Electronic Commerce Research
https://doi.org/10.1007/s10660-024-09914-6

The drivers and income effect of big data use by e-commerce farmers: evidence from China

Yiwu Zeng[1,2] · Baogang Li[3] · Lili Li[4] · Guojun Zhang[1]

Accepted: 6 October 2024

Abstract
More and more Chinese farmers have engaged in e-commerce markets, and some e-commerce farmers have used big data products developed by e-commerce platforms to improve their business decision-making. Based on the questionnaire data of 418 e-commerce farmers in China, this paper reveals how big data product attributes affect the use decisions of e-commerce farmers on big data products, and how the use of big data products affect the income of e-commerce farmers. It is found that the usefulness, ease of use, and experience of big data products all have a significantly positive impact on the e-commerce farmers' use decisions. The use of big data products has significantly increased the income of e-commerce farmers by enhancing their entrepreneurial alertness and dynamic capabilities. The findings have positive references for the academic research and policy making of rural e-commerce and big data applications in developing countries.

Keywords Big data · E-commerce farmers · Product attributes · Income effect · Taobao Villages

1 Introduction

With the increasing popularization of the Internet, e-commerce in rural areas of developing countries has developed rapidly. Especially in China, more and more farmers are using third-party e-commerce platforms to sell their

✉ Lili Li
11620059@zju.edu.cn

1 School of Economics, Hangzhou Normal University, Hangzhou, China

2 Hangzhou International Urbanology Research Center &, Zhejiang Urban Governance Studies Center, Hangzhou Normal University, Hangzhou, China

3 College of Business Administration, Northeastern University, Shenyang, China

4 School of Economics, Hangzhou Dianzi University, Hangzhou, China

Published online: 15 October 2024

products online, and they are called "e-commerce farmers". Alibaba's Taobao, with its advantages of low entry threshold, low technical difficulty and low initial capital requirements, has become a major front for farmers to participate in e-commerce. Because of this, people are used to calling the professional e-commerce village "Taobao Village". In order to count the number of Taobao Villages, the AliResearch Institute has formulated two quantitative criteria: one is the number of active online stores in the village reaches more than 100 or the number of active online stores accounts for more than 10% of the number of local households; the second is the annual total transaction amount of the village's online business reached more than 10 million yuan. According to the criteria, the earliest Taobao Villages appeared in 2009, with a total of 3. Since then, the development of Taobao Villages has gradually accelerated, and in 2022, Taobao Villages have covered 28 provinces and 180 cities, with the number reaching 7,780. The rapid development of Taobao Villages has aroused scholars' attention to the emerging group of e-commerce farmers. A large number of literatures have studied the factors that determine farmers' adoption of e-commerce and the impact of e-commerce adoption on farmers' income, consumption, employment and migration decisions [1–7].

However, with the rapid development of rural e-commerce and the continuous expansion of the scale of e-commerce farmers, the sustainability of rural e-commerce began to encounter difficulties. This is a challenge that all developing countries must face when the rural e-commerce has developed to a certain stage. And that is, the fierce competition in the online market brings e-commerce farmers difficulties in operation, and rural e-commerce clusters are under the pressure of upgrading. Along with the expansion of cluster scale, rural online businesses are prone to low-level homogeneous competition, insufficient product innovation, and declining prices, leading to operational difficulties and even the occurrence of counterfeit and shoddy phenomena [8]. At present, the bottlenecks of resources in rural e-commerce clusters are becoming increasingly prominent, including a lack of talents, tight land supply, inadequate public services, and so on. The supporting services such as training, finance, warehousing, and approval required by the development of e-commerce cannot meet the growing needs of rural online businesses, so it is urgent to seek upgrading [9]. How to improve the sustainability of rural e-commerce has become an important issue that the academic community must study and solve. Scholars have proposed measures such as increasing policy support and establishing industry associations [8–11]. Although these measures are necessary, they are only external conditions for e-commerce farmers. The core subject of rural e-commerce is e-commerce farmers. Only when the endogenous capacity of e-commerce farmers has been improved qualitatively and the entrepreneurial income has been steadily increased, can the sustainable development of rural e-commerce have a solid foundation. The endogenous capacity of e-commerce farmers is reflected in the self-management ability of the entire process of the production, supply and network marketing of the products they sell. How to improve the endogenous capacity of e-commerce farmers and continuously increase the income of e-commerce farmers is the key to strengthening the sustainability of rural e-commerce, but the existing literature lacks exploration on this.

In some rural areas of China, a new phenomenon, that is, using big data products[1] to improve the endogenous capacity and income level of e-commerce farmers, is constantly occurring. This phenomenon provides a new research perspective and feasible path for improving the sustainability of rural e-commerce. E-commerce platforms has accumulated massive data such as transaction records, comment information, and search traces. By mining these big data, it can provide important guidelines for the production and operation of online businesses [12]. E-commerce platform enterprises represented by Alibaba Group and JD Group provide professional big data services for the majority of online merchants through the development of big data products. AliResearch Institute released the "2016 Online Business Data Application Analysis Report", pointing out that many excellent online businesses are using data to gain insight into the market and find opportunities, and the era of relying on intuition and experience is becoming a thing of the past. A survey of rural households in Cao County, Shuyang County and Lin'an District, China, three regions with leading rural e-commerce development, shows that about 30% of the surveyed farmers have used big data products developed by e-commerce platform enterprises in the operation of their online stores [13].

In addition, this new phenomenon occurring in China also provides an excellent opportunity to explore how to promote the proliferation of big data technology in rural areas. Previously, quite a number of scholars have discussed the impact of big data development on enterprise management and labor share income [14–17]. The big data involved is mainly the data accumulated by enterprises themselves, and some of it is also the data opened by the government. However, these studies are not suitable for guiding the proliferation of big data technology in rural areas. Farmers are different from enterprises in production scale and management mode. Farmers do not have their own accumulated data, and currently do not have access to government public data. Although the Chinese government is trying to accelerate the construction of large databases in rural areas, how to make rural public big data close to farmers is still an unresolved issue. Big data products developed by e-commerce platform companies are rapidly spreading among e-commerce farmers. The study of this new phenomenon can undoubtedly be inspired, so as to make a positive response to how to promote the spread of big data technology in rural areas.

Behind the new phenomenon is a new pattern. There must be some theoretical logic behind this new phenomenon emerging in China. To reveal these theoretical logic is the contribution that this paper tries to make. In theory, there are at least two basic logic supporting e-commerce farmers to choose to use big data products. First, the attributes of big data products can fit the group characteristics and actual needs

[1] In a general sense, big data products refer to all kinds of tools, platforms or services based on the development and application of big data technologies. By collecting, storing, processing, and analyzing massive amounts of data, they support the decision-making of businesses and organizations. In this paper, big data product specifically refers to a system developed by e-commerce platform enterprises to use big data technology and tools to operate and manage e-commerce business. Such a system can help online merchants better understand consumer needs, optimize marketing strategies, improve user experience and recommend products accurately, thereby increasing sales and customer satisfaction. For details, see the background in Sect. 2.

of e-commerce farmers, and second, the use of big data products can promote the increase of income of e-commerce farmers. Both are indispensable. The former is a push and the latter is a pull, which together contribute to the new phenomenon of e-commerce farmers using big data products. In addition to the theoretical analysis of this mechanism, it is also necessary to collect survey data for e-commerce farmers and carry out empirical tests to prove that this is indeed the case. Therefore, the research objective of this paper is to reveals how big data product attributes drive the big data usage behavior of e-commerce farmers and how the use of big data products promotes the income increase of e-commerce farmers, through theoretical analysis and combined with empirical research methods.

Specifically, on the basis of theoretical analysis, this paper proposes five hypotheses on the relationship between big data product attributes and e-commerce farmers' big data use, as well as the relationship between big data use and e-commerce farmers' income. Then empirical tests are conducted by using the survey data of 418 e-commerce farmers in Taobao Villages in Zhejiang Province, China. It is found that the usefulness, ease of use, and experience of big data products all have a significantly positive impact on the e-commerce farmers' use decisions. The use of big data products has significantly increased the income of e-commerce farmers by enhancing their entrepreneurial alertness and dynamic capabilities. The findings confirm that the use of big data products to assist business is an effective path to improve the endogenous capacity of e-commerce farmers and promote the sustainable development of rural e-commerce, and also reveal that product attributes are the key factors driving e-commerce farmers to use big data products. Our study not only directly expand the research on the sustainability of rural e-commerce by introducing the perspective of big data application, but also have a positive reference for how to promote the diffusion of public big data in rural areas.

The rest of the paper is structured as follows: Sect. 2 reviews the background of e-commerce farmers and the e-commerce big data products in China; Sect. 3 proposes hypotheses; Sect. 4 introduces the data and methods; Sect. 5 reports the empirical results; Sect. 6 further discusses the findings; and the last section concludes.

2 Background

2.1 E-commerce farmers in China

Taobao Village is an industrial cluster phenomenon based on the entrepreneurship of e-commerce farmers. Compared with traditional farmers, e-commerce farmers are younger, have stronger computer skills, have higher frequency of internet use, attach more importance to improving the added value of products, pay more attention to consumers' experience, and have the spirit of openness and sharing [10]. E-commerce farmers evolve from simple imitation to branding innovation, and the local e-commerce ecosystem has a significant positive impact on the growth of e-commerce farmers [18]. However, there is a large gap between e-commerce farmers and urban e-merchants in cyberspace representation [19]. In terms of online store operation, e-commerce farmers are increasingly facing problems such as lack of professional talents, fierce local competition, high logistics costs, single marketing strategies, and low policy inclusiveness [8]. As the number of online merchants continues to increase, the competition in e-commerce markets becomes more and more fierce, and the technical threshold of online store operation is rapidly rising, requiring new online merchants to ensure that all aspects of online store operation reach a high level at the beginning. Old online merchants need to enhance their specialization in product presentation, picture artwork, video production, and marketing strategies so as to maintain competitiveness. At the same time, with the accumulation of online shopping experience, consumers are more and more discerning in their choices of online stores, and the professionalism of online stores must be enhanced accordingly to meet the needs of consumers [20]. Although some local governments have taken measures to support the development of Taobao Villages, these measures generally tend to support large e-commerce households [21]. In addition, some scholars have found that the strong local clan will lead to the closure of Taobao Villages, and e-commerce farmers reject the extended relationship, which is not conducive to the further development of Taobao Villages [22].

As an emerging group in rural areas, e-commerce farmers are receiving more and more attention from scholars [23]. The existing literature focusing on e-commerce farmers in Taobao Villages mainly studies the characteristics and the growth path of e-commerce farmers. Existing literature points out the bottlenecks faced by e-commerce farmers and the necessity for making changes. In reality, some e-commerce farmers have successfully overcome their business shortcomings and maintained their competitive advantages through the use of big data. The use of big data products constitutes an important influencing factor for the growth of these e-commerce farmers and the performance of their online stores [12].

2.2 The E-commerce Big Data Products in China

The big data products based on e-commerce platforms are derived from the combination of e-commerce and big data analytics technology. Big data products combine the power of data and products to empower business decisions of merchants. E-commerce platform gathers all parties, precipitates transaction data, and forms a big databases. Through mining and analyzing massive data, it can form scientific guidance for offline production and online operation of online merchants [24]. At present, China's big data products based on e-commerce platforms are mainly the "Business Advisor" (BA) developed by Alibaba Group and the "Jingdong Business Intelligence" (JBI) developed by JD Group.[2]

[2] The official website of BA is https://sycm.taobao.com/custom/login.htm?_target=http://sycm.taobao.com/, and the official website of JBI is https://sz.jd.com/sz/view/index/login.html?ReturnUrl=http%3A%2F%2Fsz.jd.com%2F.

BA, born in August 2011, is a unified data product platform developed by Alibaba for merchants, covering all platforms of Alibaba such as Taobao and Tmall, as well as all terminals such as personal computer and wireless. BA integrates data products such as data room, market situation, decoration analysis, source analysis, and competitive intelligence, involving thousands of data indicators. BA is an important platform to empower e-commerce sellers in the era of big data. The development of BA has gone through three stages. The first stage was from August 2011 to December 2012. During this stage, BA was only a value-added service project of 1688 Credit Pass, with the function of viewing wholesale procurement, market conditions and other types of data. The second stage was from January 2013 to September 2014. At this stage, BA was constantly upgraded. The first step was to launch a new version of 1688, which became the essential store operation analysis tool for the majority of suppliers on the 1688 platform. The second step was the launch of the Taobao platform, which opened up a new era of store operation analysis in the world's broadest online retail market. The third step was to launch the deluxe version of 1688, which provided more in-depth value-added services for the data of the 1688 suppliers. In the third stage from October 2014 to now, BA is upgraded to Alibaba's unified data product platform providing personalized data service solutions for all kinds of merchants. BA is committed to solving the pain points of "data cannot be found", "data cannot be understood", "data is not comprehensive", "data can not be used" and so on. According to Alibaba, in 2016, BA served more than 20 million merchants and more than 5 million merchants each month.

JBI, born in March 2017, is a platform for JD Group to provide data services for the merchants. After subscribing to JBI, a merchant can obtain data on the store's traffic, sales, customers, and commodities, as well as data on other merchants across the entire industry and in the same industry, so as to support operational decision-making. JBI also supports shopping cart marketing, customer marketing and other precision marketing to help merchants improve sales. JBI has comprehensive, accurate and professional product advantages. JBI provides all-round, all-link data solutions. All data interfaces are strictly verified to ensure that the business data is more general and accurate. JBI provides a perfect and professional data analysis solution to show the operational data and the current situation of the industry in multiple dimensions. In March 2021, on the fourth anniversary of the launch of JBI, JD Group announced a comprehensive upgrade of JBI, marking the upgrading of JBI from a data tool to a data decision center, allowing data applications to be open and innovative, and inspiring the marketing imagination of more brand merchants. It focuses more on the improvement of overall service capability and the empowerment of data technology capability, so as to better meet the diversified data needs of different types of brand businesses, and help brand businesses achieve high-quality growth with digital intelligence.

3 Hypotheses

3.1 Product attributes and big data use decisions of e-commerce farmers

Product attributes are the most direct and basic factors that affect the consumer's decision to use the product. Product attributes are the basic properties and key characteristics inherent in a product. Essentially, a product is a collection of attributes [25]. In other words, the attributes of a product are usually multidimensional. The product presented to the consumers is the joint action of multidimensional attributes of products [26–30].

The research object of this paper involves big data products developed by e-commerce platform enterprises, and is concerned with the impact of big data product attributes on e-commerce farmers' big data use decisions. Logically speaking, product attributes must be one of the key factors affecting whether e-commerce farmers use big data products. Additionally, the use of big data products by e-commerce farmers is essentially a new technology adoption behavior. Among a series of theoretical models of technology adoption behavior, Technology Acceptance Model (TAM) is a classical theory that reveals the mechanism of producers' willingness to accept new technologies from the perspective of product attributes. What's more, it applies to situations where individuals are fully autonomous in their adoption decisions without other restrictions. The decisions of e-commerce farmers to use big data products developed by e-commerce platform enterprises are purely market-based behaviors, so TAM is suitable for analysis. One of TAM's core ideas is that usefulness and ease of use reflect the adopters' subjective psychological evaluations regarding the superiority and ease of use of the technology respectively, and jointly determine their attitudes towards the new technology [31]. Because of its simplicity, ease of operation and strong explanatory power, TAM has been applied to the research of technology acceptance and adoption behavior in all walks of life, including information technology management, online education, e-commerce and other fields. Although some scholars believe that TAM has shortcomings, and try to expand and supplement it [32, 33]. However, the perception of product attributes still rests on perceived usefulness and perceived ease of use. In fact, with the advent of the digital age, the experience economy has grown even more rapidly. Growing up in the digital age, the younger generation not only demands the usefulness and ease of use of products, but also pays more attention to the experience that products bring to them. In developing countries such as China, young people are the main force of rural e-commerce development [10]. This determines that the big data products developed by e-commerce platform enterprises must be able to meet the needs of young people. In view of this, this paper defines the attributes of big data product as usefulness, ease of use, and experience of big data products psychologically perceived by online merchants. Specifically, the usefulness of big data products refers to online merchants' psychological perception of the usefulness of big data products in enabling business decisions; the ease of use of big data products refers to online merchants' psychological perception of the ease of use of big data products in enabling business decisions; and the experience of big data products refers to online merchants' psychological perception of the degree to which big data products match the risk of use, price level, the fun of use, user-friendly design, etc.

It has been shown that the psychological perception of product attributes directly influence consumers' attitude towards products, which in turn significantly influence their decisions [34]. Product attributes enable consumers to associate with products before they purchase them, thus entering the psychological simulation process of using them. Along with this psychological simulation process, consumers will have cognition and judgment, which will further influence their attitudes towards products [35]. It is demonstrated that after understanding and judging various attributes of a product, consumers will adopt a mindset to evaluate the overall image of a product, and when consumers' satisfaction with the overall image of a product reaches a certain level, it will trigger their purchase decisions [36]. The impact of big data product attributes on e-commerce farmers' big data use decisions will replicate the logic of the relationship between product attributes and consumer behaviors. In reality, e-commerce farmers will learn about the e-commerce big data product attributes through free trials, access to data product information, web search for relevant information, and listen to the experiences of other online shop sellers. If e-commerce farmers psychologically perceive that the attribute level of big data products is high, they will have positive attitudes towards e-commerce big data products, and then generate use intention and finally turn into use decisions.

Perceived usefulness and perceived ease of use are the most central variables of the technology acceptance model. In the more than thirty years since the technology acceptance model was proposed, numerous empirical studies have fully confirmed the strong explanatory power of perceived usefulness and perceived ease of use in the area of users' technology adoption behavior [37]. Regarding the usefulness of big data products, e-commerce farmers make judgments by comparing the situation before and after the use of big data products. The usefulness of big data products is reflected in the increased access to information, improved operations and incremental revenue that e-commerce farmers can achieve after using big data products. When the perceived usefulness of big data products is higher, the e-commerce farmers will be more motivated to use them and make more efforts to do so. Conversely, if e-commerce farmers feel that the data information provided by big data products is not very meaningful for their business decisions, they are more likely to choose not to use big data products. In terms of the ease of use of big data products, e-commerce farmers perceive the ease of access to data indicators, the ease of understanding data indicators and whether the data analysis process is hassle-free or laborious. When e-commerce farmers perceive the ease of use of big data products, they will be more positive and open to the use of big data products. Conversely, if big data products are not sufficiently visualized and popularized, making it difficult for e-commerce farmers to use the data, they are more likely to choose not to use the big data products. Moreover, according to the technology acceptance model, insufficient perceived ease of use will also lead to insufficient perceived usefulness as well. That is, in the case that e-commerce farmers perceive that big data products are not easy to use, it will lead to the situation that the usefulness of big data products cannot be fully understood, which eventually leads to e-commerce farmers choosing not to use big data products.

Usefulness and ease of use are key product attributes, but they do not cover the full range of product attributes. Even if a product is very useful to the user and easy to use, there is no guarantee that consumers will use the products. The reason is that consumers will take account of the risk of use, the cost of use, the fun of use, the user-friendliness of the product and other factors as well. Therefore, this paper introduces the dimension of experience by integrating the theoretical ideas of perceived risk and perceived fit, so that the framework of big data product attributes can be more complete. Perceived risk theory suggests that perceived risk stems from the uncertainty of consumption outcomes, and that users may experience some unexpected or inconsistent situations in the process of consuming products, which may lead to losses for them [38]. Specifically for big data products, e-commerce farmers, especially risk-averse ones, may choose not to use big data products due to worries and concerns about payment security, data accuracy, data timeliness, and data privacy. Perceived fit theory suggests that the extent to which new technologies are perceived by adoption subjects to fit with their own lives, experiences, preferences, needs, and values significantly affects their adoption decisions [39]. Specifically for big data products, the availability of big data products in terms of price level, interest, comprehensiveness, and user-friendly design will contrast with their own needs and preferences, creating a perceived fit and ultimately influencing the decision on big data use by e-commerce farmers. Obviously, the better the experience of the big data products, the greater the tendency of the e-commerce farmers to choose to use the big data product.

Based on the analysis above, this paper proposes the following hypotheses.

H1 The usefulness of big data products has a significant positive impact on the e-commerce farmers' big data use decisions.

H2 The ease of use of big data products has a significant positive impact on the e-commerce farmers' big data use decisions.

H3 The experience of big data products has a significant positive impact on the e-commerce farmers' big data use decisions.

3.2 Big data use, entrepreneurial alertness, and e-commerce farmers' income

In entrepreneurship theory, entrepreneurial alertness is seen as an intrinsic ability that entrepreneurs need to focus on. Entrepreneurial alertness, also known as psychological alertness of entrepreneurial opportunity, significantly influences the entrepreneurial behavior and entrepreneurial performance [40]. The Austrian economist Kirzner firstly introduced the concept of entrepreneurial alertness, defining it as the ability of entrepreneurs to respond quickly to entrepreneurial opportunities that have previously been overlooked [41]. Entrepreneurial alertness is not just a reaction to the outside world, but is always embedded in an entrepreneur's ability [42]. In reality, not everyone can identify entrepreneurial opportunities and earn profits, but individuals with higher entrepreneurial alertness are undoubtedly more likely and quickly to identify new market opportunities and develop strategic plans to drive business growth [43]. Empirical studies have shown that in order to achieve good entrepreneurial performance, entrepreneurs need to remain alert to entrepreneurial opportunities in order to accurately identify and appropriately exploit business opportunities [44, 45].

The use of big data helps to raise the entrepreneurial alertness of e-commerce farmers. In the pre-internet era, farmers' business decisions were based on subjective feelings and experience accumulation, with defects such as localization, lagging and roughness. The entire docking process from producers to consumers had long chains and many links, and the efficiency of information collection, processing and transmission was very low, with serious information lag and distortion. Information on demand side cannot be fed back to the supply side in a timely and effective manner, resulting in poor scientific and planned business decisions. The remarkable change brought about by the Internet is information, including the flow of information and information analysis. The flow of information has changed from the original inconvenient flow to the flow in nanoseconds, and from the original directional flow to non-directional flow, breaking the fortress of information asymmetry [46]. The spurt in data growth and the development of big data analysis techniques have made the ability of information analysis improve by leaps and bounds. Big data enables e-commerce farmers to obtain information in a more timely manner, and the information obtained is more comprehensive, accurate and closer to complete information, which can greatly reduce the subjective deviation of information. The Internet connects operators with consumers directly, making it easier and more effective for operators to make demand-oriented decisions and achieve user-orientation [10]. Big data products developed by e-commerce platform enterprises can help e-commerce farmers grasp dynamic information about the online market in the first time. The data will constantly create a visual and cognitive stimulus for e-commerce farmers, so that they remain highly alert to market changes, and continuously improve their abilities and probabilities to identify new opportunities that have not yet been discovered.

Based on the analysis above, this paper proposes the following hypothesis.

H4 The use of big data has a positive impact on the income level of e-commerce farmers by enhancing their entrepreneurial alertness.

3.3 Big data use, dynamic capabilities, and e-commerce farmers' income

As an expansion of the resource-based view, dynamic capabilities emphasizes that enterprises should have a keen ability to respond to the changes in the external environment and make timely reconfiguration of resources, which has a very important positive impact on the business performance of enterprises [47]. Factors such as the rapid development of digital technologies and changes in the economic situation have intensified the transformation of the entrepreneurial environment from relatively stable and orderly to highly dynamic and complex. When the change of the environment is highly discontinuous, dynamic capability is required to a large extent to develop multiple capabilities at the same time, and it has become a key capability for enterprises to cope with volatile environments [48]. For a long time, the lack of effective identification, accurate perception and timely response to external market demand has been an important reason for the difficulties in farmers' business income growth. Farmers' business decisions rely on experience, imitation and luck, which are prone to mistakes and greater risks [21]. Traditional experience can be seen as a production model in which the information element is inelastic in terms of output. That is, the information accumulated by traditional experience has a very limited role in changing the level of output. Changes in the external markets cannot be transmitted to farmers in a timely manner, resulting in very poor dynamic capabilities of farmers, a serious disconnect between supply and marketing, frequent stagnation in sales, and increased production without increased income for farmers [7]. Therefore, how to effectively enhance the dynamic capabilities of farmers is an important breakthrough to solve the problem of increasing farmers' income.

The use of big data products is beneficial to improve the dynamic capabilities of e-commerce farmers. Through the use of big data products, e-commerce farmers can access a range of optimal combinations of parameters about their own products, including colour, weight, taste, price, logistics, etc., so that they can compare and identify where there are disadvantages or deficiencies in their own operations in order to make targeted improvements [4]. This is critical as e-commerce marketplaces are highly competitive and small deviations in decision-making are magnified by the multiplier effect of the Internet, which can have a significant impact on business performance. The e-commerce market is changing rapidly. Big data products allow e-commerce farmers to keep abreast of marketing developments in real time, especially the latest developments of similar competitors and information about the fastest growing products. In addition, big data analytics technologies not only process and provide accurate information in real time, but also allow for the development of intelligent predictive functions [49]. Big data includes not only the data of events that have occurred and are occurring, but also data about events that are going to occur in the future, mined through machine learning [50], allowing operators to gain greater foresight and actively seek a first-mover advantage [12].

Based on the analysis above, this paper proposes the following hypothesis.

H5 The use of big data has a positive impact on the income level of e-commerce farmers by enhancing their dynamic capabilities.

4 Methodology

4.1 Data

The data used in this paper comes from a field household survey of e-commerce farmers conducted by our research team from July to August 2022. Zhejiang is the province with the highest level of rural e-commerce development in China and has the largest number of Taobao Villages. We selected 15 typical Taobao Villages in Zhejiang Province as survey subjects. These villages have been awarded the title of "China Taobao Village", whose e-commerce started earlier, developed rapidly, and received wide attention. These villages are distributed in Hangzhou, Jinhua, Lishui, Taizhou and other different cities, and mainly dealing in nuts and fried goods, clothing and apparel, outdoor products, tea, water heaters, shoes and other different types of products. Compared with scattered online merchants in ordinary villages, e-commerce farmers in Taobao Villages have richer experience in online shop operation and higher probability of using big data products, thus providing us with observable samples. In addition, e-commerce farmers in Taobao Villages are highly concentrated in spatial distribution, so it is convenient to investigate e-commerce farmers in Taobao Villages, and more samples can be obtained within a limited time. We investigated 30 e-commerce farmers in each village using the incidental random sampling method, and after eliminating the questionnaires with many missing values or irregularities, the final samples were 418.

4.2 Models

In order to verify the impact of big data product attributes on the big data use decisions of e-commerce farmers, considering that the big data use decisions belongs to a binary choice variable, so the binary Probit model was adopted for empirical analysis. The specific model is as follows:

$$P\left(D_i = 1|X_i\right) = \Phi\left(X_i\right) = \alpha + \beta_i T_i + \lambda_i Z_i + \mu \tag{1}$$

In Eq. (1), i represents different farmers, $P(D_i = 1\big|X_i)$ represents the probability that the farmer i chooses to use big data products, $\Phi(X_i)$ is the cumulative distribution function of the standard normal distribution, T is the core independent variable, that is the attributes of big data products, Z is the control variables, α is the constant term, and μ is the random disturbance term.

To verify the impact of big data use decisions on the income of e-commerce farmers, the following model was constructed:

$$Y_i = \alpha + \delta D_i + \beta X_i + \varepsilon_i \tag{2}$$

In Eq. (2), i represents different farmers, Y represents the income level of e-commerce farmers, D is the core independent variable of whether or not e-commerce farmers use big data products, X is the control variables, α is the constant term, and ε is the random disturbance term. If the use of big data by e-commerce farmers is random, the estimated coefficient can accurately reflect the income effect of big data use. However, in reality, whether to use big data products is a subjective decision of e-commerce farmers, not a random event. If the Equation (2) is estimated directly using OLS without considering this potential selection process, the parameter estimation results will be biased because the potential selection model and the unobserved factors are interrelated [51], resulting correlation between ε and

D. In other words, the big data use decisions of e-commerce farmers becomes an endogenous variable.

To overcome the endogeneity problem of big data use decisions, the data are preprocessed by introducing entropy balancing to control the estimation bias caused by self-selectivity to the maximum extent. Entropy balancing method was proposed by Hainmueller [52], which pre-sets a set of equilibrium constraints and normative constraints, and makes the exact matching of samples under a specific matrix for the treatment and control groups by calculating a set of optimal weights matching the constraints. The entropy balancing method can match each treatment group sample with a control group sample that is very similar to it, so as to retain the useful information of all samples, and the standardized mean difference and mean difference tests of the matched covariates are more robust and the results are more reliable [53]. Compared with propensity score matching method, entropy balancing method has the following advantages: First, it can ensure that the treatment group and the control group achieve a balance in the sample characteristics, and retain the useful information of all samples; Second, the model setting in the second stage estimation is more flexible; Thirdly, the mean difference test results of covariates matched by entropy balancing method are more robust.

4.3 Variables

When analyzing the influence of perceived big data product attributes on the big data use decisions of e-commerce farmers, this paper sets the dependent variable as the big data use decisions of e-commerce farmers, which is measured by "whether the e-commerce farmers have used big data products in the past year", so it is a binary choice variable. The core independent variables are the psychological perception of big data product attributes, which are examined from three dimensions: the usefulness of big data products, the ease of use of big data products, and the experience of big data products. The psychological perception of big data product attributes are measured in the form of a scale, as detailed in Table 1. The usefulness of big data products, the ease of use of big data products and the experience of big data products are all measured by the mean score of the corresponding measurement items with the value range from 1 to 5. In order to improve the reliability and validity, the content of the scale of big data product attributes was adapted from the existing literature. Statistics show that the Cronbach's alpha coefficients of big data product usefulness, big data product ease of use, and big data product experience are 0.944, 0.872, and 0.721 respectively, which are significantly higher than 0.60, showing good reliability. The convergent validity (AVE) of big data product usefulness, big data product ease of use, and big data product experience are 0.854, 0.694, and 0.694 respectively. The combined reliability (CR) of big data product usefulness, big data product ease of use and big data product experience are respectively 0.946, 0.872 and 0.851 which are significantly higher than 0.70, showing good combined reliability.

In terms of control variables, this paper draws on research experience from the relevant literature [4, 7, 13], so variables such as gender, age, education, CPC (the

Table 1 Content of scale measure items for big data product attributes

Product attribute dimensions	Content of measurement items	Source
Usefulness of big data products	Data products are very useful in the process of online store operation	Davis [31]
	Data products effectively help me understand market information	Venkatesh [54]
	With data products, online store business decisions become more accurate	Venkatesh and Davis [33]
Ease of use of big data products	Learning how to use data products is easy for me	Yang [55]
	Mastering the functions of the data product is easy for me	
	With the existing knowledge I can use the data products	
Experience of big data products	When I use data products, I don't worry about the risks	Wu and Wang [56]
	The data product is quite interesting, I like to use it	Elkins et al. [38]
	The interface design of the data product is quite user-friendly	Yang [55]
	The data metrics provided by the data product are quite comprehensive	
	The purchase price of data products is acceptable	
	There are too few free experience features for data products	

Communist Party of China) member, enterprise operation, registered trademark, duration of e-commerce business, main e-commerce formats, and main product types were included in the control variables. The specific definitions of the control variables are shown in Table 2. In particular, it needs to be explained that the main e-commerce business format is divided into two categories: new e-commerce business format and traditional e-commerce business format, the former including live e-commerce, social e-commerce and other new e-commerce business formats, and the latter mainly refers to the traditional e-commerce business format that is mainly based on static graphic display.

From the descriptive statistics in Table 2, we can see that about 53% of the surveyed e-commerce farmers used big data products. The attribute scores of big data products indicate that the current big data products developed by e-commerce platform enterprises are at a moderate to high level. The overall situation is good, but there is still room for improvement. Specifically for the dimensions of big data product attributes, usefulness has the highest score of nearly 4.0, followed by ease of use at about 3.6, and finally experience, only about 3.4. The surveyed e-commerce farmers are roughly evenly distributed in terms of gender, most of the surveyed e-commerce farmers are young. The average education level is mainly undergraduate and junior college, about 20% are CPC members, about 26% have conducted entrepreneurial operations, and 36% have registered trademarks. Most e-commerce enterprises have been established for less than 2 years, most of them are mainly new types of e-commerce, and mainly sell non-agricultural products.

Based on whether they use big data products to assist the business decisions of online stores, this paper classifies the e-commerce farmers into big data-using e-commerce farmers and ordinary e-commerce farmers. Table 3 shows the statistical results of the t-test of the mean values of each variable for the two groups of farmers. It can be seen that there are significant differences between big data-using e-commerce farmers and ordinary e-commerce farmers in the variables of usefulness of big data products, ease of use of big data products, experience of big data products, gender, age, education, enterprise operation, and registered trademark. Specifically, compared with ordinary e-commerce farmers, big data-using e-commerce farmers have higher perceptions of usefulness, ease of use, and experience of big data products. Meanwhile, big data-using e-commerce farmers are younger, more educated, have been operating online stores for a longer period of time, are more inclined to new e-commerce business, have more brand awareness, and pay attention to trademark registration and its importance for business operation.

In the study of the impact of big data use decisions on the income of e-commerce farmers, the dependent variable is the income of e-commerce farmers, which is measured in the form of a scale, as shown in Table 4. The reason for not considering the direct income measurement in this paper is that in reality farmers do not know exactly what their income is. Additionally, the direct income measurement relies on farmers' memory to estimate, which is inaccurate. Therefore, this paper adopts the psychological perception method to measure the income level of e-commerce farmers. Specifically, the income level of e-commerce farmers is measured by the mean score of all measurement items, and the value ranges from 1 to 5. In order to improve the reliability and

Table 2 Variable descriptions and descriptive statistics

Variable types	Variable names	Variable definition	Mean	S.D
Dependent variable	Whether or not to use big data products	Yes = 1, No = 0	0.526	0.538
Independent variables	Usefulness of big data products	The mean scores of all measurement items ranged from 1 to 5	3.951	0.851
	Ease of use of big data products	The mean scores of all measurement items ranged from 1 to 5	3.631	0.862
	Experience of big data products	The mean scores of all measurement items ranged from 1 to 5	3.404	0.591
Control variables	Gender	Male = 1, female = 0	0.505	0.501
	Age	Under 30 = 1, 30–50 = 2, over 50 = 3	1.660	0.536
	Education	Senior high school/technical secondary school or below = 1, junior college/bachelor degree = 2, bachelor degree or above = 3	1.914	0.684
	Whether a CPC member or not	Yes = 1, No = 0	0.199	0.399
	Whether enterprise operation or not	Yes = 1, No = 0	0.259	0.439
	Whether registered trademark or not	Yes = 1, No = 0	0.364	0.482
	Duration of e-commerce business	Less than 2 years = 1, 2–5 years = 2, more than 5 years = 3	1.742	0.774
	Main e-commerce formats	New e-commerce business formats = 1, traditional e-commerce business formats = 0	0.586	0.844
	Main product types	Agricultural products = 1, non-agricultural products = 0	0.273	0.446

Table 3 The comparisons between big data-using e-commerce farmers and ordinary e-commerce farmers

Variables	Big data-using e-commerce farmers		Ordinary e-commerce farmers		Difference in mean (t-test)
	Mean	S.D	Mean	S.D	
Usefulness of big data products	4.193	0.747	3.779	0.854	0.415***
Ease of use of big data products	3.830	0.768	3.490	0.910	0.340***
Experience of big data products	3.554	0.540	3.284	0.601	0.270***
Gender	0.650	0.478	0.389	0.489	0.260***
Age	1.595	0.522	1.739	0.541	-0.144***
Education	2.050	0.678	1.803	0.668	0.247***
Whether a CPC member or not	0.250	0.434	0.162	0.370	0.087**
Whether enterprise operation or not	0.340	0.475	0.182	0.387	0.158***
Whether registered trademark or not	0.455	0.499	0.280	0.450	0.174***
Duration of e-commerce business	1.821	0.767	1.674	0.779	0.147**
Main e-commerce formats	0.525	0.770	0.455	0.624	0.140**
Main product types	0.260	0.440	0.286	0.453	0.026

validity, the content of the scale of e-commerce farmers' income is adapted from the literature on farmers' entrepreneurial performance. Statistics showed that the Cronbach's alpha coefficient of the e-commerce farmers' income scale was 0.936, significantly higher than 0.60, showing good reliability. The convergent validity (AVE) was 0.715, significantly higher than 0.50, showing good convergent validity. The combined reliability (CR) was 0.938, significantly higher than 0.70, showing good combined reliability.

The mechanism variables are entrepreneurial alertness and dynamic capabilities, which are measured in the form of scales, as shown in Table 4, and the mean scores of the corresponding items are used, ranging from 1 to 5. In order to improve the reliability and validity, the content of the entrepreneurial alertness and dynamic competence scales are adapted from the existing literature. Statistics show that the Cronbach's alpha coefficients of entrepreneurial alertness and dynamic capabilities are respectively 0.895 and 0.907, which are significantly higher than 0.60, showing good reliability. The convergent validity (AVE) are respectively 0.743 and 0.651, significantly higher than 0.50, showing good convergent validity. The combined reliability (CR) ares respectively 0.896 and 0.912, significantly higher than 0.70, showing good combined reliability.

From the descriptive statistics in Table 5, we can see that the average income score of the e-commerce farmers is 3.33, which shows that the subjective judgment of the interviewed e-commerce farmers on their income level is in the middle to upper position. In terms of the mechanism variables, the mean scores of entrepreneurial alertness and dynamic capabilities are very close to each other, about 3.5, which is above the medium level.

Table 4 The content of the scale measurement items of the dependent variable and the mechanism variables

Variables	Content of measurement items	Sources
E-commerce farmers' income	At present, e-commerce profits are very good	Covin et al. [57]
	E-commerce orders are growing fast	Cooper and Artz [58]
	E-commerce operations have not encountered financial difficulties	Wall et al. [59]
	The goals set at the beginning have been achieved	
	Social status has been greatly improved than before engaging in e-commerce	
	I am satisfied with the current income level	
Entrepreneurial alertness	Even on a vacation, I always think about e-commerce	Ardichvili et al. [40]
	I would spend an evening talking about e-commerce	Guo and Zhou [60]
	When I'm not at work, I'm always thinking about e-commerce	Hu and Wang [61]
Dynamic capabilities	If I know about a new technology or product, I will look for an opportunity to experience it	Yi et al. [62]
	Among my friends and family, I was one of the first to try out new technology	Lan [63]
	I am willing to participate in the use of new technology, even if it costs a little money and time	Chien and Tsai [64]
	I will acquire external cutting-edge knowledge in a timely manner	Wilden et al. [65]
	I will quickly and accurately understand and master new knowledge acquired from outside	Lee et al. [66]
	I will digest and absorb the new knowledge acquired from outside	
	I am good at responding quickly to market changes and external opportunities	
	I will update the management mode as needed	

Table 5 Variable descriptions and descriptive statistics

Variable names	Variable definition description	Mean	S.D
E-commerce farmers' income	The mean scores of all measurement items ranged from 1 to 5	3.333	0.846
Entrepreneurial alertness	The mean scores of all measurement items ranged from 1 to 5	3.510	0.886
Dynamic capabilities	The mean scores of all measurement items ranged from 1 to 5	3.545	0.822

5 Empirical results

5.1 The impact of product attributes on e-commerce farmers' big data use decisions

Table 6 reports the results of regressions on the big data product attributes that influence whether or not e-commerce farmers use big data products. In particular, regression (1), regression (2), and regression (3) take whether e-commerce farmers use big data products as dependent variables and adopt binary Probit model for estimation. The difference is that regression (1), regression (2), and regression (3) take usefulness of big data products, ease of use of big data products and experience

Table 6 Regression results for product attributes influencing e-commerce farmers' use of big data products

Variables	Dependent variable: Whether or not to use big data products		
	(1)	(2)	(3)
Usefulness of big data products	0.120*** (0.033)	–	–
Ease of use of big data products	–	0.106*** (0.033)	–
Experience of big data products	–	–	0.191*** (0.049)
Gender	0.195*** (0.058)	0.234*** (0.058)	0.216*** (0.058)
Age	−0.208*** (0.060)	−0.179*** (0.061)	−0.195*** (0.060)
Education	0.096** (0.042)	0.088** (0.042)	0.106** (0.043)
Whether a CPC member or not	0.073 (0.078)	0.093 (0.078)	0.072 (0.078)
Whether enterprise operation or not	0.144** (0.072)	0.178** (0.071)	0.209*** (0.072)
Whether registered trademark or not	0.113 (0.073)	0.093 (0.077)	0.059 (0.078)
Duration of e-commerce business	0.009 (0.042)	−0.008 (0.043)	0.007 (0.043)
Main e-commerce formats	0.132*** (0.037)	0.127*** (0.037)	0.118*** (0.037)
Main product types	−0.682 (0.061)	−0.050 (0.060)	−0.077 (0.061)
Observations	418	418	418
R^2	0.156	0.156	0.161

(i) *** and ** denote 1% and 5% significance levels respectively; (ii) robust standard errors are reported in brackets; (iii) binary Probit models were used for estimation and to report the marginal effects of the independent variables

Table 7 OLS regression results for the impact of big data use on e-commerce farmers' income and its mechanism

Dependent variables	Income	Entrepreneurial alertness	Dynamic capabilities
Whether or not to use big data products	0.302*** (0.093)	0.577*** (0.099)	0.421*** (0.089)
Gender	0.042 (0.098)	0.102 (0.102)	0.193** (0.094)
Age	0.031 (0.106)	0.027 (0.107)	−0.004 (0.096)
Education	0.101 (0.663)	0.145** (0.069)	0.141** (0.070)
Whether a CPC member or not	0.021 (0.112)	−0.185 (0.118)	−0.074 (0.106)
Whether enterprise operation or not	−0.373*** (0.086)	−0.165 (0.103)	−0.863 (0.099)
Whether registered trademark or not	0.365*** (0.098)	0.077 (0.123)	0.217** (0.110)
Duration of e-commerce business	0.041 (0.065)	0.093 (0.075)	0.011 (0.066)
Main e-commerce formats	0.154*** (0.058)	−0.073 (0.068)	0.082 (0.079)
Main product types	0.148 (0.099)	0.042 (0.092)	0.034 (0.092)
Observations	418	418	418
R^2	0.149	0.149	0.151

(i) *** and ** denote 1% and 5% significance levels respectively; (ii) robust standard errors are reported in brackets

of big data products as the core independent variables respectively. The estimation results show that big data product usefulness, big data product ease of use and big data product experience are all significantly influence e-commerce farmers' big data use decisions at the 1% level. It indicates that the psychological perception of big data product attributes are important factors to drive e-commerce farmers to use big data products, and thus the hypotheses of H1, H2, and H3 are verified. In terms of the marginal effects of the estimated coefficients, each unit increase in the score of usefulness, ease of use and experience of big data can increase the average probability of using big data products among e-commerce farmers by 12.0%, 10.6% and 19.1% respectively. In comparison, the marginal effect of big data product experience is greater in driving the use of big data products by e-commerce farmers. This demonstrates the necessity of introducing the dimension of perceived experience into the technology acceptance model in this paper.

5.2 The impact of big data use decisions on e-commerce farmers' income

Table 7 reports OLS regression[3] results of the impact of big data use on the income of e-commerce farmers and its mechanism. It can be seen that big data

[3] Since cross-section data is used in this study, the model test of OLS regression is mainly heteroscedasticity and multicollinearity. White test results show that the model has heteroscedasticity, and we use robust standard error to deal with it. The results of VIF show that there is no multicollinearity in the model.

products use significantly contribute to the income increase of e-commerce farmers, with e-commerce farmers using big data scoring on average 0.302 more than ordinary e-commerce farmers. In addition, the use of big data products significantly increased the entrepreneurial alertness and dynamic capabilities of e-commerce farmers. However, OLS regressions need to be treated with caution for their estimation results do not take the endogeneity of self-selection into account.

Since OLS regression cannot effectively control the differences in covariates between the treatment group (i.e. big data-using e-commerce farmers) and the control group (i.e. ordinary e-commerce farmers), the two groups of samples are adjusted by further adopting an entropy balancing approach to set constraints on the covariates such as first-order moments (mean), second-order moments (variance) and third-order moments (skewness) and using their automatically calculated optimal weights as equilibrium weights, so that the two groups of samples are precisely matched under the constraints, and the selective bias of the samples is controlled to the maximal extent. Table 8 demonstrates the mean and variance of the covariates before and after the entropy balancing treatment and the results of the matching test. As can be seen, before matching, the means and variances of the covariates in the treatment and control groups were significantly different, and after the entropy balancing treatment, the differences in the means and variances of the covariates were significantly reduced. To further test the reliability of the entropy equalization results, the standardized mean differences (SMD) between the treatment and control groups can be further calculated and a t-test of the differences in means can be conducted. The results show that the standardized mean differences between the two groups after treatment are all zero and the p-values of the t-tests of the differences in means of the covariates are all one, indicating that the data of each covariate in the treatment and control groups have been matched exactly.

Table 9 reports the results of the entropy balancing method of estimating the impact of big data products use on the income of e-commerce farmers and its mechanism. It can be seen that after removing the sample selection bias, the big data use decisions still significantly and positively affects the income, entrepreneurial alertness and dynamic capabilities of e-commerce farmers at the 1% level. Thus the hypotheses of H4 and H5 are verified. By comparing the magnitude of the estimated coefficients in Tables 7 and 9, it can be seen that OLS regression leads to an overestimation of the income-increasing effect of big data use and produce exaggerated judgments due to the problem of self-selection. Theoretically, the use of big data products can significantly improve the business capabilities of e-commerce farmers and promote income growth by bringing them information services and decision-making guidance. However, in practice, the extent to which the use of big data can increase income is closely related to the ability of platform enterprises to develop big data products and the depth of use by e-commerce farmers. On the one hand, in China, big data products based on e-commerce platforms are still at an early stage, and it takes time for big data technology to reach a more mature level in terms of information collection, processing and analysis. On the other hand, the use of big data products by e-commerce farmers is still in its infancy, and it will take some time for them to make targeted improvements in their production operations based

Table 8 Matching tests for covariates after treatment

Covariates		Mean		Variance		Standardized Mean Difference (SMD)	Difference in means t-test p-value
		Treatment group	Control group	Treatment group	Control group		
Gender	Before	0.650	0.389	0.228	0.239	0.039	0.000
	After	0.653	0.653	0.228	0.228	0	1
Age	Before	1.595	1.739	0.272	0.293	-0.019	0.007
	After	1.590	1.590	0.275	0.285	0	1
Education	Before	2.050	1.803	0.459	0.446	0.026	0.000
	After	2.037	2.037	0.470	0.478	0	1
Whether a CPC member or not	Before	0.250	0.162	0.188	0.137	0.014	0.030
	After	0.226	0.226	0.176	0.176	0	1
Whether enterprise operation or not	Before	0.340	0.182	0.226	0.150	0.024	0.000
	After	0.347	0.347	0.228	0.228	0	1
Whether registered trademark or not	Before	0.455	0.280	0.249	0.203	0.025	0.000
	After	0.463	0.463	0.250	0.250	0	1
Duration of e-commerce business	Before	1.821	1.674	0.588	0.607	0.014	0.059
	After	1.832	1.832	0.596	0.635	0	1
Main e-commerce formats	Before	0.525	0.455	0.593	0.389	0.006	0.047
	After	0.584	0.584	0.604	0.620	0	1
Main product types	Before	0.260	0.286	0.194	0.205	-0.004	0.560
	After	0.253	0.253	0.190	0.190	0	1

$SMD = (\overline{X}_t - \overline{X}_c)/\sqrt{[S_t^2(n_t-1) + S_c^2(n_c-1)/(n_t+n_c-2)]}$, where $\overline{X}_t$ and $\overline{X}_c$ are the means of the variables in the treatment and control groups respectively,S_t^2 and S_c^2 are the variances of the variables in the treatment and control groups respectively, and n_t and n_c are the sample sizes of the treatment and control groups respectively.

Table 9 Estimation results of the entropy balancing method

Dependent variables	Income level	Entrepreneurial alertness	Dynamic capabilities
Whether or not to use big data products	0.264*** (0.098)	0.552*** (0.126)	0.352*** (0.122)
Control variables	Yes	Yes	Yes
Observations	418	418	418
R^2	0.036	0.089	0.042

(i) *** and ** denote 1% and 5% significance levels respectively; (ii) robust standard errors are reported in brackets

on the analysis results of big data products. The results of the entropy balancing method are more reasonable than the OLS regression results.

6 Discussion

The findings of this paper firstly contribute to enrich the field of rural e-commerce sustainability and provide a new way to promote the sustainability of rural e-commerce in developing countries. The rise of rural e-commerce has brought surprises to farmers' income in the beginning stage, and has aroused the research excitement of scholars. A large number of previous studies have consistently confirmed the positive effect of rural e-commerce development on farmers' income [67–70]. However, these studies can only prove the positive role of the rise of rural e-commerce, and do not pay attention to the phased changes in the development of rural e-commerce. With the continuous evolution of rural e-commerce, e-commerce farmers will gradually face various challenges. In general, the development of rural e-commerce can be divided into three stages: scale expansion period, quality improvement period and brand innovation period. During the scale expansion period, the rural social network plays the role of information diffusion, stimulates the imitation learning effect of farmers, and a large number of farmers join the camp of e-commerce entrepreneurship. However, with the rapid expansion of the number of e-commerce farmers and the high degree of product homogeneity, price wars will inevitably break out, and the profit margin of products will continue to decline. At this time, it is very easy to produce an endogenous quality crisis, that is, some e-commerce farmers begin to sacrifice product quality and implement low-price competitive strategies. In this regard, local governments need to prevent the decline of the industry caused by race to the bottom, pay attention to strengthening product quality supervision, and make rural e-commerce enter a period of quality improvement. At this stage, the number of e-commerce farmers will stabilize, and the improvement of product quality will increase the output value. With the further development of rural e-commerce, the price of land rent, labor and other factors will gradually rise, coupled with the intensification of external competition from other regions, so that e-commerce farmers have to upgrade. In general, there are three ways to solve the problems of rural e-commerce. The first idea is to enhance policy supports. The government needs to provide public goods and services in a timely manner according to the changing needs of rural e-commerce development. Supporting services such as land index, skill training, talent introduction, finance, warehousing and administrative approval required for the development of e-commerce need to be provided systematically [9, 21]. The second idea is to take collective actions. E-commerce farmers must actively participate in collective actions, such as joining e-commerce associations, strengthening industry self-discipline together with other e-commerce farmers, and jointly improving collective competitive advantages [10, 12]. And the improvement of collective efficiency will also help to improve individual efficiency. The third idea is to improve the endogenous capacity of e-commerce farmers, so that the online operation level of e-commerce farmers can be improved. The solutions that scholars mainly mention focus on the first and the second ideas, but ignore the third one. There is no denying that e-commerce farmers can't successfully cope with various challenges without the assistance of the government and industry associations. However, the assistance of the government and industry associations ultimately belongs to the common external force, and the upgrade of e-commerce farmers needs their own personalized improvement. In the face of profit decline, no matter it is caused by product homogenization or factor cost rise, the coping strategy of e-commerce farmers must be to improve their alertness to new opportunities and dynamic capabilities of innovation. It seems that scholars have not found any good way to improve the endogenous capacity of e-commerce farmers. In this regard, this paper paid attention to the empowerment of big data. This paper introduces a new perspective on the use of big data products to improve the endogenous capacity of e-commerce farmers and the sustainability of rural e-commerce. The existing literature pays little attention to the role of big data products developed by e-commerce platform enterprises in improving the sustainability of rural e-commerce, and this paper fills this gap. We consider the new phenomenon which first occurred in China theoretically and test it empirically. The empirical study confirms that the use of big data has a significant positive effect on the income of e-commerce farmers. Compared to ordinary e-commerce farmers, those e-commerce farmers who use big data products have seen a qualitative change in their access to information. With more timely, comprehensive and accurate information, e-commerce farmers can improve their weaknesses in a more targeted manner, become more demand-driven and user-oriented, and be more proactive in seeking first-mover advantages, thereby continuously improving their performances. The sustainability problem of rural e-commerce is ultimately the problem of sustainable income increase of e-commerce farmers. The latter is the micro-foundation of the former. The key to promoting the successful evolution of rural e-commerce in developing countries from the period of scale expansion to the period of brand innovation is to enhance the entrepreneurial alertness and dynamic ability of e-commerce farmers. China's practice shows that it is an effective way to give play to the empowering role of big data, that is, e-commerce farmers use big data products developed by e-commerce enterprises to guide their own product innovation, technological innovation, marketing innovation and brand construction. Therefore, this paper provides micro-evidence for data-enabled rural e-commerce improvement, and thus opens more horizons for rural e-commerce research in developing countries.

Another contribution of this paper is that it provides practical contribution for the construction and applications of rural public database, including agricultural big data. Existing literature on the application of rural public database mainly discusses the impact of big data technology on agricultural and rural development and the promotion of big data in developing countries from the macro level [49, 71–73]. However, the research on farmers from the micro perspective is rare, especially the survey research and empirical analysis are quite scarce [74]. There is also a part of relevant literature focused on the impact of big data development on the labor income share of enterprises [8–11], and these studies cannot provide effective guidance for the diffusion of big data technology in rural areas. Farmers and enterprises, as well as rural entrepreneurs and enterprises in the labor force are different. To study the impact of big data on farmers' income, it is necessary to conduct empirical research specifically for farmers. The application of big data in rural China is still in the early stage, and existing studies have not yet observed the new phenomenon, let alone from the perspective of product attributes. The application of big data products on e-commerce platforms by e-commerce farmers is the first practice of big data application in rural China, which has important demonstration value. The research of this paper on the product attributes of big data developed by e-commerce platform enterprises is beneficial to provide references for the construction of rural public database from the product attributes perspective. This paper proposes a big data product attributes framework with three dimensions of usefulness, ease of use and experience. Perceived usefulness and perceived ease of use are the two core variables of TAM [31]. In the more than 30 years since the technology acceptance model was proposed, numerous empirical studies have confirmed the powerful explanatory power of perceived usefulness and perceived ease of use in the field of technology adoption [74]. However, this article argues that usefulness and ease of use do not cover the full perspective of product attributes. Even if a product is useful and easy to use, there is no guarantee that consumers will use the product. Especially the younger generation of consumers, they attach great importance to the humanized and personalized design of products. Thus, this paper constructs a three-dimensional conceptual framework of big data product attributes in terms of usefulness, ease of use and experience, which makes the understanding of big data product attributes more complete. The empirical results show that the experience of big data products plays the largest marginal role in driving the use of big data products by e-commerce farmers. This demonstrates the necessity of incorporating experience into the dimensionalized framework of big data product attributes. TAM's extended study was later used to analyze older adults and was once seen as an innovation in the subject matter [75]. This paper expands the research from the perspective of the younger generation. TAM's explanatory power for the elderly is still strong, but it may not be fully applicable to the younger generation. Farmers are different from enterprises, and if the government wants to further promote the spread of big data technology in rural areas, it must make the product attributes of public big databases meet the needs and characteristics of farmers. In this process, the usefulness of big data products is the foundation, the ease of use of big data products is the driving force, and the experience of big data products is the guarantee. In other words, governments should focus on the three aspects of usefulness, ease of use and experience in the process of promoting the construction of rural public database. Only in this way, the promotion of rural public database in the later stage can produce real diffusion effect, so that the data can truly empower the modernization of the rural areas.

7 Conclusion

With the rapid development of e-commerce in rural China, more and more farmers have become specialized in e-commerce. Big data products developed by e-commerce platform enterprises can assist e-commerce farmers in making more scientific and accurate decisions on online store operation. Based on the survey data of 418 e-commerce farmers collected from 15 typical Taobao Villages in Zhejiang Province, China, this paper adopts Probit model and entropy balancing method to make empirical studies. It is found that the perception of big data product attributes is an important factor driving e-commerce farmers to use big data products. Specifically, the usefulness, ease of use and experience of big data products all have a significantly positive impact on the e-commerce farmers' big data use decisions. Additionally, the use of big data significantly promotes the income increase of e-commerce farmers, and the improvement of entrepreneurial alertness and dynamic capabilities is important mechanisms for the use of big data to affect the income of e-commerce farmers. It is indicated that the use of big data products is an effective way for e-commerce farmers to enhance their competitive advantages, and improving e-commerce farmers' perception of big data product attributes is an important prerequisite for the proliferation of big data products.

The research conclusions have important implications for governments, e-commerce platform enterprises, and e-commerce farmers. Governments should enhance the attention of big data, increase the support for rural e-commerce skills training, and provide low-income farmers with training in big data application. In addition, governments should pay attention to improving the data product attributes of databases, including usefulness, ease of use and experience, in the process of promoting the construction of rural public database such as big data in the whole industry chain of agricultural products. It is suggested that the government should stimulate the extensive participation of farmers in the process of construction of rural public database. E-commerce platform enterprises need to continue to strengthen product innovation and further improve the usefulness, ease of use and experience of big data products, especially the experience of big data products. E-commerce farmers should improve their awareness of learning and using big data products, actively participate in e-commerce skills training, and constantly improve their professional knowledge. Especially those e-commerce farmers who have not yet contacted and used big data products should change their psychological cognition and try to accept new things. With the continuous development of digital economy, data will play an increasingly important role in rural development. E-commerce farmers must recognize the nature of digital economy, and learn to use big data products to improve their entrepreneurial alertness and dynamic capabilities. Only in this way, can they maintain their competitiveness and avoid elimination.

The use of big data products by e-commerce farmers is a new phenomenon, and this paper opens the research on this topic. This paper reveals the driving mechanism of e-commerce farmers' use of big data products from the perspective of product attributes, but product attributes are certainly not the only factor affecting e-commerce farmers' decisions on big data use. In the future, scholars can explore other possible driving factors, such as the social network of e-commerce farmers and participation in skill training. In addition, this paper discusses the positive effect of the use of big data on the income of e-commerce farmers. In the future, scholars can further study the impact of the use of big data on other aspects of e-commerce farmers (e.g. innovative learning, innovative ability, growth), and pay attention to the impact of the big data use by e-commerce farmers on the evolution of industrial clusters and community governance through case studies.

Acknowledgements The authors acknowledges the supports in data collection, article editing and suggesting revisions from Zhuoer Mao, Hongtao Wang, Kaida Wang, Shiyu Ying, and Tianhe Hu.

Funding This study was funded by the Philosophy and Social Science Planning Zhejiang Youth Special Project of Zhejiang Province (Grant No. 24ZJQN046Y), the Natural Science Foundation of Zhejiang Province (Grant No. LQ23G030007), and the Special Fund for Basic Scientific Research of Universities in Zhejiang Province (Grant No. GK229909299001-216).

Declarations

Conflict of interest The authors state that they have no conflicts of interest.

References

1. Qi, J., Zheng, X., & Guo, H. (2019). The formation of taobao villages in China. *China Economic Review, 53*, 106–127.
2. Ma, W., Zhou, X., & Liu, M. (2020). What drives farmers' willingness to adopt e-commerce in rural China? The role of Internet use. *Agribusiness, 36*(1), 159–163.
3. Guo, G., Liang, Q., & Luo, G. (2014). Effects of clusters on China's e-commerce: Evidence from the Junpu Taobao Village. *International Journal of Business Management, 9*(6), 180–186.
4. Zeng, Y., Guo, H., & Jin, S. (2018). Is e-commerce good for farmers' income? Evidence from Shuyang. *Jiangsu Province. Chinese Rural Economy, 2*, 49–64. (in Chinese).
5. Luo, X., & Niu, C. (2019). E-commerce participation and household income growth in Taobao Villages. World Bank Policy Research Working Papers.
6. Qi, J., Zheng, X., Cao, P., & Zhu, L. (2019). The effect of e-commerce agribusiness clusters on farmers' migration decisions in China. *Agribusiness, 35*(1), 20–35.
7. Li, X., Guo, H., Jin, S., Wang, M., & Zeng, Y. (2021). Do farmers gain internet dividends from e-commerce adoption? *Evidence from china. Food Policy, 101*, 102024.
8. Tang, W., & Zhu, J. (2020). Informality and rural industry: Rethinking the impacts of e-commerce on rural development in China. *Journal of Rural Studies, 75*, 20–29.
9. Gao, P., & Liu, Y. (2020). Endogenous inclusive development of e-commerce in rural China: A Case Study. *Growth and Change, 51*(4), 1611–1630.
10. Zeng, Y., Guo, H., Yao, Y., & Huang, L. (2019). The formation of agricultural e-commerce clusters: A case from China. *Growth and Change, 50*(4), 1356–1374.

11. Yu, H., & Cui, L. (2019). China's e-commerce: Empowering rural women. *The China Quarterly, 238*, 418–437.
12. Li, L., Zheng, L., Zhang, Z., & Song, Y. (2024). How to trigger and strengthen the positive impact of the internet on the income of farmers in the region? A case from China. *Electronic Commerce Research, 4*, 1407–1433.
13. Zeng, Y., Zhang, Z., Fang, H., & Guo, H. (2019). The driving forces and income-added effects of big data usage by rural e-commerce households. *Chinese Rural Economy, 12*, 29–47. (in Chinese).
14. Graetz, G., & Michaels, G. (2018). Robots at work. *Review of Economics and Statistics, 100*(5), 753–768.
15. Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review, 108*(6), 1488–1542.
16. Goldfarb, A., & Tucker, C. (2019). Digital economics. *Journal of Economic Literature, 57*(1), 3–43.
17. Nordhaus, D. (2021). Are we approaching an economic singularity? Information technology and the future of economic growth. *American Economic Journal: Macroeconomics, 13*(1), 299–332.
18. Zhang, Q., Sun, S., Wu, S., & Li, D. (2019). Study on the continuous growth and evolution of online business groups of farmers in Taobao Village. *Journal of Agro-technical Economics, 1*, 121–134. (in Chinese).
19. Shao, Z. (2019). Limitation and reflection on the ideation ability of farmers' online merchants in cyberspace. *Journal of Northwest A&F University (Social Science Edition), 19*(2), 90–98. (in Chinese).
20. Shao, Z. (2017). Rules and the logic of capital: The shaping mechanism of farmers' online stores in taobao village. *Journal of Northwest A&F University (Social Science Edition), 17*(4), 74–82. (in Chinese).
21. Zeng, Y., Jia, F., Wan, L., & Guo, H. (2017). E-commerce in agri-food sector: A systematic literature review. *International Food and Agribusiness Management Review, 20*(4), 439–459.
22. Lin, G., Xie, X., & Lv, Z. (2016). Taobao practices, everyday life and emerging hybrid rurality in contemporary China. *Journal of Rural Studies, 47*, 514–523.
23. Leong, C., Pan, S. L., Newell, S., & Cui, L. (2016). The emergence of self-organizing e-commerce ecosystems in remote villages of China: A tale of digital empowerment for rural development. *MIS Quarterly, 40*(2), 475–484.
24. Niebel, T., Rasel, F., & Viete, S. (2019). Big data - big gains? Understanding the link between big data analytics and innovation. *Economics of Innovation and New Technology, 28*(3), 296–316.
25. Kim, K., & Chhajed, D. (2002). Product design with multiple quality-type attributes. *Management Science, 48*(11), 1502–1511.
26. Anderson, E. W., Claes, F., & Donald, R. L. (1994). Consumer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing, 58*(3), 53–66.
27. Leclerc, F., Schmitt, B. H., & DubÉ, L. (1994). Foreign branding and its effects on product perceptions and attitudes. *Journal of Marketing Research, 31*(2), 263–270.
28. Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and utilitarian dimensions of consumer attitude. *Journal of Marketing Research, 40*(3), 310–320.
29. Sun, M. (2011). Disclosing multiple product attributes. *Journal of Economics & Management Strategy, 20*(1), 195–224.
30. Kwark, Y., Chen, J. Q., & Raghunathan, S. (2014). Online product reviews: Implications for retailers and competing manufacturers. *Information Systems Research, 25*(1), 93–110.
31. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.
32. Dishaw, M. T., & Strong, D. M. (1999). Extending the technology acceptance model with task-technology fit constructs. *Information & Management, 36*(1), 9–21.
33. Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204.
34. DubÉ, L., Cervellon, M. C., & Han, J. (2003). Should consumer attitudes be reduced to their affective and cognitive bases? Validation of a hierarchical model. *International Journal of Research in Marketing, 20*(3), 259–272.
35. Pham, L. B., & Taylor, S. E. (1999). From thought to action: Effects of process-versus outcome-based mental simulations on performance. *Personality and Social Psychology Bulletin, 25*(2), 250–260.

36. Mittal, V., & Kamakura, W. A. (2001). Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. *Journal of Marketing Research, 38*(1), 131–142.
37. Zhang, P. W. (2017). The impact of social capital on farmers' willingness to participate in credit union: An analysis based on 393 survey data of farmers in Xuzhou City. *Journal of Hunan Agricultural University (Social Science Edition), 18*(4), 24–29. (in Chinese).
38. Elkins, A. C., Dunbar, N., Adame, B., & Nunamaker, J. F. (2013). Are users threatened by Credibility assessment systems? *Journal of Management Information Systems, 29*(4), 249–262.
39. Oh, J., & Yoon, S. (2014). Validation of haptic enabling technology acceptance model (HE-TAM): Integration of IDT and TAM. *Telematics and Informatics, 31*(4), 585–596.
40. Ardichvili, A., Cradozo, R., & Ray, S. (2003). A theory of entrepreneurial opportunity Identification and development. *Journal of Business Venturing, 18*(1), 105–123.
41. Kirzner, M. (1978). *Competition and Entrepreneurship*. University of Chicago Press.
42. Valliered, D. (2013). Towards a schematic theory of entrepreneurial alertness. *Journal of Business Venturing, 28*(3), 430–442.
43. Gaglio, C. M., & Katz, J. A. (2001). The psychological basis of opportunity identification: Entrepreneurial alertness. *Small Business Economics, 16*(2), 95–111.
44. Ko, S., & Butler, J.E. (2003). Alertness, bisociative thinking ability, and discovery of entrepreneurial opportunities in Asian hi-tech firms. In *Frontiers of Entrepreneurship Research,* Proceedings of the Babson-Kauffman Entrepreneurship Research Conference, Wellesley, MA: Babson College, 421–429.
45. Allinson, C. W., Chell, E., & Hayes, J. (2000). Intuition and entrepreneurial behaviour. *European Journal of Work and Organizational Psychology, 9*(1), 31–43.
46. Hajli, M., Sims, J., & Ibragimov, V. (2015). Information technology productivity paradox in the 21st century. *International Journal of Productivity and Performance Management, 64*(4), 457–478.
47. Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning, 51*(1), 40–49.
48. Newbert, S. L. (2005). New firm formation: A dynamic capability perspective. *Journal of Small Business Management, 43*(1), 55–77.
49. Lokers, R., Knapen, R., Janssen, S., Randen, Y. V., & Janse, J. (2016). Analysis of big data technologies for use in agro-environmental science. *Environmental Modelling & Software, 84*, 494–504.
50. Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. *Artificial Intelligence in Agriculture, 2*, 1–12.
51. Shaver, J. M. (1998). Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival ? *Management Science, 44*(4), 571–585.
52. Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis, 20*(1), 25–46.
53. Hainmueller, J., & Xu, Y. (2013). E-balance: A stata package for entropy balancing. *Journal of Statistical Software, 54*(7), 1–18.
54. Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research, 11*(4), 342–365.
55. Yang, S. (2016). Role of transfer-based and performance-based cue on initial trust in mobile shopping services: A cross-environment perspective. *Information System and e-Business Management, 14*(1), 47–70.
56. Wu, J. H., & Wang, S. C. (2005). What drives mobile commerce? An empirical evaluation of the revised technology acceptance model. *Information & Management, 42*(5), 719–729.
57. Covin, J. G., Mcdougal, P. P., Robinson, J., & Herron, L. (1994). The effects of industry growth and strategic breadth on new venture performance and strategy content. *Strategic Management Journal, 15*(7), 537–554.
58. Cooper, A. C., & Artz, K. W. (1995). Determinations of satisfaction for entrepreneurs. *Journal of Business Venturing, 10*(6), 439–457.
59. Wall, T. D., Michie, J., Patterson, M., Wood, S. J., & West, M. A. (2004). On the validity of subjective measures of company performance. *Personnel Psychology, 57*(1), 95–118.
60. Guo, H., & Zhou, H. (2013). Previous experience, entrepreneurial alertness and farmers' entrepreneurial opportunity recognition: A mediating effect model and its implications. *Journal of Zhejiang University (Humanities and Social Sciences), 43*(4), 17–27. (in Chinese).

61. Hu, H., & Wang, C. (2013). Frontier analysis and future prospect of entrepreneurial alertness research. *Foreign Economics and Management, 35*(12), 11–19. (in Chinese).
62. Yi, M. Y., Fiedler, K. D., & Park, J. S. (2006). Understanding the role of individual innovativeness in the acceptance of IT-based innovations: Comparative analyses of models and measures. *Decision Sciences, 37*(3), 393–426.
63. Lan, P. (2017). Research on tourists' adoption behavior of mobile tourism applications. *Doctoral dissertation of Xiamen University*. (in Chinese)
64. Chien, S. Y., & Tsai, C. H. (2012). Dynamic capabilities, knowledge, learning, and firm performance. *Journal of Organizational Change Management, 125*(3), 434–444.
65. Wilden, R., Gudergan, S. P., Nielsen, B. B., & Lings, L. (2013). Dynamic capabilities and performance, strategy, structure and environment. *Long Range Planning, 46*(1–2), 72–96.
66. Lee, P., Joseph, L. C., & Wu, M. (2018). The roles of cross-cultural adjustment and social capital formation in the dynamic capabilities development of multi-unit organizations. *Asian Pacific Management Review, 23*(1), 20–29.
67. Galloway, L., Sanders, J., & Deakins, D. (2011). Rural small firms' use of the internet: From global to local. *Journal of Rural Studies, 27*(3), 254–262.
68. Zapata, S. D., Carpio, C. E., Isengildina-Massa, O., & Lamie, D. R. (2013). The economic impact of services provided by an electronic trade platform: The case of Marketmaker. *Journal of Agricultural and Resource Economics, 38*(3), 359–378.
69. Fan, Q., & Garcia, V. B. S. (2018). Information access and smallholder farmers' market participation in Peru. *Journal of Agricultural Economics, 69*(2), 476–494.
70. Ma, W., & Wang, X. (2020). Internet use, sustainable agricultural practices and rural incomes: Evidence from China. *Australian Journal of Agricultural and Resource Economics, 64*(4), 1087–1112.
71. Nandyala, C. S., & Kim, H. K. (2016). Big and meta data management for u-agriculture mobile services. *International Journal of Software Engineering & Application, 10*(1), 257–270.
72. Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy, 40*(1), 79–96.
73. Turland, M., & Slade, P. (2020). Farmers' willingness to participate in a big data platform. *Agribusiness, 36*(1), 20–36.
74. Chen, L., & Chen, D. (2019). Research on farmers' use behavior of agricultural big data services under the strategy of rural revitalization: Based on the micro data of 973 farmers. *Journal of Nanjing Normal University (Social Sciences Edition), 6*, 123–132. (in Chinese).
75. He, J., & Huang, X. (2020). The smartphone use and well-being of urban elderly: Based on intergenerational support and TAM. *Journal of International Communication, 3*, 49–73. (in Chinese).