provided by Zahra and George (2002)[7], absorptive capacity can be split into sub-dimensions, such as acquisition, assimilation, transformation, and exploitation. For the case of procurement, it measures the procurement ability to acquire and assimilate external supply and customer market knowledge, transform it for the company's purposes, and use it to contribute to the creation of competitive advantage.

Figure 1 summarizes the main statistics of the different components of the absorptive capacity for the companies included in the sample.

Figure 1. Value of absorptive capacity of procurement departments

[6] https://www.oxford-review.com/oxford-review-encyclopaedia-terms/encyclopaedia-absorptive-capacity/

[7] Zahra, S. A., & George, G. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of management review*, *27*(2), 185-203.

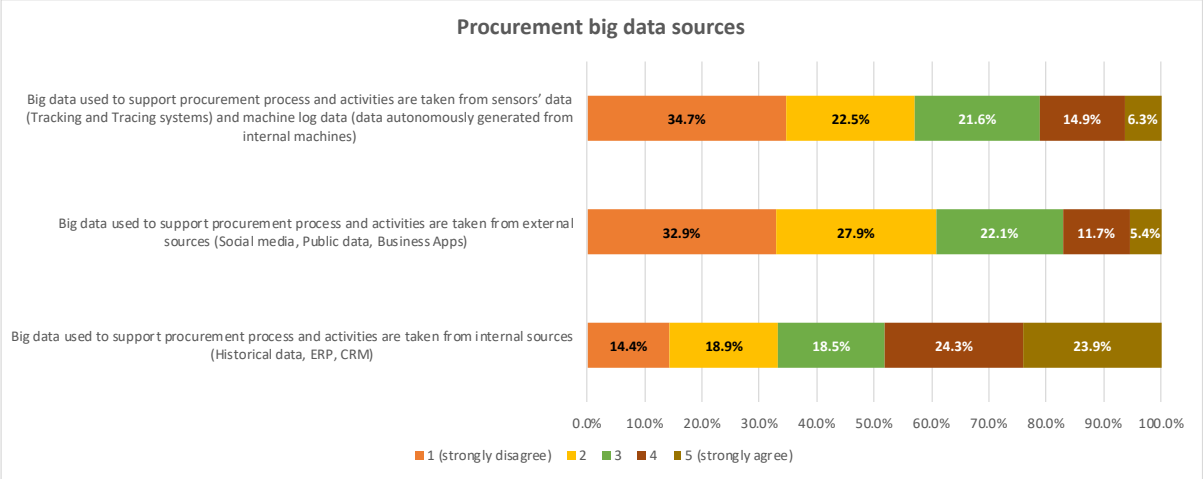


As we can see, the sample includes companies with relatively high values of absorptive capacity, particularly for what concerns **transformation** aspects, which includes communication and integration mechanisms that are essential for appropriate knowledge management and use.

3.2 Purpose and sources of big data in procurement

The questionnaire first aimed to capture the main sources of data used by procurement for the purpose of big data analytics. *Figure 2* reports the distribution of the sample.

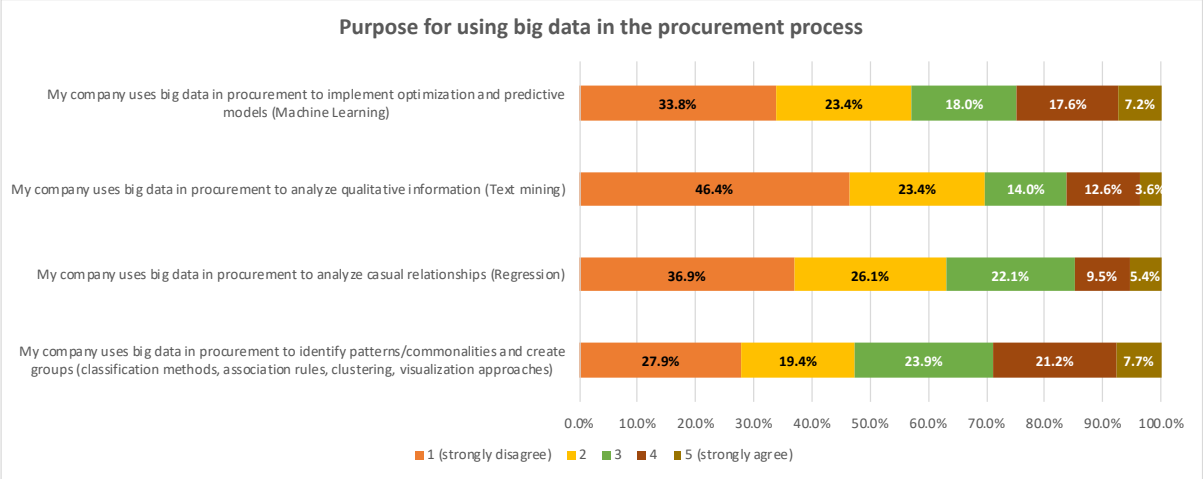
Figure 2. Sources of procurement big data



According to our respondents, most of the sources of procurement big data are **internal** (i.e., data retrieved from internal information systems or historical data included in worksheets); this is an expected result, as these are the sources easier to be accessed, managed, and understood. Instead, much lower is the percentage of respondents that operate in companies that are able to use external sources, sensors, and machine log data – which represent areas for future investments by the companies.

A pattern is also evident for what concerns the main purposes and applications of big data in procurement (*Figure 3*).

Figure 3. Purposes and techniques for analyzing procurement big data

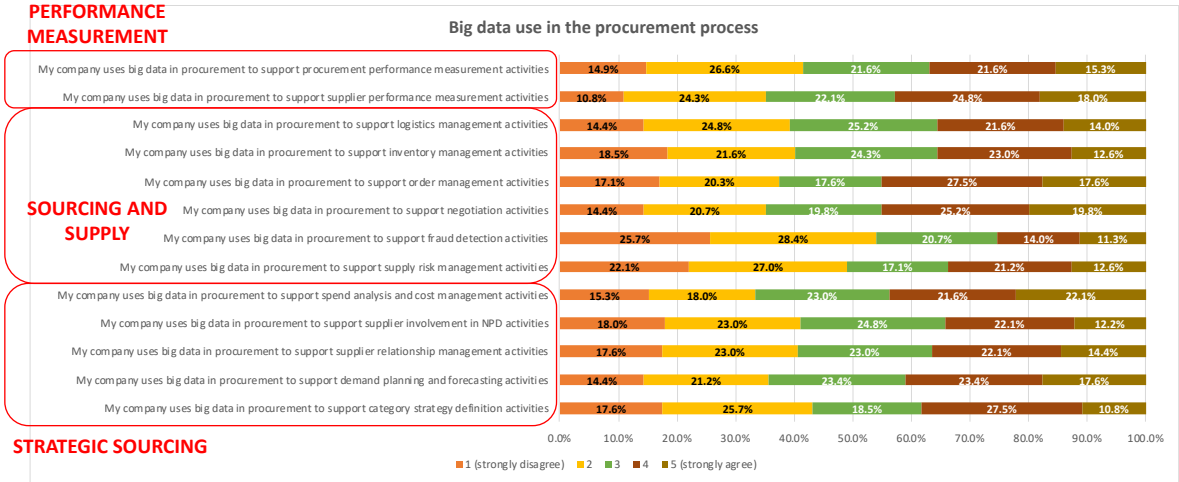


As we can see, the big data use and analysis in procurement are primarily performed with the purpose of creating groups and identify patterns between relevant procurement items (e.g., suppliers, commodities). Cluster analyses and classification approaches are being increasingly adopted in the procurement area, consistently with category management principles, which promotes the idea that procurement and supply management strategies should be consistent across similar groups of items and suppliers.
Instead, analyses that use techniques such as regression, text mining, and machine learning are much less diffused.

3.3 Use of big data in the procurement process

The questionnaire specifically asked the respondents to rate how much big data are used to support the different decisions and activities that characterize the procurement process – which cannot be grouped into 1) strategic sourcing (that includes all the strategic aspects of procurement, such as category strategy definition, supplier relationship management activities, specific supplier collaborations, etc.); 2) sourcing and supply (that includes all the tactical and operational aspects of procurement, such as negotiation and selection, order management, risk management, fraud detection, etc.); 3) performance measurement (that includes all the aspects related to performance measurement and management, for both the procurement process and the supplier). *Figure 4* summarizes the distribution of the sample.

Figure 4. Use of big data for procurement activities

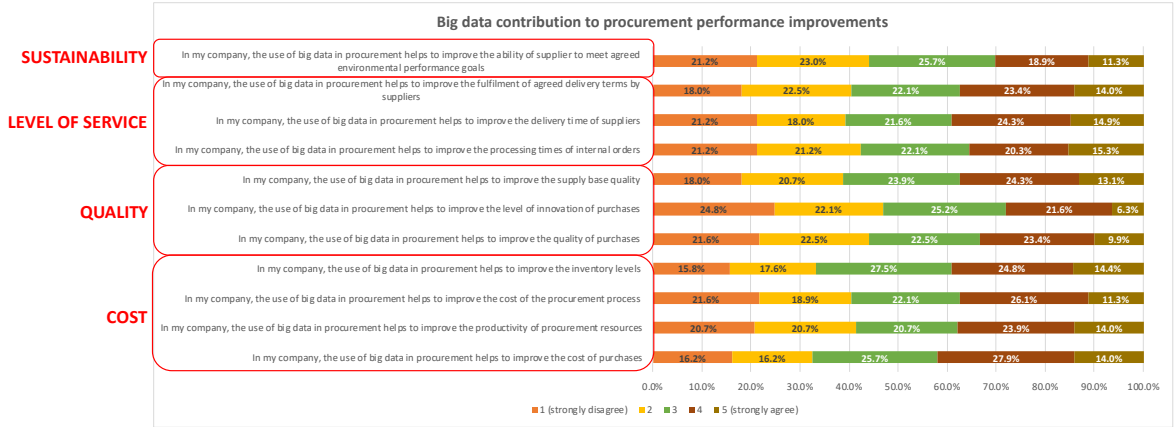


Our sample shows that the more substantial use of big data happens at the strategic sourcing level (e.g., for planning and forecasting, spend analysis, and cost management) and support activities such as negotiation, order management, and supplier performance measurement.

3.4 Big data contribution to procurement performance improvement

Last, the survey also captured the contribution that the use of big data in their companies had in improving specific procurement performance (*Figure 5*).

Figure 5. Big data contribution to performance improvements



Our data shows that the use of big data mostly contributes to improving costs (in terms of reduction of prices paid and cost of inventory through optimization of inventory levels), supply base composition (in terms of reach out to "best-in-class" suppliers), and suppliers' level of service (delivery times and compliance to delivery terms).

4. Procurement organization and big data use

Overall, the previous statistics show that, although not widespread, big data diffusion in procurement is progressing toward higher maturity.

This poses the issues of understanding what factors might favor, or prevent, stronger use of big data in this supply chain area; in this sense, the characteristics of the procurement organization, and its absorptive capacity, could help to explain possible patterns.

This idea is reinforced if we consider the Pearson's correlation between these organizational aspects and the extent of use of big data in the different procurement activities (*Table 3*).

Table 3. Big data use in the procurement process and correlation with procurement organization characteristics

Big data use in the procurement process	***Correlation with the level of centralization***	***Correlation with the level of absorptive capacity***[a]
My company uses big data in procurement to support **category strategy definition** activities	0.197*	0.297**
My company uses big data in procurement to support **demand planning and forecasting** activities	0.053^{NS}	0.214**
My company uses big data in procurement to support **supplier relationship management** activities	0.134*	0.309**

My company uses big data in procurement to support **supplier involvement in NPD** activities	0.07^{NS}	0.225**
My company uses big data in procurement to support **spend analysis and cost management** activities	0.201**	0.207**
My company uses big data in procurement to support **supply risk management** activities	0.078^{NS}	0.343**
My company uses big data in procurement to support **fraud detection** activities	0.031^{NS}	0.259**
My company uses big data in procurement to support **negotiation** activities	0.176*	0.281**
My company uses big data in procurement to support **order management** activities	0.185*	0.274**
My company uses big data in procurement to support **inventory management** activities	0.061^{NS}	0.214**
My company uses big data in procurement to support **logistics management** activities	0.019^{NS}	0.237**
My company uses big data in procurement to support **supplier performance measurement** activities	0.191*	0.311**
My company uses big data in procurement to support **procurement performance measurement** activities	0.018^{NS}	0.242**

[a]Calculated as the average of the items reported in *Figure 1*
*p< 0.05; **p<0.01 level

A more intense use of big data for several procurement activities is positively correlated to a higher level of centralization and/or absorptive capacity.

To further explore this potential relationship, we decided to classify our sample according to these two procurement organizational characteristics, discriminating between:

- High (i.e., centralized organizations), medium (i.e., hybrid organizations), and low (i.e., decentralized organizations) centralization;
- High (i.e., higher than the average), medium (i.e., around the average), and low (i.e., lower than the average) absorptive capacity.

This generated nine clusters, whose characteristics are summarized in *Table 4*, together with the ANOVA tests for group mean differences.

Table 4. Characteristics of the clusters

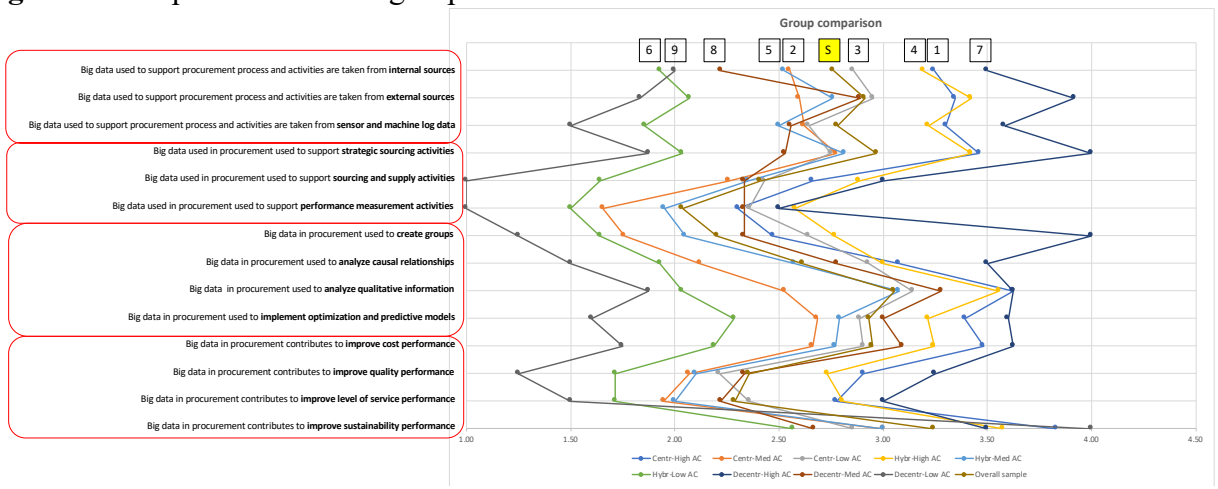
		Group 1	*Group 2*	*Group 3*	*Group 4*	*Group 5*	*Group 6*	*Group 7*	*Group 8*	*Group 9*	**Overall sample**	**Mean difference significance (p-value)**
		Centr-High AC	*Centr-Med AC*	*Centr-Low AC*	*Hybr-High AC*	*Hybr-Med AC*	*Hybr-Low AC*	*Decentr-High AC*	*Decentr-Med AC*	*Decentr-Low AC*		
	Number of observations	*53*	*58*	*14*	*26*	*40*	*14*	*4*	*9*	*4*	*222*	
Big data sources	Big data used to support procurement process and activities are taken from **internal sources**	3.83	3.00	2.86	3.58	3.00	2.57	3.50	2.67	4.00	3.24	0.005**
	Big data used to support procurement process and activities are taken from **external sources**	2.77	1.95	2.36	2.81	2.00	1.71	3.00	2.22	1.50	2.29	0.000***
	Big data used to support procurement process and activities are taken from **sensors' data** and **machine log data**	2.91	2.07	2.21	2.73	2.10	1.71	3.25	2.33	1.25	2.36	0.001**
Big data use	Big data in procurement is used to support **strategic sourcing** activities	3.48	2.66	2.91	3.24	2.77	2.19	3.63	3.09	1.75	2.95	0.000***
	Big data in procurement is used to support **sourcing and supply** activities	3.39	2.68	2.89	3.22	2.79	2.29	3.60	3.00	1.60	2.93	0.000***
	Big data in procurement is used to support **performance measurement** activities	3.62	2.53	3.14	3.56	3.08	2.04	3.63	3.28	1.88	3.05	0.000***
Big	Big data in procurement is used to **create groups**	3.08	2.12	2.93	3.00	2.58	1.93	3.50	2.78	1.50	2.61	0.000***

	Big data in procurement is used to **analyze causal relationships**	2.47	1.76	2.64	2.77	2.05	1.64	4.00	2.33	1.25	2.20	0.000***
	Big data in procurement is used to **analyze qualitative information**	2.30	1.66	2.36	2.58	1.95	1.50	2.50	2.33	1.00	2.04	0.004**
	Big data in procurement is used to **implement optimization and predictive models**	2.66	2.26	2.43	2.89	2.35	1.64	3.00	2.33	1.00	2.41	0.032*
Big data contribution to performance improvement	Big data in procurement contributes to improving **cost performance**	3.46	2.78	2.75	3.42	2.81	2.04	4.00	2.53	1.88	2.97	0.000***
	Big data in procurement contributes to improving **quality performance**	3.30	2.62	2.64	3.22	2.50	1.86	3.58	2.56	1.50	2.78	0.000***
	Big data in procurement contributes to improving the **level of service performance**	3.35	2.60	2.95	3.42	2.76	2.07	3.92	2.89	1.83	2.91	0.000***
	Big data in procurement contributes to improve sustainability **performance**	3.25	2.55	2.86	3.19	2.53	1.93	3.50	2.22	2.00	2.76	0.002**

*p< 0.05; **p<0.01 level; ***p<0.001

The tests show that the clusters differ significantly across all the big data variables included in the questionnaire.

Figure 6. Comparison between groups



The hybrid and decentralized organizations characterized by low absorptive capacity (Group 6 and 9) are also those that have a lower profile for what concerns big data use and contribution to performance. In situations where the procurement responsibilities are spread across different offices, it is hard to implement systematic approaches and shared procedures; if this is paired with a low absorptive capacity, the possibility to use big data strategically is reduced.

Conversely, groups including companies with high levels of absorptive capacity (1, 4, 7) are those with higher use and impact of big data in procurement, no matter the level of centralization. Surprisingly, it is the case of decentralization that is able to provide the strongest profile.

Figures 7, *8*, and *9* provide a more detailed view of the characteristics of the clusters that adopt a more-or-less centralized approach to procurement at different levels of absorptive capacity.

Figure 7. Big data in centralized procurement organizations

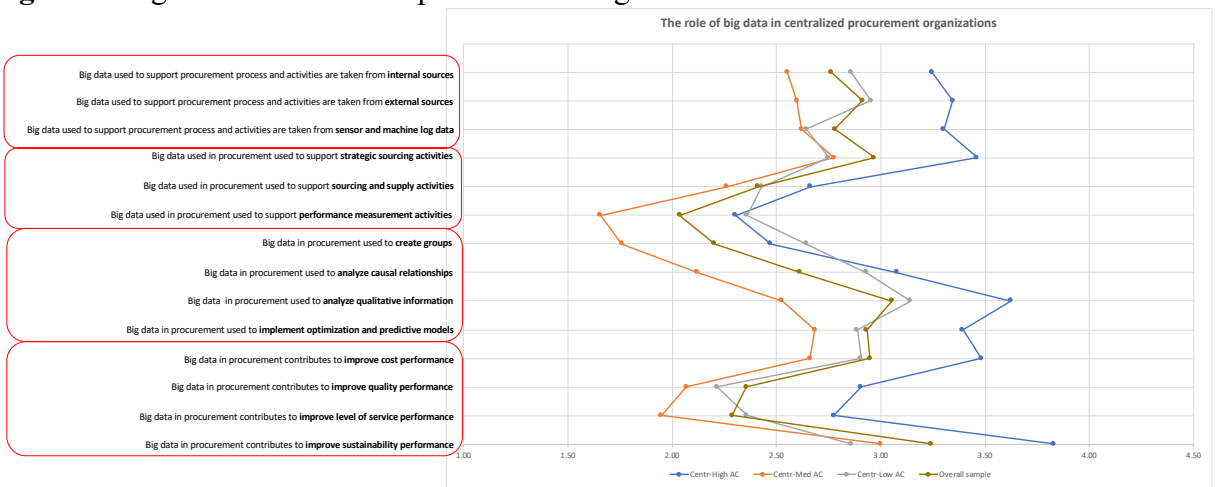


Figure 8. Big data in hybrid procurement organizations

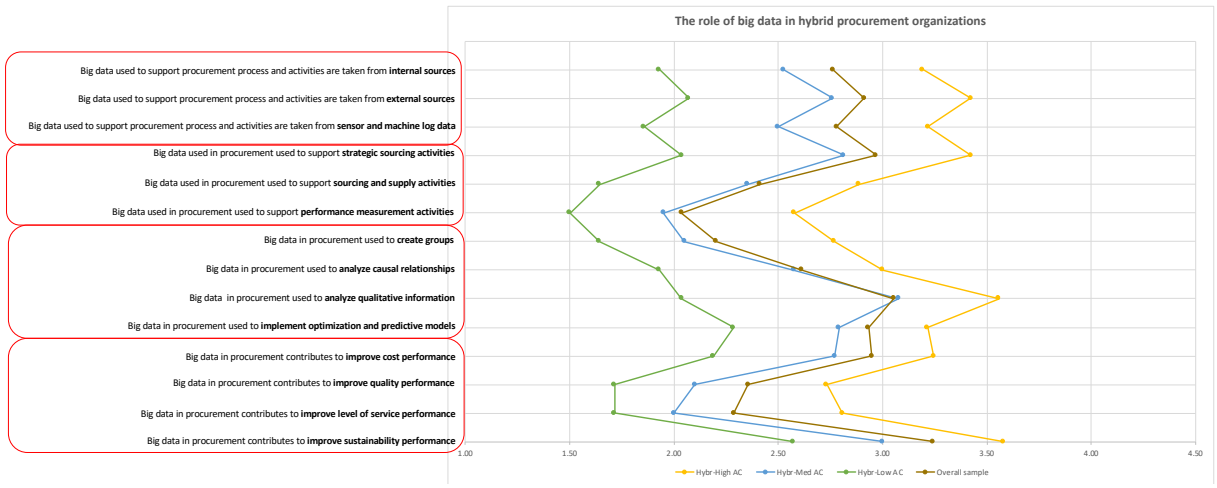
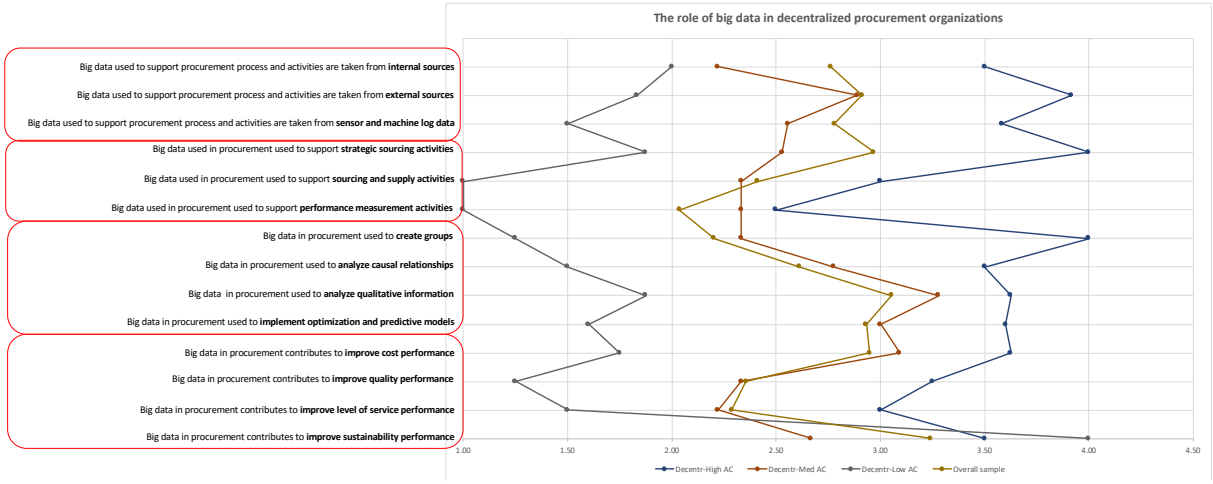


Figure 9. Big data in decentralized procurement organizations



In most cases, higher absorptive capacity in procurement leads to a more substantial big data use profile. We only have one exception in *Figure 7*. Centralized procurement organizations characterized by high absorptive capacity are those reporting the highest values for all the dimensions, except for the use of big data to support performance measurement activities and the use of techniques to create commonalities and groups. In both cases, these aspects are stronger in a centralized procurement organization with low absorptive capacity. Surprisingly, this organizational setting dominates the case of centralized organizations characterized by medium absorptive capacity.

5. Big data in procurement: reality or illusion?

Procurement teams are under intense pressure to reduce costs, maximize efficiency and enhance profit margins while establishing relationships with reliable and trustworthy suppliers.

Harnessing the power of big data is leading to significant opportunities for companies to predict market trends, spending, consumer behavior, and supply chain needs in real-time. From a procurement perspective, there is the possibility to use big data to move away from making reactive decisions on fiscal data and instead take a forward-looking approach that is insight-led.

How do companies use big data to support procurement decisions and activities?

Although there is a common feeling that procurement represents one of the supply chain areas that still have to grow from this perspective, our results show that big data in procurement are being used to support more and more activities and decisions.
Many of these activities are **strategic**; with the help of data, procurement teams can become a more strategic advisor to steer the direction of the business (which confirms the high value of big data in supporting more effective decision-making).
This information is mainly retrieved from **internal sources**: data analytics can process complex data sets from a broad range of ERP systems, consolidating relevant information from multiple streams onto single dashboards in order to inform business decisions.
There is a lot of space for improvement for what concerns the purpose of application of big data: many of the responses declared to use big data in procurement to **create groups** and establish commonalities between commodities and/or suppliers – which represents the natural use for procurement. However, other techniques, more diffused in different supply chain areas (such as regression or optimization models), could represent further approaches procurement professionals should look at in order to extract even more value from big data analysis.
Most importantly, our respondents confirm the big data contribution to **improve performance areas**, particularly for what concerns costs (price and cost of inventory) and supplier level of service (delivery times and conditions).

Are there procurement organizational aspects that influence a more or less intense adoption of big data in the procurement area?

Our survey also looked at procurement organization characteristics that could enable more intense big data use and adoption.
Results show that the **level of procurement centralization** is not associated with a higher maturity in big data use in terms of support to more activities and/or types of data analysis techniques.
Instead, it seems key the role of **absorptive capacity**: procurement organizations characterized by the higher ability to acquire, assimilate, transform and exploit knowledge and information are also those that show a higher level of big data use.