Investigation and Empirical Analysis of Personal Information Collected by Digital Business Platforms

From the Perspective of Big Data Discriminatory Pricing (BDDP)

Huali Chen[1] and Fei Wang[2(✉)]

[1] School of Law, Guangzhou Xinhua University, Tianhe District, Guangzhou, China
[2] School of Management, Guangzhou Xinhua University, Tianhe District, Guangzhou, China
1318382963@qq.com

Abstract. The rapid development of big data and artificial intelligence is changing the traditional economic model. By collecting a wealth of personal information from the consumer manipulated by algorithms and then using algorithmic technology to accurately profile them, digital business platforms provide greater scope for their pricing activities, thus giving rise to the phenomenon of "big data discriminatory pricing (BDDP)". The research revealed that the consumer manipulated by algorithms are not fully aware of exactly of what personal details are being collected from them and are not enjoying the full rights of informed consent, resulting in a lack of consent to price discrimination. The correlation analysis of the data suggests that digital business platforms should disclose the personal details collected from individual the consumer manipulated by algorithms to safeguard their rights of informed consent, so as to build a platform of trust between the two parties and adopt reasonable price discrimination under the premise of legal compliance.

Keywords: Algorithmic Consumer · Information Collection · Big Data Discriminatory Pricing (BDDP) · Informed Consent · Price Discrimination

1 Introduction

Intelligent technologies with the core of data and algorithms have accelerated the development of the digital economy. The 49th Statistical Report on the Development of the Internet in China, released in February 2022, showed that "as of December 2021, the scale of online shopping users in China had reached 842 million, with an increase of 59.68 million compared to December 2020, which accounted for 81.6% of the overall number of Internet users". The China Mobile Internet Development Report (2022) showed that "in 2021, national online retail sales reached 13.1 trillion yuan, with an increase of 14.1% over the previous year, which accounted for 24.5% of the total retail sales of consumer goods". As an extremely significant business entity in the digital economy, these technologies collect and use the personal information of customers. Then with the help of data algorithms, they automatically identify consumer needs, build persona for

G. Guan et al. (Eds.): ICBDSS 2022, AHCS 8, pp. 410–420, 2023.
https://doi.org/10.2991/978-94-6463-064-0_44

consumers and make precise recommendations. At the same time, these technologies also adjust marketing strategies in a suitable manner to enhance both users' experience and profitability, thus building a brand-new business operation model. Therefore, the consumers on the Big Data platform become the algorithmic consumers.

While the digital economy has brought convenience, it has also brought new problems. Consumers have found that operating platforms sell the same goods or services to different consumers at different prices, thus capturing more consumer surplus. Judgment of (2018) Xiang 0102 Min Chu No.13515 showed that the plaintiff, Liu Quan, ordered a meal through the Meituan takeout platform operated by Beijing Science and Technology Co., Three Fast at 11:55 p.m. on 19 July 2018. At 12:08 that day, another colleague of Liu Quan ordered the same meal from the same merchant on the platform with the same delivery address, but the delivery fee was 1 yuan less than Liu Quan. The price discrimination by Beijing Science and Technology Co., Three Fast has infringed upon Liu Quan's legitimate rights and interests by taking advantage of the monopoly of the industry and the technical means of "big data discriminatory pricing (BDDP)". Judgment of (2020) Hu 0105 Min Chu No.9010 showed that on 22 August 2018, the plaintiff Zheng Yugao had a similar experience to Liu Quan when he bought two tickets on Ctrip for the same flight on the same day, but with different prices around RMB 500 between the two purchases.

Do digital business platforms fulfil their obligation to inform while collecting personal information and algorithmic recommendations? Are consumers aware of what personal information has been collected from them? Are consumers aware of whether they are enjoying the right to consent? Are consumers informed enough to accept a certain level of price discrimination, thereby reducing consumer surplus and achieving simultaneous growth in profits and sales volume?

2 Questionnaire on the Collection of Personal Information by Digital Business Platforms

2.1 Design of the Questionnaire

The purpose of this survey is to understand the collection of personal information of the consumer manipulated by algorithms by digital business platforms. To be specific, this survey includes whether the algorithmic consumer is informed, consents to the personal in-formation collection, and accepts a certain amount of price discrimination on the basis of informed. The questionnaire was designed using a 12-question 5-point Likert scale, including "strongly agree, agree, neutral, disagree and strongly disagree", with five statistical options of "1, 2, 3, 4 and 5". The details are shown in Table 1.

2.2 Distribution of the Questionnaire

The questionnaire was distributed randomly through the questionnaire website from 29 May 2022 to 11 Jun-e 2022, a total of 362 questionnaires were collected over a period of 2 weeks, with a total of 336 valid questionnaires. In terms of the age of respondents, 94.94% were aged 18–40. In terms of the gender of the respondents, 29.76% were male

Table 1. Design of the questionnaire

Purpose	Question No.	Content of the Question
Whether informed	Q1	I read the Guidelines for the Protection of Personal Information when registering on online shopping platforms.
	Q2	I read the Privacy Policy when I registered on the online shopping platform.
	Q3	I know exactly what personal information is collected and used by the shopping platform.
	Q4	I clearly know the personalized recommendation function of the shopping platform.
Whether to agree to collect	Q5	The shopping platform gets my browsing information to recommend products that are useful to me.
	Q6	The shopping platform makes it easier for me to buy the products I want through getting my browsing information and recommending them to me.
	Q7	I think the online shopping platform will protect the safe use of my personal information.
Whether to accept price discrimination on the basis of informed	Q8	I am able to accept that shopping platforms selling the same product at different prices for different customers with different purchasing power in the same situation when I had been informed.
	Q9	I am able to accept that the shopping platform will only issue a certain number of vouchers to new users when I am informed.
	Q10	I am able to accept that the shopping platforms determine my purchasing power by accessing my personal usage information, such as my shopping history, in order to increase the price for me in a targeted manner.

(*continued*)

Table 1. (*continued*)

Purpose	Question No.	Content of the Question
	Q11	I am able to accept that shopping platforms determine my preferences by accessing my personal usage information, such as product browsing, in order to increase prices for me in a targeted manner.
	Q12	A shopping platform is knowingly the big data discriminatory pricing (BDDP), which sells products to me at a higher price than other customers in the same situation, and I will still buy products or services from this shopping platform.

and 70.21% were female. In terms of the educational qualifications of the respondents, 94.34% had a bachelor's degree or above. In terms of the income of the respondents, 74.7% had a monthly income of RMB 3,000 and above.

3 Questionnaire Analysis of the Collection of Personal Information by Digital Business Platforms

3.1 Validity and Reliability Analysis of the Questionnaire

According to the data analysis through the software of SPSSAU, the reliability coefficient of the questionnaire, the Cronbach alpha coefficient, was 0.809, which is greater than 0.8, indicating the high quality of the reliability of the study data; the validity coefficient of the questionnaire, the KMO value, was 0.752, which ranged from 0.7 - 0.8, reflecting the good validity from a side perspective, which indicates that the study data is suitable for extracting information.

3.2 Descriptive Analysis

According to the data analysis through the software of SPSSAU, on the three purposes of the research on "whether informed", "whether to agree to collect" and "Whether to accept price discrimination on the basis of informed", the consumer manipulated by algorithms have a combined mean of around 2.5, which is generally neutral. Details are shown in Table 2.

On both Q7 (I think the online shopping platform will protect the safe use of my personal information.) and Q12(A shopping platform is knowingly the big data discriminatory pricing (BDDP), which sells products to me at a higher price than other customers in the same situation, and I will still buy products or services from this shopping platform.) issues, respondents were neutral in agreement. From this data it is objectively clear

Table 2. Mean and combined mean for each question in the questionnaire

Purpose	Question No.	Sample Sizes	Minimum Value	Maximum value	Mean	Combined mean
Whether informed	Q1	336	1.000	5.000	2.464	2.541
	Q2	336	1.000	5.000	2.420	
	Q3	336	1.000	5.000	2.509	
	Q4	336	1.000	5.000	2.774	
Whether to agree to collect	Q5	336	1.000	5.000	2.506	2.421
	Q6	336	1.000	5.000	2.485	
	Q7	336	1.000	5.000	2.271	
Whether to accept price discrimination on the basis of informed	Q8	336	1.000	5.000	2.634	2.523
	Q9	336	1.000	5.000	2.512	
	Q10	336	1.000	5.000	2.664	
	Q11	336	1.000	5.000	2.577	
	Q12	336	1.000	5.000	2.229	

that consumers do not show a tendency to disagree or strongly disagree with the price discrimination if the digital business platform is able to establish a stable relationship of trust with the algorithmic consumer and achieve the consumer's right to informed consent. "The fairness of trading conditions should be judged by relatively objective criteria, such as production costs, supply and demand and other market factors, which allow for a certain range of fluctuation" [1]. "The act of price discrimination itself is reasonable for businesses, and there is no way to prohibit it, which is also a market rule" [2]. When we evaluate the big data discriminatory pricing (BDDP), if we focus on the same price for the same person, this is actually a violation of the Internet's right to operate autonomously. We should not be concerned with the objective outcome of pricing, but the act itself that produces the pricing. Under the premise of protecting trade secrets, if digital business platforms can be as open as possible to consumers about their accurate profiling acts and improve their trust in their platforms. As far as consumers are concerned, they may not consider that price discrimination violates their right to fair exchange.

3.3 Correlation Analysis

Data from this research, which was analyzed by SPSSAU software for correlations, showed some correlations between informed and consent to collection, and between consent to collection and acceptance of the price discrimination. The correlation coefficients and p-values were analyzed as follows.

3.3.1 Informed (Q1, Q2, Q3, Q4) and Consent Collection (Q5, Q6, Q7) Show a Significant Positive Correlation

The correlation coefficient values between Q1, Q2, Q3, Q4 and Q5 were 0.154, 0.224, 0.249 and 0.331 respectively and all showed statistical significance at the 0.01 level, thus indicating that there was significant positive correlation between Q5 and Q1, Q5 and Q2, Q5 and Q3 and Q5 and Q4. This is shown in Table 3.

The correlation coefficient values between Q1, Q2, Q3, Q4 and Q6 were 0.240, 0.270, 0.239 and 0.332 respectively, and all showed statistical significance at the 0.01 level, thus indicating that there was a significant positive correlation between Q6 and Q1, Q6 and Q2, Q6 and Q3, and Q6 and Q4. This is shown in Table 4.

The correlation coefficient values between Q1, Q2, Q3, Q4 and Q7 were 0.264, 0.290, 0.357 and 0.330 respectively and all showed statistical significance at the 0.01

Table 3. Correlation analysis between informed and consent collection1

Pearson Correlation - Detailed Format		
		Q5
Q1	Correlation coefficient	0.154**
	p-value	0.005
Q2	Correlation coefficient	0.224**
	p-value	0.000
Q3	Correlation coefficient	0.249**
	p-value	0.000
Q4	Correlation coefficient	0.331**
	p-value	0.000

Table 4. Correlation analysis between informed and consent collection2

Pearson Correlation - Detailed Format		
		Q6
Q1	Correlation coefficient	0.240**
	p-value	0.000
Q2	Correlation coefficient	0.270**
	p-value	0.000
Q3	Correlation coefficient	0.239**
	p-value	0.000
Q4	Correlation coefficient	0.332**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

Table 5. Correlation analysis between informed and consent collection3

Pearson Correlation - Detailed Format		
		Q7
Q1	Correlation coefficient	0.264**
	p-value	0.000
Q2	Correlation coefficient	0.290**
	p-value	0.000
Q3	Correlation coefficient	0.357**
	p-value	0.000
Q4	Correlation coefficient	0.330**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

level, thus indicating that there was a significant positive correlation between Q7 and Q1, Q7 and Q2, Q7 and Q3, Q7 and Q4. This is shown in Table 5.

3.3.2 Consent Collection (Q5, Q6, Q7) Shows a Significant Positive Correlation with Acceptance of Price Discrimination (Q8, Q9, Q10, Q11, Q12)

The correlation coefficient values between Q5, Q6, Q7 and Q8 were 0.279, 0.237 and 0.264 respectively and all showed statistical significance at the 0.01 level, thus indicating that there was a significant positive correlation between Q8 and Q5, Q8 and Q6, Q8 and Q7. This is shown in Table 6.

The correlation coefficient values between Q5, Q6, Q7 and Q9 were 0.125, 0.147 and 0.276 respectively and showed statistical significance at the 0.05 level, 0.01 level and 0.01 level respectively, thus indicating that there was a significant positive correlation bet-ween Q9 and Q5, Q9 and Q6, Q9 and Q7. This is shown in Table 7.

Table 6. Correlation analysis of consent collection and acceptance of price discrimination1

Pearson Correlation - Detailed Format		
		Q8
Q5	Correlation coefficient	0.279**
	p-value	0.000
Q6	Correlation coefficient	0.237**
	p-value	0.000
Q7	Correlation coefficient	0.264**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

Table 7. Correlation analysis of consent collection and acceptance of price discrimination2

Pearson Correlation - Detailed Format		
		Q9
Q5	Correlation coefficient	0.125*
	p-value	0.022
Q6	Correlation coefficient	0.147**
	p-value	0.007
Q7	Correlation coefficient	0.276**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

Table 8. Correlation analysis of consent collection and acceptance of price discrimination3

Pearson Correlation - Detailed Format		
		Q10
Q5	Correlation coefficient	0.311**
	p-value	0.000
Q6	Correlation coefficient	0.345**
	p-value	0.000
Q7	Correlation coefficient	0.308**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

The correlation coefficient values between Q5, Q6, Q7 and Q10 were 0.311, 0.345 and 0.308 respectively, and showed a statistical significance at the 0.01 level respectively, thus indicating that there was a significant positive correlation between Q10 and Q5, Q10 and Q6, and Q10 and Q7. This is shown in Table 8.

The correlation coefficient values between Q5, Q6, Q7 and Q11 were 0.229, 0.232 and 0.245 respectively, and showed a statistical significance at the 0.01 level respectively, thus indicating that there were significant positive correlations between Q11 and Q5, Q11 and Q6, and Q11 and Q7. This is shown in Table 9.

The correlation coefficient values between Q5, Q6, Q7 and Q12 were 0.218, 0.186 and 0.262 respectively and showed a statistical significance at the 0.01 level respectively, thus indicating that there was a significant positive correlation between Q12 and Q5, Q12 and Q6, Q12 and Q7. This is shown in Table 10.

From the above correlation analysis, we can visualize that in this survey, the more the right to be informed about the collection of their personal information by digital business platforms is guaranteed to the consumer manipulated by algorithms, the higher the level of the right to know to the use of their personal information. Once a sense of trust

Table 9. Correlation analysis of consent collection and acceptance of price discrimination4

Pearson Correlation - Detailed Format		
		Q11
Q5	Correlation coefficient	0.229**
	p-value	0.000
Q6	Correlation coefficient	0.232**
	p-value	0.000
Q7	Correlation coefficient	0.245**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

Table 10. Correlation analysis of consent collection and acceptance of price discrimination5

Pearson Correlation - Detailed Format		
		Q12
Q5	Correlation coefficient	0.218**
	p-value	0.000
Q6	Correlation coefficient	0.186**
	p-value	0.001
Q7	Correlation coefficient	0.262**
	p-value	0.000

* $p < 0.05$ ** $p < 0.01$

has been established between the two parties, the consumer manipulated by algorithms are not completely excluded from price discrimination. The more consumers are aware of the type of personal information being collected, the higher the acceptance of price discrimination. "This is partly because the consumer manipulated by algorithms can significantly reduce the cost of search and transactions and help consumers overcome bias, thus enabling more rational and sophisticated choices" [3].

4 Conclusions and Recommendations

4.1 It is Appropriate for Digital Business Platforms to Disclose the Content of the Information They Collect from Individuals

"Instead of making purchase decisions directly, consumers will delegate these tasks to algorithms, thus minimizing the direct role they play in purchase decisions." [3] This is the consumer's adaptation to the time, but digital business platforms should also implement protection of the consumer's right to know. Through wholesale notification models such as privacy policies or personal information collection checklists, which do

not allow the consumer manipulated by algorithms to visualize when, where, to what extent, with what frequency, and for what purposes their personal information is collected by the platform, it is appropriate for algorithmic controllers to disclose the above acts to the consumers themselves, "following the social responsibility to be fair and equitable, open and transparent, and scientific and reasonable" [4].

4.2 Agree to the Choice of a Diverse, Sub-scene Model

From the correlation of the research data, we can see that the extent of protection of the rights of informed consent by the consumer manipulated by algorithms directly influences their acceptance of price discrimination by digital business platforms. The current model we are using is a one-click consent model. Some scholars have suggested that information processors "should provide different versions of the consent rules to information subjects, setting up 'negotiable' and 'one-click consent' versions respectively" [5], but there needs to be some discussion as to whether the "negotiable" model can maximize the protection of consumers' right to know. The author believes that differentiated authorized consent can be given depending on the different situations in which information is collected. However, this mode of operation has some disadvantages. Consumers are likely to be required to carry out cumbersome authorization consent activities when information collection situations are switched.

4.3 The Game of Consent Authorization vs Price Discrimination

Algorithmic pricing offers the advantage of lower operating costs and higher profits. "High-quality customers with low price sensitivity are often the target of competition between major platforms. Once the data information of such customers is shared synchronously, the customers will be faced with continuous promotional bombardment and deceive acquaintance of price by major platforms" [6]. However, as mentioned in the previous correlation analysis, the consumer manipulated by algorithms are able to accept a certain degree of "price discrimination", assuming that the operator optimizes the services agreed to be authorized. Article 33 in the Constitution of the People's Republic of China also specifies the protection of the right of equality. While the essence of the right of equality allows for reasonable differential treatment. In the future, while operators carry out algorithmic differentiated pricing, they should think about the reasonable scope of that differentiation from a consumer perspective in order to retain regular customers. At the same time, they should develop new customers and finally achieve a dynamic balance between personal information protection and algorithmic operation.

Acknowledgments. This paper was supported by Department of Education of Guangdong Province (Project No. 2022WTSCX154) and Dongguan Social Science Association (Project No. 2022CG82).

References

1. Tian Mengchi, Hou Liyang. The pros and cons of consumers being deceived by algorithms[J]. Procuratorial View,2019(14):16.
2. Wang Han. The price discrimination behind "deceive acquaintance" by big data[N]. Democracy and Law Times, 2018–03–29 (006).
3. Gal M S, Elkin-Koren N. Algorithmic Consumers[J]. Social Science Electronic Publishing, 2017(30):311.
4. Liu Ying. On the relationship between algorithm and legal behavior: Institutional influence and legal response[J]. Journal of Chongqing University (Social Science Edition):2021(11):11.
5. Zhao Jingwei. The normative construction of the rule of informed consent in the protection of personal information[J]. Journal of North Minzu University, 2022(02):148.
6. Zhu Changgen, Chen Yeqi. Research on the regulation of differentiated pricing behavior of e-commerce platforms in the context of big data[J]. Prices Monthly,2021(12):91.