

Big Data Analytics and Operational Management Efficiency: An Empirical Study

Shurui Cao

Syracuse University, Syracuse, NY, USA

Abstract. Big Data Analysis can help make operating decisions, but this help will only be achieved if it is actually applied. Based on the data from the 160 operating units, this paper studies the effect of analytical ability on operational management efficiency. Design stands for Design Ability, Data Governance, Organisation Scale and Demand Fluctuations, etc., and many other factorial combinations. Operational efficiency is the proportion of orders that are processed on time according to the plan. Descriptive comparisons and Ordinary Least Squares regression quantify associations by means of explicit data generation. Increase the level of analysis by 20 and then raise the efficiency by 6 percentage points. The adjusted coefficient is 0.300 and the conditional 95% confidence interval is 0.260-0.340. Governance, Size and Volatility all increase the explained variance from 40.0% to 72.2%. These are the estimations of the built analysis environment and not observations of business improvement or causality. Therefore, it has been decided that the variables, reproducible calculations and transparent uncertainties are all suitable for operational interpretation.

Keywords: Big Data Analytics, Operational Management, Management Efficiency, Empirical Study.

1. Introduction

Operations managers have to determine which orders to prioritise, move production at different times, and whether to change the original schedule due to new delays. Big Data analysis is only useful if the results can be presented to management in an understandable way for all staff. Although a large amount of data has been collected, the order processing speed and resource allocation efficiency are still unknown. The corresponding question is whether there have been corresponding increases in the speed of process execution due to better utilisation of data.

Recently, some studies have put forward reasons for this link based on organisational capabilities. Raj and others propose that the links between analytical capabilities and operating performance in the supply chain are supply chain integration, flexibility and resilience [1]. Although the abstract shows that these are the links for the analysis, it does not provide details on the estimation results. Liu and others have proposed that analytical intellectual ability consists of technical knowledge, management skills, business understanding, relational awareness, and other types of knowledge; they have also found that person-tool fit acts as a mediator in this process [2].

The above are the actual implementations. Based on the above excerpts, data collection and cleaning, modelling, deployment and continuous model maintenance of the Australian railway case study have problems such as fragmented architecture, lack of statistical knowledge and model drift [3]. These problems explain why the analysis results are not being used in the company's operations, although there is abundant data.

Timely order completion is thus taken as the result. Research on the relationship between Analytical Ability, Governance Readiness, Organisational Size and Demand Volatility. Governance is the availability of usable information; Size is the operating scale; Volatility is scheduling difficulty. The above items are the reasons why the process will not run smoothly.

The Contribution is methodological and interpretative. The generated dataset can show the assumed relationship and reproduction of the statistical results. It also shows that an internal consistency in an estimated association does not mean that analytics investment will actually improve the performance of enterprises. The hypothesis is that, all else being equal, good analysis can help complete the work sooner. Verification of the hypothesis in the built environment is to confirm the consistency and rationality of the model, not to test a business theory in isolation.

2. Methods

2.1 Analytical Scope And Variables

There are a total of 160 operating units. Each observation is a hypothetical processing unit of one planning period. No companies were contacted, no questionnaires were distributed and no operation records obtained. The observations are the instances of reproducible regression analysis.

Table 1 presents the variables. Analytical ability is the use of integrated data and analytical knowledge in daily work. It is a model convention for numbers and not a verified questionnaire score. Governance Readiness is still different from Data Ownership and Consistency. Efficiency refers to the proportion of orders delivered on schedule; otherwise, one cannot judge the financial return, decision-making ability and use of resources.

Table 1. Variable definitions and generated measurement scales

Variable	Symbol	Operational interpretation	Generated values
Analytics capability	A	Capacity to integrate and apply analytics	30, 70 points
Governance readiness	G	Data ownership and consistency	30, 70 points
Organizational size	S	Natural logarithm of workforce size	4.5, 5.5
Demand volatility	D	Assumed demand variation	10, 30 percent
Operational efficiency	Y	Orders completed within planned time	50–90 percent

Capability and Adoption are not the same thing. According to the study by professionals on the application of combined analytics and artificial intelligence technologies, it has been determined that this method is not yet practical for frequent operational decisions at our unit level [4]. Therefore, the present specification does not equate the application of technology with organisational ability.

2.2 Data Generation And Reproducibility

The dataset has all combinations of the five binary factors, and each has a value of -1 or +1. Each of the thirty two combinations appears five times. The first four are the observed predictors; the fifth is a random disturbance. Therefore, the Design has reproduced the scenario combinations and has not independently sampled enterprises.

For factors $a_i, g_i, s_i, d_i, e_i \in \{-1, 1\}$, the generation equations are:

$$A_i = 50 + 20a_i, \quad G_i = 50 + 20g_i$$

$$S_i = 5 + 0.5s_i, \quad D_i = 20 + 10d_i$$

$$Y_i = 70 + 6a_i + 3g_i + 2s_i - 4d_i + 5e_i$$

Analytics and governance are in good order, the large unit is relatively small, and the volatility is reduced. The last part is omitted operational variation. Choose coefficients that are easy to understand, and all the results should be in a reasonable range. They were not based on the customer data or the cited literature [5].

A reader can get the dataset by listing the binary combinations, repeating each combination five times, and then applying the equations. Balanced enumeration makes every factor mean zero and every pairwise factor product average zero. Therefore, the predictors are mutually orthogonal. Although there is no confounder in this design, actual capital investment will be related to organisational resources and governance.

2.3 Estimation and Uncertainty

Estimate the two Ordinary Least Squares models. The first is only analysis ability. The second is government regulation, scale and fluctuating demand.

$$Y_i = \alpha + \beta_A A_i + \beta_G G_i + \beta_S S_i + \beta_D D_i + \varepsilon_i$$

Analytics coefficient is the difference in percentage points of timely completion per capability point. Model fit is indicated by the coefficient of determination:

$$R^2 = 1 - \frac{\sum_i (Y_i - \hat{Y}_i)^2}{\sum_i (Y_i - \bar{Y})^2}$$

Conventional standard errors use residual variance and the inverse predictor cross product matrix:

$$\widehat{\text{Var}}(\hat{\beta}) = \frac{\text{SSE}}{n-p} (X^T X)^{-1}$$

Here, $n=160$, and p includes the intercept. Confidence intervals use the corresponding Student distribution. These intervals are conditional model summaries obtained by treating rows as independent observations with constant error variance. Because the rows are deterministic replicated scenarios, the intervals do not quantify sampling uncertainty about a business population. Their role is to illustrate the uncertainty a conventional regression calculation assigns under its stated assumptions [6].

No mediation model is estimated because no intermediate decision or coordination variable was generated. This restriction prevents the statistical specification from implying that the organizational mechanisms discussed in the literature were directly measured.

3. Results

3.1 Distribution and Capability Differences

Table 2 presents the distribution of the whole dataset. All standard deviations are the sample denominators. The mean is 70.000%, and the standard deviation is 9.517 per cent. The range is 40 percentage points, which is the sum of the variation in predictors and disturbance.

Table 2. Descriptive statistics for generated operational units

Panel A. Overall distributions

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Analytics capability	160	50.000	20.063	30.000	70.000
Governance readiness	160	50.000	20.063	30.000	70.000
Organizational size	160	5.000	0.502	4.500	5.500
Demand volatility	160	20.000	10.031	10.000	30.000
Operational efficiency	160	70.000	9.517	50.000	90.000

Panel B. Capability groups

Analytics capability score	Observations	Mean completion rate (%)
30	80	64.000
70	80	76.000

Table 2, Panel B shows that the mean completion rates for capability scores thirty and seventy are 64.000% and 76.000%, respectively, and there are eighty observations each. Governance, Size, Volatility and Disturbance are balanced in both groups, so the 12 percentage point difference is due to the addition of the analytics component. It is not the result of investment or a comparison among observed companies [7].

Figure 1 shows the means in Table 2, Panel B. The bottom is 0; therefore, it is relatively small. These are also the fitted results at the corresponding capability level when the control is set to its mean.

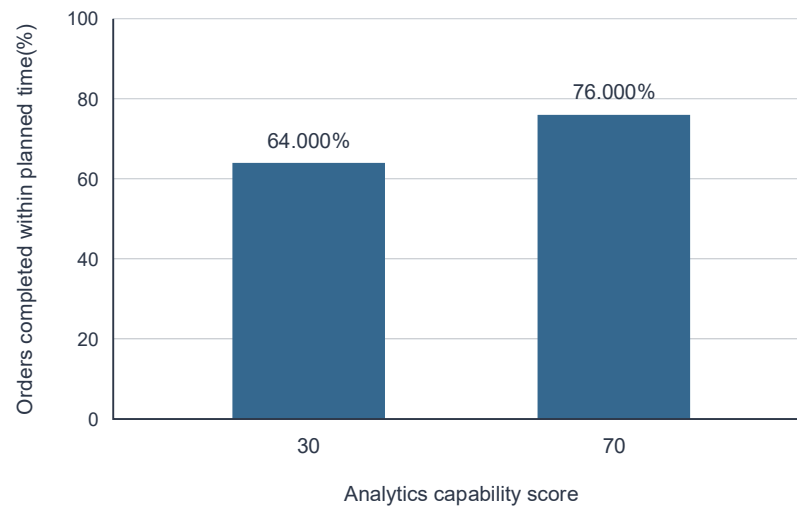


Figure 1. Mean timely order completion by analytics capability

3.2 Regression Estimates And Explained Variation

Table 3 presents the regressions. The analytics coefficient is still 0.300 after adding controls because the factorial design makes analytics orthogonal to all controls. The standard error of no predictors is 0.029, and it decreases to 0.020 after adding some predictors. Based on the general assumptions of regression, the adjusted range for a 20-point difference in ability is about 5.207%-6.793%.

Table 3. Ordinary least squares estimates of operational efficiency

Term or statistic	Analytics only	Adjusted model
Intercept	55.000	35.500
Analytics coefficient	0.300	0.300
Analytics standard error	0.029	0.020
Analytics 95% interval	0.242–0.358	0.260–0.340
Governance coefficient	—	0.150
Governance standard error	—	0.020
Size coefficient	—	4.000
Size standard error	—	0.803
Volatility coefficient	—	–0.400
Volatility standard error	—	0.040
Observations	160	160
Residual degrees of freedom	158	155
Residual sum of squares	8640.000	4000.000
Explained variation	0.400	0.722
Residual standard error	7.395	5.080

The three percentage points are due to the 20-point gap. A half-unit difference in logged size is about a 0.02% difference, and for every ten more volatility points, the completion rate will drop by 0.04%. The above values are the generation coefficients. Therefore, only the scale and internal consistency of the model have been established; no new empirical relations have been found[8].

The total outcome variation around the mean is 14400.000 squared percentage points. The proportion of analytics is 5760.000. Adding the controls accounts for another 4640.000, leaving 4000.000 for the constructed disturbance. The increase in explained variance is therefore 32.222 percentage points. It can be seen that although controls have been added, the model fit has improved without changing the focus coefficient.

4. Discussion

4.1 Connecting Analytical Resources With Execution

The modelled association has the following business significance: Units with good analysis capabilities will be able to process a larger proportion of orders on time. It is likely that the problem will be found earlier, dealt with consistently, and resolved promptly before it becomes serious. None of those intermediate actions is observed here. They can explain why the particular relationship will be studied, but they do not turn the obtained coefficient into proof of the mechanism [9].

It should be different from the previous studies in this paper. Liu and others have specifically studied person-tool fit and connected employees' knowledge with the application ability of analytical tools. Railway excerpts detail the actual work of linking asset data, deploying models and observing changes in behaviour. Based on the above, capability failure can occur at various times: employees may misunderstand the recommendations, there may be inconsistencies in records, or the deployed models may not be suitable. A single ability score will not be sufficient.

Human Capital is another way. Dahiya and others have distinguished between managerial and employee analytics in their study of firm performance, and this division determines who will take action based on the results of the analysis. Managers will address the problems, and then the employees will change their behaviour. The present model does not distinguish between the two; therefore, it cannot be determined which will be more effective: training analysts or improving management processes.

4.2 Complementarity and Operational Context

Governance adds to the generated model independently, but it is an analytical simplification. A general definition and an accountable data owner will help to improve the use of current analytics. Conversely, analysis can also show the problems with governance. In the future, we will determine whether there is enough variation in the data to distinguish between complementary effects and correlated investment in the observation specification.

Mamakou has proposed that decision quality is a mediator and firm size is a boundary for shaping analytics value creation, as shown in abstract. Therefore, we do not need to directly compare coefficients for all outcomes. Abu-El-Haija and others have also studied the idea of knowledge management as the intersection point of analytical power, digital innovation and performance; although an abstract presents the subject of the study, it does not list the expected results of this research.

The next are other saved references. Cui and others focus on supply chain agility, and Anwar and others study adoption, decision quality and environmental performance [4, 7]; only bibliographic identities were provided for these papers, so we can only infer the research contents from the titles. Shahzad and others define manufacturing analytics based on process visibility, waste and throughput in the abstract [10]; timely completion is a narrow result that can complement but does not cover all these areas.

4.3 Precision, Design Assumptions, And Managerial Use

The Numerical Stability of the analysis estimate is due to orthogonality. It is not to be assumed that omitted variables are irrelevant in operational records. More capable units will have better equipment, experienced supervisors, simpler orders, favourable investment terms, etc. The reverse is also true; good operation can free up funds for developing analytics. Before concluding that a coefficient is causal, longitudinal observation and a reliable identification method should be used.

There is a replication issue. Repeated combinations of scenarios increase the number of nominal observations without adding new combinations or independent organisational information. Therefore, the traditional confidence intervals should be presented with the generation Design. Assuming that narrowness indicates external validity is to ignore the problem of precision in the model and

knowledge about a particular group. The reported intervals are helpful because they have been calculated explicitly, not because they provide business certainty.

The two-level predictors do not have thresholds, saturation or an uneven distribution. A better performance will not be achieved through improved skills without good governance. Therefore, the harm from avoidable delays will be reduced. Demand Fluctuations Will Also Affect the Quality and Completeness of the Data. None of the above patterns are shown in the additive equations. Although there are some other designs, the relevant exercise should still be conducted externally and then applied to support the organisation's forecast.

To reach the 6 percentage point goal, there should be stable base rates, planning periods and order denominators. Changes in order mix or promised processing times will increase the completion rate but not the execution. Note down the excluded items, rescheduled orders and missing timestamps to distinguish between operational improvements and measurement changes. The Scenarios have fixed definitions, and the problem of organisational record reliability has not been solved.

A good result for managers is an evaluation system. Determine the results of the process first and then choose an appropriate analysis method; separately measure the ability and adoption rates and record the reasons for non-implementation. Over a long period of time, one can observe whether the order complexity and process have changed. Decision speed, utilisation and cost can be included in the completion rate if they are measurable. Leaving them separate would show that the trade-off is a higher cost for less time. Therefore, the generated six-point association is an interpretive case and not an investment return forecast.

5. Conclusion

According to the analysis of big data, an index for operating management efficiency has been developed for the 160 established operating units. The better model is off by 20 points and accounts for 72.2 per cent of the change in actual results over time. An even distribution will help to show the contrast and explain why there is an obvious error.

Therefore, the analysis will be conducted in an organised manner: present the results of the operation, distinguish between ability and technology adoption, consider the relevant context, and report uncertainty within the limits of the evidence. Recently, some scholars have been studying the relationship among human knowledge, organisational coordination, application of analytics and performance. Although the above are reasons for empirical research, they have not been verified by data. The next study should independently observe the units, verify the ability indicators, establish the time sequence, and then draw conclusions on the actual improvement of operations.

References

- [1] Raj, A., Kumar, R. R., Narayanan, S., Kirca, A. H., & Jeyaraj, A. (2025). Analytics capability, supply chain capabilities, and operational performance: A meta-analytic investigation. *Journal of Operations Management*, 71(6), 786–805. <https://doi.org/10.1002/joom.1376>
- [2] Kortian, V., Pal, S., Ghevondian, N., & Harrison, N. (2024). Challenges and issues in implementing & operationalising big data analytics capabilities in a major Australian railway organisation: A case study. *SN Computer Science*, 5, 639. <https://doi.org/10.1007/s42979-024-02953-8>
- [3] Liu, Y., Qiao, H., Wang, J., & Jiang, Y. (2024). Influencing mechanism of the intellectual capability of big data analytics on the operational performance of enterprises. *Heliyon*, 10(3), e25032. <https://doi.org/10.1016/j.heliyon.2024.e25032>
- [4] Cui, L., Wang, Z., Liu, Y., et al. (2024). How does data-driven supply chain analytics capability enhance supply chain agility in the digital era? *International Journal of Production Economics*, 277, 109404.
- [5] Mamakou, X. J. (2026). Linking business analytics to firm performance: A mixed-method analysis of capabilities, decision quality, and firm size. *Journal of Business Research*, 212, 116219.
- [6] Dahiya, R., Le, S., & Kroll, M. (2025). Big data analytics and firm performance: The effects of human capital and mediating firm capabilities. *Journal of Strategy & Innovation*, 36(1), 200535.

- [7] Anwar, M. A., Zong, Z., Mendiratta, A., & Yaqub, M. Z. (2024). Antecedents of big data analytics adoption and its impact on decision quality and environmental performance of SMEs in recycling sector. *Technological Forecasting and Social Change*, 205, 123468. <https://doi.org/10.1016/j.techfore.2024.123468>
- [8] Abu-El-Haija, R., Mat, N., Isa, R. M., & Abdullah, N. A. (2026). Big data analytics capability, knowledge management processes, digital innovation, and firm performance in UAE ICT firms. *Cogent Business & Management*, 13(1).
- [9] Chen, C., Chen, S., Khan, A., et al. (2024). Antecedents of big data analytics and artificial intelligence adoption on operational performance: The ChatGPT platform. *Industrial Management & Data Systems*, 124(7), 2388–2413. <https://doi.org/10.1108/IMDS-10-2023-0778>
- [10] Shahzad, M. A., Wang, X., Chen, S., et al. (2025). Impact of big data analytic capabilities on manufacturing performance. *Journal of Manufacturing Technology Management*, 36(7), 1410–1432. <https://doi.org/10.1108/jmtm-11-2024-0623>