Achieving Sustainable New Product Development by Implementing Big Data-Embedded New Product Development Process

Yufan Wang and Haili Zhang *

School of Economics and Management, Xi’an Technological University, Xi’an 710021, China; wangyufan@xatu.edu.cn
* Correspondence: zhanghaili@xatu.edu.cn

Received: 4 May 2020; Accepted: 5 June 2020; Published: 8 June 2020

Abstract: Literature suggests that new product development (NPD) has an impact on sustainable organizational performance. Yet, previous studies in NPD have mainly been based on “experience-driven”, not data-driven, decision-making in the NPD process. We develop a research model to examine how the big data-embedded NPD process affects the sustainable innovation performance of NPD projects. We test the proposed model and conduct the cross-national comparison using data collected on 1858 NPD projects in the United States of America (USA), the United Kingdom (UK), and Australia. The research findings suggest that big data-embedded business analysis, product design, and product testing increase sustainable innovation performance in all three countries. The study findings also reveal several surprising results: (1) in the USA, big data-embedded product testing has the highest effect on sales growth and gross margin, (2) in Australia, big data-embedded commercialization has the highest effect on sales growth and gross margin, and (3) in the UK, big data-embedded commercialization has the highest effect on second-year sales growth, first-year, and third-year gross margin; in addition, big data-embedded product testing has the highest effect on third-year sales growth and second-year gross margin.

Keywords: big data; big data-embedded new product development process; sustainable innovation performance

1. Introduction

The role of the new product development (NPD) process in driving sustainable growth and performance has been the central focus of many innovation scholars [1–11]. The NPD process can be characterized as a series of stages, comprising idea development, business analysis, product design, product testing, and commercialization [3,4,6]. Previous research has demonstrated that the “experience-driven” NPD process is a positive effect on innovation performance [2–7,9,12,13]. Yet, few extant studies have examined how the big data-embedded NPD process affects sustainable innovation performance. Therefore, it is important to investigate the relationship between big data-embedded NPD processes and sustainable innovation performance [14,15].

Prior studies suggest that big data have effects on the NPD process [14,16–19]. Big data can help firms build a data-driven decision-making environment and increase the effectiveness of decision-making in NPD [14]. Moreover, big data can help firms identify customer needs, analyze business opportunities, optimize innovation processes, and increase innovation performance [16–19]. Since firms increasingly see big data as an opportunity rather than a cost, many now embed big data into various aspects of their NPD process. Unfortunately, studies suggest that embedding big data into the NPD process suffers the “Big Data Productivity Paradox” [19,20]. Furthermore, no studies
have examined what big data-embedded NPD process is and how a big data-embedded NPD process affects sustainable innovation performance. Therefore, we develop the following research questions for this study:

**RQ1:** Does a big data-embedded NPD process increase sustainable innovation performance?

**RQ2:** Do different stages of the big data-embedded NPD process have different effects on sustainable innovation performance?

**RQ3:** Which stage of big data-embedded NPD process has the highest effect on sustainable innovation performance?

Based on the NPD theory, we propose a research model and focus on five key stages of big data-embedded NPD process: big data-embedded idea development, business analysis, product design, product testing, and commercialization. We also develop new measurement scales for measuring big data-embedded NPD process. We hypothesize that big data-embedded NPD process has a positive relationship with sustainable innovation performance. We collect empirical data from 497 NPD projects in the United States of America (USA), 510 NPD projects in the United Kingdom (UK), and 851 NPD projects in Australia. The data collected include NPD projects in telecommunications, high-tech technology, automotive, pharmaceutical, and healthcare systems and services.

We find that big data-embedded business analysis, product design, and product testing increase sustainable innovation performance in the USA, UK, and Australia. However, in the USA, big data-embedded commercialization has no impact on sustainable innovation performance, and in Australia, big data-embedded idea development has a partial impact on sustainable innovation performance. Furthermore, the most important stage with regard to the impact on sales growth and gross margin is big data-embedded product testing in the USA, and big data-embedded commercialization in Australia. The most important stage with regard to the impact on second-year sales growth, first-year and third-year gross margin is big data-embedded commercialization in the UK, but the most important stage with regard to the impact on third-year sales growth and second-year gross margin is big data-embedded product testing in the UK.

We make three contributions to sustainable NPD literature. First, we enrich the NPD theory by identifying big data-embedded NPD processes and developing new measurement scales. Second, we advance extant NPD literature by empirically testing the effect of big data-embedded NPD process on sustainable innovation performance. Third, we enrich cross-national comparative literature by exploring similarities and differences in big data-embedded NPD. Finally, we also provide managerial implications for creating a big data-embedded NPD process.

### 2. Literature Review

#### 2.1. The NPD Process

The NPD literature suggests that the NPD process often comprises five stages, starting with idea development, then moving through business analysis, product design, product testing, and commercialization [1,3,4]. Each stage of the NPD process involves several specific NPD activities and tasks [7]. During the idea development stage, innovative ideas are proposed by evaluating customer needs and identifying business objectives. Good ideas are developed into full product concepts and poor ideas are removed from this stage [3,4]. During the business analysis stage, the NPD team further evaluates the market potential, competitive advantage, investment required, and reward potentials of the product concept. The purpose of the business analysis stage is to identify potentials and evaluate the necessary resources to fully developed the project concept. During the product design stage, the NPD team develops product prototypes, production models, or marketable products [3,4]. During the product testing stage, the NPD team adjusts product prototype according to the market and customer feedback and tests marketing and commercialization strategies. During the commercialization stage, the NPD team develops the final manufacturing and marketing plans, selects
business partners and suppliers, creates the overall direction of the commercialization, and introduces the products into the marketplace.

2.2. The NPD Process and Innovation Performance

Although there are many studies on the relationship between the NPD process and innovation performance, the findings are mixed. For example, many studies find that the NPD process helps firms embrace business opportunities, meet customer needs, and establish competitive advantages, which lead to improving innovation performance [1,3–6,9,11,13]. Other studies suggest that some stages of the NPD process do not increase innovation performance. For example, Barczak [21] finds that only the idea development stage significantly increases innovation performance. Rubera, Chandrasekaran, and Ordanini [22] find that the two early stages of the NPD process (idea development and business analysis) enhance innovation performance.

2.3. The Big Data-Embedded NPD Process

Most current research studies focus on how “experience-driven” NPD process affects innovation performance [1,3–6,9,11,13]. However, recent studies have found that the “experience-driven” NPD process has some problems such as higher development cost and risk, lower efficiency of decision-making, and longer development cycle [14]. Thus, scholars suggest that firms need to establish a big data-driven NPD process to solve these problems [14,23–25].

Big data are profoundly influencing the “experience-driven” NPD process [17]. On the one hand, embedding big data into the NPD process helps firms build “data-driven” decision-making environments, which lead to higher quality NPD decisions [14]. On the other hand, embedding big data into the NPD process helps firms accurately identify business opportunities, analyze customer needs, optimize the NPD process, and improve innovation performance [23–25]. Although the importance of big data usage in NPD is a significant issue in building competitive advantages, few studies have investigated what big data-embedded NPD is and how embedding big data into the NPD process can help companies achieve superior innovation performance.

3. Theoretical Model and Research Hypotheses

3.1. Theoretical Model

To address these existing literature gaps, we develop a research model based on the NPD theory in Figure 1 and suggest that the five stages of the big data-embedded NPD process enhance sustainable innovation performance. Big data-embedded NPD process refers to using big data in the five stages of the NPD process, including big data-embedded idea development, big data-embedded business analysis, big data-embedded product design, big data-embedded product testing, and big data-embedded commercialization.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Big data-embedded new product development (NPD) process and sustainable innovation performance.
3.2. Big Data-Embedded Idea Development and Sustainable Innovation Performance

Big data-embedded idea development refers to using big data to identify innovative ideas and key business opportunities, develop multiple product concepts, and develop NPD project proposals [3,4,15]. We argue that big data-embedded idea development positively affects sustainable innovation performance. On the one hand, big data-embedded idea development can help firms obtain comprehensive market and customer needs information. By analyzing the information, the NPD project team can better generate creative ideas [13]. On the other hand, big data-embedded idea development can help firms identify key business opportunities and provide a data-driven NPD proposal to guide the NPD project team to develop new product concepts, which are vital for NPD projects to profit and succeed [13,15,26,27]. Thus, the first formal hypothesis is proposed:

**Hypothesis 1.** Big data-embedded idea development increases sustainable innovation performance.

3.3. Big Data-Embedded Business Analysis and Sustainable Innovation Performance

Big data-embedded business analysis refers to using big data to conduct market opportunity analyzes (market potential, customer preference, purchase process, and buyer adoption), appraise current and future products and services, determine desired innovative features and development feasibility, assess costs, time, risks, and business implications, and evaluate internal and external innovation capability and resources [3,4,15]. We argue that big data-embedded business analysis positively affects sustainable innovation performance. On the one hand, big data-embedded business analysis can help firms accurately analyze business information related to market potential, customer preferences, purchase process, and buyer adoption [28–30]. The market and customer knowledge provided by analyzing big data help firms select good product concepts and determine potential innovative attributes and characteristics. On the other hand, big data-embedded business analysis can help firms effectively assess NPD’s costs, time, and risks [30]. Therefore, the second formal hypothesis is developed:

**Hypothesis 2.** Big data-embedded business analysis increases sustainable innovation performance.

3.4. Big Data-Embedded Product Design and Sustainable Innovation Performance

Big data-embedded product design refers to using big data to conduct engineering, technological, and manufacturing assessments, develop technologies, product features, and product prototypes, evaluate prototypes against performance specifications, determine the final product design, and perform cost reduction and quality control tasks [3,4,15]. We argue that big data-embedded product design positively affects sustainable innovation performance. First, big data-embedded product design can help firms better understand the existing engineering, technological, and manufacturing capabilities and thereby provide effective guidance for new product prototype development [31]. Second, big data-embedded product design enables customers to participate in product design through data. Thus, firms can adjust product prototypes in time to better meet customer needs [3,4,15]. Third, big data-embedded product design can help firms integrate digital technologies in the NPD process, which is lead to increasing differentiation characteristics and competitive advantages for the product prototype. Finally, a higher quality of big data-embedded product design also resolves the trade-offs between product costs and product quality, which will lead to lower cost and higher value for customers. Therefore, the third formal hypothesis is proposed:

**Hypothesis 3.** Big data-embedded product design increases sustainable innovation performance.
3.5. Big Data-Embedded Product Testing and Sustainable Innovation Performance

Big data-embedded product testing refers to using big data to evaluate prototypes, market performance, and different product and service offers, test alternative technologies, simulate customer acceptance and customer use, and assess feedback from technology and market tests [3,4,15]. We argue that big data-embedded product testing positively affects sustainable innovation performance. First, big data-embedded product testing can help firms obtain customer feedback information on the use of new product prototypes. Therefore, firms can find the deficiencies and defects in product prototypes and then adjust product attributes and functions to solve these problems [3,4,15]. Second, big data-embedded product testing can help firms simulate consumer behavior and better predict customer buy decisions. Third, big data-embedded product testing can help firms predict possible successful marketing and commercialization strategies. Therefore, the fourth formal hypothesis is developed:

**Hypothesis 4.** Big data-embedded product testing increases sustainable innovation performance.

3.6. Big Data-Embedded Commercialization and Sustainable Innovation Performance

Big data-embedded commercialization refers to using big data to evaluate suppliers and commercialization partners, complete the final plans for production, evaluate and complete the final commercialization plans (the timing of product launch), develop pricing strategies and tactics, and launch the innovation in the marketplace (sales, promotion, and distribution) [3,4,15]. We argue that big data-embedded commercialization positively affects sustainable innovation performance. First, big data-embedded commercialization can help firms accurately assess and select suppliers and business partners, resulting in providing sufficient guarantee for the final production and sale [32]. Second, big data-embedded commercialization can help firms better understand customer’s price expectations, as well as their purchase channels and promotion preferences, which lead to accurately located target market segments, determined appropriate product prices, selected appropriate distribution and promotion strategies. Third, big data-embedded commercialization can help firms develop digital marketing methods (e.g., email marketing, social network marketing, viral marketing, and mobile marketing) [13]. Thus, customers can easily compare new products and thereby reach a quick decision about which new products to purchase. Finally, a higher quality of big data-embedded commercialization increases customer experience with new product purchases and uses. Therefore, the fifth formal hypothesis is proposed:

**Hypothesis 5.** Big data-embedded commercialization increases sustainable innovation performance.

4. Methodology

In order to empirically test five research hypotheses, we collected project-level data which include 1858 NPD projects with the following characteristics: (1) 497 NPD projects were from the USA, 510 NPD projects were from the UK, and 851 NPD projects were from Australia, (2) the projects include five product industries (telecommunications, high-tech technology, automotive, pharmaceutical, and healthcare systems and services), and (3) to be included in this study, all project-level data must include sales growth and gross margin for the first three years after the commercialization.

4.1. Overall Research Design

We executed this study in three phases. In the first phase, we followed the methods in Douglas and Craig [33] and Song and Montoya-Weiss [5] to develop new measurement scales to assess the levels of the big data-embedded NPD process and to select methods to implement the cross-national comparative study. In the second phase, we administered the survey to collect data regarding five
stages of the big data-embedded NPD process. In the third phase, we collected three-year sales growth and gross margin data.

4.2. Measuring Scale Development

The current NPD and management literature did not have existing measures for assessing the big data-embedded NPD process. Therefore, we followed a three-step procedure reported by Song and Montoya-Weiss [5] to develop a specific measurement scale for measuring big data-embedded NPD process.

First, we conducted in-depth case studies of 18 big data-embedded NPD projects as well as focus-group interviews with NPD project team members in the USA, UK, and Australia. The focus-group interviews are divided into three parts. The first part was designed to examine the conceptual equivalence of the constructs by asking team members to define each of the variables [5,33]. The second part was designed to assess the functional equivalence of the constructs by asking team members to evaluate the relationship between the constructs [5,33]. The third part was designed to evaluate whether or not there is a category equivalence of the theoretical variables by examining the relevance and completeness of the measurement scale items [5,33]. These interviews suggested some modifications to the scale items.

Second, we adopted Churchill’s method to develop big data-embedded NPD scales [34] by involving academic experts in the USA, UK, and Australia, and 18 NPD project leaders. Experts were asked to evaluate measurement items and provide recommendations for improvement [5]. Minor revisions and a further review by the experts and leaders resulted in measurement scales with high consistency and face validity.

Third, we conducted two pretests to further validate the measurement scales [5]. The first pretest was on-site to review the questionnaire to determine whether any items were unclear. The second pretest was through a professionally drafted survey that was administered to 32 project team members. We used these pretests to further modify the measurement scales.

4.3. Variable Measurement

To measure sustainable innovation performance, we used objective performance data (sales growth and gross margin) adopted by Hao et al. [19] and Hu et al. [35].

To measure the big data-embedded NPD process, we used five 4-item scales for the five stages of the NPD process. Respondents rated each measure from 0 to 10 (big data was extensively used). The big data-embedded idea development includes (a) using big data to identify innovative ideas, (b) using big data to develop product concepts, (c) using big data to identify key business opportunities, and (d) using big data to develop written project proposals.

The big data-embedded business analysis measures the extent of (a) using big data to conduct market opportunity analyses, (b) using big data to appraise current and future products and services, (c) using big data to determine the desired innovative features and development feasibility, and (d) using big data to measure costs, time, risks, and business implications.

The measures for the big data-embedded product design are: (a) using big data to conduct engineering, technological, and manufacturing assessments, (b) using big data to develop technologies and product features, (c) using big data to develop product prototypes, and (d) using big data to evaluate prototypes against performance specifications.

The big data-embedded product testing is measured with four measures: (a) using big data to evaluate prototype market performance, (b) using big data to test alternative technologies, (c) using big data to simulate customer acceptance and customer use tests, and (d) using big data to evaluate different product and service offers.

The measurement items for the big data-embedded commercialization include four measures: (a) using big data to evaluate suppliers and commercialization partners, (b) using big data to complete
the final plans for production, (c) using big data to evaluate, and (d) using big data to determining the
timing of product launch and pricing strategies.

In addition to the independent variable and dependent variable, past studies suggest that industry
characteristics may influence sustainable innovation performance [19]. Thus, we included five industry
dummies as control variables: IND1 = telecommunications industry, IND2 = high-tech technology
industry, IND3 = automotive industry, IND4 = pharmaceutical industry, and IND5 = healthcare
systems and services industry.

4.4. The Data Collection Method

In the first step, we sent express mail and/or e-mail to all selected companies with a cover letter,
a one-page presurvey, a data confidential agreement, and a list of free research reports for participating
organizations. In the second step, we adopted a survey methodology to collect the five stages of
the big data-embedded NPD process data from the companies from all companies that responded
during the first step [36]. Following the survey methodology, we sent a business card, a cover letter,
the survey, a prepaid express mail return envelope, and a list of free research reports selected by the
companies. We asked each company to provide four NPD projects if possible: a recently completed
project, a typical project, a successful project, and a failed project. We followed up twice and made
multiple phone calls to project managers to increase the participation rate.

4.4.1. The USA Data

We selected 1000 firms randomly from the companies in the Russell 3000 Index and the Dun &
Bradstreet database. We strictly followed the above data collection method and was able to get 176 USA
firms to provide data for at least one NPD project. The participation rate at the firm level was 17.6%.

From 176 firms, we collected 558 big data-embedded NPD projects, including 106 from
telecommunications, 96 from automotive, 138 from pharmaceutical, 92 from healthcare systems
and services, and 126 from high-technology. We also tracked three years and collected sales growth
and gross margin but were unable to get sales and gross margin data for 61 projects. Therefore, we ended
up with complete data for a sample of 497 big data-embedded NPD projects in the USA.

Among the 497 respondents, 67.03% was male and 32.97% was female; 19.18% has bachelor
degrees and 80.82% has master degrees or above; 10.22% was president, 15.59% was division manager,
21.51% was strategic business manager, 32.97% was new business project manager, and 19.71% was
R&D director; the average age was 39.87 years; the average industry experience was 14.09 years; 39.25%
developed less than 5 NPD projects, 53.05% developed 6–10 NPD projects, 3.58% developed 11–15 NPD
projects, 2.69% developed 16–20 NPD projects, and 1.43% developed more than 20 NPD projects.

4.4.2. The UK Data

From the listings of the Dun & Bradstreet database and the World Business Directory, we identified
companies that develop products in the same industries as in the USA sample. We randomly selected
1500 firms from the final list for conducting data collection. We used the same data collection protocol
as described above for the USA and collected 595 big data-embedded NPD projects from 247 UK firms.
The participation rate at the firm level was 16.5%.

The NPD projects were divided 73 from telecommunications, 115 from automotive, 153 from
pharmaceutical, 98 from healthcare systems and services, and 156 from high-technology. Unfortunately,
85 projects of the 595 projects failed to provide sales and gross margin data. Therefore, the final sample
for this study included only 510 big data-embedded NPD projects in the UK.

Among the 510 respondents, 67.56% was male and 32.44% was female; 17.82% has bachelor
degrees and 82.18% has master degrees or above; 9.92% was president, 16.47% was division manager,
21.51% was strategic business manager, 31.60% was new business project manager, and 20.50% was
R&D director; the mean value for the age was 41.28 years; the mean working experience was 19.19 years;
2.35% developed 5 NPD projects, 16.98% developed 6–10 NPD projects, 39.50% developed 11–15 NPD projects, 26.55% developed 16–20 NPD projects, and 14.64% developed more than 20 NPD projects.

4.4.3. Australia Data

We used the same sample selection protocol of the UK sample to select Australian firms. Following the sample selection criteria as in the UK sample, we randomly selected 1500 firms from the Dun & Bradstreet database and the World Business Directory. We collected data on 924 big data-embedded NPD projects from 341 Australian firms. The participation rate at the firm level was 22.7%.

The 924 NPD projects included 187 from telecommunications, 197 from automotive, 229 from pharmaceutical, 144 from healthcare systems and services, and 167 from high-technology. Unfortunately, 73 projects failed to provide performance data, we ended up with complete data for a sample of 851 big data-embedded NPD projects in Australian.

The data providers’ characteristics are: 66.67% was male and 33.33% was female; 17.86% has bachelor degrees and 82.14% has master degrees or above; 10.28% was president, 15.80% was division manager, 21.65% was strategic business manager, 32.90% was new business project manager, 21.12% was departmental manager, 21.12% was project team leader, 17.86% was R&D director; the average age was 39.61 years; the mean industry working experience was 5.12 years; 31.62% was senior manager, 45.35% was manager, 23.03% was team leader; 43.62% was middle age; 42.01% was over 40 years old; 53.33% was R&D; 36.95% was business; 50.00% was pharmaceutical; 25.86% was healthcare systems and services; 24.04% was high-technology.

In Table 1, we report means, standard deviations, and correlations for all variables in the three samples. Results from Table 1 provide some initial results of the variables. We also reported Cronbach’s alpha in Table 1 to evaluate reliability. The Cronbach’s alpha ranged from 0.739 to 0.914 in the USA sample, 0.758 to 0.908 in the UK sample, and 0.702 to 0.904 in the Australian sample, indicating that all constructs’ reliability is acceptable.

Table 1. Descriptive statistics of the sample.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|---|---|---|---|---|---|---|---|---|----|
| Mean     | 74.61 177.141 29.127 45.047 84.784 4.743 6.282 5.498 5.350 5.276 |
| S.D.     | 58.311 99.220 12.296 14.772 28.549 2.493 1.983 2.857 2.148 2.190 |
| 1.SALEG1 | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. |
| 2.SALEG2 | 0.964*** | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. |
| 3.GMY1   | 0.895*** | 0.884*** | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. |
| 4.GMY2   | 0.828*** | 0.835*** | 0.933*** | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. |
| 5.GMY3   | 0.841*** | 0.837*** | 0.943*** | 0.935*** | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. |
| 6.BDIDEAA | 0.425*** | 0.414*** | 0.383*** | 0.328*** | 0.362*** | 0.651 | n.a. | n.a. | n.a. | n.a. |
| 7.BDBMOA | 0.443*** | 0.434*** | 0.420*** | 0.395*** | 0.400*** | 0.345*** | 0.657 | n.a. | n.a. | n.a. |
| 8.BDDES | 0.230*** | 0.229*** | 0.266*** | 0.274*** | 0.248*** | 0.122*** | 0.130*** | 0.791 | n.a. | n.a. |
| 9.BDTEST | 0.524*** | 0.519*** | 0.489*** | 0.475*** | 0.485*** | 0.451*** | 0.345*** | 0.100*** | 0.757 | n.a. |
| 10.BDCOMM | -0.022 -0.021 -0.004 0.000 -0.019 -0.066 -0.167*** -0.099*** -0.040 0.875 |
| Cronbach’s α | 0.779 0.739 0.828 0.833 0.914 | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. | n.a. |

5. Results

5.1. Basic Statistics

In Table 1, we report means, standard deviations, and correlations for all variables in the three samples. Results from Table 1 provide some initial results of the variables. We also reported Cronbach’s alpha in Table 1 to evaluate reliability. The Cronbach’s alpha ranged from 0.739 to 0.914 in the USA sample, 0.758 to 0.908 in the UK sample, and 0.702 to 0.904 in the Australian sample, indicating that all constructs’ reliability is acceptable.
Table 1. Cont.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|---|---|---|---|---|---|---|---|---|----|
| 1.SALEG1 | n.a. |  |  |  |  |  |  |  |  |  |
| 2.SALEG2 | 0.927*** | n.a. |  |  |  |  |  |  |  |  |
| 3.GMY1   | 0.814*** | 0.797*** | n.a. |  |  |  |  |  |  |  |
| 4.GMY2   | 0.683*** | 0.684*** | 0.817*** | n.a. |  |  |  |  |  |  |
| 5.GMY3   | 0.717*** | 0.709*** | 0.849*** | 0.849*** | n.a. |  |  |  |  |  |
| 6.BDIDEAA | 0.321*** | 0.307*** | 0.354*** | 0.294*** | 0.325*** | 0.716 |  |  |  |  |
| 7.BDBMOA | 0.410*** | 0.406*** | 0.365*** | 0.303*** | 0.320*** | 0.250*** | 0.627 |  |  |  |
| 8.BDDESI | 0.197*** | 0.174*** | 0.191*** | 0.157*** | 0.143*** | 0.008 | 0.140*** | 0.813 |  |  |
| 9.BDTEST | 0.458*** | 0.442*** | 0.455*** | 0.365*** | 0.396*** | 0.394*** | 0.358*** | 0.056*** | 0.778 |  |
| 10.BDCOMM | 0.462*** | 0.463*** | 0.458*** | 0.414*** | 0.435*** | 0.530*** | 0.220*** | 0.105*** | 0.405*** | 0.735 |
| Mean     | 81.015 | 185.684 | 82.469 | 116.646 | 138.099 | 4.802 | 6.350 | 5.792 | 5.665 | 5.216 |
| S.D.     | 60.533 | 92.343 | 34.400 | 49.110 | 56.527 | 2.098 | 1.850 | 2.843 | 2.121 | 2.121 |
| Cronbach's α | 0.819 | 0.702 | 0.904 | 0.858 | 0.923 |

Note: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$ (all significant levels are based on two-tailed test). SALEG1 = Second-year sales growth; SALEG2 = Third-year sales growth; GMY1 = First-year gross margin; GMY2 = Second-year gross margin; GMY3 = Third-year gross margin; BDIDEAA = Big data-embedded idea development; BDBMOA = Big data-embedded business analysis; BDDESI = Big data-embedded product design; BDTEST = Big data-embedded product testing; BDCOMM = Big data-embedded commercialization. The numbers on diagonal are square root of AVE; n.a. indicates the non-multiitem scales.

5.2. Measurement Model Validation Using CFA

We performed a confirmative factor analysis (CFA) to validate the measurement model fit using AMOS software [37]. The initial measurement model for each country was run by using all measures. The fit indices and modification indices suggested that the initial models can be significantly improved by adding four correlations among four errors of the four measures (as shown in Figure 2): (1) the error term of BDIDEAA2 should be correlated with the error term of BDIDEAA1 and BDIDEAA4 respectively, (2) the error terms of BDDESI1 and BDDESI2 should be correlated, and (3) the error term of BDDESI3 and BDDESI4 should be correlated. Table 2 presents the relevant statistics of the final measurement models for all three countries.
Consistent with prior studies [5,38], Table 2 suggests that the empirical data have good fit measurement models: (1) for the USA sample, RMSEA = 0.070; χ²/degrees of freedom (df) = 3.761; CFI = 0.928; NFI = 0.905; GFI = 0.907. (2) for the UK sample, RMSEA = 0.063; and χ²/degrees of freedom (df) = 3.346; CFI = 0.960; NFI = 0.943; GFI = 0.923. (3) for Australian sample, CFI = 0.943; NFI = 0.931; GFI = 0.921; RMSEA = 0.069; χ²/degrees of freedom (df) = 5.350; given the large sample in Australian sample, the χ²/degrees of freedom (df) is acceptable [5,38].

To assess convergent validity, we also used Cronbach’s alphas. As shown in Table 1, the lowest alpha is 0.739 for the USA sample, 0.758 for the UK sample, and 0.702 for Australia sample. We then computed the average variance explained (AVE) for each construct to evaluate discriminant validity (as shown in the last column of Table 2).

### Table 2. Statistics for the measurement model.

| Construct Items | Unstandardized Estimate | Standard Error | Z-Value | p-Value | Standardized Estimate | AVE  |
|-----------------|-------------------------|----------------|---------|---------|-----------------------|------|
| BDIDEAA1        | 1                       | 0.539          | 0.424   |         |                       |      |
| BDIDEAA2        | 1.205                   | 0.094          | 12.796  | ***     | 0.678                 |      |
| BDIDEAA3        | 1.665                   | 0.163          | 10.192  | ***     | 0.864                 |      |
| BDIDEAA4        | 0.785                   | 0.163          | 8.38    | ***     | 0.447                 |      |
| BDIMOA1         | 1                       | 0.758          |         |         |                       |      |
| BDIMOA2         | 0.989                   | 0.073          | 13.577  | ***     | 0.743                 |      |
| BDIMOA3         | 0.793                   | 0.077          | 10.324  | ***     | 0.499                 |      |
| BDIMOA4         | 0.785                   | 0.065          | 12.084  | ***     | 0.59                  |      |
| BDDES1          | 1                       | 0.596          |         |         |                       |      |
| BDDES2          | 1.192                   | 0.074          | 16.067  | ***     | 0.695                 | 0.625|
| BDDES3          | 1.172                   | 0.048          | 10.324  | ***     | 0.755                 |      |
| BDDES4          | 1.6                     | 0.635          | 2.521   | 0.015   | 0.755                 |      |
| BDTEST1         | 1                       | 0.796          |         |         |                       |      |
| BDTEST2         | 1.201                   | 0.056          | 21.58   | ***     | 0.892                 | 0.573|
| BDTEST3         | 0.765                   | 0.052          | 14.54   | ***     | 0.618                 |      |
| BDTEST4         | 0.938                   | 0.056          | 16.619  | ***     | 0.693                 |      |
| BDCOMM1         | 1                       | 0.929          |         |         |                       |      |
| BDCOMM2         | 0.988                   | 0.023          | 42.831  | ***     | 0.765                 |      |
| BDCOMM3         | 1.011                   | 0.019          | 52.342  | ***     | 0.79                  |      |
| BDCOMM4         | 0.564                   | 0.035          | 16.047  | ***     | 0.581                 |      |
| Overall Model   |                         | χ²/degrees of freedom (df) = 586.764/156 = 3.761; CFI = 0.928; NFI = 0.905; GFI = 0.907; IFI = 0.929; RMSEA = 0.070. |

| Construct Items | Unstandardized Estimate | Standard Error | Z-Value | p-Value | Standardized Estimate | AVE  |
|-----------------|-------------------------|----------------|---------|---------|-----------------------|------|
| BDIDEAA1        | 1                       | 0.89           |         |         |                       |      |
| BDIDEAA2        | 0.996                   | 0.029          | 34.115  | ***     | 0.885                 | 0.697|
| BDIDEAA3        | 1.041                   | 0.03           | 34.785  | ***     | 0.993                 |      |
| BDIDEAA4        | 0.507                   | 0.04           | 12.571  | ***     | 0.474                 |      |
| BDIMOA1         | 1                       | 0.755          |         |         |                       |      |
| BDIMOA2         | 1.013                   | 0.074          | 13.673  | ***     | 0.751                 | 0.449|
| BDIMOA3         | 0.969                   | 0.078          | 12.404  | ***     | 0.591                 |      |
| BDIMOA4         | 0.792                   | 0.064          | 12.3    | ***     | 0.559                 |      |
| BDDES1          | 1                       | 0.537          |         |         |                       |      |
| BDDES2          | 1.991                   | 0.142          | 13.99   | ***     | 1.065                 | 0.672|
| BDDES3          | 1.483                   | 0.137          | 10.782  | ***     | 0.881                 |      |
| BDDES4          | 1.197                   | 0.118          | 10.185  | ***     | 0.699                 |      |
| BDTEST1         | 1                       | 0.79           |         |         |                       |      |
| BDTEST2         | 1.153                   | 0.052          | 22.068  | ***     | 0.878                 | 0.577|
| BDTEST3         | 0.798                   | 0.05           | 15.933  | ***     | 0.657                 |      |
| BDTEST4         | 0.907                   | 0.053          | 17.094  | ***     | 0.694                 |      |
| Overall Model   |                         | χ²/degrees of freedom (df) = 521.980/156 = 3.346; CFI = 0.960; NFI = 0.943; GFI = 0.923; IFI = 0.960; RMSEA = 0.063. |
Table 2. Cont.

c. Results of Measurement Model Validation Using CFA for Australia Sample (N = 924).

| Construct Items | Unstandardized Estimate | Standard Error | Z-Value | p-Value | Standardized Estimate | AVE |
|-----------------|--------------------------|----------------|---------|---------|------------------------|-----|
| BDIDEAA1        | 1                        |                |         |         |                        |     |
| BDIDEAA2        | 0.689                    | 0.05           | 17.64   | ***     | 0.657                  | 0.513|
| BDIDEAA3        | 1.144                    | 0.054          | 20.949  | ***     | 0.825                  |     |
| BDIDEAA4        | 0.777                    | 0.046          | 16.712  | ***     | 0.603                  |     |
| BDDESI1         | 1                        |                |         |         |                        |     |
| BDDESI2         | 2.094                    | 0.274          | 7.644   | ***     | 1.08                   | 0.661|
| BDDESI3         | 1.499                    | 0.169          | 8.875   | ***     | 0.861                  |     |
| BDDESI4         | 1.203                    | 0.139          | 8.633   | ***     | 0.666                  |     |
| BDTEST1         | 1                        |                |         |         |                        |     |
| BDTEST2         | 1.055                    | 0.041          | 25.887  | ***     | 0.842                  | 0.605|
| BDTEST3         | 0.911                    | 0.042          | 21.945  | ***     | 0.732                  |     |
| BDTEST4         | 0.969                    | 0.043          | 22.561  | ***     | 0.754                  |     |
| BDCOMM1         | 1                        |                |         |         |                        | 0.671|
| BDCOMM2         | 1.143                    | 0.061          | 18.654  | ***     | 0.725                  | 0.54 |
| BDCOMM3         | 1.239                    | 0.063          | 19.826  | ***     | 0.828                  |     |
| BDCOMM4         | 1.054                    | 0.057          | 18.369  | ***     | 0.707                  |     |
| Overall Model   |                         | χ²/degrees of freedom (df) = 834.635/156 = 5.350; |          |         |                        |     |
|                 |                          | CFI = 0.943; NFI = 0.931; IFI = 0.943; RMSEA = 0.069. |          |         |                        |     |

Note: *** p < 0.001 (two-tailed test). BDIDEAA = Big data-embedded idea development; BDBMOA = Big data-embedded business analysis; BDDESI = Big data-embedded product design; BDTEST = Big data-embedded product testing; BDCOMM = Big data-embedded commercialization.

As shown in the lower left off-diagonal of Table 1, the lowest square root of the AVE (USA: 0.750; UK: 0.670; Australian: 0.627) was greater than the highest correlation between that construct and other constructs (USA: 0.524; UK: 0.528; Australian: 0.530). Thus, prior studies suggest that these results indicate acceptable discriminant validity [39].

5.3. Hypothesis Testing

The research model has following five equations:

\[
\text{SALEG1} = \alpha + \beta_1 \times \text{BDIDEAA} + \beta_2 \times \text{BDBMOA} + \beta_3 \times \text{BDDESI} + \beta_4 \times \text{BDTEST} + \beta_5 \times \text{BDCOMM} + \beta_6 \times \text{IND1} + \beta_7 \times \text{IND2} + \beta_8 \times \text{IND3} + \beta_9 \times \text{IND4} + \varepsilon \tag{1}
\]

\[
\text{SALEG2} = \alpha + \beta_1 \times \text{BDIDEAA} + \beta_2 \times \text{BDBMOA} + \beta_3 \times \text{BDDESI} + \beta_4 \times \text{BDTEST} + \beta_5 \times \text{BDCOMM} + \beta_6 \times \text{IND1} + \beta_7 \times \text{IND2} + \beta_8 \times \text{IND3} + \beta_9 \times \text{IND4} + \varepsilon \tag{2}
\]

\[
\text{GMY1} = \alpha + \beta_1 \times \text{BDIDEAA} + \beta_2 \times \text{BDBMOA} + \beta_3 \times \text{BDDESI} + \beta_4 \times \text{BDTEST} + \beta_5 \times \text{BDCOMM} + \beta_6 \times \text{IND1} + \beta_7 \times \text{IND2} + \beta_8 \times \text{IND3} + \beta_9 \times \text{IND4} + \varepsilon \tag{3}
\]

\[
\text{GMY2} = \alpha + \beta_1 \times \text{BDIDEAA} + \beta_2 \times \text{BDBMOA} + \beta_3 \times \text{BDDESI} + \beta_4 \times \text{BDTEST} + \beta_5 \times \text{BDCOMM} + \beta_6 \times \text{IND1} + \beta_7 \times \text{IND2} + \beta_8 \times \text{IND3} + \beta_9 \times \text{IND4} + \varepsilon \tag{4}
\]

\[
\text{GMY3} = \alpha + \beta_1 \times \text{BDIDEAA} + \beta_2 \times \text{BDBMOA} + \beta_3 \times \text{BDDESI} + \beta_4 \times \text{BDTEST} + \beta_5 \times \text{BDCOMM} + \beta_6 \times \text{IND1} + \beta_7 \times \text{IND2} + \beta_8 \times \text{IND3} + \beta_9 \times \text{IND4} + \varepsilon \tag{5}
\]

where: SALEG1 = Second-year sales growth; SALEG2 = Third-year sales growth; GMY1 = First-year gross margin; GMY2 = Second-year gross margin; GMY3 = Third-year gross margin; BDIDEAA = Big data-embedded idea development; BDBMOA = Big data-embedded business analysis; BDDESI = Big data-embedded product design; BDTEST = Big data-embedded product testing; BDCOMM = Big data-embedded commercialization; IND1 = Telecommunications industry; IND2 = High-tech...
Technology industry; IND3 = Automotive industry; IND4 = Pharmaceutical industry; \( \alpha \) = intercept; \( \beta \) = Coefficient estimate; \( \varepsilon \) = disturbance terms for the equation.

Since we used objective performance data (three-year sales and gross margin data), we followed Zhang et al. [40] and used ordinary least squares (OLS) estimation to estimate these five equations. Prior studies suggest that OLS estimation procedure is more appropriate when the dependent variables are objective data with wide ranges [40]. The estimation results for Equations (1) and (2) are presented in Table 3 and the estimations for Equations (3)–(5) are presented in Table 4. To summarize the results of hypothesis testing, we present in Table 5.

Table 3. Regression results of big data-embedded NPD process on sale growth.

| dependent Variable: Sale Growth | USA Sample (N = 497) | UK Sample (N = 510) | Australia Sample (N = 851) |
|--------------------------------|----------------------|---------------------|---------------------------|
|                                | \( \beta \) (S.E.)  | \( \beta \) (S.E.)  | \( \beta \) (S.E.)  |
| Intercept                      | −53.947*** (11.489) | −45.959 (33.932)  | −59.549*** (16.004)  |
|                               | 4.890*** (1.172)    | 13.532*** (3.461) | 2.273** (0.987)     |
| BDIDEAA                       | 8.272*** (1.249)    | 23.714*** (3.688) | 5.380*** (1.132)    |
| BDBMOA                        | 3.253*** (0.831)    | 9.522*** (2.456)  | 2.225*** (0.717)    |
| BDDESI                        | 10.757*** (1.167)   | 31.399*** (3.446) | 6.620*** (1.205)    |
| BDTEST                        | 0.475 (0.907)       | 1.442 (2.678)     | 7.190*** (1.146)    |
| IND1                          | 3.277 (7.641)       | 7.502 (22.566)    | 4.691 (7.691)       |
| IND2                          | 10.517 (7.289)      | 22.751 (21.527)   | 9.375 (6.385)       |
| IND3                          | 4.060 (7.758)       | 4.981 (22.912)    | 3.783 (6.803)       |
| IND4                          | 12.114 (7.241)      | 25.047 (21.387)   | 0.207 (6.353)       |
| F-value \( R^2 \)             | 37.630*** (5.608)   | 35.520*** (5.531) | 37.430*** (7.626)   |
| Adjusted \( R^2 \)            | 0.410 (0.572)       | 0.396 (0.564)     | 0.403 (0.778)       |

Note: * \( p < 0.10; ** \( p < 0.05; *** \( p < 0.01. \) \, \text{SALEG1} = \text{Second-year sales growth}; \text{SALEG2} = \text{Third-year sales growth}; \text{BDIDEAA} = \text{Big data-embedded idea development}; \text{BDBMOA} = \text{Big data-embedded business analysis}; \text{BDDESI} = \text{Big data-embedded product design}; \text{BDTEST} = \text{Big data-embedded product testing}; \text{BDCOMM} = \text{Big data-embedded commercialization}. \text{IND}1 = \text{Telecommunications industry}; \text{IND}2 = \text{High-Tech technology industry}; \text{IND}3 = \text{Automotive industry}; \text{IND}4 = \text{Pharmaceutical industry}; \text{IND}5 = \text{Healthcare Systems and Services industry}. \text{IND}5 \text{ is the baseline; } \beta = \text{Coefficient estimate; S.E. = Standard error.}

Table 4. Regression results of big data-embedded NPD process on gross margin.

| dependent Variable: Gross Margin | USA Sample (N = 497) | UK Sample (N = 510) | Australia Sample (N = 851) |
|---------------------------------|----------------------|---------------------|---------------------------|
|                                | \( \beta \) (S.E.)  | \( \beta \) (S.E.)  | \( \beta \) (S.E.)  |
| Intercept                      | 4.856*** (5.608)    | 19.401*** (5.331)  | 20.624*** (7.626)  |
|                               | 1.878*** (0.572)    | 1.192*** (0.564)   | 2.136*** (0.778)    |
| BDIDEAA                        | 1.878*** (0.572)    | 1.192*** (0.564)   | 2.136*** (0.778)    |
|                               | 0.425*** (0.209)    | 0.538*** (0.265)   | 1.210** (0.498)     |
|                               | 0.370 (0.561)       | 0.373 (0.858)      | 0.366 (0.969)       |
Sustainability 2020, 12, 4681

Table 4. Cont.

| USA Sample (N = 497) | UK Sample (N = 510) | Australia Sample (N = 851) |
|----------------------|---------------------|-----------------------------|
| GMY1  | GMY2  | GMY3  | GMY1  | GMY2  | GMY3  | GMY1  | GMY2  | GMY3  |
| β (S.E.)  | β (S.E.)  | β (S.E.)  | β (S.E.)  | β (S.E.)  | β (S.E.)  | β (S.E.)  | β (S.E.)  | β (S.E.)  |
| BDBMOA | 3.758*** | 3.382*** | 4.639*** | 0.667*** | 0.758*** | 1.180*** | 3.298*** | 3.974*** | 4.773*** |
| (0.609) | (0.601) | (0.829) | (0.240) | (0.303) | (0.571) | (0.571) | (0.974) | (0.986) |
| BDDESI | 2.000*** | 2.083*** | 2.492*** | 0.503*** | 0.347*** | 0.746*** | 1.587*** | 1.796*** | 1.740*** |
| (0.406) | (0.400) | (0.552) | (0.152) | (0.192) | (0.361) | (0.343) | (0.525) | (0.592) |
| BDTEST | 4.766*** | 4.704*** | 6.561*** | 1.524*** | 1.884*** | 3.563*** | 4.085*** | 4.158*** | 5.289*** |
| (0.570) | (5.562) | (0.775) | (0.255) | (0.323) | (0.608) | (0.528) | (0.808) | (0.912) |
| BDCOMM | 0.383 | 0.330 | 0.219 | 1.723*** | 1.672*** | 3.808*** | 4.406*** | 6.434*** | 7.493*** |
| (0.443) | (0.437) | (0.602) | (0.243) | (0.307) | (0.578) | (0.556) | (0.850) | (0.960) |
| IND1 | −1.000 | 0.128 | −2.882 | 0.022 | −0.433 | −1.670 | 0.367 | −1.032 | 1.460 |
| (3.730) | (3.679) | (5.072) | (1.628) | (2.061) | (3.878) | (3.245) | (4.966) | (5.605) |
| IND2 | 1.824 | 1.824 | 2.445 | 2.153 | 1.658 | 3.407 | 2.949 | 1.262 | 1.301 |
| (3.558) | (3.509) | (4.838) | (1.352) | (1.711) | (3.220) | (3.334) | (5.103) | (5.760) |
| IND3 | −0.423 | 0.771 | −0.119 | 1.198 | 1.727 | 2.011 | 2.026 | −1.814 | −0.233 |
| (3.787) | (3.735) | (5.149) | (1.441) | (1.823) | (3.430) | (3.195) | (4.090) | (5.519) |
| IND4 | 1.732 | 1.925 | 1.982 | 1.408 | 1.890 | 2.600 | 0.530 | 0.789 | 2.741 |
| (3.353) | (3.486) | (4.807) | (1.345) | (1.703) | (3.203) | (3.309) | (4.743) | (5.394) |

Note: *p < 0.10; **p < 0.05; ***p < 0.01. GMY1 = First-year gross margin; GMY2 = Second-year gross margin; GMY3 = Third-year gross margin; BDDEAI = Big data-embedded idea development; BDBMOA = Big data-embedded business analysis; BDDESI = Big data-embedded product design; BDTEST = Big data-embedded product testing; BDCOMM = Big data-embedded commercialization. IND1 = Telecommunications industry; IND2 = High-tech technology industry; IND3 = Automotive industry; IND4 = Pharmaceutical industry. IND5 = Healthcare Systems and Services. IND5 is set to be the baseline; β = Coefficient estimate; S.E. = standard error.

Table 5. Summary of hypotheses testing in USA, UK, and Australia samples.

| Hypothesis | USA Sample (N = 497) | UK Sample (N = 510) | Australia Sample (N = 851) |
|------------|---------------------|---------------------|-----------------------------|
| H1: Big data-embedded idea development increases sustainable innovation performance. | Supported | Supported | Partial supported |
| H2: Big data-embedded business analysis increases sustainable innovation performance. | Supported | Supported | Supported |
| H3: Big data-embedded product design increases sustainable innovation performance. | Supported | Supported | Supported |
| H4: Big data-embedded product testing increases sustainable innovation performance. | Supported | Supported | Supported |
| H5: Big data-embedded commercialization increases sustainable innovation performance. | Not supported | Supported | Supported |

5.3.1. USA Results

For the USA sample, the numbers in the first two columns in Table 3 indicate that the effects of big data-embedded idea development ($\beta_{SALEG1} = 4.890$, $\beta_{SALEG2} = 13.532$, $p < 0.01$), business analysis ($\beta_{SALEG1} = 8.272$, $\beta_{SALEG2} = 23.714$, $p < 0.01$), product design ($\beta_{SALEG1} = 3.253$, $\beta_{SALEG2} = 9.522$, $p < 0.01$), and product testing ($\beta_{SALEG1} = 10.757$, $\beta_{SALEG2} = 31.399$, $p < 0.01$) on sales growth are positive and significant.
The numbers in the first two columns in Table 4 show that big data-embedded idea development ($\beta_{GMY1} = 1.878, p < 0.01; \beta_{GMY2} = 1.192, p < 0.05; \beta_{GMY3} = 2.136, p < 0.01$), business analysis ($\beta_{GMY1} = 3.758, \beta_{GMY2} = 3.382, \beta_{GMY3} = 4.639, p < 0.01$), product design ($\beta_{GMY2} = 2.000, \beta_{GMY3} = 2.083, \beta_{GMY2} = 2.492, p < 0.01$), and product testing ($\beta_{GMY1} = 4.766, \beta_{GMY2} = 4.704, \beta_{GMY3} = 6.561, p < 0.01$) significantly increase gross margin. However, big data-embedded commercialization does not increase sales growth or gross margin ($p > 0.10$). Therefore, H1-H4 were supported, but H5 was not supported.

5.3.2. UK Results

For the UK sample, the numbers in the third and fourth columns of Table 3 reveal that big data-embedded idea development ($\beta_{SALEG1} = 2.273, \beta_{SALEG2} = 3.475, p < 0.05$), business analysis ($\beta_{SALEG1} = 5.380, \beta_{SALEG2} = 8.834, p < 0.01$), product design ($\beta_{SALEG1} = 2.225, p < 0.01; \beta_{SALEG2} = 3.170, p < 0.05$), product testing ($\beta_{SALEG1} = 6.620, \beta_{SALEG2} = 12.770, p < 0.01$), and commercialization ($\beta_{SALEG1} = 7.190, \beta_{SALEG2} = 9.943, p < 0.01$) significantly increase sales growth.

The numbers in the third and fourth columns of Table 4 suggest that big data-embedded idea development ($\beta_{GMY1} = 0.425, \beta_{GMY2} = 0.538, \beta_{GMY3} = 1.210, p < 0.05$), business analysis ($\beta_{GMY1} = 0.667, p < 0.01; \beta_{GMY2} = 0.758, p < 0.05; \beta_{GMY3} = 1.180, p < 0.05$), product design ($\beta_{GMY1} = 0.503, p < 0.01; \beta_{GMY2} = 0.347, p < 0.10; \beta_{GMY3} = 0.746, p < 0.05$), product testing ($\beta_{GMY1} = 1.524, \beta_{GMY2} = 1.884, \beta_{GMY3} = 3.563, p < 0.01$), and commercialization ($\beta_{GMY1} = 1.723, \beta_{GMY2} = 1.672, \beta_{GMY3} = 3.808, p < 0.01$) significantly increase gross margin. Therefore, H1-H5 were supported.

5.3.3. Australian Results

For Australian sample, the numbers in the fifth and sixth columns of Table 3 indicate that the effects of big data-embedded business analysis ($\beta_{SALEG1} = 7.685, \beta_{SALEG2} = 11.988, p < 0.01$), product design ($\beta_{SALEG1} = 2.714, \beta_{SALEG2} = 3.367, p < 0.01$), product testing ($\beta_{SALEG1} = 7.113, \beta_{SALEG2} = 10.152, p < 0.01$), and commercialization ($\beta_{SALEG1} = 8.409, \beta_{SALEG2} = 13.591, p < 0.01$) on sales growth were significantly positive.

The numbers in the fifth and sixth columns of Table 4 show that big data-embedded business analysis ($\beta_{GMY1} = 3.298, \beta_{GMY2} = 3.974, \beta_{GMY3} = 4.773, p < 0.01$), product design ($\beta_{GMY1} = 1.587, \beta_{GMY2} = 1.796, \beta_{GMY3} = 1.740, p < 0.01$), product testing ($\beta_{GMY1} = 4.085, \beta_{GMY2} = 4.158, \beta_{GMY3} = 5.289, p < 0.01$), and commercialization ($\beta_{GMY1} = 4.406, \beta_{GMY2} = 6.434, \beta_{GMY3} = 7.493, p < 0.01$) significantly increase gross margin.

Tables 3 and 4 reveal some interesting results for Australian sample. For sales growth, big data-embedded idea development did not significantly increase sales growth ($p > 0.10$). Yet, big data-embedded idea development significantly increases first-year gross margin ($\beta_{GMY1} = 0.961, p < 0.10$), but it did not significantly increase second-year and third-year gross margin ($p > 0.10$). Therefore, H2-H5 were supported, but H1 was partial supported.

5.4. Cross-National Comparison

5.4.1. Cross-National Similarities

From the regression results reported in Tables 3 and 4, we found some interesting cross-national similarities. In the USA, UK, and Australia, the effects of big data-embedded business analysis, product design, and product testing on sales growth and gross margin were all significantly positive.

5.4.2. Cross-National Differences

There are some interesting results by comparing the results among the three countries. In the USA, big data-embedded commercialization does not increase sales growth and gross margin. In Australia, big data-embedded idea development does not increase sales growth, second-year, and third-year gross margin. Because the development stage of big data technologies differs between the USA and Australia, a plausible explanation for the surprising results may be that big data have been more widely
used in companies in the USA. Thus, the value of big data-embedded commercialization has already
evidence. A second plausible explanation for the surprising findings may be that Australian companies
may not yet fully understand the importance of big data at the beginning of the NPD process. The idea
development may still be based on experience-driven decision-making. Therefore, firms benefit less
from big data-embedded idea development.

A second interesting cross-national difference is the rank ordering of the effect of each stage of big
data-embedded NPD process on sales growth and gross margin. We summarized the standardized estimates in Tables 6 and 7.

Table 6 presents the rank ordering of the effects of each stage of the big data-embedded NPD
process on sales growth. The results suggest the following:

(1) For the USA, big data-embedded product testing increases sales growth the most.
(2) For the UK, big data-embedded commercialization increases second-year sales growth the most,
but big data-embedded product testing increases third-year sales growth the most.
(3) For Australia, big data-embedded commercialization increases sales growth the most.

| Table 6. Relative ranking of regression coefficients: sale growth as the dependent variable. |
|---------------------------------------------------|----------------|----------------|----------------|
| Dependent Variable: Sale Growth                   | USA Sample (N = 497) | UK Sample (N = 510) | Australia Sample (N = 851) |
| Standardized Estimate (Rank)                      | SALEG1 | SALEG2 | SALEG1 | SALEG2 | SALEG1 | SALEG2 |
| BDIDEAA                                           | 0.167 (3) | 0.158 (3) | 0.097 (5) | 0.087 (5) | −0.001 (ns) | −0.017 (ns) |
| BDBMOA                                            | 0.255 (2) | 0.250 (2) | 0.186 (3) | 0.179 (3) | 0.233 (3) | 0.238 (2) |
| BDDESI                                            | 0.139 (4) | 0.139 (4) | 0.109 (4) | 0.091 (4) | 0.127 (4) | 0.103 (4) |
| BDTEST                                            | 0.363 (1) | 0.363 (1) | 0.245 (2) | 0.278 (1) | 0.251 (2) | 0.235 (3) |
| BDCOMM                                            | 0.019 (ns) | 0.019 (ns) | 0.267 (1) | 0.217 (2) | 0.295 (1) | 0.312 (1) |

Note: 1 = The most important stage; 5 = The least important stage; ns = Not significant; SALEG1 = Second-year sales
growth; SALEG2 = Third-year sales growth; BDIDEAA = Big data-embedded idea development; BDBMOA = Big
data-embedded-business analysis; BDDESI = Big data-embedded product design; BDTEST = Big data-embedded
product testing; BDCOMM = Big data-embedded commercialization.

Table 7 presents the rank ordering of the effects of each stage of the big data-embedded NPD
process on gross margin. The results suggest the following:

(1) For the USA, big data-embedded product testing has the biggest effect on gross margin.
(2) For the UK, big data-embedded commercialization has the biggest effect on first-year and
third-year gross margin and big data-embedded product testing has the biggest effect on
second-year gross margin.
(3) For Australia, big data-embedded commercialization has the biggest effect on gross margin.

| Table 7. Relative ranking of regression coefficients: gross margin as dependent variable. |
|---------------------------------------------------|----------------|----------------|----------------|
| Dependent Variable: Gross Margin                  | USA Sample (N = 497) | UK Sample (N = 510) | Australia Sample (N = 851) |
| Standardized Estimate (Rank)                      | GMY1 | GMY2 | GMY3 | GMY1 | GMY2 | GMY3 | GMY1 | GMY2 | GMY3 |
| BDIDEAA                                           | 0.135 (4) | 0.089 (4) | 0.115 (4) | 0.086 (5) | 0.091 (5) | 0.105 (3) | 0.059 (5) | 0.035 (ns) | 0.054 (ns) |
| BDBMOA                                            | 0.245 (2) | 0.229 (2) | 0.227 (2) | 0.109 (4) | 0.103 (3) | 0.083 (4) | 0.176 (3) | 0.148 (3) | 0.155 (3) |
| BDDESI                                            | 0.181 (3) | 0.196 (3) | 0.169 (3) | 0.117 (3) | 0.067 (4) | 0.075 (5) | 0.131 (4) | 0.104 (4) | 0.087 (4) |
| BDTEST                                            | 0.341 (1) | 0.349 (1) | 0.351 (1) | 0.267 (2) | 0.275 (1) | 0.269 (2) | 0.254 (2) | 0.181 (2) | 0.200 (2) |
| BDCOMM                                            | 0.032 (ns) | 0.028 (ns) | 0.014 (ns) | 0.303 (1) | 0.245 (2) | 0.289 (1) | 0.272 (1) | 0.278 (1) | 0.281 (1) |

Note: 1 = The most important stage; 5 = The least important stage; ns = Not significant; GMY1 = First-year gross
margin; GMY2 = Second-year gross margin; GMY3 = Third-year gross margin; BDIDEAA = Big data-embedded
idea development; BDBMOA = Big data-embedded business analysis; BDDESI = Big data-embedded product
design; BDTEST = Big data-embedded product testing; BDCOMM = Big data-embedded commercialization.
To explain these differences, we consider the actual levels of using big data in the NPD process. In the USA, firms have widely used big data to simulate consumer behavior and collect customer feedback information. Therefore, big data-embedded product testing can help firms accurately improve new products and increase the possibility of successful product sales. In Australia, we speculate that the Australians may indicate a willingness to learn and adopt the successful experience of big data application in the USA, this may reflect a willingness to widely use big data into commercialization stage, which lead to better commercialization strategies and superior sustainable innovation performance. In the UK, we speculate that Britisher may better copy the experience of big data applications in the USA. Thus, the most significant impact on sustainable innovation performance is big data-embedded commercialization, followed by big data-embedded product testing. Our studies reveal many similarities and differences between the USA, UK, and Australia. Some of the differences may be due to cultural factors.

6. Discussion

6.1. Results

Based on the NPD theory, we defined big data-embedded NPD processes and developed new measurement scales. We then proposed a theoretical model that investigated how the big data-embedded NPD process affects sustainable innovation performance. Using empirical data from the USA, UK, and Australia, we verified the research hypothesis and explored cross-national similarities and differences.

First, we identified the connotation of the big data-embedded NPD process and developed new measurement scales. Previous research has suggested that an “experience-driven” NPD process should be transformed into a big data-driven NPD process [14,17]. Consistent with this logic, our study first developed related theories about big data-embedded NPD process.

Second, we empirically tested the relationship between big data-embedded NPD process and sustainable innovation performance using data from the USA, UK, and Australia. Previous studies have demonstrated that the “experience-driven” NPD process had a significant positive effect on innovation performance [1,3–6,9,11,13]. Consistent with prior studies, we found that big data-embedded business analysis, product design, and product testing significantly increase sustainable innovation performance in the USA, UK, and Australia.

Third, we also presented some differences for the USA, UK, and Australia. On the one hand, in the USA, big data-embedded commercialization did not affect sales growth and gross margin. In Australia, big data-embedded idea development did not affect sales growth and second-year and third-year gross margin. On the other hand, in the USA, big data-embedded product testing has the biggest effect on sales growth and gross margin. In Australia, big data-embedded commercialization has the biggest effect on sales growth and gross margin. In the UK, big data-embedded commercialization has the biggest effect on second-year sales growth, first-year and third-year gross margin. In addition, big data-embedded product testing has the biggest effect on third-year sales growth and second-year gross margin. Therefore, American firms, British firms, and Australian firms should have some differences in the big data-embedded NPD process.

6.2. Theoretical Implications

We make three contributes to the literature of sustainable NPD. First, our study extends the NPD theory by defining big data-embedded NPD processes and developing new measurement scales. Prior studies have highlighted the role of big data for the NPD process and sustainable innovation performance [14,15,17,23–25]. However, the existing studies do not develop big data-embedded NPD process theory. We identified the connotations of the big data-embedded NPD process and developed new measurement scales. Our findings provide a more comprehensive understanding of the big data-embedded NPD process.
Second, this study extends the NPD literature by empirically testing the effects of big data-embedded NPD process on sustainable innovation performance. Previous research has confirmed that the “experience-driven” NPD process has a significant positive impact on innovation performance [1,3–6,9,11,13]. Consistent with this logic, we find that some stages of big data-embedded NPD process increase sustainable innovation performance while others don’t. Our findings add to the research findings of McAfee and Brynjolfsson [14] and George, Haas, and Pentland [17] that big data-driven innovation has superior innovation performance. In addition, we find that in the USA, big data-embedded commercialization does not enhance sustainable innovation performance, and in Australia, big data-embedded idea development only has a partial effect on sustainable innovation performance. Our findings indicate that different stages of the big data-embedded NPD process have different impacts on sustainable innovation performance. These findings add to the findings of Barczak [21] and Rubera, Chandrasekaran, and Ordanini [22].

Third, this study supplements cross-national comparative study by exploring similarities and differences in the USA, UK, and Australia. Previous research has emphasized that the effect of the “experience-driven” NPD process on innovation performance has cross-national similarities and differences [3,4]. Consistent with this logic, we revealed that the effects of five stages of big data-embedded NPD process on sustainable innovation performance had cross-national similarities and differences. Our findings add to new research results of how the big data-embedded NPD process affects sustainable innovation performance in different countries [41–45].

6.3. Insights for NPD Managers

Our findings offer three new insights for NPD project managers regarding developing a big data-embedded NPD process. First, when resources are limited, it is suggested that firms in the USA should not invest in big data-embedded commercialization, and firms in Australia should not invest in big data-embedded idea development. Since investing in these stages does not improve sustainable innovation performance but increases the innovation cost.

Second, if the objective is to improve sales growth, American firms should first invest in big data-embedded product testing, while Australian firms should first invest in big data-embedded commercialization, and British firms should first invest in big data-embedded commercialization and then invest in big data-embedded product testing to improve longer-term sales growth (i.e., third-year sales growth).

Third, if the objective is to improve gross margin, American firms should first invest in big data-embedded product testing, while Australian firms should priority invest in big data-embedded commercialization, and British firms should priority invest in big data-embedded commercialization and then invest in big data-embedded product testing.

Firms should do the following to implement big data-embedded product testing:

1. using big data to evaluate prototype market performance,
2. using big data to test alternative technologies,
3. using big data to simulate customer acceptance and customer use tests, and
4. using big data to evaluate different product and service offers.

Firms should do the following to implement big data-embedded commercialization:

1. using big data to evaluate suppliers and commercialization partners,
2. using big data to complete the final plans for production,
3. using big data to evaluate and complete the final commercialization plans including the timing of product launch, and
4. using big data to develop pricing strategy and tactics.
6.4. Future Research Directions

This study has some research limitations that offer some future research directions. First, we only included five stages of the big data-embedded NPD process in this study. There are other possible stages (e.g., big data-embedded post-commercialization, big data-embedded operations management). These stages may also affect sustainable innovation performance. Future studies should investigate what additional stages should be included in the big data-embedded NPD process and how the new stages affect sustainable innovation performance. Second, because the objective of this study was to evaluate the effects of the big data-embedded NPD process on sustainable innovation performance, we did not make any formal hypotheses regarding cross-cultural differences. It would be interesting for future research studies to develop cross-cultural hypotheses and test them. The initial findings of this study may serve as starting points for developing cross-cultural hypotheses. Third, we only conducted three empirical studies in the USA, UK, and Australia. This limits the generalizability of the research results to other countries. Although cross-national data are expensive and difficult to obtain, additional studies in other countries will provide additional insights and help future theoretical development.

7. Conclusions

In this study, we defined big data-embedded NPD processes, developed new measure scales for big data-embedded NPD process, and investigated the effect of big data-embedded NPD process on sustainable innovation performance. Using data collected from 497 NPD projects in the USA, 510 NPD projects in the UK, and 851 NPD projects in Australia, we make the following conclusions: (1) For USA, UK, and Australia, big data-embedded business analysis, product design, and product testing are positive effects on sustainable innovation performance, but big data-embedded commercialization has no significant impact on sustainable innovation performance in the USA and big data-embedded idea development has a partial impact on sustainable innovation performance in Australia, (2) Big data-embedded product testing is the most important stage for sustainable innovation performance improvement in the USA, (3) Big data-embedded commercialization is the most important stage for sustainable innovation performance improvement in Australia, and (4) Big data-embedded commercialization is the most important stage for second-year sales growth and first-year and third-year gross margin improvement and big data-embedded product testing is the most important stage for third-year sales growth and second-year gross margin improvement in the UK. Our findings enrich extant NPD literature and deepen our understanding of the effect of big data-embedded NPD processes on sustainable innovation performance in the USA, UK, and Australia.

Author Contributions: Y.W. and H.Z. share equally for the first authorship and contributed equally to the development of this article. Conceptualization, Y.W. and H.Z.; data curation, H.Z.; formal analysis, H.Z.; methodology, H.Z.; writing-original draft, Y.W.; writing-review and editing, Y.W. and H.Z.; funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Humanities and Social Science Project of the China Ministry of Education under the grant with project title: “Breakthrough service innovation: effects of big data analytics and AI capability”, grant number 20YJCZH224.

Acknowledgments: The authors thank three anonymous reviewers for their useful suggestions to an earlier version which improved the quality of this research. The authors also thank Michael Song for his contributions in this research.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Cooper, R.G.; Kleinschmidt, E. New Products: The Key Factors in Success; American Marketing Association: Chicago, IL, USA, 1990.
2. Song, X.M.; Parry, M.E. What separates Japanese new product winners from losers. *J. Prod. Innov. Manag.* 1996, 13, 422–439. [CrossRef]
3. Song, X.M.; Parry, M.E. A cross-national comparative study of new product development processes: Japan and the United States. *J. Mark.* 1997, 61, 1–18. [CrossRef]
4. Song, X.M.; Parry, M.E. The determinants of Japanese new product success. *J. Mark. Res.* 1997, 34, 64–76. [CrossRef]
5. Song, M.; Montoya-Weiss, M.M. The effect of perceived technological uncertainty on Japanese new product development. *Acad. Manag. J.* 2001, 44, 61–80. [CrossRef]
6. Cooper, R.G.; Edgett, S.J. Overcoming the crunch in resources for new product development. *Res. Manag.* 2003, 46, 48–58. [CrossRef]
7. Swink, M.; Song, M. Effects of marketing-manufacturing integration on new product development time and competitive advantage. *J. Oper. Manag.* 2006, 25, 203–217. [CrossRef]
8. Crawford, J.C. Product innovation: Leading change through integrated product development by David, L. Rainey. *J. Prod. Innov. Manag.* 2007, 24, 188–190. [CrossRef]
9. Reid, M.; Brady, E. Improving firm performance through NPD: The role of market orientation, NPD orientation and the NPD process. *Australas. Mark. J.* 2012, 20, 235–241. [CrossRef]
10. García-Alcaraz, J.; Maldonado-Macias, A.A.; Hernández-Hernández, S.I.; Hernandez-Arellano, J.L.; Blanco-Fernandez, J.; Muro, J.C.S.-D.; Diez-Muro, J.C.S. New product development and innovation in the maquiladora industry: A causal model. *Sustainability* 2016, 8, 707. [CrossRef]
11. Dubiel, A.; Banerjee, S.; Ernst, H.; Subramaniam, M. International-market-information use across new-product-development stages. *Int. Mark. Rev.* 2018, 35, 760–784. [CrossRef]
12. Paladino, A. Investigating the drivers of innovation and new product success: A comparison of strategic orientations. *J. Prod. Innov. Manag.* 2007, 24, 534–555. [CrossRef]
13. Song, L.Z.; Song, M.; Di Benedetto, C.A. Resources, supplier investment, product launch advantages, and first product performance. *J. Oper. Manag.* 2010, 29, 86–104. [CrossRef]
14. McAfee, A.; Brynjolfsson, E. Big data: The management revolution. *Harv. Bus. Rev.* 2012, 90, 60–68. [PubMed]
15. Mikalef, P.; Pappas, I.O.; Krogstie, J.; Giannakos, M. Big data analytics capabilities: A systematic literature review and research agenda. *Inf. Syst. e-Bus. Manag.* 2017, 16, 547–578. [CrossRef]
16. Gobble, M.M. Big data: The next big thing in innovation. *Res. Manag.* 2013, 56, 64–67. [CrossRef]
17. George, G.; Haas, M.R.; Penland, A. Big data and management. *Acad. Manag. J.* 2014, 57, 321–326. [CrossRef]
18. Akter, S.; Wamba, S.F.; Gunasekaran, A.; Dubey, R.; Childe, S.J. How to improve firm performance using big data analytics capability and business strategy alignment? *Int. J. Prod. Econ.* 2016, 182, 113–131. [CrossRef]
19. Hao, S.; Zhang, H.; Song, M. Big data, big data analytics capability, and sustainable innovation performance. *Sustainability* 2019, 11, 7145. [CrossRef]
20. Gupta, M.; George, J.F. Toward the development of a big data analytics capability. *Inf. Manag.* 2016, 53, 1049–1064. [CrossRef]
21. Barczak, G. New product strategy, structure, process, and performance in the telecommunications industry. *J. Prod. Innov. Manag.* 1995, 12, 224–234. [CrossRef]
22. Rubera, G.; Chandrasekaran, D.; Ordanini, A. Open innovation, product portfolio innovativeness and firm performance: The dual role of new product development capabilities. *J. Acad. Mark. Sci.* 2015, 44, 166–184. [CrossRef]
23. Johnson, J.S.; Friend, S.B.; Lee, H.S. Big data facilitation, utilization, and monetization: Exploring the 3Vs in a new product development process. *J. Prod. Innov. Manag.* 2017, 34, 640–658. [CrossRef]
24. Tan, K.H.; Zhan, Y. Improving new product development using big data: A case study of an electronics company. *R D Manag.* 2016, 47, 570–582. [CrossRef]
25. Zhan, Y.; Tan, K.H.; Li, Y.; Tse, Y.K. Unlocking the power of big data in new product development. *Ann. Oper. Res.* 2016, 270, 577–595. [CrossRef]
26. Markus, M.L. New games, new rules, new scoreboards: The potential consequences of big data. *J. Inf. Technol.* 2015, 30, 58–59. [CrossRef]
27. Beyer, M.A.; Laney, D. The importance of ‘big data’: A definition. *Stanford CT Gart.* 2012, 21, 2014–2018.
28. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* 2015, 35, 137–144. [CrossRef]
29. Leeflang, P.S.H.; Verhoef, P.C.; Dahlström, P.; Freundt, T. Challenges and solutions for marketing in a digital era. *Eur. Manag. J.* 2014, 32, 1–12. [CrossRef]
30. Sivarajah, U.; Kamal, M.; Irani, Z.; Weerakkody, V. Critical analysis of big data challenges and analytical methods. *J. Bus. Res.* 2017, 70, 263–286. [CrossRef]

31. Raguseo, E. Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *Int. J. Inf. Manag.* 2018, 38, 187–195. [CrossRef]

32. Gunasekaran, A.; Papadopoulos, T.; Dubey, R.; Wamba, S.F.; Childe, S.J.; Hazen, B.; Akter, S. Big data and predictive analytics for supply chain and organizational performance. *J. Bus. Res.* 2017, 70, 308–317. [CrossRef]

33. Douglas, S.P.; Craig, C.S. *International Marketing Research*; Englewood Cliffs: Bergen County, NJ, USA, 1983.

34. Churchill, G.A., Jr. A paradigm for developing better measures of marketing constructs. *J. Mark. Res.* 1979, 16, 64–73. [CrossRef]

35. Hu, C.; Zhang, H.; Song, M.; Liang, D. Past performance, organizational aspiration, and organizational performance: The moderating effect of environmental jolts. *Sustainability* 2019, 11, 4217. [CrossRef]

36. Dillman, D.A. *Mail and Telephone Surveys: The Total Design Method*; Wiley: New York, NY, USA, 1978.

37. Anderson, J.C.; Gerbing, D.W. Structured equation modeling in practice: A review and recommended two step approach. *Psychol. Bull.* 1988, 103, 411–423. [CrossRef]

38. Brown, T.A. *Confirmatory Factor Analysis for Applied Research (Methodology in the Social Sciences)*; The Guilford Press: New York, NY, USA, 2006.

39. Larcker, D.F.; Fornell, C. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* 1981, 18, 39. [CrossRef]

40. Zhang, H.; Song, M.; He, H. Achieving the success of sustainability development projects through big data analytics and artificial intelligence capability. *Sustainability* 2020, 12, 949. [CrossRef]

41. Yang, D.; Park, S. Too much is as bad as too little? Sources of the intention-achievement gap in sustainable innovation. *Sustainability* 2016, 8, 712. [CrossRef]

42. Chen, S. The influencing factors of enterprise sustainable innovation: An empirical study. *Sustainability* 2016, 8, 425. [CrossRef]

43. Can, U.; Alatas, B. Big social network data and sustainable economic development. *Sustainability* 2017, 9, 2027. [CrossRef]

44. Syafrudin, M.; Fitriyani, N.L.; Li, D.; Alfian, G.; Rhee, J.; Kang, Y.-S. An open source-based real-time data processing architecture framework for manufacturing sustainability. *Sustainability* 2017, 9, 2139. [CrossRef]

45. Choy, M.; Park, G. Sustaining innovative success: A case study on consumer-centric innovation in the ICT industry. *Sustainability* 2016, 8, 986. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).