Journal of Applied Finance & Banking, vol. 7, no. 3, 2017, 121-135
ISSN: 1792-6580 (print version), 1792-6599 (online)
Scienpress Ltd, 2017

Rationality of Investors in P2P Online Lending Platform with Guarantee Mechanism: Evidence in China

Nanfei Zhang[1]

Abstract

This paper investigates whether investors in P2P online lending platforms in China are rational. In China, most P2P platforms run a guarantee mechanism using loan loss provision. I take into account the effect of the guarantee mechanism on loan's cash flow and calculate expected internal rate of return of each loan. The empirical results show evidence against rationality assumption. Firstly, expected return calculated under guarantee mechanism of a loan in China is not only affected by systematic risk, but also by idiosyncratic risk. Secondly, China P2P investors do not maximize their expected return. They take into account other variables although their influence on default and prepayment risk is already reflected in the expected return. Conclusively, China P2P investors are not rational. The guarantee mechanism might contribute to some of the findings.

JEL classification numbers: G110 G120 G140 G170 G210
Keywords: Rational expectation, Expected return, CAPM, Guarantee mechanism

1 Introduction

Whether investors have rational expectation is an important topic in finance research. Rational expectation implies investors' rationality, in which case they are able to value securities in a rational way. The intrinsic value of a security is the sum of the discounted cashflows it generates in each period. Fama proposes the famous efficient market hypothesis, and investors' rationality was one of its core assumptions [1]. The capital asset pricing model (CAPM) is also based upon the assumption that investors are rational. A major branch of empirical finance researches pursue to justify or refute the rationality assumption. This paper also aims to analyze if investors are rational, using an exclusive dataset, a major peer-to-peer (P2P) online lending platform in China, an emerging market, while standard approach utilizes data from conventional securities in a developed market such as stock market in the United States. This paper provides evidence that investors in China P2P platform do not form rational expectation.

Risk of an asset is divided into two parts, systematic risk and idiosyncratic risk. CAPM implies that only systematic risk is priced. In other words, expected return is only

[1]School of Economics and Management, Tsinghua University.

Article Info: *Received :* February 2, 2017. *Revised :* March 6, 2017.
Published online : May 1, 2017

determined by systematic factors. Therefore, if it is found that the expected return of an asset is not only affected by individual factors other than systematic ones, then investors do not form rational expectation.

Rational expectation assumption is the core debate between traditional finance and behavioral finance. Numerous researches test whether investors form rational expectation. Jensen applies CAPM to calculate portfolio's risk adjusted return for the first time [2]. Fama finds that asset price follows a random walk, and market is weakly efficient [3]. Furthermore, Fama et al. uses event study to show that US stock market is in the form of semi-strong efficiency [4]. Fama firstly proposes efficient market hypothesis[1]. This hypothesis brings about a complete new branch of empirical finance. However, some researchers do not believe investors are rational, and this opens up a whole new field in finance, namely behavioral finance. In behavioral finance, investors are not rational, and the irrationality is systematic in the sense that it has long term effect on asset prices. Shleifer reviews the theories of behavioral finance comprehensively [5].

Most researches make use of stock market to test if investors are rational. In late 2000s, a new form of market emerged. Online lending platform, usually known as peer-to-peer (P2P), injects fresh blood into rationality studies. Unlike stock markets, where the information investors observe cannot be controlled, in P2P, all information investors can see is listed on the website, consisting of clearly divided information on loan characteristics (amount, interest rate, term, etc.), systematic risk relevant characteristics (credit rating of borrowers, etc.) and idiosyncratic risk relevant characteristics (sex, age, etc.)

Researches in P2P mainly focus on the impact of information provided with the loan on investors' decision. Michels finds that larger number of pieces of voluntarily disclosed information leads to significantly lower borrowing cost [6]. Barasinska studies the role of information of borrowers' sex [7]. Other papers study whether investors are smart in the process of lending. Freedman and Jin finds that investors systematically underestimates credit risk of borrowers, but can learn [8]. Duarte et al. find that the average expected return in US P2P market is negative, yet investors still invest in it [9]. Iyer et al., on the contrary, show that investors' decision is even more accurate than credit score in identifying borrowers' risk [10]. Among the aforementioned literature, Duarte et al. tests if investors form rational expectation in US P2P market and finds that they do not[9].

However, P2P in China tremendously differs from P2P in the US. In the US, P2P platform serves only as information intermediary. The platform is not responsible for investors' loss incurred by borrowers' default. Oppositely, in China, P2P platform plays the dual role of information and credit intermediary. The platform supplies loan information and simultaneously promises a full principal refund to investors who encounter a default, namely a guarantee mechanism.

This paper calculates the expected internal rate of return(EIRR) of each loan and finds that in China investors do not form a rational expectation, just like the case in the US as shown by Duarte et al. [9]. In a behavioral finance perspective, the loan EIRR is affected by systematic risk as well as borrowers' individual risk, which means the loans are not properly priced. On the other hand, in a P2P research perspective, no paper has studied the China-specific guarantee mechanism's impact on investors' rationality and the EIRR both at loan level and market level. Empirical results show that all loans in China P2P platform have a positive EIRR, opposite to the case in the US. Furthermore, investors in China do not maximize their expected return. All of the above results provide strong evidence against rational expectation assumption.

The structure of this paper is as follows. Section 2 introduces the model. Section 3 is data and variable description. Section 4 is empirical results and discussion. Section 5 concludes.

2 Preliminary Notes

For each loan, there is one nominal interest rate specified in the contract. However, nominal interest rate is neither the expected nor realized rate of return, both of which reflect more importance. What investors place more weights on should be the nominal interest rate adjusted by risks inherent in loans. Two risks exist in lending process, default risk and prepayment risk. Both affect the contingent cash flow generated by the loan in each period. An intuitive illustration is given below. Assume there are two loans both having only one balloon payment. Loan A has a nominal interest rate of 10% and a default probability of 0, while loan B's interest rate is 100% but is sure to default. Under the circumstance of no insurance or guarantee, it is easy to see that loan A's expected interest rate is 10% and loan B's is -100%, since its borrower will surely default. So a rational investor should choose loan A. Nominal interest rate barely means anything, only providing a benchmark.

The above case is too simple. We need to specify one return measure that can cover more general cases. Internal rate of return(IRR) is a reasonable choice. IRR is usually implemented in calculating the net present value of cash flow generating projects. It is the discount rate that equates a project's net present value to zero. A loan is essentially a fixed income asset or project that has a fixed cash flow payment in each period, justifying the usage of IRR to symbolize a loan's performance.

However, the cash flow of a loan is random. In each period, the loan might end up defaulted or prepaid. Therefore before calculation of IRR, the probability of default and prepayment in each period should be estimated, under which expected IRR is calculated. The following introduces the estimation model.

2.1 Expected Internal Rate of Return Model

In this section, I introduce a model calculating risk-adjusted expected IRR in detail. The loan can terminate any time due to different risks as described above, from initiation to maturity. Each case is one realization of a path. I estimate for every loan the probability of every possible path the loan can take. Freedman and Jin and Duarte et al. use different models to calculate IRR [8][11][12][9].

Termination of a loan before maturity might be due to two reasons, prepayment and default. Prepayment refers to the case the borrower pays back the remaining principal and interests before maturity date once and for all, while default the case that borrower fails to timely pay the obligatory interests and ceases to pay anything afterwards.

In the P2P platform we choose, Renrendai, payment guarantee or insurance mechanism is present, typical in China P2P market. In the case of prepayment, borrower pays off the remaining principal as well as an additional 1% of remaining principal as penalty. In the case of default, the cash flow in that period is zero. But the platform guarantees to pay back the remaining principal one month period(month) later to the investor using the loan loss provision. This loan loss provision is raised from borrowers. Each borrower is charged an administrative fee every month.

I define P_{tp} to be probability that a loan is prepaid in month t conditional on the loan not terminating before month t, P_{td} probability that the borrower defaults in month t conditional on the loan not terminating before month t, and P_{tc} probability otherwise conditional on the loan not terminating before month t. It is easily seen that:

$$P_{tp} + P_{td} + P_{tc} = 1$$

A loan can terminate in every month before maturity T because of default or prepayment, totaling $2(T-1)$ paths. In the last month, or month T, it will either terminate naturally or default, totaling 2 paths. Therefore, there are $2(T-1)+2 = 2T$ possible paths for a loan with maturity T.

For each path, we will use the conditional probability to recursively calculate the unconditional probability. I denote the unconditional probability for a loan to terminate in month t (t<T) $Q_{tk}(k = p, d)$, where p stands for prepayment and d default. Using probability theory,

$$Q_{tk} = (\prod_{u=1}^{t} P_{uc}) P_{tk}, (t = 1,2,...,T-1)$$

In the final month,

$$Q_{Tk} = (\prod_{u=1}^{T} P_{uc}) P_{Tk}$$

Each path corresponds to a net present value of loan at month 0. There are three scenarios:

a. Prepayment in month t (t<T).

$$NPV_{tp} = -AMOUNT + \sum_{u=1}^{t} \frac{PAYMENT}{(1+r)^u} + \frac{REMAININGPRINCIPAL * 1.01}{(1+r)^t}$$

The last term is the penalty of prepayment.

b. Default in month t.

$$NPV_{tp} = -AMOUNT + \sum_{u=1}^{t-1} \frac{PAYMENT}{(1+r)^u} + \frac{REMAININGPRINCIPAL}{(1+r)^{t+1}}$$

In month t, there is no cash flow, whereas Renrendai will refund the remaining principal in month t+1.

c. Natural termination in month T.

$$NPV_{Tc} = -AMOUNT + \sum_{u=1}^{T} \frac{PAYMENT}{(1+r)^u}$$

Combined with Q_{tk},

$$E(NPV) = \sum_{t,k} Q_{tk} NPV_{tk}$$

Equating *E(NPV)* zero, *r* can be solved numerically. Annualizing r by multiplying *r* by 12, I get the expected IRR. I argue expected IRR should be the core rate of return a rational and sophisticated investor pays attention to, since it already incorporate all relevant systematic and unsystematic information in the process of calibrating the probabilities, as will be covered later.

2.2 Probability Estimation Model

In this section, I describe the calibration of the conditional probabilities. It is trivial to calculate the unconditional probabilities given conditional ones. More specifically, I estimate the conditional probability for each of three cases (c=censored, p=prepayment, d=default) in month *t* given the loan survives to month *t*. Multinomial logistic regression is applied in the following form:

$$\ln(\frac{P_{itp}}{P_{itc}}) = \alpha_{11} + \alpha_{12} loancharacteristics_i + \alpha_{13} progress_{it} + \alpha_{14} loancharacteristics_i^{et} \cdot progress_{it} + \varepsilon_{1it}$$

$$\ln(\frac{P_{itd}}{P_{itc}}) = \alpha_{21} + \alpha_{22} loancharacteristics_i + \alpha_{23} progress_{it} + \alpha_{24} loancharacteristics_i^{et} \cdot progress_{it} + \varepsilon_{2it}$$

The dependent variables are a pair of odds ratio, which are ratios of conditional probability of prepayment or default in month *t* to a benchmark conditional probability that loan *i* is censored or normally progresses in month *t*. All loan characteristics available are included in the model. I also add a progress variable that is month *t* divided by term *T* to account for the possible time varying model structure, and interaction terms between progress and loan characteristics.

I run the regression to get all the coefficients and plug them back in the model to generate all conditional probabilities:

$$\ln(\frac{\hat{P}_{itp}}{\hat{P}_{itc}}) = \hat{\alpha}_{11} + \hat{\alpha}_{12} loancharacteristics_i + \hat{\alpha}_{13} progress_{it} + \hat{\alpha}_{14} loancharacteristics_i \cdot progress_{it} = \hat{A}_{itp}$$

$$\ln(\frac{\hat{P}_{itd}}{\hat{P}_{itc}}) = \hat{\alpha}_{21} + \hat{\alpha}_{22} loancharacteristics_i + \hat{\alpha}_{23} progress_{it} + \hat{\alpha}_{24} loancharacteristics_i^{et} \cdot progress_{it} = \hat{A}_{itd}$$

And using simple math, I have

$$\hat{P}_{itp} = \frac{e^{\hat{A}_{itp}}}{e^{\hat{A}_{itp}} + e^{\hat{A}_{itd}} + 1}$$

$$\hat{P}_{itd} = \frac{e^{\hat{A}_{itd}}}{e^{\hat{A}_{itp}} + e^{\hat{A}_{itd}} + 1}$$

$$\hat{P}_{itc} = \frac{1}{e^{\hat{A}_{itp}} + e^{\hat{A}_{itd}} + 1}$$

2.3 Data and Variables Description

2.3.1 Data description

This paper's data comes from a large P2P platform, Renrendai, in China. Renrendai was established in 2010. The sample period starts in January 2011 and ends in February 2015. There are 233679 loans in the sample. I use all the sample to calibrate the multinomial logit model and estimate the expected IRR for each loan. However, only data before August 2012 are used to study whether investors are rational, since in August 2012 Renrendai launched a manual loan screening mechanism. The platform invalidates some loans based on some unknown criteria. As a result, a certain amount of loans contained in the data are not really seen by investors. To study whether investors are rational when

making investing decisions, I have to make sure that all data I use are actually seen by investors.
In the 233679 loans, 212000 were not fully bid and 21679 loans were fully bid. Some of the fully bid loans were invalidated by the platform, leaving 16063 loans for model calibration. As for the rational expectation analysis part, we use all loans, fully bid or not, from January 2011 to August 2012, totaling 32168 loans.

2.3.2 Variable description

Core dependent variable:
FundingSuccess: 1 if a borrower gets the money and 0 otherwise.

Independent Variables:
Loan specific characteristics:
Amount (in 1000s): Amount of the loan divided by 1000.
Interest: Annualized nominal interest.
Term: Maturity of a loan in months.
Borrower-specific characteristics:
Gender: 1 if male and 0 if female.
Age: Age of borrower.
IsVoc: 1 if borrower has a vocational college degree and 0 otherwise.
IsUnGrad: 1 if borrower has a bachelor's degree and 0 otherwise.
IsPoGrad: 1 if borrower has a master or PhD degree and 0 otherwise.
IsDivorced: 1 if borrower is divorced and 0 otherwise.
IsMarried: 1 if borrower is married and 0 otherwise.
Line of work relevant dummies: A group of dummies that are 1 if borrower is in a corresponding line of work and 0 otherwise. Line of work includes catering, real estate, public affairs, charity, construction, transportation, education, finance or law, retail, media, energy, agriculture, sport or art, medical, entertainment, government, and manufacture.
WorkExp: Years of working of borrower. 1 if no more than 1 year. 2 if more than 1 year and no more than 3 years. 3 if more than 3 years and no more than 5 years. 4 if more than 5 years.
Income: 1 if borrower's income is lower than RMB 1000. 2 if between RMB 1000 and 2000. 3 if between RMB 2000 and 5000. 4 if between RMB 5000 and 10000. 5 if between RMB 10000 and 20000. 6 if between RMB 20000 and 50000. 7 if higher than 50000.
House: 1 if borrower has at least one house and 0 otherwise.
Mortgate: 1 if borrower is in a mortgage contract and 0 otherwise.

Credit specific characteristics:
CR: Credit rating of borrower by Renrendai, including AA, A, B, C, D, E, HR in a descending order. CR take the value of 6 for AA, 5 for A, 4 for B, 3 for C, 2 for D, 1 for E and 0 for HR.
Application: Historical total number of loans borrower has applied.
Success: Historical number of successes in borrowing.
PaidLoan: Historical number of loans fully paid off by borrower.
TotalLending (in 1000s): Historical amount of loans borrowed.

Termination relevant:
Prepayment: 1 if terminated due to prepayment and 0 otherwise.
Default: 1 if terminated due to default and 0 otherwise.
Tertime: Months before termination.
Terprogress: Months before termination as a percentage of loan term.

Table 1 is the summary statistics for two samples, panel A for multinomial logit sample and panel B for rational expectation sample.

Table 1: Summary statistics

First	second	third
45.09	35.00	35.00
45.09	35.00	35.00

Panel A

Variable Name	N	Mean	Standard Deviation	Min	Medium	Max	Skewness
amount	16063	28.43	50.46	3	12	500	4.57
term	16063	11.24	8.57	3	9	36	1.44
age	16063	33.54	7.02	22	32	72	0.97
applied	16063	3.16	9.31	0	1	147	8.78
success	16063	2.28	8.99	0	0	143	9.27
cleared	16063	2.26	8.98	0	0	143	9.3
totallending	16063	29.62	97.73	0	0	2050	6.89
gender	16063	0.87	0.34	0	1	1	-2.2
isvoc	16063	0.4	0.49	0	0	1	0.41
isungrad	16063	0.31	0.46	0	0	1	0.8
ispograd	16063	0.03	0.17	0	0	1	5.47
isdivorced	16063	0.03	0.17	0	0	1	5.69
ismarried	16063	0.68	0.47	0	1	1	-0.79
isit	16063	0.09	0.29	0	0	1	2.86
isres	16063	0.02	0.15	0	0	1	6.38
isreales	16063	0.02	0.15	0	0	1	6.18
ispublic	16063	0.02	0.14	0	0	1	6.93
ischarity	16063	0	0.03	0	0	1	30.69
isconstr	16063	0.04	0.2	0	0	1	4.69
istrans	16063	0.04	0.2	0	0	1	4.67
isedu	16063	0.04	0.2	0	0	1	4.45
isfinlaw	16063	0.04	0.21	0	0	1	4.44
isretail	16063	0.15	0.35	0	0	1	2
ismedia	16063	0.02	0.15	0	0	1	6.58
isenergy	16063	0.04	0.2	0	0	1	4.69
isagri	16063	0.01	0.12	0	0	1	8.25
issportart	16063	0	0.06	0	0	1	16.7
ismed	16063	0.03	0.18	0	0	1	5.32
isent	16063	0.01	0.11	0	0	1	9.11
isgov	16063	0.09	0.28	0	0	1	2.97
ismanuf	16063	0.24	0.43	0	0	1	1.21
workexp	16063	2.9	0.99	1	3	4	-0.21
CR	16063	1.44	2.09	0	0	6	1.13
prepayment	16063	0.16	0.37	0	0	1	1.87
default	16063	0.06	0.25	0	0	1	3.53
censored	16063	0.78	0.42	0	1	1	-1.33
tertime	16063	5.54	4.33	0	4	36	2.27
terprogress	16063	0.64	0.36	0	0.67	1	-0.31

Panel B

Variable Name	N	Mean	Standard Deviation	Min	Medium	Max	Skewness
Amount	32168	20.65	57.86	1	3	500	5.42
Term	32168	8.85	7.1	3	6	36	1.61
Age	32168	30.53	6.29	15	29	69	1.34
Income	32168	3.54	1.2	1	3	7	1.15
House	32168	0.38	0.48	0	0	1	0.5
mortgage	32168	0.11	0.31	0	0	1	2.49
#applied	32168	1.03	4.3	0	0	102	12.4
#success	32168	0.58	3.99	0	0	100	14.15
#cleared	32168	0.58	3.98	0	0	100	14.2
Totallending	32168	4.62	30.02	0	0	830	11.33
Gender	32168	0.87	0.34	0	1	1	-2.14
Isvoc	32168	0.35	0.48	0	0	1	0.62
Isungrad	32168	0.17	0.38	0	0	1	1.72
Ispograd	32168	0.01	0.12	0	0	1	8.09
Isdivorced	32168	0.00	0.05	0	0	1	18.42
Ismarried	32168	0.46	0.5	0	0	1	0.16
Isit	32168	0.09	0.28	0	0	1	2.92
Isres	32168	0.04	0.2	0	0	1	4.64
Isreales	32168	0.03	0.18	0	0	1	5.16
Ispublic	32168	0.02	0.12	0	0	1	7.86
Iscarity	32168	0.00	0.05	0	0	1	19.49
Isconstr	32168	0.05	0.23	0	0	1	3.92
Istrans	32168	0.04	0.19	0	0	1	4.73
Isedu	32168	0.04	0.19	0	0	1	4.74
Isfinlaw	32168	0.04	0.2	0	0	1	4.51
Isretail	32168	0.15	0.35	0	0	1	2.01
Ismedia	32168	0.01	0.1	0	0	1	9.69
Isenergy	32168	0.03	0.17	0	0	1	5.42
Isagri	32168	0.02	0.13	0	0	1	7.28
Issportart	32168	0.01	0.08	0	0	1	11.94
Ismed	32168	0.03	0.17	0	0	1	5.58
Isent	32168	0.04	0.19	0	0	1	4.81
Isgov	32168	0.04	0.2	0	0	1	4.62
Ismanuf	32168	0.20	0.4	0	0	1	1.52
Workexp	32168	2.31	1.01	1	2	4	0.43
CR	32168	0.44	1.36	0	0	6	3.22
Fundingsuccess	32168	0.13	0.33	0	0	1	2.23
Interest	32168	15.28	4.75	3	15	24.4	-0.32

3 Main Results

3.1 Distribution of Expected IRR

Expected IRR of each loan is calculated as illustrated above. In the default/prepayment rate estimation model, loan's nominal interest rate is not included, due to endogeneity concern, for interest rate is considered to reflect default/prepayment risk. As robustness check, I also repeat the procedure including nominal interest rate. And the results are similar. The results of logit regression are not the main focus in this paper thus not presented. But they are available upon request.

The relationship between expected IRR and nominal interest rate in China is of main interest in this section. A large disparity between China and US P2P lending platform is the existence of a guarantee mechanism in the former. In the US, without the guarantee, if

borrower defaults, investor will lose everything, having a return of -100%. Therefore, it is possible for a loan to have a negative expected return. However in China online lending market, since each loan is protected against default by the platform, it is impossible for a loan's expected return to be negative, zero at least. In the case of default, expected IRR of a loan is smaller than its nominal interest rate undoubtedly. Investor only loses one interest payment at the month of default, for the remaining principal refund is one month later. Whereas in the case of prepayment, expected IRR can even be larger than its nominal interest rate thanks to the penalty of 1%. It is natural to infer that the expected IRR in China P2P lending market to be positive and highly correlated with nominal interest rate.

Figure 1 plots the distribution of expected IRR and nominal interest rate. The relationship is as expected. The empirical probability density function of expected IRR is more leftward than that of nominal interest rate. Figure 2 plots the distribution of nominal interest rate in excess of expected IRR, and Table 2 presents some statistics of the difference. In most cases, the difference is positive, implying a larger nominal interest rate than expected IRR. But sometimes prepayment penalty drives expected IRR higher than nominal interest rate. On average, the difference is 0.52%. And the difference is skewed to the right.

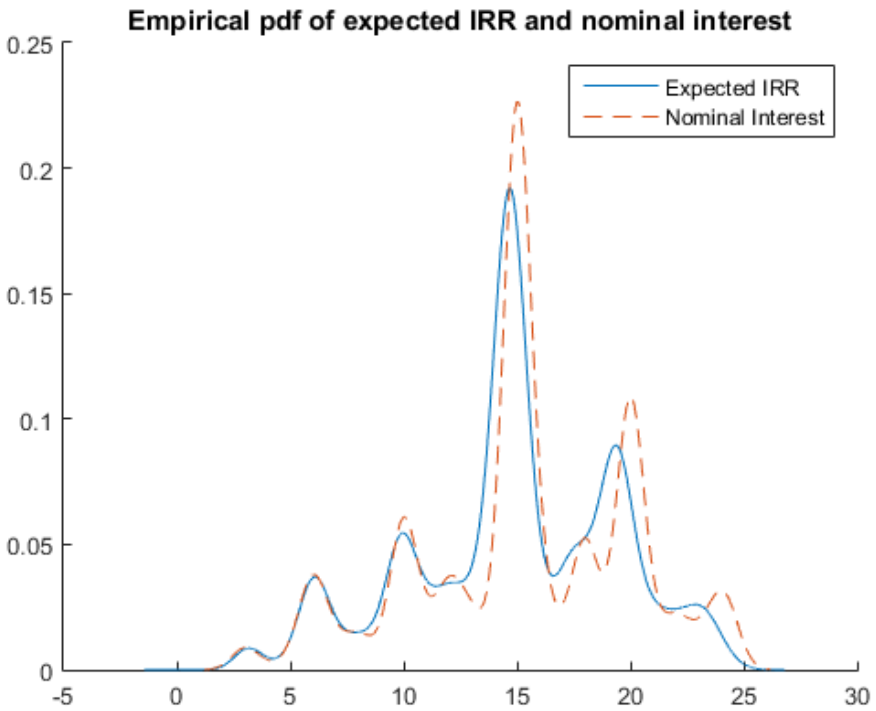

Figure 1: Empirical distribution of expected IRR and nominal interest

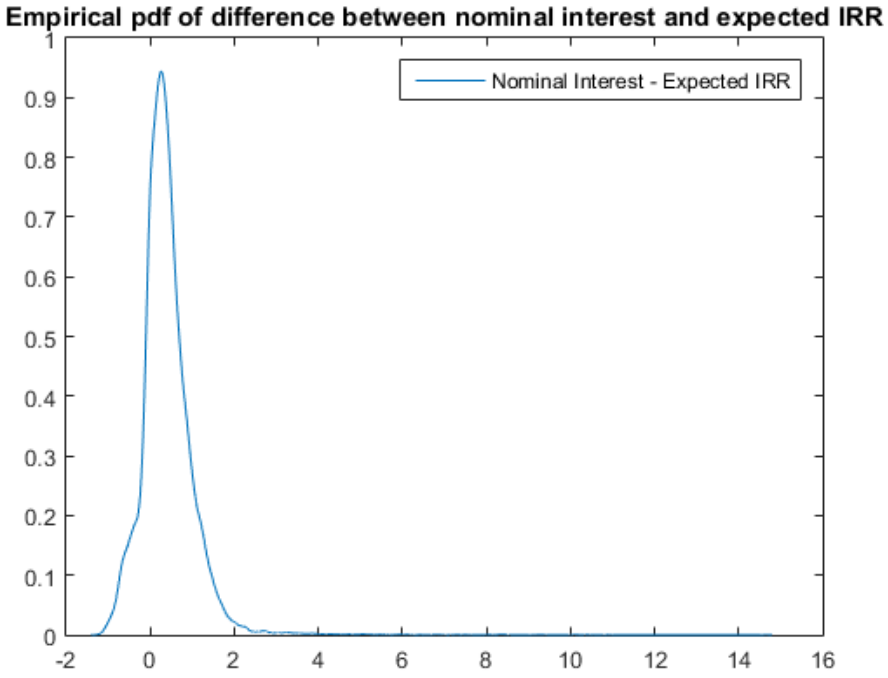

Figure 2: Empirical distribution of difference between nominal interest and expected IRR

Also, the expected IRR in China P2P market as a whole is positive, on the contrary to Duarte et al. (2014)'s finding of a negative overall expected IRR in US market. I have to point out that my finding does not implicate higher quality in China online loan assets than in the US. The positive expected IRR in China is simply a direct result of guarantee mechanism. This places more unpredictability in investors' behavior, in that they do not have that high an incentive as in the US market to do the risk return analysis. Everything they lose will be compensated. Therefore, rational expectation in China market remains a puzzle even though that in the US is already studied by many researchers.

Table 2: Statistics for difference of interests

N	Mean	Min	Medium	Max	Skewness
233679	0.52	-1.48	0.35	20.23	5.8

3.2 Influence Factors of Expected IRR

Based on the assumption that investors have rational expectation, according to the renowned capital asset pricing model, risk-adjusted expected return only depends on the systematic risk of the portfolio instead of idiosyncratic risk. I aim to use China P2P market to study if China P2P investors form rational expectation. First I have to identify proxies for systematic risks. Taking into account Duarte et al. (2014)'s choice, I use borrower's credit rating, historical borrowing records, education, income, working experience and loan specific characteristics to proxy systematic risk of a loan. Other variables represent idiosyncratic risk. Expected IRR is then regressed on both systematic factors and idiosyncratic factors. The regression model is expression (1).

$$ExpectedIRR_i = \beta_0 + \beta_1 SystematicFactors_i + \beta_2 IdiosyncraticFactors_i + \varepsilon_i \quad (1)$$

If investors are rational, when controlling all systematic factors from above, other variables should be insignificant. What individual risk has impact on should be the difference between nominal interest rate and expected IRR. Expected IRR should be solely driven by systematic factors. Table 3 displays the regression result.

Table 3: Expected returns determinants under guarantee mechanism

	(A)		(B)		(C)	
Variable	**Coefficient**	**P Value**	**Coefficient**	**P Value**	**Coefficient**	**P Value**
Intercept	14.97647	<.0001	14.43182	<.0001	14.36617	<.0001
age	-0.01844	<.0001	-0.00035	0.781	-0.00324	0.0158
gender	0.09402	<.0001	0.112	<.0001	0.10358	<.0001
isdivorced	-0.28655	<.0001	-0.29332	<.0001	-0.31434	<.0001
ismarried	-0.3363	<.0001	-0.26375	<.0001	-0.26816	<.0001
isit	-0.31034	<.0001	-0.2712	<.0001	-0.23486	<.0001
isres	-0.16293	0.0002	-0.19399	<.0001	-0.19055	<.0001
isreales	-0.09763	0.0317	-0.16592	0.0002	-0.13103	0.0036
ispublic	0.00193	0.9733	-0.13208	0.0206	-0.13795	0.0161
ischarity	-0.11659	0.4956	-0.20073	0.2355	-0.15693	0.3536
isconstr	-0.05953	0.1157	-0.05719	0.1264	-0.04845	0.1958
istrans	-0.04836	0.2418	-0.13104	0.0013	-0.1503	0.0002
isedu	0.12855	0.0032	0.0115	0.7899	0.04874	0.2686
isfinlaw	-0.32451	<.0001	-0.33871	<.0001	-0.26205	<.0001
isretail	-0.18975	<.0001	-0.13918	<.0001	-0.1397	<.0001
ismedia	-0.18608	<.0001	-0.19837	<.0001	-0.17934	0.0002
isenergy	-0.2048	<.0001	-0.27511	<.0001	-0.28496	<.0001
isagri	-0.46491	<.0001	-0.31117	<.0001	-0.31227	<.0001
issportart	-0.10278	0.3097	-0.11926	0.2332	-0.11097	0.2671
ismed	0.00294	0.9498	-0.06572	0.1543	-0.05305	0.2518
isent	0.23505	<.0001	0.19746	<.0001	0.19169	<.0001
isgov	-0.20654	<.0001	-0.30238	<.0001	-0.27865	<.0001
ismanuf	-0.26209	<.0001	-0.26881	<.0001	-0.28328	<.0001
CR			-0.03489	0.0008	-0.02829	0.0075
applied			-0.0008	0.8236	2.37E-05	0.9948
success			-1.77593	<.0001	-1.78145	<.0001
cleared			1.73029	<.0001	1.73647	<.0001
totallending			-0.00057	0.0295	-0.00067	0.01
amount			-0.00503	<.0001	-0.005	<.0001
term			0.01879	<.0001	0.01853	<.0001
income					0.01113	0.1114
house					-0.06783	<.0001
mortgage					-0.13934	<.0001
workexp					0.09122	<.0001
isvoc					-0.05314	0.0012
isungrad					-0.14238	<.0001
ispograd					-0.56705	<.0001
Adjusted R^2	**0.0069**		**0.0292**		**0.0305**	
Num. of Obs.	**233678**		**233678**		**233678**	
F test	**74.94(p<0.0001)**		**243.15(p<0.0001)**		**204.99(p<0.0001)**	

Panel A includes only idiosyncratic factors, panel B adds hard information systematic factors from the platform and panel C adds all other systematic factors. After controlling for systematic factors, expected IRR is still significantly affected by individual factors

cross-sectionally. Elder borrowers imply lower expected IRR. Loan of a male borrower has lower expected IRR than that of a female one. Loan of a single borrower has higher expected IRR than that of a married or divorced borrower. These are just a few examples of significant influence of idiosyncratic risk. The empirical evidence is against rational expectation assumption. Expected IRR is not only correlated with systematic risk, but also idiosyncratic risk. Loans are not properly priced due to investors irrationality. The reason behind the result might be the guarantee mechanism reducing the difference of nominal interest rate and expected IRR, blurring the risk return structure of the loan. In other words, no matter how much risk is inherent in a loan, the expected IRR will be close to the nominal interest rate because of guarantee mechanism.

3.3 Investors' Decision

Although the loans in China P2P market are not properly priced, it does not mean that investors are irrational when making investment decision. For a expected return maximizing, rational and sophisticated investor, she can still only use systematic information to make decision. Therefore, I regress the dummy variable of a successful loan transaction on expected IRR as well as other variables. The equation is expression (2) and (3), and it is designed to study when controlling expected IRR, whether other variables still significantly affect the probability that a representative investor invests in the loan.

$$\log\left(\frac{\Pr(FundingSuccess_i)}{1-\Pr(FundingSuccess_i)}\right) = \alpha_{10} + \alpha_{11} EIRR_i + \eta_{1i} \tag{2}$$

$$\log\left(\frac{\Pr(FundingSuccess_i)}{1-\Pr(FundingSuccess_i)}\right) = \alpha_{20} + \alpha_{21} EIRR_i + \alpha_{22} Variables + \eta_{2i} \tag{3}$$

If investors in China have rational expectation and maximize their expected return, then expected IRR should have a positive impact on the probability of loan being funded. Furthermore, controlling for expected IRR, other variables should be jointly insignificant. And if it is the most important factor in investor's decision, regression coefficient of expected IRR should be robust and remain significantly positive whichever other factors controlled. More specifically, comparing to the benchmark model merely including expected IRR as explanatory variable, models that include expected IRR as well as other variables do not have a significant increase in explanatory power. The result of the regression is given in Table 4. Panel A only includes expected IRR as independent variable, panel B adds idiosyncratic variables and panel C contains both systematic and idiosyncratic variables.

Table 4: Investors' decision under guarantee mechanism

	(A)		(B)		(C)	
Variable	**Coefficient**	**P value**	**Coefficient**	**P value**	**P value**	**P value**
Intercept	-1.5128	<.0001	-4.7369	<.0001	-7.3469	<.0001
irr	-0.0277	<.0001	-0.0243	<.0001	0.0417	<.0001
age			0.0717	<.0001	0.041	<.0001
gender			-0.00039	0.994	-0.1946	0.01
isdivorced			0.541	0.0588	0.2573	0.4765
ismarried			0.9405	<.0001	0.2566	<.0001
isit			0.7633	<.0001	-0.0173	0.889
isres			0.1044	0.3527	-0.0937	0.604
isreales			-0.1563	0.2398	-0.1248	0.5008
ispublic			0.3853	0.0054	-0.3331	0.1781
ischarity			-0.2558	0.5565	0.3223	0.5068
isconstr			-0.0248	0.8083	-0.5304	0.0008
istrans			-0.3234	0.0082	-0.3583	0.0388
isedu			0.4607	<.0001	0.1374	0.345
isfinlaw			0.5415	<.0001	-0.3638	0.0238
isretail			0.5294	<.0001	0.3284	0.0011
ismedia			0.4814	0.0071	0.2203	0.3709
isenergy			-0.0467	0.7055	-0.2912	0.1014
isagri			-0.286	0.0699	-0.5821	0.0142
issportart			-0.817	0.0197	-2.1704	<.0001
ismed			0.403	0.0004	0.1111	0.5234
isent			-0.7622	<.0001	-0.6743	0.0026
isgov			0.8085	<.0001	0.6854	<.0001
ismanuf			0.36	<.0001	0.4361	<.0001
CR					0.7643	<.0001
applied					0.0459	0.0395
success					1.8154	<.0001
cleared					-0.961	0.0041
totallending					-0.00293	0.0123
amount					-0.00046	0.2644
term					-0.0402	<.0001
income					0.4533	<.0001
house					-0.0492	0.4129
mortgage					-0.1761	0.0251
workexp					0.2729	<.0001
isvoc					0.3392	<.0001
isungrad					0.3142	<.0001
ispograd					0.0724	0.6855
-2 L	24557.818		21830.113		11615.664	
Likelihood ratio test			**2727.705**		**12942.15**	
	D.F.		**22**		**36**	
	Confidence		**95%**	**99%**	**95%**	**99%**
	Chi-square		**33.92**	**30.81**	**51**	**58.62**

From panel A, expected IRR has a significantly negative impact on the probability a loan being funded. Investors do not aim for high expected return. When controlling for all other variables, as in panel C, the influence of expected IRR turns significantly positive. For example, for two loans having the same expected IRR, one with a male borrower will be less likely to be fully funded than one with a female borrower. This is evidence that investors' irrationally prefer female borrowers to male ones. By 'irrationally', I mean the preference does not rise from return difference but from investors' taste. As a whole, expected IRR is not even robust across different model specifications. Investors' still take into account other factors, despite the risks of other factors already incorporated into

expected IRR.
Furthermore, by comparing the explanatory power of different models with panel (A) model being the benchmark, it is easily seen that adding other variables significantly decrease -2*log-likelihood, in that the statistics of likelihood ratio test are both larger than the chi-square critical values. The null hypothesis that none of the other variables except expected IRR have any impact on investors' decision are rejected. That means investors in China consider a great deal of information besides expected return of the loan. To conclude, investors in China do not form rational expectation.

3.4 Discussion

The lack of robustness in expected IRR in China P2P market might result from the guarantee mechanism. Investors might get confused when they make their decision. Even if they are rational, investors might not be clear about how to interpret the impact of the guarantee on future cash flows. Some of them might choose to believe that the platform will actually fulfill its responsibility when borrower defaults, but some might not. Those who could not completely trust the platform will put little weight on the guarantee effect when making decision, in other words a complete different expected IRR will be observed. Therefore, it is natural for the expected IRR effect to be not robust given that some investors believe in it whereas others do not. But given the good record of Renrendai, where all defaulted loans are refunded by the platform, it might still be more rational for investors to give credit to the guarantee mechanism.

5 Conclusion

Whether investors for rational expectation is an important issue both in traditional and behavioral finance research. If investors are rational, then according to CAPM model, the expected IRR of loan should only be affected by systematic risk instead of idiosyncratic risk. This paper is especially interested in China P2P investors' rationality. In China, P2P is one of the hottest topic in web finance. Furthermore, unlike the case in the US, China P2P platforms have a guarantee mechanism that pays back all remaining principal for a defaulted loan to investors. What effect the mechanism may have on investors is unknown. No other research has used a market with guarantee mechanism to study investors' rationality.
This paper builds a multinomial logistic model to incorporate the default/prepayment risk in each month for each loan and computes the expected IRR, or risk-adjusted return, of every loan. The result shows that in a platform with guarantee mechanism, the difference of expected IRR and nominal interest rate is particularly small. Expected return of a loan is not only affected by systematic risk but also idiosyncratic risk, not in accordance with CAPM. Besides, investors take into consideration other risk factors along with expected IRR, although the return-relevant influence of the risk factors is already incorporated in the latter. All empirical results give evidence to the fact that China P2P investors are not rational.

References

[1] E. Fama, Efficient capital markets: a review of theory and empirical work, *Journal of Finance*, **25**(2), (1970), 383-417.
[2] M. Jensen, The performance of mutual funds in the period 1945-1964, *Journal of Finance*, **23**(2), (1967), 389-416.
[3] E. Fama, The behavior of stock market prices, *Journal of Business*, **38**(1), (1965), 34-105.
[4] E. Fama, L. Fisher, M. Jensen and R. Roll, The adjustment of stock prices to new information, *International Economic Review*, **10**(1), (1969), 1-21.
[5] A. Shleifer, *Inefficient markets*, Oxford University Press, Oxford, 2000.
[6] J. Michels, Do unverifiable disclosures matter? Evidence from peer-to-peer lending, *Accounting Review*, **87**(4), (2012), 1385-1413.
[7] N. Barasinska, Does gender affect investors' appetite for risk? Evidence from peer-to-peer lending, DIW *Discussion Paper* 1125, (2011).
[8] S. Freedman and G. Jin, Learning by doing with asymmetric information: evidence from prosper.com, NBER *Working Paper* **W16855**, (2011).
[9] J. Duarte, S. Siegel and L. Young, To lend or not to lend: revealed attitudes towards gender, ethnicity, weight, and age in the U.S., SSRN *Working Paper* **2473240**, (2015).
[10] R. Iyer, A. Khwaja, E. Luttmer and K. Shue, Screening peers softly: inferring the quality of small borrowers, *Management Science*, **62**(6), (2015), 1554-1577.
[11] S. Freedman and G. Jin, Do social networks solve information problems for peer-to-peer lending? Evidence from Prosper.com, SSRN *Working Paper* **1936057**, (2008).
[12] S. Freedman and G. Jin, The signaling value of online social networks: lessons from peer-to-peer lending, NBER *Working Paper* **W19820**, (2014).