HIGH SCHOOL EDITION
Journal of Student Research

Volume 11 Issue 4 (2022)

A Study on The Effect of Big Data Analytics on Product Innovation Performance

Chih-Hsuan Ashley Cheng[1], Millie Y.S. Hong[1] and Colin C. J. Cheng[#]

[1] Kang Chiao International School Taipei Campus, Taiwan, Province of China
#Advisor

ABSTRACT

Despite the rising interests in the use of big data analytics (BDA), it is unclear whether the use of BDA is beneficial to product innovation. Building on the Resource-Based View, this study aims to examine the relationship between BDA and product innovation performance. Conducting survey from 163 firms, this study provides compelling evidence indicating that the use of BDA has a significant positive impact on firms' product innovation. This study discusses theoretical implications for advancing BDA research and suggests actionable steps for managers to get benefits from using BDA.

Introduction

The recent emergence of big data analytics (BDA) has been heralded as the next revolution that will transform the way firms' business operations are conducted and, eventually, bring the firms benefits (Mikalef et al., 2020; Wamba et al., 2017; Akter et al., 2016). As a result, there is much attention from practitioners and academics on the value that firms can create through the use of BDA (Mikalef et al., 2018). While BDA have the potential to significantly create business value, like with all innovations, many firms have encountered many challenges in identifying the merits and limits of BDA. One of the most critical challenges for firms is making sense of and responding to BDA (Crittenden et al., 2019).

Specifically, most previous studies have focused on the role of BDA in driving firm performance in many aspects, such as profitability, market share, customer retention, and financial performance (Mikalef et al., 2019; Raguseo and Vitari, 2018; Wamba et al., 2017). However, they have paid less attention to whether firms can rely on BDA to enhance product innovation performance, and particularly how firms transfer their existing technology-specific know-how to product innovation development (Crittenden et al., 2019). Thus, the use of BDA is critical for firms to understand its merits and limits as well as to find the growth opportunities. Accordingly, it is important to assess how firms use BDA to enhance product innovation performance.

We aim to contribute to BDA research as follows. Because the literature is limited in answering whether firms can extract value from their BDA in the product innovation development (Caesarius and Hohenthal, 2018), this study extends previous research to address this critical question. Using survey and archival data of firms with longitudinal design, we find that firms can rely on BDA to enhance product innovation performance, over time. In this way, the findings add to the understanding of the importance of BDA for product innovation outcomes.

In the following sections, we first discuss theoretical background and hypotheses development. Next, we describe research method, followed by analyses and results of this study. Finally, we present the findings and conclude with a discussion and implications for both academics and practitioners.

Theoretical Background

We draw on the resource-based view (RBV) to explain the relationship between BDA and product innovation performance.

RBV is useful in understanding innovation-related performance through the theoretical perspectives of organizational resources (Barney, 2001). RBV portrays technology competence as intangible resources derived from combinations of internal investments and external appraisals, which create difficult-to-replicate knowledge assets (Teece, 1998). Given its prominent role, prior studies have examined various types of technology competence, including the use of marketing analytics, social media analytics, and BDA (e.g., Mikalef et al., 2019; Stieglitz et al., 2018; Germann et al., 2013). For the purpose of this study, we mainly focus on BDA, which refer to the collection of data, analytical tools, computer algorithms and techniques to derive meaningful insights and patterns from the collected large data sets (Mikalef et al., 2020).

According to prior studies, strengths in such intangible resource assets grows overall BDA of a firm and this can lead to competitive advantages and ultimate performance superiority (Mikalef et al., 2018; Barney, 2001). BDA thus play important roles in affording a firm's competitive advantage as an imperfectly imitable and non-substitutable resource. In addition, the use of BDA enables firms to achieve heterogeneity and hence afford higher value and awareness in securing sustainable advantages. In this study, we particularly focus on the impact of BDA on firms' product innovation performance, defined as the extent to which firms can achieve their stated objectives of product innovation development (Liu and Atuahene-Gima, 2018).

Hypothesis

Previous studies have provided empirical evidence of the effect of BDA on various indicators of firm performance. For example, using data from two Delphi studies and 152 online surveys of business analysts in the U.S., Akter et al. (2016) confirm that the value of BDA has a significant impact on firm performance. Wamba et al. (2017), using an online survey to collect data from 297 Chinese IT managers, demonstrate that BDA has a direct impact on firm performance. Drawing on the resource-based view, Raguseo and Vitari (2018) find that the business value achieved from investments in BDA leads to superior financial performance of a firm. Using survey data from 175 chief information officers and IT managers working in Greek firms, and three case studies, Mikalef et al. (2020) show that BDA is crucial to firm performance.

We propose that the positive effect of BDA will enhance product innovation performance for firms, over time. First, the literature suggests that through focused deployment of BDA, firms are able to sense emerging opportunities and threats, generate critical insight, and adapt their operations based on trends observed in the competitive environment (Mikalef et al., 2019). As a result, the major competitive differentiator that BDA provides lies in the fact that it facilitates better informed decision-making (Mikalef et al., 2019; Abbasi et al., 2016). The increased interest in BDA has been particularly evident in firms operating in complex and changing environments (Hajli et al., 2020; Constantiou and Kallinikos, 2015). Managers nowadays are basing their decisions more and more on real-time insight generated from BDA, and are directing a growing number of initiatives in this direction (Jabbar et al., 2020). Several research studies demonstrate that BDA can offer substantial value, when applied to problems of specific domains such as product innovation (Johnson et al., 2017), service provision (Lehrer et al., 2018), and business model innovation (Ciampi et al., 2021). As echoed by a study of the MIT Sloan Management Review indicates that BDA can be a source of innovation, with those companies that are leaders in adoption being more likely to deliver new products in comparison to the laggards (Ransbotham et al., 2017). Overall, BDA can be expected to have a positive effect on product innovation performance.

Second, as firms can learn innovative knowledge from multiple sources, BDA can exert a positive effect on product innovation performance. Firms can learn innovations from observing the strategic moves of their counterparts (Obal, 2013). They can also learn from their own successful and failed innovation experiences to accumulate innovative knowledge (Eggers and Kaul, 2018; Kapoor and Adner, 2012), which in turn guides future product innovation development and helps them through using BDA more effectively to enhance product innovation performance. With more innovative knowledge accumulated over time, firms can be more capable of applying their BDA to adjust their product innovation development operations.

Third, researchers argue that knowledge is an important asset that helps firms develop innovation and gain competitive advantage (Eggers and Kaul, 2018; Rehm and Goel, 2015; Kapoor and Adner, 2012). From a marketing perspective, De Luca and Atuahene-Gima (2007) argue that firms need to see beyond the technology and focus on how to manage their knowledge to gain superior product innovation. Examining innovation success from a strategic decision making perspective, Zhou and Li (2012) point out that knowledge plays an important role in driving product innovation success. We then take the perspective that using BDA helps firms to benefit their product innovations because the essence of BDA is to turn the vast amount of raw data into meaningful and actionable information (Mikalef et al., 2020; Raguseo and Vitari, 2018), in which BDA can help firms turn it valuable knowledge for product innovation development.

Accordingly, we expect that firms that employ BDA are more likely to achieve better product innovation performance, over time. Therefore, we hypothesize the following:

Hypothesis: BDA has a positive effect on firms' product innovation performance, over time.

Methods

Sample and Data Collection

To test our hypothese, we used the data set of respondents from China Credit Information Service (2017), a leading business database in Taiwan, which lists firms, their financial data, and contact information. Accordingly, we obtained the initial sample frame of 500 firms.

To correctly collect the data, we contacted new product development (NPD) manager of each firm by telephone to determine whether the firm had used BDA in its product innovation process. Since the unit of analysis refers to the project level, we then asked NPD managers to identify one recently completed product innovation project (within the past three years) for which they had the best knowledge to respond to the items as they related to that particular project. This resulted in a target population of 359 firms, which were asked to participate in this study.

To reduce the potential for common method bias associated with single sourcing (Podsakoff et al., 2003, p. 898), we measured the different variables by obtaining different data from multiple respondents in each firm. Specifically, chief information officers (or similar personnel) were asked to provide information regarding BDA, while NPD managers provided information for product innovation performance. Importantly, to corroborate the reliability of some survey measures, we also collected archival data related to each firm's product innovation performance.

In early-2022, we obtained archival data of product innovation performance and control variables (firm size, firm age, and R&D intensity), and also asked chief information officers to rate BAD and NPD managers to rate product innovation performance and the control variable. After matching the various data sets, we successfully obtained the matched 163 firms.

We compared participating and nonparticipating firms in terms of firm size, firm age, and R&D intensity and found no significant difference ($p < 0.05$), suggesting non-response bias is not a major concern (Armstrong and Overton, 1977).

Measurement

We measured the perceptual items with a seven-point scale. For items adapted from the literature and written in English, a double-translation method was used to translate them into Chinese. This process included: (1) our initially translating the items into Chinese; (2) two other academics then translating the Chinese version back into English; and (3) this translation being checked by a third academic to ensure conceptual equivalence.

To measure *BDA*, we adapted the 12 items from Johnson et al. (2017), which reflects a firm captures data and derives insights by using BDA approaches.

For *product innovation performance*, we adapted from De Luca and Atuahene-Gima (2007) that measure the extent to which in the last three years firms have achieved the product development performance objectives. Further, we follow Kostopoulos et al. (2011) to measure product innovation performance by using archival data as the percent change in the ratio of the annual sales from Y_0 to Y_1 that originated from the particular new product development project, such that we obtained the data (2017) in 2018 and the data (2020) in 2021. Specifically, we measured product innovation performance as (annual $sale_{y2020}$ – annual $sale_{y2017}$)/annual $sale_{y2017}$ × 100. The variable was expressed as a percentage going from zero to 100%. This type of variable has been widely used in previous innovation studies (Berchicci 2013; Laursen and Salter, 2006).

Finally, we included four control variables based on their relevance to firm characteristics. *Firm size* was measured as the number of employees. *Firm age* was measured on the number of years the firm had been established. Following Marano and Kostova (2016), we controlled *R&D intensity* using the ratio of R&D expenditures to sales. Finally, *prior innovation performance* was controlled because if firms have good innovation performance in past years, they more likely to achieve higher product innovation performance (Cheng and Shiu, 2020). A two-item scale adapted from Yanadori and Cui (2013) measures prior innovation performance.

Construct Reliability and Validity

We used the MPlus Exploratory Structural Equation Modeling technique to examine the internal consistency of the scale because it combines exploratory factor analysis and confirmatory factor analysis (CFA) in one procedure and avoids the problems related to the traditional two-step process (Muthén and Muthén, 2017).

We first conducted a CFA to verify that the indicators reflected their intended latent variables. We then assessed the overall model fit using the goodness-of-fit index (GFI), comparative fit index (CFI), and normed fit index (NFI) based on the chi-square statistic, and evaluated these fit indexes with the cutoff value of 0.90 recommended by (Hair et al., 2019). We also used the root mean square error of approximation (RMSEA) with the cutoff value of 0.08, recommended by Hu and Bentler (1999), to assess the potential lack of model fit. The fit statistics indicate our model fits the data well (x^2/d.f. = 1.94, $p < 0.001$; GFI = 0.93, CFI = 0.94, NFI = 0.93, RMSEA = 0.04).

The results indicate that convergent validity is met (Hair et al., 2019; Fornell and Larcker, 1981). As shown in Table 1; (1) all standardized factor loadings of the latent variables on their indicators are significant ($p < 0.01$) and well above the recommended value of 0.65, between 0.70 and 0.85; (2) the composite reliability of all constructs (ranging from 0.81 to 0.95) exceeds the 0.80 threshold; and (3) average variance extracted (AVE) ranges from 0.57 to 0.68, exceeding the 0.50 threshold.

The results also indicate that discriminant validity is achieved. As shown in Table 2, all diagonal elements representing the square root of the AVE are greater than the highest shared variance (Fornell and Larcker, 1981). In addition, the value of the unconstrained model is significantly lower than that of the constrained model for all possible pairs of constructs, with all chi-square differences ranging from 26.31 to 47.29 ($p < 0.001$), which supports discriminant validity (O'Leary-Kelly and Vokurka, 1998). Table 2 shows descriptive statistics, correlations, and the square root of AVEs of constructs.

Results

Our hypotheses testing relies on hierarchical multiple regressions. Before testing the moderating effects, we mean-centered the variables that constitute the interaction terms (Aiken and West, 2011). To check for multicollinearity, we assessed the diagnostic statistics of the variance inflation factor (VIF) of the variables and find VIF values ranging from 1.02 to 1.23. We further calculated the model-dependent VIF values of the full regression models and obtained 1.33 and 1.39, respectively (Craney and Surles, 2002). All the VIF values of the variables included were below their corresponding model-dependent VIF values, suggesting that the independent variables in our models have better explanatory power than that when regressing any variable on the remainder of the independent variables. Thus, multicollinearity is not a serious concern in our study.

Table 3 reports the estimated effects on product innovation performance. Model 1 contains the effects of the control variables. Model 2 tests the direct effects of BDA in addition to the control variables. We use Models 2, 4, and 6 to discuss the hypotheses.

In Hypothesis, we expect that BDA enhances firms' product innovation performance over time. Model 2 reveals that BDA is positively related to product innovation performance ($\beta = 0.32$, $p < 0.001$), in support of Hypothesis.

We also used the archival data of product innovation performance to test the hypotheses. The results shown in Table 4 are highly consistent with those presented in Table 3 of the survey data in terms of Hypothesis (Model 2, $\beta = 0.32$; $p < 0.001$).

Discussion

Drawing on the RBV, we show that BDA leads to positive product innovation performance. As such, this study contributes to the research fields of big data in following aspects. First, it contributes to big data literature by extending the growing research on BDA to the context of product innovation. The predominant view in prior research is that BDA plays an important role in determining firm performance (Mikalef et al., 2019; Wamba et al., 2017; Akter et al., 2016). However, researchers have devoted limited efforts to examining whether firms can use BDA to enhance product innovation performance (Crittenden et al., 2019). This study thus augments extant big data literature through confirming that firms can utilize their BDA to enhance product innovation performance.

Second, the deployment of BDA presents critical challenges for firms. This new emerging technique has been undergoing fundamental changes, and the transition is far from accomplishment. Managers, thus, must adjust their existing data analytics strategies to adapt to this ongoing developing technique. Our findings suggest that managers can enhance product innovation by using BDA. Specifically, managers of BDA can use a mountain of data about market information in order to find the correlations that allow managers to develop deep market insights and make relatively accurate predictions on product innovation development. Therefore, employing BDA is a valid strategic choice for managers to create points of difference from competition.

Limitations and Future Research

This study has several limitations, which in turn suggest fruitful directions for future research. First, future research could study whether our results will hold in other types of innovation such as service innovation or organizational innovation.

Second although our sample size and the ratio of sample to variables are comparable with those of previous BDA research (e.g., Mikalef et al., 2019), our sample size may raise concerns about the generalizability of the findings. Future research should use larger samples to help verify the relationship among the focal constructs of this study.

Finally, we acknowledge that our measurement of BDA (based on informants' knowledge) has limitations. Although the measurements are reasonable and have been used in prior research (e.g., Gupta and George, 2016), further research should use finer-grained measurements to reexamine our findings. For example, tapping into objective sources of BDA measures would help validate our findings.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the association for information systems*, *17*(2), 3.

Aiken, L. S., & West, S. G. (2011). Multiple regression: Testing and interpreting interactions. SAGE, Thousand Oaks, California, USA: Sage Publications, Inc.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment?. *International Journal of Production Economics*, *182*, 113-131.

Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of marketing research*, *14*(3), 396-402.

Bao, Y., Chen, X., & Zhou, K. Z. (2012). External learning, market dynamics, and radical innovation: Evidence from China's high-tech firms. *Journal of Business Research*, *65*(8), 1226-1233.

Barney, J. B. (2001). Resource-based theories of competitive advantage: A ten-year retrospective on the resource-based view. *Journal of management*, *27*(6), 643-650.

Berchicci, L. (2013). Towards an open R&D system: Internal R&D investment, external knowledge acquisition and innovative performance. *Research policy*, *42*(1), 117-127.

Caesarius, L. M., & Hohenthal, J. (2018). Searching for big data: How incumbents explore a possible adoption of big data technologies. *Scandinavian Journal of Management*, *34*(2), 129-140.

Caesarius, L. M., & Hohenthal, J. (2018). Searching for big data: How incumbents explore a possible adoption of big data technologies. *Scandinavian Journal of Management*, *34*(2), 129-140.

Cheng, C., & Shiu, E. C. (2020). What makes social media-based supplier network involvement more effective for new product performance? The role of network structure. *Journal of Business Research*, *118*, 299-310.

China Credit Information Service (2017). *2017 Business Groups in Taiwan*. China Credit Information Service.

Ciampi, F., Demi, S., Magrini, A., Marzi, G., & Papa, A. (2021). Exploring the impact of big data analytics capabilities on business model innovation: The mediating role of entrepreneurial orientation. *Journal of Business Research*, *123*, 1-13.

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44-57.

Craney, T. A., & Surles, J. G. (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, *14*(3), 391-403.

Crittenden, A. B., Crittenden, V. L., & Crittenden, W. F. (2019). The digitalization triumvirate: How incumbents survive. *Business Horizons*, *62*(2), 259-266.

De Luca, L. M., & Atuahene-Gima, K. (2007). Market knowledge dimensions and cross-functional collaboration: Examining the different routes to product innovation performance. *Journal of marketing*, *71*(1), 95-112.

Eggers, J. P., & Kaul, A. (2018). Motivation and ability? A behavioral perspective on the pursuit of radical invention in multi-technology incumbents. *Academy of Management Journal*, *61*(1), 67-93.

Eggers, J. P., & Park, K. F. (2018). Incumbent adaptation to technological change: The past, present, and future of research on heterogeneous incumbent response. *Academy of Management Annals*, *12*(1), 357-389.

Fornell, C. & Larcker, D. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 382-388.

Germann, F., Lilien, G. L., & Rangaswamy, A. (2013). Performance implications of deploying marketing analytics. *International Journal of Research in Marketing*, *30*(2), 114-128.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information & Management*, *53*(8), 1049-1064.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European business review*.

Hajli, N., Tajvidi, M., Gbadamosi, A., & Nadeem, W. (2020). Understanding market agility for new product success with big data analytics. *Industrial Marketing Management*, *86*, 135-143.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Jabbar, A., Akhtar, P., & Dani, S. (2020). Real-time big data processing for instantaneous marketing decisions: A problematization approach. *Industrial Marketing Management*, *90*, 558-569.

Johnson, J. S., Friend, S. B., & Lee, H. S. (2017). Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process. *Journal of Product Innovation Management*, *34*(5), 640-658.

Kapoor, R., & Adner, R. (2012). What firms make vs. what they know: how firms' production and knowledge boundaries affect competitive advantage in the face of technological change. *Organization Science*, *23*(5), 1227-1248.

Kostopoulos, K., Papalexandris, A., Papachroni, M., & Ioannou, G. (2011). Absorptive capacity, innovation, and financial performance. *Journal of Business Research*, *64*(12), 1335-1343.

Laursen, K., & Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. *Strategic management journal*, *27*(2), 131-150.

Lehrer, C., Wieneke, A., Vom Brocke, J. A. N., Jung, R., & Seidel, S. (2018). How big data analytics enables service innovation: materiality, affordance, and the individualization of service. *Journal of Management Information Systems*, *35*(2), 424-460.

Li, Y., Thomas, M. A., & Osei-Bryson, K.-M. (2016). A snail shell process model for knowledge discovery via data analytics. *Decision Support Systems, 91*, 1-12.

Liu, W., & Atuahene-Gima, K. (2018). Enhancing product innovation performance in a dysfunctional competitive environment: The roles of competitive strategies and market-based assets. *Industrial Marketing Management*, *73*, 7-20.

Marano, V., & Kostova, T. (2016). Unpacking the institutional complexity in adoption of CSR practices in multinational enterprises. *Journal of Management Studies*, *53*(1), 28-54.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*, 261-276.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics capabilities and innovation: the mediating role of dynamic capabilities and moderating effect of the environment. *British Journal of Management*, *30*(2), 272-298.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2020). The role of information governance in big data analytics driven innovation. *Information & Management*, *57*(7), 103361.

Mikalef, P., Krogstie, J., Pappas, I. O., & Pavlou, P. (2020). Exploring the relationship between big data analytics capability and competitive performance: The mediating roles of dynamic and operational capabilities. *Information & Management*, *57*(2), 103169.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, *16*(3), 547-578.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, *16*(3), 547-578.

Muthén, B., & Muthén, L. (2017). *Mplus*. Chapman and Hall/CRC.

Obal, M. (2013). Why do incumbents sometimes succeed? Investigating the role of interorganizational trust on the adoption of disruptive technology. *Industrial Marketing Management*, *42*(6), 900-908.

O'Leary-Kelly, S. W., & Vokurka, R. J. (1998). The empirical assessment of construct validity. *Journal of operations management*, *16*(4), 387-405.

Raguseo, E., & Vitari, C. (2018). Investments in big data analytics and firm performance: an empirical investigation of direct and mediating effects. *International Journal of Production Research*, *56*(15), 5206-5221.

Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017). Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review*, *59*(1).

Rehm, S. V., & Goel, L. (2015). The emergence of boundary clusters in inter-organizational innovation. *Information and Organization*, *25*(1), 27-51.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. *International journal of information management*, *39*, 156-168.

Teece, D. J. (1998). Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets. *California management review*, *40*(3), 55-79.

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J. F., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356-365.

Yanadori, Y., & Cui, V. (2013). Creating incentives for innovation? The relationship between pay dispersion in R&D groups and firm innovation performance. *Strategic Management Journal*, *34*(12), 1502-1511.

Zhou, K. Z., & Li, C. B. (2012). How knowledge affects radical innovation: Knowledge base, market knowledge acquisition, and internal knowledge sharing. *Strategic management journal*, *33*(9), 1090-1102.