International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 11 Issue 1 ǁ January 2023 ǁ PP. 471-484

Regression analysis of user level based on Netease cloud user data set

Xinyu Wang, Feiran Gao, Haoyao He, Qingke Cai
School of Management, Tianjin University of Technology, Tianjin, China

Abstract
In recent years, the level of users' level well reflects the activity of users, and can help the music platform better classify and recommend users, algorithm implementation and platform operation management. In the new wave of merger and integration, digital music platform has gradually clarified its own positioning and actively formulated differentiated development strategy. Firstly, this study combs the characteristics of the user set of digital music platform. Secondly, through data modeling, taking the user characteristics of Netease cloud music as the starting point, this study defines its main user group as young and highly educated people outside the system, and analyzes the product characteristics of shrimp music. Finally, the researchers put forward three strategic development strategies, including enhancing user identity, paying attention to product emotional design and improving product social attributes, in order to provide user research basis for the development of digital music platform.

Keywords: *Stata analysis*、*Big data*、*Music platform*、*Linear regression*

Date of Submission: 14-01-2023 Date of acceptance: 29-01-2023

I. INTRODUCTION

In the era of big data, various types of digital music platforms are developing rapidly and emerging in an endless stream. The huge social changes brought by each wave of technological innovation have become the fuse to open a new era and awaken new lifestyles. Big data technology provides platform operators with convenient means of user data processing. The huge data hides users' music preferences. Through data calculation and analysis, the digital music platform will more accurately push music information to users, thus improving user utilization and enhancing user stickiness. The Internet provides new business opportunities for the music industry, especially for both independent artists and record companies. [11]In the heyday of network sites (e.g. Weibo, Wechat, and QQ), a music platform called NetEase Cloud Music unprecedentedly dominated in competitive markets in China. The success of NetEase Cloud Music is related to its social function.[12]By analyzing the online music user portrait collection, this study conducts empirical research on various factors affecting digital music users, so as to help the platform formulate targeted industrial management and operation strategies. At the same time, the research has important theoretical significance and value for understanding the user behavior of Internet applications in the new media era.

II. RESEARCH SUMMARY

2.1 The development of digital music

Initial stage: since 1999, a series of free online music platforms have sprung up in China. The early music platform has the functions of online music listening and downloading, which marks the official entry of domestic music into the digital music era. In November 2002, Baidu, which started the search engine, officially released the MP3 search function. Its huge traffic has driven the development of MP3 function. MP3 search subsequently became the second largest search after conventional web search, promoting the development of digital music. However, free music auditions and downloads have low commercial value, do not involve clients, players or payment interfaces, etc., have weak liquidity, and constitute an infringement of intellectual property rights from the legal level, which leads to the emergence of RBTs with higher commercial value on the historical stage. Since 2003, communication operators in Chinese Mainland have successively launched RBT services, such as mobile "RBT", Unicom "Xuanling", Netcom and Telecom "Yueling", to provide users with single song download services. The popularity of RBT has broken the original bundling consumption of the entire album. Consumers can freely choose single music consumption, which greatly reduces the cost of music consumption and greatly reduces the distribution cost of digital music. At the same time, the RBT service of communication operators has changed the traditional mode of singers making money solely by relying on records and channels, and has provided more profit channels. Through cooperation with Internet service providers, singers have launched a single to share RBT revenue, thus realizing the rapid realization of single.

During the period of rapid development: the free sharing music market relies on advertising for profit. Due to technical limitations, RBT music often provides segment music to consumers, and the long tail effect is not obvious, which can not meet the normal demand of the market for music. Therefore, the rise of PC-end music software that can download free whole music gradually replaces the segmented RBT music of operators. After 2003, P2P music sharing websites represented by Kuro, ExPee, Kugou and Kuwo were established successively, rapidly expanding the market scale of digital music. For example, Kugou is a representative P2P music download tool, with an average of more than 5 million independent IPs per day, and up to 680000 online downloads, with about 5 million downloads. At the same time, a large amount of capital has also entered the relevant fields because of the optimistic development prospects of the digital music market. For example, record companies such as Taihe Maitian and Haidie received a large amount of capital injection around 2004. After 2008, with the development of 3G technology, the music market stored in the form of player has risen rapidly, the habit of consumers downloading music has been further consolidated, and advertising has also had a better display carrier. The revenue at this stage is mainly download. The major operators have laid out the music download business in the 3G field, and continue to expand and consolidate the accumulated market through various value-added services. However, due to the imperfection of the legal system and the lack of awareness of intellectual property protection, people are still used to consuming free music for a long time.

Mature and stable period: intellectual copyright gradually improved, and the awareness of payment initially formed. With the gradual improvement of the legal system related to intellectual property, around 2013, a large number of music websites were forced to close because they did not obtain music copyright. Subsequently, the digital music industry began to move towards a commercial model. Some mainstream record companies cooperated with Internet music service providers to explore providing differentiated services for users. At the technical level, achieving technical compatibility between different copyright protection or implementing unified digital rights management technology is a problem that needs to be further solved. Capital flows to the copyright part that is more easily controlled. [1]

Mature and stable period: intellectual copyright gradually improved, and the awareness of payment initially formed. With the gradual improvement of the legal system related to intellectual property, around 2013, a large number of music websites were forced to close because they did not obtain music copyright. Subsequently, the digital music industry began to move towards a commercial model. Some mainstream record companies cooperated with Internet music service providers to explore providing differentiated services for users. At the technical level, achieving technical compatibility between different copyright protection or implementing unified digital rights management technology is a problem that needs to be further solved. Capital flows to the copyright part that is more easily controlled.[1]

2.2 Purpose and significance of the study

2.2.1 Purpose of the study

（1）Analyze the relevant characteristics of domestic online music platform users through data survey

This paper obtains relevant user profiles through the analysis of Netease Cloud Music user data set, and obtains the impact of each feature on user level through multiple linear regression. It aims to help Netease Cloud App in operation and management.

（2）Broaden the relevant theories of mobile music application platform on marketing strategy

This paper analyzes the marketing models and strategies used by the current music platform from multiple perspectives, infuses the connotation system of the new era into relevant marketing theories, and further provides academic and theoretical knowledge for the future research direction of this paper.

(3)Through a case to guide the practice, provide feasible suggestions for the subsequent promotion of mobile music APP

The research on domestic mobile music platform, such as NetEase Cloud Music, found the advantages and disadvantages of NetEase Cloud products themselves, analyzed the demand points and consumption tendencies of platform users, and provided universal and practical suggestions for product function setting and product marketing promotion methods related to other similar platforms.

2.2.2 Significance of the Research

Studying the marketing strategy of mobile music platforms with a serious and cautious attitude will not only make us have a more comprehensive and relatively rational understanding of these similar APPs, but also greatly help the future operation direction and mode of mobile application platforms and the sensory improvement of platform users, that is, users.

This paper is based on the ways and methods adopted by mobile music application companies to improve user perception, and makes a comprehensive and multi-angle survey, analysis and reflection on user data sets of mobile online music platforms in the current market. What we need to know is how to make the marketing strategies adopted by these platforms more effective? What new direction will mobile music apps extend in the

context of a more developed Internet environment in the future?[20]
The main purpose of my research on this topic is to hope to put forward suggestions for the future development of all domestic music platforms of the same type and improve the sense of APP users, and finally guide the practice.

III. DOCUMENT SUMMARY

3.1 Research status at home and abroad

3.1.1 Related research topics of foreign researchers

With the development of faster internet speed andbetter mobile connections, many people are now streaming music through servicessuchasSpotify and Apple Music instead of purchasing the hard-copy music CDs. In fact, a recent report from the Recording Industry Association of America (RIAA) reveals that streaming accounted for 80% of the U.S. music market in 2019, compared with 7% in 2010. 1 According to the same report from RIAA, the number of music streaming subscribers in the United States rose from 1.5 million to around 61 million from 2010 to the first half of 2019.[7]

What is more,the paper "Business Model Comparison Between Spotify and NetEase Cloud Music" also compare the business model between Spotify and NetEase Cloud.With the development of technologies, the importance of music has been increased. Based on one of the researches about music, it pointed out that music somehow has the ability to release our pain, help release our stress or give us confidence. In one sentence, music has a powerful therapeutic effect on the human psyche. In the past of the music market, the control was generally dispersed and hard to be gathered together. However, now the market of music in China is mainly dominating by NetEase Cloud Music while Spotify owns a lot of markets in the world range. The main focus of this article will be illustrating the business model for both companies and put together for comparison to get a conclusion about if the business model works for both countries or not.[9]

Generally speaking, the research on mobile music in foreign countries is earlier than that in China, so the survey and research on the marketing strategy of mobile music APP platform in foreign countries are relatively more comprehensive and rich. From the perspective of the development direction of mobile online music APP product features, a foreign researcher took the non-Chinese mobile music platform "AOL" with social characteristics as an example in the article "AOLHitsMilestones with SocialMusicApp". The article mentioned that music platform products play an important role in the current Internet market, and also emphasized the important role that music products can not be missing in life. For another example, scholars Aharony and Noa, in the article "What's App: social capital perspective", drew a conclusion of "social capital class" by studying what's app, which focuses on making friends. Although it does not study the application of music types, this article puts forward the link between the rapid development of the current network era and the role of the marketing strategy it adopts, Therefore, Aharony and Noa have made a bold guess about many kinds of application apps, including music products - all mobile products will eventually inevitably join in the product marketing socialization.

The paper "A Report on Netease Cloud Music Business" written by Yang Zihan found out that despite the platform giving free music to its customers, other music genres may be difficult to be accessed by customers until they subscribe to premium services. This could be a discouraging factor because other platforms offering the same services could use it as a merit to compete NetEase and take away some of its customers. The paper recommended that the platform should improve the quality of freemium users' music to help receive many users annually. In line with this, the rate of premium users on the platform should be reduced in favor of the customers. [6]

Some foreign researchers have investigated the emergence and development of their own national online music platform, and obtained the general rules and context of the development of the music platform, providing the literature resources and direction for this paper.However, if we want to apply it to our country, we need to take some of the knowledge and theories from them to their essence and discard their dross before they can be used to guide domestic research practice. This is also the academic attitude advocated by the sinicization of Marxist philosophy, which is based on national conditions and specific analysis of specific problems.

3.1.2 Related research topics of domestic researchers

China's music market started and developed relatively late compared with foreign countries, but its development momentum is extremely rapid. Up to now, it has been able to compete with western markets. And in terms of marketing strategy, there are not a few papers in China.
Music sharing and social functions span the boundaries of language and cultural communication between people,and through the development of the Internet,the choice of music has been upgraded from the traditional record player mode to the discovery music mode of the music social platform.The article "The Users Emotional Study of Netease Cloud Music Based on LDA Model" found something.Based on the emotional needs users

in "music social",this study selects Netease cloud music reviews as data set.Through the optimized LDA model,Word2vec word vector model and sentiment dictionary,the topics are quantified from the perspectives of semantics and emotions respectively.In this way,we can explore the distribution characteristics of users' emotions under various music.At the same time,the music platform can conduct targeted emotional marketing and personalized recommendation through the emotional needs of users,which can promote user retention and improve marketing efficiency. [8]

The scholar Lou Junqiao also analyzed the important operational role of socialization in the mobile music platform in the article "Marketing Planning Plan for the Xiami Mobile Music Client", and proposed that the Xiami Music APP is trying to continuously attract some users from a more comprehensive perspective and make them rely on and improve their loyalty, expand the user base of their products, and maximize their benefits.

He Xinru, the author of"Netease Cloud Music Marketing planning-building a platform for music performances and playback", argues that social networking is a unique marketing model for netease Cloud APP, and from this infer netease cloud music of this marketing proposition can lead other similar type of music platform.

With the upgrading of consumer consumption structure, music platforms have been unable to meet consumer needs only from the music attribute, and the industry needs new market growth points. NetEase Cloud Music, as a young force, stands out in the industry with its "music social" positioning. Starting from the emotional marketing of NetEase Cloud Music, this paper explores the influence mechanism of emotional marketing on users from the perspective of social needs and emotional identity. Research shows that comment interaction is positively correlated with users' platform loyalty, and the social attributes of the platform are positively correlated with users' loyalty to the platform. Emotional marketing can arouse emotional resonance among consumers and then generate users' dependence on the platform. Online music platforms can use emotional marketing strategies to improve their market competitiveness in the future. In the field of emotional marketing, this paper supplements the content of the influence mechanism of social needs on user loyalty, and enriches the emotional marketing content of the music platform.[13]Emotional marketing has gradually become a new development direction in the personalized service market, especially in the market of online music platforms. We also get that NetEase Cloud Music has helped its market expansion through emotional influence, and users are gradually increasing.[14]

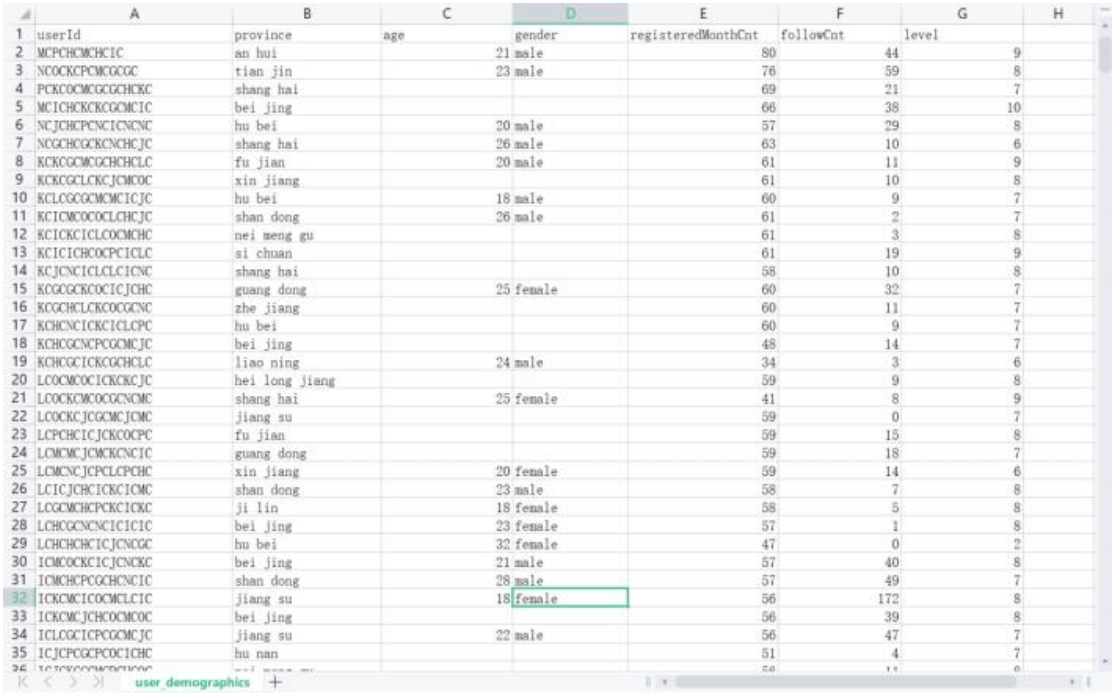
Various industries have begun to merge in the era of new media, and many offline industries have started to carry out brand new marketing by combining online methods. Through extensive data analysis, it will also study whether people attach importance to emotional value.[15]

There is no shortage of research on this type in China. Without exception, these articles all show that proper and correct marketing strategies can have both psychological and physical effects on platform users.Because I was using a domestic application as an example, I was not limited by language and data, so it was relatively easy to get some of my research results, it can also provide me with the research perspective and research method of this paper paradigm.[2]

IV. DATA SAMPLES AND VARIABLES

4.1 Basic characteristics of a dataset

Several important concepts need to be understood before the study begins. The first is the provinces, provinces have an impact on the population density, development level and so on, so it also has an impact on the user class of music platform,According to the P value, the eastern part is regarded as the reference group, and the Central and western part is regarded as the study group. From my perspective,province not only represents the economic development situation, also represents the culture.This exploratory research focuses on analysing the impact of marketing-mix, culture and experience on purchase intentions and purchase decision for online music products in Indonesia. The study draws inference from consumer behaviour theories - theory of reasoned action and theory of planned behaviour as foundation for this research. Culture and experience from these theories are considered as the moderating variables to purchase intention and decision. Marketing-mix or the 7Ps are also added as variables to the research model to identify their individual and collective impact on the intention to purchase and decision to buy. The research deploys structural equation modelling (SEM) using confirmatory factor analysis (CFA) to analyse the impact. The study finds that price, distribution and physical evidence have an impact on purchase intention, while the product and promotion have impact on purchase decision. Culture and experience as moderating variables also augment the consumer purchase intentions for online music product. [17]Next is the registration month time, reflects the user to use the platform time length.Then there is the number of followers, reflecting the number of platform creators that users are following. Let's start with the data set:



	A	B	C	D	E	F	G	H
1	userId	province	age	gender	registeredMonthCnt	followCnt	level	
2	MCPCHCMCHCIC	an hui	21	male	80	44	9	
3	NCOCKCPCMCGCGC	tian jin	23	male	76	59	8	
4	PCKCOCMCGCGCHCKC	shang hai			69	21	7	
5	MCICHCKCKCGCMCIC	bei jing			66	38	10	
6	NCJCHCPCNCICNCNC	hu bei	20	male	57	29	8	
7	NCGCHCGCKCNCHCJC	shang hai	26	male	63	10	6	
8	KCKCGCMCGCHCHCLC	fu jian	20	male	61	11	9	
9	KCKCGCLCKCJCMCOC	xin jiang			61	10	8	
10	KCLCGCGCMCMCICJC	hu bei	18	male	60	9	7	
11	KCICMCOCOCLCHCJC	shan dong	26	male	61	2	7	
12	KCICKCICLCOCMCHC	nei meng gu			61	3	8	
13	KCICICHCOCPCICLC	si chuan			61	19	9	
14	KCJCNCICLCLC1CNC	shang hai			58	10	8	
15	KCGCGCKCOCICJCHC	guang dong	25	female	60	32	7	
16	KCGCHCLCKCOCGCNC	zhe jiang			60	11	7	
17	KCHCNCICKCICLCPC	hu bei			60	9	7	
18	KCHCGCNCPCGCMCJC	bei jing			48	14	7	
19	KCHCGCICKCGCHCLC	liao ning	24	male	34	3	6	
20	LCOCMCOC1CKCKCJC	hei long jiang			59	9	8	
21	LCOCKCMCOCGCNCMC	shang hai	25	female	41	8	9	
22	LCOCKCJCGCMCJCMC	jiang su			59	0	7	
23	LCPCHCICJCKCOCPC	fu jian			59	15	8	
24	LCMCMCJCMCKCNCIC	guang dong			59	18	7	
25	LCMCNCJCPCLCPCHC	xin jiang	20	female	59	14	6	
26	LCICJCHCICKC1CMC	shan dong	23	male	58	7	8	
27	LCGCMCHCPCKC1CKC	ji lin	18	female	58	5	8	
28	LCHCGCNCNCICICIC	bei jing	23	female	57	1	8	
29	LCHCHCHCICJCNCGC	hu bei	32	female	47	0	2	
30	ICMCOCKC1CJCNCKC	bei jing	21	male	57	40	8	
31	ICMCHCPCGCHCNCIC	shan dong	28	male	57	49	7	
32	ICKCMCICOCMCLC1C	jiang su	18	female	56	172	8	
33	ICKCMCJCHCOCMCOC	bei jing			56	39	8	
34	ICLCGC1CPCGCMCJC	jiang su	22	male	56	47	7	
35	ICJCPCGCPCOC1CHC	hu nan			51	4	7	

user demographics

Figure1

Then look at the overall description and the statistical description:

```
Contains data
  obs:     2,085,533
 vars:             7
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
userid          str20   %20s                  userId
province        str14   %14s
age             byte    %8.0g
gender          str7    %9s
registeredmon~t int     %8.0g                 registeredMonthCnt
followcnt       int     %8.0g                 followCnt
level           byte    %8.0g
-------------------------------------------------------------------------------
Sorted by:
     Note: Dataset has changed since last saved.

.
```

Figure2

```
. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
userid	0				
province	0				
age	1,344,042	21.5078	3.756435	11	57
gender	0				
registered~t	2,085,124	23.79156	15.13285	0	608
followcnt	2,085,124	10.63215	26.15071	0	2465
level	2,085,124	5.669522	2.169045	0	10

Figure3

We can see a large amount of data, about 2,000,000 data, but there are many missing values, so in order to avoid the follow-up study bias, to delete the missing values in the record bar.It is one of the main goals of personalized music recommendation system that how to accurately recommend the songs in line with users'

interests in the huge music library. [16]So it is an significant problem that how to use the big data better.

4.1.1.Int variable

1.Age: the distribution was mainly between 15 and 26, with a mean of 21 but more missing values, so data were deleted to prevent impact on follow-up studies

Research Direction: Age has a great impact on whether active, need to focus on

2.RegisteredMonthCnt: approximately distributed between 0-38 months, 38-42 about a sudden drop, there are outliers 600, considering netease cloud on April 23,2013, there will be at most 108 months, therefore, the outlier data is deleted

Research Direction: Does the time of registration month affect user activity

3. followCnt: probably between 0-20, basically in line with the normal distribution

Research Direction: does the number of followers affect user activity

4.Level: level for ordinal variables, 1-10 level, can be a good reflection of user activity

4.1.2.Str Variable quantity

1. Gender: 37 percent male, 27 percent female, for the full data available

Methods: A follow-up study was carried out by dividing gender into 1 for men and 2 for women through the new variable new gender

2.Provinces: in the existing data, a total of 34 values, including Hong Kong and Macao, of course, there are unknown areas that are missing values, so choose to delete the missing data, to prevent the impact of follow-up research

Research methods: the province variable is redefined by west_1、middle_1、east_1, and the reference group is east_1 to prevent collinearity.

Specific analysis: position level as a Y-dependent variable, since the level directly reflects the activity of the user; other factors are X

V. MULTIPLE LINEAR REGRESSION ANALYSIS

5.1.1 Using stata software to have a constant term of multivariate linear regression[19]

```
. reg level age newgender registeredmonthcnt followcnt east_1 west_1 middle_1
```

Source	SS	df	MS	Number of obs	= 1,313,681
				F(7, 1313673)	= 96643.31
Model	1630059.06	7	232865.579	Prob > F	= 0.0000
Residual	3165343.07	1,313,673	2.40953652	R-squared	= 0.3399
				Adj R-squared	= 0.3399
Total	4795402.13	1,313,680	3.65035787	Root MSE	= 1.5523

level	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-0.024	0.000	-64.41	0.000	-0.024	-0.023
newgender	-0.212	0.003	-76.86	0.000	-0.217	-0.206
registeredmonthcnt	0.075	0.000	761.07	0.000	0.075	0.075
followcnt	0.008	0.000	145.03	0.000	0.008	0.008
east_1	-0.001	0.004	-0.18	0.857	-0.009	0.007
west_1	0.009	0.004	2.05	0.040	0.000	0.017
middle_1	-0.053	0.004	-13.00	0.000	-0.061	-0.045
_cons	4.565	0.010	453.33	0.000	4.545	4.585

Figure4

The available data are about 1,300,000, and the F statistic is 96643.31, and the corresponding p value (Prob > F) is 0.0000, which indicates that the regression equation is highly significant.R-squared is about 0.3399, which means that the interpretable variable is about 33.99% ; only east is not significant in the p > Abs (t) column, so the east is taken as the reference group.The constant term P value 0.000 is very significant, so we ignore the constant term and that leads to the estimation deviation, and can not get the consistent estimation.And there is no time series, so there is no need for autocorrelation testing. Here is the final regression:

```
. reg level age newgender registeredmonthcnt followcnt west_1 middle_1
```

Source	SS	df	MS	Number of obs	= 1,313,681
				F(6, 1313674)	> 99999.00
Model	1630058.98	6	271676.496	Prob > F	= 0.0000
Residual	3165343.15	1,313,674	2.40953475	R-squared	= 0.3399
				Adj R-squared	= 0.3399
Total	4795402.13	1,313,680	3.65035787	Root MSE	= 1.5523

level	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-0.024	0.000	-64.41	0.000	-0.024	-0.023
newgender	-0.212	0.003	-76.86	0.000	-0.217	-0.206
registeredmonthcnt	0.075	0.000	761.67	0.000	0.075	0.075
followcnt	0.008	0.000	145.04	0.000	0.008	0.008
west_1	0.009	0.003	2.71	0.007	0.002	0.015
middle_1	-0.052	0.003	-17.08	0.000	-0.058	-0.046
_cons	4.564	0.010	475.55	0.000	4.545	4.583

Figure5

5.1.2In order to guarantee the accuracy, considering and sample test

Use the east as the sub-sample：

```
. reg level age newgender registeredmonthcnt followcnt if east_1
```

Source	SS	df	MS	Number of obs	= 605,836
				F(4, 605831)	= 80084.17
Model	783836.228	4	195959.057	Prob > F	= 0.0000
Residual	1482416.18	605,831	2.44691371	R-squared	= 0.3459
				Adj R-squared	= 0.3459
Total	2266252.41	605,835	3.74070895	Root MSE	= 1.5643

level	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-0.022	0.001	-40.52	0.000	-0.023	-0.021
newgender	-0.224	0.004	-54.64	0.000	-0.232	-0.216
registeredmonthcnt	0.076	0.000	525.36	0.000	0.075	0.076
followcnt	0.008	0.000	96.92	0.000	0.008	0.008
_cons	4.524	0.014	319.33	0.000	4.496	4.551

Figure6

Use the west as the sub-sample：

```
. reg level age newgender registeredmonthcnt followcnt if west_1
```

Source	SS	df	MS	Number of obs	= 279,965
				F(4, 279960)	= 34102.53
Model	315971.083	4	78992.7708	Prob > F	= 0.0000
Residual	648480.193	279,960	2.3163316	R-squared	= 0.3276
				Adj R-squared	= 0.3276
Total	964451.277	279,964	3.44491176	Root MSE	= 1.5219

level	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-0.025	0.001	-32.41	0.000	-0.027	-0.024
newgender	-0.200	0.006	-34.19	0.000	-0.212	-0.189
registeredmonthcnt	0.073	0.000	344.45	0.000	0.072	0.073
followcnt	0.006	0.000	62.81	0.000	0.006	0.006
_cons	4.653	0.020	230.67	0.000	4.613	4.692

Figure7

Based on the above results, it is shown that the return rate of the variables in the east is higher than that of the users in the west. It is proved that the closer to the coast, the higher the return rate of the variables, and the more active the users of the music platform are, but in order to prevent collinearity take the east as the reference group.

VI. TEST

6.1 Heteroscedasticity testing

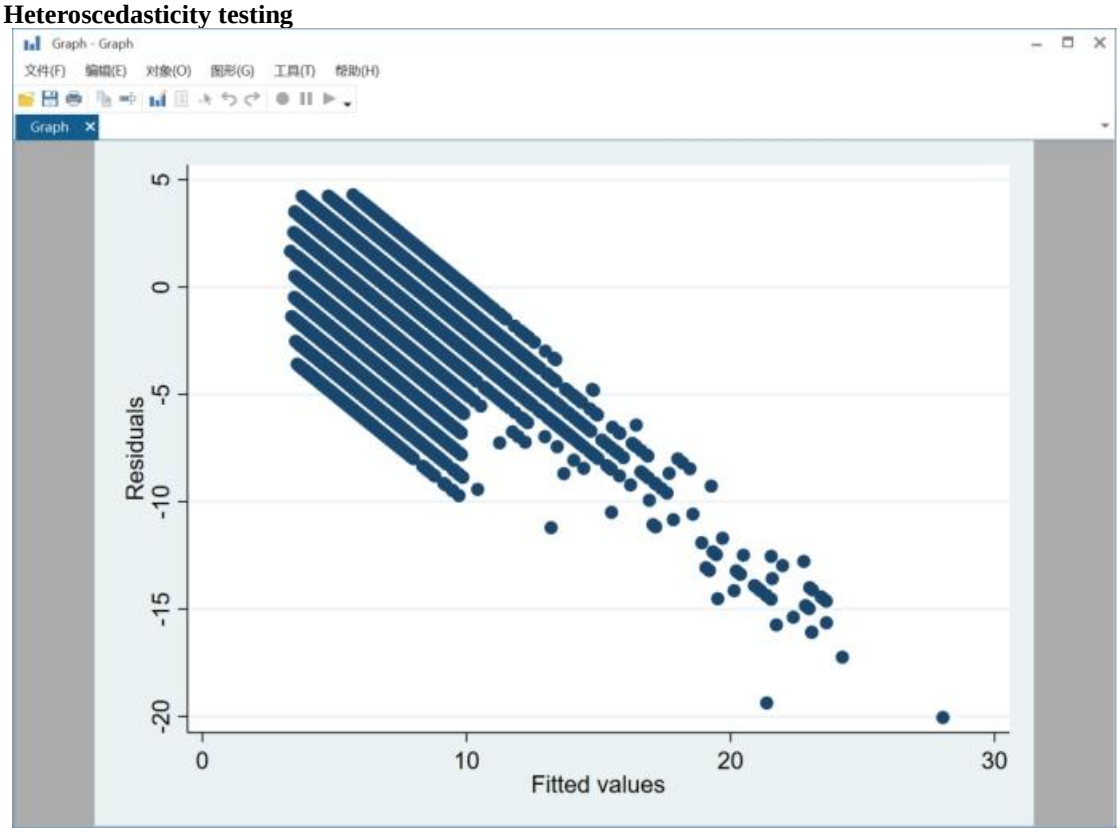


Figure8

According to the previous regression results, the variance of the perturbation term is larger when the rank is larger, and the scatter plot of the number of registered months and residuals need to be further considered.

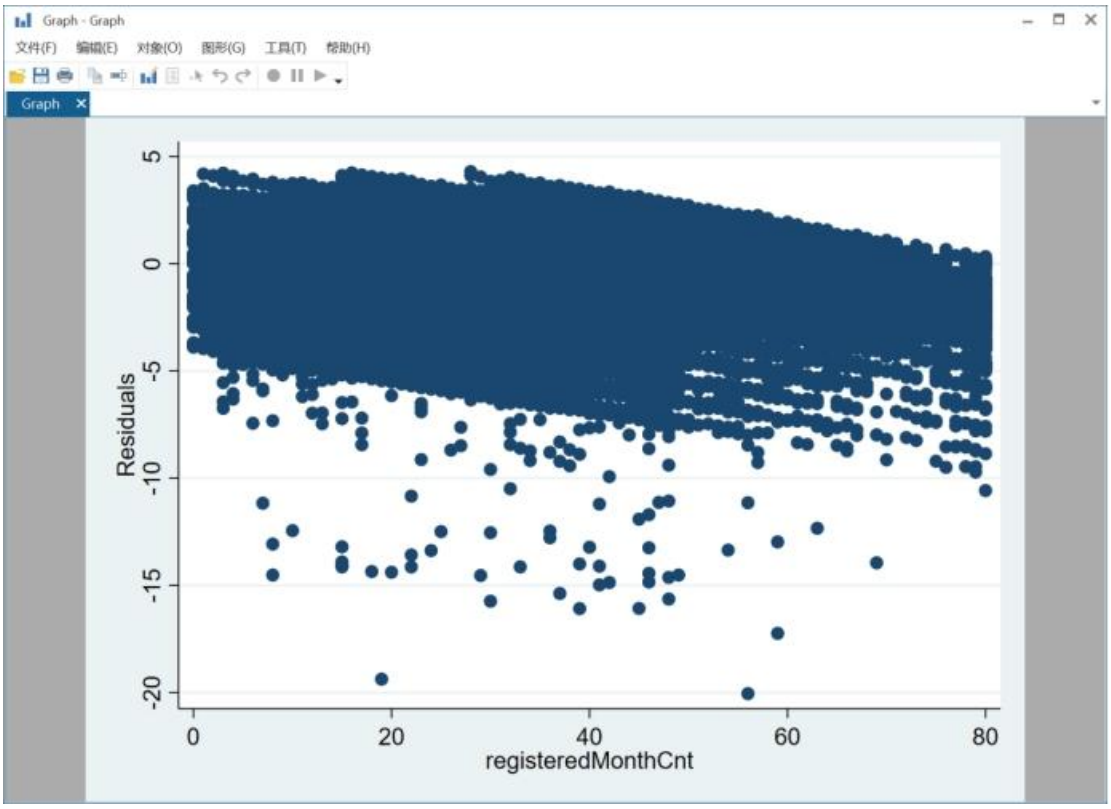


Figure9

The heteroscedasticity may be smaller if the difference between two graphs is larger.

6.2 Standard error test

```
. reg level age newgender registeredmonthcnt followcnt west_1 middle_1,r

Linear regression                               Number of obs     =  1,313,681
                                                F(6, 1313674)     >   99999.00
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3399
                                                Root MSE          =     1.5523
```

level	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-0.024	0.000	-59.72	0.000	-0.024	-0.023
newgender	-0.212	0.003	-76.87	0.000	-0.217	-0.206
registeredmonthcnt	0.075	0.000	606.09	0.000	0.075	0.075
followcnt	0.008	0.000	31.06	0.000	0.008	0.009
west_1	0.009	0.003	2.71	0.007	0.002	0.015
middle_1	-0.052	0.003	-17.20	0.000	-0.058	-0.046
_cons	4.564	0.010	448.08	0.000	4.544	4.584

Figure10

From the graph above, the regression coefficients of the robust standard error WLS are the same whether or not the robust standard error WLS is applied, only the standard error of the constant term is slightly different.

6.3 RESET Test

```
. estat ovtest,rhs
(note:  newgender dropped because of collinearity)
(note:  newgender^3 dropped because of collinearity)
(note:  newgender^4 dropped because of collinearity)

Ramsey RESET test using powers of the independent variables
       Ho:  model has no omitted variables
            F(10, 1313665) =   8650.46
                  Prob > F =      0.0000
```

Figure11

As shown in the table above, the P value is 0.0000, rejecting the original assumption that the higher-order term is missing. This paper only tries multiple linear regression, there are still many deficiencies. There are many binary variables in the variables, the next will try Logistics regression experiments, observe the results.

6.4 Correlation test

```
. pwcorr level age newgender registeredmonthcnt followcnt west_1 middle_1

             |    level      age newgen~r regist~t follow~t   west_1 middle_1
-------------+------------------------------------------------------------------
       level |   1.0000
         age |   0.0136   1.0000
   newgender |  -0.0958  -0.1185   1.0000
registered~t |   0.5690   0.1091  -0.0819   1.0000
   followcnt |   0.1975  -0.0714  -0.0091   0.1609   1.0000
      west_1 |   0.0015  -0.0019  -0.0025  -0.0096   0.0403   1.0000
    middle_1 |  -0.0175  -0.0627   0.0087  -0.0131  -0.0043  -0.0375   1.0000

.
```

Figure12

The order of correlation size was as follows: number of registered months > number of followers > age > whether it was west or not. The lower the level of users in the middle, the lower the level of boys, with a decrease of 0.0958 per level

VII. LOGISTICS REGRESSION

7.1 Variable handling

Since logistics regression requires variables to be discrete, consider dividing age, number of registration months, and number of followers into discrete variables by quartile.
0-25% for 0,25% -50% for 1,50% -75% for 2,75% -100% for 3, and then logistics regression was performed.

```
gen newregisteredmonthcnt = 0
replace newregisteredmonthcnt=0 if registeredmonthcnt<13
replace newregisteredmonthcnt=1 if registeredmonthcnt>=13&registeredmonthcnt<23
replace newregisteredmonthcnt=2 if registeredmonthcnt>=23&registeredmonthcnt<34
replace newregisteredmonthcnt=3 if registeredmonthcnt>=34
gen newfollowcnt=0
replace newfollowcnt=0 if followcnt<3
replace newfollowcnt=1 if followcnt>=3&followcnt<5
replace newfollowcnt=2 if followcnt>=5&followcnt<10
replace newfollowcnt=3 if followcnt>=10
gen newage=0
replace newage=0 if age<19
replace newage=1 if age>=19&age<22
replace newage=2 if age>=22&age<24
replace newage=3 if age>=24
```

Figure13

7.2 logistics Regression

```
. logit level newage newgender newregisteredmonthcnt newfollowcnt west_1 middle_1,r

Iteration 0:   log pseudolikelihood = -271483.96
Iteration 1:   log pseudolikelihood = -235079.82
Iteration 2:   log pseudolikelihood = -206430.47
Iteration 3:   log pseudolikelihood =  -204903.1
Iteration 4:   log pseudolikelihood = -204871.26
Iteration 5:   log pseudolikelihood =  -204871.2
Iteration 6:   log pseudolikelihood =  -204871.2

Logistic regression                             Number of obs     =  2,052,424
                                                Wald chi2(6)      =   76529.76
                                                Prob > chi2       =     0.0000
Log pseudolikelihood =  -204871.2               Pseudo R2         =     0.2454
```

level	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
newage	-0.406	0.007	-62.25	0.000	-0.419	-0.393
newgender	0.495	0.008	63.48	0.000	0.480	0.510
newregisteredmonthcnt	1.437	0.008	174.71	0.000	1.421	1.453
newfollowcnt	0.673	0.005	130.64	0.000	0.663	0.683
west_1	-0.020	0.011	-1.86	0.063	-0.040	0.001
middle_1	-0.013	0.010	-1.29	0.197	-0.032	0.007
_cons	2.230	0.021	105.33	0.000	2.188	2.271

Figure14

Similarly, in order to prevent collinearity, the eastern part was taken as the reference group and the regression was performed. The results showed that the sample passed the Wald chi-square test and the result was 76529.76, with a corresponding p value (Prob > F) of 0.0000, it shows that the regression equation is highly significant.Pseudo R 2 was about 0.2454, indicating that about 24.54% of the variables could be interpreted, with p > middle being less significant in the ABS (t) column.The constant term P value 0.000 is very significant, so we ignore the constant term and that leads to the estimation deviation, and can not get the consistent estimation.

VIII. REFRESSION ANALYSIS

8.1 Variables and models

Interpreted variable: user level（level）

Explain variables：age、gender、province、registeredMonthCnt、followCnt

Multivariate regression statistical model：

$$\mathbf{level} = \beta_1 + \beta_2 \bullet \mathbf{age} + \beta_3 \bullet \mathbf{newgender} + \beta_4 \bullet \mathbf{province} + \beta_5 \bullet \mathbf{registered\,month}\mathit{Cnt} + \beta_6 \bullet \mathbf{follow}\mathit{Cnt} + \varepsilon$$

8.2 Conclusion

```
. reg level age newgender registeredmonthcnt followcnt east_1 west_1 middle_1
```

Source	SS	df	MS		
Model	1630059.06	7	232865.579	Number of obs	= 1,313,681
Residual	3165343.07	1,313,673	2.40953652	F(7, 1313673)	= 96643.31
				Prob > F	= 0.0000
				R-squared	= 0.3399
				Adj R-squared	= 0.3399
Total	4795402.13	1,313,680	3.65035787	Root MSE	= 1.5523

level	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age	-0.024	0.000	-64.41	0.000	-0.024	-0.023
newgender	-0.212	0.003	-76.86	0.000	-0.217	-0.206
registeredmonthcnt	0.075	0.000	761.07	0.000	0.075	0.075
followcnt	0.008	0.000	145.03	0.000	0.008	0.008
east_1	-0.001	0.004	-0.18	0.857	-0.009	0.007
west_1	0.009	0.004	2.05	0.040	0.000	0.017
middle_1	-0.053	0.004	-13.00	0.000	-0.061	-0.045
_cons	4.565	0.010	453.33	0.000	4.545	4.585

Figure15

Through two regression results can be seen, multiple linear regression is suitable for the processing of the data set, but does not rule out I will be treated as continuous variables for away

The results showed that the influence of each factor on the rank was in the following order: number of registered months > number of followers > age > whether it was west or not. And users in the middle of the lower level, the lower the male level, the older the lower level, each year of age decreased by 0.024 grade.The regression coefficients were -0.024 for age, -0.212 for male, 0.075 for number of registered months, 0.008 for number of followers, and 0.009 for Western.The regression results are credible. The number of months of registration, the number of attention, whether the west has a positive impact on the level of users.Age, gender, whether the Middle East has a negative effect on user rank.

Some foreign scholars said that "Major labels are no longer superior to minor labels in survival of streaming services." in the paper "A survival analysis of songs on digital music platform"[10]The integration of digital music services and social networking has become a major trend, in the digital music services platform, and this trend will become more and more obvious in the future, and become the main capital of platform competition.Predictably, the future of the development of digital music service will take on characteristics as follows: (1) offline and online mode of social integration.Through social networks, digital music services platforms will gradually evolve into virtual social networks of shared interests and purposes, with the popularity of virtual communities, the virtualization of entity organization relationship and the localization of virtual relationship on the Internet make the tradition.There is an overlap of"Online" and"Offline" in the interpersonal network, resulting in a complex social network environment.This integration will create more opportunities for online singers to earn more income. At present, some online musicians, after successfully utilizing the music form and becoming influential, break away from the original service platform and create their own off-line product and product marketing channels, this is actually bad for digital music service providers.And when the service platform is able to integrate this set of processes, it can either generate revenue, or it can use the expanded channels to organize some offline activities and train new people.In the future, the development of live broadcasting technology will lead to more and more online and offline events. Offline events can be disseminated through online live broadcasting, while the original online events, such as the purchase of products, etc. can also use the convenient transaction and delivery system to quickly achieve the line, under delivery.(2) based on the new interactive device social experience, according to Moore's law, in the coming period of time, the computer's computing power will continue to improve.At the same time, computing equipment will continue to shrink in size. With the development of VR and AR technology, the consumption of digital music products will give users more and more strong generation.Constantly, in the future, appreciate the music of more and more shock to the user experience of music.And the development of Internet technology will also greatly promote this process. Today's users don't have to carry around a lot of devices to get that information.At present, in most cases, intelligent terminals can meet people's daily entertainment needs. I believe that in the future, there will be some hearing, visual integration of the device, through more developed Internet technology, for us to create more ways of music appreciation.(3) the deep integration of social and

life.In the future, the development of computing equipment and network technology will make life and the network more closely connected.Topic of music features make it have better points, on this basis, the exchange will breakthrough the limitation of cultural background.The social network provided by the digital music service platform, in line with online and offline interactions, will not only affect the process, of online communication, but will also make this kind of social relationship directly enter people's daily lives and form a life, live can come into contact with the same group, like comic-con, auto show and other existing same-sex activities, like the same.Concerts and music festivals will no longer be high above, but rather a form of daily interaction in which people can take part, and many star singers will no longer be"Superheroes" in captivity in the entertainment industry, but in our daily life can often contact, exchange friends.With the development of Internet technology, digital music service presents a trend of standardization, socialization and routine.Over the past few years, we've seen digital music services change, from the initial phase of unfettered, unrestricted sharing through piracy, to content recommendation by the Internet industry, and eventually into syndicates, content promotion network in the form of.In this series of changes, we can observe, digital music services from disorder to order, from unconscious to directional evolution process. I believe that in the future, digital music services will become more universal and human-oriented, greatly, to enrich our lives[5]

(4) through the combination of big data and recommendation algorithms, digital music platforms will unleash greater energy, deliver more accurate and more three-dimensional user portraits, and promote the development of digital music in China in the future, will be closely integrated with the Internet, blockchain, big data.

REFERENCES

[1]. Yang Zihan.(2022).A Report on Netease Cloud Music Business. Tourism Management and Technology Economy(1). doi:10.23977/TMTE.2022.050106.

[2]. Zhang Dennis J.,Hu Ming,Liu Xiaofei,Wu Yuxiang & Li Yong.(2020).NetEase Cloud Music Data. Manufacturing & Service Operations Management(1). doi:10.1287/msom.2020.0923.

[3]. Hong Zhang,Jiayi Nie & Zenan Ruan.(2019).The Users Emotional Study of Netease Cloud Music Based on LDA Model..(eds.)Proceedings of 2019 4th International Conference on Cloud Computing and Internet of Things(2019 CCIOT)(pp.32-35).

[4]. Z. Liu, R. Gao, et al. Two-period pricing and strategy choice for a supply chain with dual uncertain information under different profit risk levels. Computers & Industrial Engineering, 2019, 136: 173-186.

[5]. Li Yonghao,Zhang Qirui,Li Anna Rita & Cheng Jun.(2022).Business Model Comparison Between Spotify and NetEase Cloud Music..(eds.)Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)(pp.).Atlantis Press.

[6]. Hyunsuk Im,Haeyeop Song & Jaemin Jung.(2018).A survival analysis of songs on digital music platform. Telematics and Informatics(6). doi:10.1016/j.tele.2018.04.013.

[7]. Puerta M.P..(2017).Colombian artists and digital music platforms: Some difficulties [Los artistas colombianos y las plataformas de música digitales: Algunas dificultades]. Revista de Derecho Privado(33). doi:10.18601/01234366.1133.05.

[8]. C. Zhou, M. Leng, Z. Liu, et al. The impact of recommender systems and pricing strategies on brand competition and consumer search. Electronic Commerce Research and Applications, 2022, 53: 1-15.

[9]. Zhu Qinxuan.(2022).Negative Impact of Platform Framework on User Experience -- A Case Study of NetEase Cloud Music..(eds.)Proceedings of the 2022 5th International Conference on Humanities Education and Social Sciences (ICHESS 2022)(pp.).Atlantis Press.

[10]. C. Zhou, N. Ma, X. Cui, Z. Liu. The impact of online referral on brand market strategies with consumer search and spillover effect. Soft Computing, 2020, 24(4): 2551-2565.

[11]. Wang Mengyu.(2022).Analysis of Online Music Platform based on Emotional Marketing Theory: Taking NetEase Cloud Music as an Example..(eds.)Proceedings of the 2022 4th International Conference on Economic Management and Cultural Industry (ICEMCI 2022)(pp.).Atlantis Press.

[12]. C. Zhou, G. Xu, Z. Liu. Incentive contract design for internet referral services: Cost per click vs cost per sale. Kybernetes, 2020, 49(2): 601-626.

[13]. Liu Kaijie.(2022).The Usage of Emotional Marketing in the Online Music Industry – A Case Analysis of NetEase Cloud Music..(eds.)Proceedings of the 2022 4th International Conference on Economic Management and Cultural Industry (ICEMCI 2022)(pp.).Atlantis Press.

[14]. Teng Ge.(2022).Research on Brand Building and Communication Strategy of New Media APP:Take NetEase Cloud Music as an Example..(eds.)Proceedings of the 2022 7th International Conference on Social Sciences and Economic Development (ICSSED 2022)(pp.).Atlantis Press.

[15]. J. Yu, J. Zhao, et al. Strategic business mode choices for e-commerce platforms under brand competition. Journal of Theoretical and Applied Electronic Commerce Research, 2022, 17(4): 1769-1790.

[16]. M. Chu. The impact of online referral services on cooperation modes between brander and platform. Journal of Industrial and Management Optimization, 2022, DOI: 10.3934/jimo.2022174.

[17]. Gao Hong.(2022).Automatic Recommendation of Online Music Tracks Based on Deep Learning. Mathematical Problems in Engineering. doi:10.1155/2022/5936156.

[18]. C. Li, M. Chu, et al. Is it always advantageous to add-on item recommendation service with a contingent free shipping policy in platform retailing? Electronic Commerce Research and Applications, 2019, 37: 1-11.

[19]. Reni Diah Kusumawati, Teddy Oswari, Tristyanti Yusnitasari, Saurabh Mittal & Vikas Kumar.(2021).Impact of marketing-mix, culture and experience as moderator to purchase intention and purchase decision for online music product in Indonesia. International Journal of Business Innovation and Research(4). doi:10.1504/IJBIR.2021.117089.

[20]. Z. Liu, et al. Impact of cost uncertainty on supply chain competition under different confidence levels. International Transactions in Operational Research, 2021, 28(3): 1465-1504.

[21]. C. Zhou, W. Tang, R. Zhao. Optimal consumer search with prospect utility in hybrid uncertain environment. Journal of Uncertainty Analysis and Applications, 2015, 3(6): 1-20.

[22]. Perdana Adityo Cakra,Rachman Rais & Palinggi Monika Datu Mirring.(2021).An analysis of population using Multiple Linear Regression Analysis. IOP Conference Series: Materials Science and Engineering(1). doi:10.1088/1757-899X/1088/1/012091.
[23]. Cerulli Giovanni.(2022).Machine learning using Stata/Python. The Stata Journal(4). doi:10.1177/1536867X221140944.
[24]. C. Zhou, W. Tang, R. Zhao. Optimal consumption with reference-dependent preferences in on-the-job search and savings. Journal of Industrial and Management Optimization, 2017, 13(1): 503-527.