"A new conceptualization of investor sophistication and its impact on herding and overconfidence bias"

AUTHORS	Ashutosh Yadav (iD)
	Deepshikha Yadav (iD)
	Ishan Kashyap Hazarika (iD)

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
61	**1**	**7**

Ashutosh Yadav (India), Deepshikha Yadav (India), Ishan Kashyap Hazarika (India)

A NEW CONCEPTUALIZATION OF INVESTOR SOPHISTICATION AND ITS IMPACT ON HERDING AND OVERCONFIDENCE BIAS

Abstract

Despite the success of behavioral finance, the question of whether behavioral biases persist in the face of expertise is an oft-expressed concern. It becomes pertinent to explore if investor sophistication is associated with behavioral biases, as traders gain sophistication with experience and knowledge. The current study explores this relationship by proposing a new conceptualization of investors' sophistication via the processes of learning and competition. The study empirically explores if herding and overconfidence biases are related to learning and competition, and thus, with investors' sophistication via these aspects. Using data from equity investors from India (n = 257), the study employs ANOVA and multiple regression analysis through indicator function to form dummy variables for different categories. The results of the study conclude that diversification is significantly related to both the biases using ANOVA (F(3,253) = 3.081; p < 0.05) as well as multiple regression (p < 0.05). The other variables considered are found to be non-significant (p > 0.05) for both the biases. The study controls for all the other observed variables of the conceptual model to find out the effect of the change in the observed variables on the level of investor sophistication, making this study a novel and a distinct attempt.

Ashutosh Yadav, Assistant Professor, Department of Commerce, Hansraj College, University of Delhi, India.

Deepshikha Yadav, Assistant Professor, Department of Commerce, Hansraj College, University of Delhi, India. (Corresponding author)

Ishan Kashyap Hazarika, Student, Delhi School of Economics, University of Delhi, India.

INTRODUCTION

It is currently a widely accepted claim that human decisions in most fields, including financial decisions by investors, are subject to irrational biases. The Behavioral Finance literature is replete with studies that establish the presence and influence of such biases in different domains. The abundance of evidence in support of the existence of systematic biases is, quite famously, a significant force in the dethronement of the formerly dominant Efficient Market Hypothesis and the Rational Actor Model in Finance. Despite the advancement in the field of behavioral finance, how biases would emerge in the presence of different investors' characteristics still remains unraveled to a great extent. It is argued that commoners or students, who are often the subjects of behavioral experiments, may display systematic biases or errors in decision making due to their inexperience or lack of knowledge. Experts would not exhibit such biases, or at least, exhibit them to an extent that is not harmful. This concern has two distinct foundations: *learning* and *competition*. There is a need to discuss both these concerns to develop the contextual background for further research, in which the relevance of the current study becomes evident.

Investment decision-making of individual investors can be considered as an intricate behavioral process which is influenced by ration-

al and irrational factors that contribute to inefficiency of security markets (Shanmugham & Ramya, 2012). The susceptibility of an investor to a particular bias may be considered as a function of unique physiognomies of investors ranging from demographic, psychological, and personality factors (Yadav & Narayanan, 2021).

Overconfidence and herding biases are two of the most prominent behavioral biases that investors get susceptible to while making financial decisions. Odean (1998) defined the concept of overconfidence as investors' tendency to overestimate the precision of their knowledge about the value of a security. Prosad et al. (2013) proved the presence of overconfidence bias in the Indian equity market. Herding propensities, on the other hand, may develop not just because people imitate the actions of others as they conclude that others have better knowledge about the fundamental long-term values of goods and assets, but also because assenting to a group bequeaths a utility that is independent of the information implicit in the decisions of others (Baddeley et al., 2010).

This paper proposes an entirely new conceptual framework to define and understand investor sophistication. This study also demonstrates an empirical application of this conceptualization by relating investor sophistication to herding and overconfidence, two highly important biases concerning learning and competition.

1. LITERATURE REVIEW

The first strand of literature arguing the irrelevance of behavioral aspects, termed as '*learning*', is based on the idea that rational behavior can be learned by training or experience. Smith (2009) notes that the 'principal findings' of Experimental Economics show that even in experimental settings, under repeated interactions, impersonal exchange in markets converge to the equilibrium states predicted by economic theory. Interestingly, the information conditions under which this occurs in experiments can be much weaker than what is demanded by theory. That is, on repeated interaction, the predictions of rational economic theory are often found to be true, even when some of the assumptions of the model are not satisfied. Real markets operate daily, and a large amount of trade takes place every day. It can thus be expected that most traders would soon learn to behave rationally, by imitation if not consciously. This raises serious questions on the relevance of the findings of Behavioral Finance. While many such results may be true in the laboratory, they may be irrelevant in real markets. As a response, their validity can be tested in real markets through field studies (Pope & Sydnor, 2015). But this approach is like a black box that only affirms the presence of biases, not explain why they exist despite learning. Another approach is to relate investors' sophistication (which results partly from learning) to the extent of biases in real markets, which gives us a clearer picture. For now, we only define 'investor sophistication' in an admittedly hand-wavy fashion, as the quality of having more experience and knowledge, engaging in relatively complex activities and possibly having high net-worth. This is an unsatisfactory definition, but it is clear that it relates to the investors' experience, how frequently the investor trades, professional expertise, portfolio diversification, etc. This shall suffice for the discussion that follows. Numerous biases affect human decision making (Davis, 2018). But certain biases have received more attention due to their prominence in this field. Baker et al. (2018), in particular, note that four biases are majorly important here: overconfidence, herding, disposition and mental accounting.

According to Gigerenzer et al. (1991), the overconfidence bias occurs when the confidence judgments are larger than the relative frequencies of the correct answers. Block and Harper (1991) calls it a 'cognitive deceit'. Fellner and Krügel (2012) redefines it in terms of single-cue signals. This captures the qualitative understanding better. On the other hand, discussions of herding in the context of economic and financial decision making can be traced back at least to studies by Keynes (Baddeley, 2010). Asch (1952) is an early psychological study on the topic. A fairly simple definition is found in Bikhchandani et al. (1992), where it is defined

as imitation behavior resulting from individual factors and leading to market inefficiencies. The connection of herding to learning is especially important for new investors. They are often not sophisticated and face a dearth of information and may in fact find it rational to actively observe other participants and try to elicit information on possible better strategies. This is rational herding. Irrational herding, on the other hand, involves a passive imitation of others. Zhang and Liu (2012) report rational herding among lenders in the microloan market. Toyokawa et al. (2019) find that for challenging tasks, copying may lead to maladaptive herding, while for less-challenging tasks, it may lead to 'wisdom of the crowd'. Similarly, the very definition of the overconfidence bias relates it to learning. Overconfidence entails mistakes in learning from signals: is a loss or reward due to luck or strategy? Gervais and Odean (2015) explain overconfidence using a multi-period market model. In their model, traders are initially unaware of their abilities. They try to infer it from their successes and losses and update their beliefs with each success or loss. A flaw in this learning process, whereby the traders take too much credit for their successes, leads to overconfidence. This possibly relates to self-attribution, which is found to be an important factor in Allen and Evans (2005) too.

The second strand of literature arguing the irrelevance of behavioral laboratory results in real markets relies on what we term the logic of *'competition'*: if rational behaviors are optimal, then regardless of whether the market participants 'learn' to behave rationally or not, over time markets would tend to have only rational participants. This happens because market activities involve risks, and profit and loss calculations. In the long run, participants who act irrationally would be eliminated from markets as they would face suboptimal outcomes frequently, and would be outcompeted by more rational participants. Thus, markets act as sieves. There can be several variants of this argument. An interesting finding is in Gode and Sunder (1993), where even in the presence of computerized zero-intelligence traders who make random bids and asks subject to a zero-profit budget constraint, the markets converge to the competitive equilibrium allocation. A related argument is that rational agents will drive the irrational agents

from the market because rational agents make higher profits (Fehr & Tyran, 2005). Finally, it may be argued that even if irrational participants exist in extreme ends of the supply and demand curves, the actions of more rational agents who are marginal buyers and sellers will determine the market equilibrium. A market where there is a continuous entry and exit of agents would have both rational and irrational agents, but irrational agents would not last long. They may of course be replaced by other irrational agents who enter the market. But even if irrational agents constitute a majority in the market, their experience and amount of trading would be limited, compared to rational agents. For the purpose of this study, this means that a negative association between an investor's sophistication and the extent of her biases can be expected. This highlights the role of the market mechanism in the rationality or irrationality that can be observed in it. If the market is good enough to weed out irrational agents quickly, a strong negative relationship between sophistication and biases can be found, but only a weak relationship otherwise. Thus, in this paper, the term 'competitive' has been used, to specifically mean the extent to which the process discussed above is taking place, unlike the convention in microeconomic theory.

This also means that extremely competitive and extremely non-competitive markets may not offer enough variability in terms of vulnerability to biases. Extremely competitive markets would have very few irrational agents, while in extremely non-competitive markets, there may be very few rational ones. The markets that are in-between the two extremes can offer interesting insights. It can be argued that markets in emerging-market economies offer this balance. Shusha and Tounny (2016) and Christie and Huang (1995) note that herding bias is usually common in emerging markets. Chen et al. (2004) focus on emerging markets too. Vo and Phan (2016) affirm the presence of herding in Vietnam. Prosad et al. (2015) note that this is because emerging markets are less 'mature'. They are more likely to sustain the coexistence of rationality and irrationality. Thus, Kawshala et al. (2020) urge future researchers to focus on these markets. Mushinada and Veluri (2018) note the presence of self-attribution and overconfidence bias and show that a large part of excessive and asymmetric volatility in Indian stock market is

explained by these biases. Prosad et al. (2017) also affirm the presence of overconfidence bias in the emerging market of India.

Extant literature also highlights the different aspects of investor sophistication: experience, frequency of trade, profession and diversification, with overconfidence and herding biases. Experience is expected to be negatively related to the propensity to exhibit behavioral biases. More experience implies more opportunities to learn, and also that the investor has survived in the market for long. But this is not unanimously found empirically. Menkhoff et al. (2006) and Kawshala et al. (2020) find a negative relationship, as expected, in the Colombo Stock Exchange. In a similar vein, Prosad et al. (2015) find that investors with less than one year of experience are more likely to engage in herding behavior. They also find that highly experienced investors also engage in more herding, however. Herding is less in the middle. Bodnaruk and Simonov (2015) find what apparently is to the contrary. They find that inexperienced investors are interested in investing on their own rather than being advised by consultants. While this may mean that they are less dependent on others, this may also mean that they prefer to imitate other investors instead of taking formal help. This can also be related to overconfidence. Inexperienced investors may be overconfident of their abilities and choose to invest on their own. But a distinction must be made between confidence and overconfidence, which is the confidence that exceeds reason. It cannot be ascertained which one it is. Many studies do not make such a distinction in clear terms. One possible answer, theoretically, can be that in efficient markets, it is not really possible to beat the market, and thus any confidence in one's abilities is overconfidence. But that is a debatable proposition. Glaser et al. (2004) and Baker et al. (2018) find that more experienced investors are more overconfident. Barber and Odean (2001) find that more experienced investors are less overconfident.

Similarly, the profession of the investor is also related to sophistication. It can be expected that people with professions related to the relevant market would exhibit biases less. Beatrice et al. (2021) and Elizabeth et al. (2020) find that investors whose professions are related to finance en-

gage less in herding. Kawshala et al. (2020), however, report that they engaged more in herding. Tekçe et al. (2016) and Sarkar and Sahu (2018) find significantly different levels of herding across professions. Prosad et al. (2015) find no significant relationship between occupation and herding. Coming to overconfidence, Chandra et al. (2017) find that traders with finance-related professions are more overconfident. Lin (2011), Kumar and Goyal (2016) and Elizabeth et al. (2020), however, find no significant relationship between the two.

The frequency of trade is related too. The higher the frequency, given the same amount of experience, the more chances for the investor to learn and the more trades she has survived in the competition. Kawshala et al. (2020), as expected, report that investors with a higher frequency of trade engage less in trading. Prosad et al. (2015) find the opposite. As per the study, intraday traders are found to be prone to overconfidence and herding. The investors who trade less frequently are more cautious in comparison to intraday traders. Barber and Odean (2001) also record that perhaps it is overconfidence that leads to excessive trading, as well as poor performance.

Finally, diversification is also a sign of sophistication. Even Feng and Seasholes (2005), who argue that factors like experience do not constitute investor sophistication, relate diversification to sophistication. The relationship between herding and diversification is complicated. Filiz et al. (2018), unlike the discussion, consider the question in reverse. Filiz et al. (2018) and the current study are equally concerned with the effects of herding on the optimality of decisions. But, while this acts as a motivation for this paper, and it seeks to use characteristics of investors such as sophistication to explain the propensity to herd, Filiz et al. (2018) instead ask how does herding affect optimal diversification. The answer is that it possibly generates non-optimality. Kim and Pantzalis (2003) find that herding is actually negatively associated with diversification. Fuertes et al. (2012) note the possibility that poorer diversification is a reflection of overconfidence. Cervellati et al. (2013) note this in the case of entrepreneurs. Merkle (2017) states that overconfidence can lead to less diversification. These studies look at diversification as an outcome of overconfidence, while the current

study considers diversification as a component of investors' sophistication, which in turn may be associated with overconfidence.

From the literature, it is clear that there is no consensus on what exactly is meant by investor sophistication, what variables ideally should be there, or what should be the proxies for those variables. Some studies attempt to properly define investor sophistication as a state of expertise but fall short of explaining the process of sophistication. Researchers are left to their own means to characterize and measure sophistication. The Corporate Finance Institute (n.d.) definition is useful but vague. The classification of the Security and Exchange Board is legal and not behavioral. A possibly accurate characterization is by Kacperczyk et al. (2019). But it ends up using income as the proxy, which is too general a variable and includes several other unrelated influences in it. Calvet et al. (2009) provide a rigorous definition, terming unsophistication as 'mistakes'. They do not relate it to both learning and competition, however, to shed light on the dynamics that this paper delves on. Other papers choose proxy variables in an ad hoc manner, without properly characterizing investors' sophistication.

On the basis of the literature, a new characterization of investors' sophistication can be proposed by imagining two states: unsophistication and sophistication (see Figure 1). Unsophistication can perhaps be defined as trading at random: choosing any security with uniform probability and investing a random sum of money in it. Sophistication, on the other hand, is the state where the investor satisfies all the axioms of rationality. The level of sophistication of an investor is: how far they are from unsophistication or how close to sophistica-

tion. Obviously, this cannot be directly measured, and one is forced to look at imperfect indicators.

In Figure 1, the distance of investors from unsophistication is due to either learning or competition. Learning happens either through training (termed here as knowledge) or through experience. Competition also selects learned and experienced participants. The boxes in the middle, experience, knowledge and successful strategies can be directly or indirectly captured through observable variables. Learning and competition are processes and cannot be observed. There is a crucial difference between the roles of learning and competition in this process. Two major differences can be highlighted. Firstly, there is a difference in causality relating them to sophistication. Learning causes individual investors to become more sophisticated. Competition, on the other hand, does not cause individual investors to be more sophisticated but eliminates unsophisticated investors. One may, however, argue that competition thus causes the average sophistication of the market to rise. Secondly, there is a difference in causality linking them to experience, knowledge and strategies. Experience and knowledge directly cause learning. But surviving competition does not necessarily cause experience or knowledge. Competition simply eliminates unsophisticated investors, and thus on average, the experiences of unsophisticated investors in the market is less. Similarly, the market may eliminate investors who have less knowledge and perform poorly. The point is that while in some cases competition may cause individual experience or knowledge, this is not necessary for the relationship in the diagram to hold. Finally, while learning may cause successful strategies, successful strategies cause survival in the market. Experience, knowledge and strategies are

Source: Based on authors' understanding of literature.
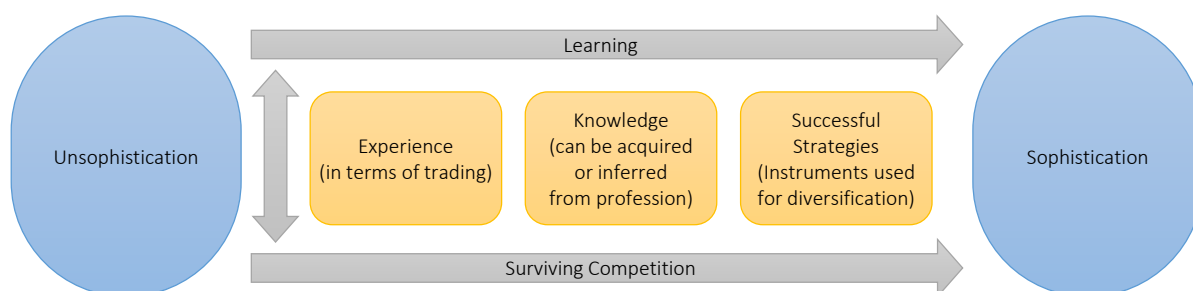
Figure 1. Authors' conceptualization of Investor Sophistication

all positively related conceptually to both learning and survival in competition, whether causally or not. Therefore, higher scores in these may indicate higher levels of sophistication. Also note, that this characterization is not a Structural Equation Model, as learning and competition are not latent variables, but processes. The characterization presented in this study is a qualitative conceptual model that can be used to generate a multitude of quantitative models.

Investor sophistication is relevant in both the arguments provided for *Learning* and *Competition*. The more an investor engages in trade, the more the opportunity to learn to behave rationally. Similarly, more sophisticated agents are more likely to survive competition. A sophisticated investor is a high networth investor who has sophisticated knowledge and investment experience that makes her capable of increasing returns and lowering risks in more advanced opportunities, according to the Corporate Finance Institute (n.d.). This is an industry-oriented definition. It also has a related legal meaning, for example, the Securities and Exchange Commission defines the categories of investors in the US. In the academic literature, the exact characterization of investor sophistication is missing, but the definitions used are similar to the idea described above. In Kacperczyk et al. (2019), sophisticated investors have access to better information, which allows them to earn higher income on the assets they hold. This notion is also found in Arrow (1987). According to the Carlin and Manso's (2011) model, when non-expert investors become informed through access to experts or public signals, they pick the optimal fund and are termed sophisticated. Operationalizing this idea is another issue. There are different approaches in the literature. Kacperczyk et al. (2019) simply take wealth as a proxy. Bartov et al. (2000) and Collins et al. (2003) use the proportion of firm shares held by institutional investors as a measure. Feng and Seasholes (2005) differentiate between sophistication and experience. The former constitutes static differences between investors, such as the level of diversification, while the latter is an evolving behavior of the same investor. Dhar and Zhu (2006) use investor literacy about financial markets and trading frequency.

This new conceptual model may help include all aspects of sophistication in research. Based on the literature review and the conceptual model, the study aims to establish if experience, frequency, profession and diversification are associated with the herding effect and overconfidence bias. Accordingly, the hypotheses for the study can be stated as:

H_a1: *Experience as a proxy of investor sophistication impacts herding and overconfidence.*

H_a2: *Profession of investors as a proxy of investor sophistication impacts herding and overconfidence.*

H_a3: *Frequency of trade as a proxy of investor sophistication impacts herding and overconfidence.*

H_a4: *Diversification as a proxy of investor sophistication impacts herding and overconfidence.*

2. METHODS

The proposed conceptualization can be used to choose variables for empirically analyzing the relationship between investor sophistication and herding and overconfidence, but data in perfect alliance with it are not available. As noted earlier, this study uses the number of years an investor has been trading (*time-experience*, which is only one component of experience, but we shall simply call it experience from now on for brevity), the frequency of trading, the profession and the number of securities in the portfolio. These are only imperfect proxies for the variables in our conceptual model, but it is argued that these shall be correct at least on average.

None of the variables of the concern is directly observable. Accordingly, the quantification of these variable is based on the reported degrees of various variables. For both the biases and the components of investor sophistication, data from 257 equity investors have been collected. As the biases involved manifest themselves in multiple ways, multi-question questionnaire has been used to triangulate the extent of these biases in individual investors. The questionnaire included reflective statements about different aspects of the biases. An example would be, 'discussing my investment

decisions with colleagues reduces my pressure'. The respondent is asked to rate a series of such questions, on a 5-point Likert scale. This paper uses a 7-item questionnaire for overconfidence and a 4-item questionnaire for herding.

For the components of investor sophistication, categorical variables have been used. The investors are divided into 'less than one month', 'one – three times per month', 'four – eight times per month' and 'more than eight times per month' based on the frequency of trading. Similarly, we have 'less than three years', 'three – 10 years' and 'more than 10 years' as categories of experience and 'zero – three', 'four – seven' and 'eight and more' securities in the portfolio as the categories of diversification. The occupational categories are 'unemployed', 'employed', 'profession' and 'business'.

Analysis has been performed based on these investor sophistication variables to examine if the behavioral biases differ amongst different categories of these sophistication variables. ANOVA is used to test the difference of biases across the sample for Occupation, Trading Frequency, Experience and Number of Securities. The null hypothesis of the ANOVA test is:

H_0:　*There is no significant difference between/ among the mean responses of various groups.*

Levene's statistics is used to test whether the assumption of homogeneity is met, with the null hypothesis of no difference in variances across the different categories, to ascertain the validity of our inferences based on ANOVA. In cases where this assumption is violated, the analysis is performed using the Brown-Forsythe Test. Post-hoc analysis is also conducted to see which exact categories differ, in case significant differences are found in the means of various categories.

The results from ANOVA have been further ingrained through multiple regression analysis. The categorical regressors (each component of investor sophistication) are converted into dummy variables, indicating the presence or absence of the membership of each particular category in each individual investor. One of the categories or the levels is taken as the reference level and dummy variables are constructed using the indicator function i.e.,

$$I(A) = \begin{cases} 1 & A\,is\,true \\ 0 & A\,is\,false \end{cases}, \quad (1)$$

for every level/category of the categorical variable such that there are $X_{i,l}$ variables formed for each categorical variable. l here represents the levels or categories in a categorical variable. Accordingly, the regression equation takes the following form:

$$Y_i \,\widetilde{ind}\, N\left(\beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3}\sigma^2\right), \quad (2)$$

where l is 1, 2 and 3 for three different levels of the categorical variable given that there are total of four categories in the variable

- β_0 is the expected response for the reference level;

- β_1, β_2 and β_3 are the expected responses for the level 1, level 2 and level 3, respectively.

Two such regressions, one for each bias, are conducted, with the dummies for the categories of each component of investors' sophistication as the regressors.

3. RESULTS

Before the analysis, the standard deviations of each respondent's responses across all statements are calculated, and those with standard deviation less than 0.5 are excluded from the sample as unengaged responses. The study followed Micceri (1989) and omit normality tests, as is the custom for categorical variables in psychometric studies. This situation is faced given the subjective nature of the questionnaire. For ordinal data sets, assessment of normality becomes difficult because the interval between scale points cannot be said to be equal, and so it is not strictly possible to regard an ordinal data set to be normal. Table 1 presents the frequency distribution of the respondents in different groups of sophistication variables.

As multiple reflective statements have been used for herding and overconfidence, internal consistency of the measurement instruments has been assessed using Cronbach's Alpha, Spearman Brown's Coefficient, and Guttman's Split Half Coefficient. A Cronbach's alpha score exceeding

Table 1. Frequency distribution of investors in different categories

Occupation	N	Percent	Frequency of trade	N	Percent
Unemployed	44	17.1	Less than one per month	32	12.5
Employed	122	47.5	1-3 times per month	116	45.1
Profession	37	14.4	4-8 times per month	37	14.4
Business	54	21.0	More than 8 times per month	72	28.0
Experience	**N**	**Percent**	**No. of securities in portfolio**	**N**	**Percent**
Less than 3 years	130	50.6	0-3	55	21.4
3-10 years	77	30.0	4-7	86	33.5
More than 10 years	50	19.5	8 and more	116	45.1

Table 2. Reliability of questionnaire statements

Measures of reliability		Overconfidence	Herding	Overall
Cronbach's Alpha		0.814	0.641	0.853
Spearman-Brown coefficient	Equal Length	0.881	0.715	0.833
	Unequal Length	0.833	0.715	0.883
Guttman Split-Half coefficient		0.832	0.713	0.871

0.70 is used as an indication of strong item co-variance as Adadan and Savasci (2011). The results of the tests are shown in Table 2.

Internal consistency is found for both herding and overconfidence using both Spearman-Brown's Coefficient and the Guttman split-half Coefficient. Using Cronbach's Alpha, strong internal consistency is found in overconfidence, but herding missed the mark. However, this borderline case of herding, which has a reliability score of less than 0.70, but more than 0.60 has also been considered for further analysis (Hair et al., 2010) as the two other parameters of reliability present acceptable figures.

To test the homogeneity assumption for ANOVA, Levene's statistics is used. The results are presented in Table 3.

Levene's test suggested that the homogeneity assumption is not violated for all the categorical variables for both herding and overconfidence biases except in the case of experience for herding and diversification for overconfidence. Welch and Brown-Forsythe tests are used to analyze the significant differences between the categories.

The analysis of the equality of means is carried out using ANOVA results for the relevant variables. The summarized results are presented in Table 4.

Table 3. Levene's statistics: test of homogeneity of variances

Biases	Occupation	Frequency	Experience	Diversification
Herding	1.459 (0.226)	1.032 (0.379)	3.701* (0.026)	1.121 (0.328)
OC	0.735 (0.532)	0.744 (0.527)	1.748 (0.176)	3.296* (0.039)

Note: * Figures in parentheses indicate the level of significance, p-values indicate significance at the 5 percent level.

Table 4. ANOVA results

Biases	Overconfidence			Herding		
Test statistics	Occupation	Frequency	Experience	Occupation	Frequency	Diversification
Mean square	9.39	12.74	38.28	2.13	15.22	39.64
F	.28	.38	1.15	.16	1.17	3.08*
Sig.	.84	.77	.32	.92	.32	.05

Note: * Significant at the 5 percent level.

Table 5. Results of the Welch and Brown-Forsythe tests

Parameters	Test	Test statistics	df1	df2	Sig.
Overconfidence and diversification	Welch	4.82*	2	145	.01
	Brown-Forsythe	5.36*	2	227	.01
Herding and experience	Welch	0.67	2	114	0.51
	Brown-Forsythe	0.62	2	170	0.54

Note: * Significant at the 5 percent level.

The study proceeds to analyze the equality of means for experience for herding and diversification for overconfidence using Welch and Brown-Forsythe tests. The results of both these analyses are presented in Table 5.

For overconfidence bias, occupation, frequency of trade and experience are found to be insignificant with $F(3,253) = 0.281$; $F(3.253) = 0.380$; and $F(2,254) = 1.153$, respectively, with p-values greater than 0.05. It can be inferred that significant differences do not exist in susceptibility towards overconfidence bias among investors based on occupation, frequency of trade and experience of the investors. However, diversification is found to have significant differences in the susceptibility to overconfidence bias using both Welch and Brown-Forsythe as test statistics are significant at the 5 percent level.

For herding bias, occupation and frequency of trading are found to be insignificant with $F(3,253) = 0.162$ and $F(3,253) = 1.66$, respective-ly. Diversification, however, is found to have a significant impact on herding behavior with $F(3,253) = 3.081$. Significant differences do not exist in herding behavior based on experience as both Welch and Brown-Forsythe test statistics are insignificant.

To determine the categories among which significant differences exist, the results of the post hoc analysis are displayed in Table 6 for diversification for both overconfidence and herding bias. The results of the post-hoc analysis indicated that significant differences exist between investors with moderate diversification and high diversification for herding behavior and between investors with no diversification and high diversification in case of overconfidence bias.

The results of the multiple regression analysis for the questions of difference of means are similar to the ANOVA results and are presented in Table 7. Accordingly, H_a4 is not rejected and H_a1, H_a2, and H_a3 are rejected.

Table 6. Post hoc analysis (multiple comparisons: Hochberg)

Dependent variable	(I) Number of securities	(J) Number of securities	Mean difference (I-J)	Std. error	Sig.
Herding	No diversification	Moderate diversification	.000	.619	1.000
		High diversification	1.116	0.587	0.165
	Moderate diversification	No diversification	0.000	0.619	1.000
		High diversification	1.116	0.510	.086**
	High diversification	No diversification	−1.116	0.587	0.165
		Moderate diversification	−1.116	0.510	0.086**
Overconfidence	No diversification	Moderate diversification	−1.355	0.981	0.424
		High diversification	−2.831	0.930	0.008*
	Moderate diversification	No diversification	1.355	0.981	0.424
		High diversification	−1.476	0.808	0.193
	High diversification	No diversification	2.831	0.930	0.008*
		Moderate diversification	1.476	0.808	0.193

Table 7. Multiple-regression analysis results

Investor Sophistication variables		Model 1 (Overconfidence)		Model 2 (Herding)	
		Beta	**p-value**	**Beta**	**p-value**
Frequency (*Low Frequency* = 0)	(Constant)	23.594	0.000	11.781	0.000
	FreqDum1	0.122	0.916	−0.497	0.492
	FreqDum2	0.595	0.670	−1.43	0.102
	FreqDum3	0.962	0.435	−0.976	0.205
Occupation (*Unemployed* = 0)	(Constant)	23.773	0.000	11.159	0.000
	OccDum1	0.572	0.575	0.005	0.994
	OccDum2	−0.232	0.857	−0.456	0.574
	OccDum3	−0.032	0.978	−0.085	0.908
Experience (*Less than 3 years* = 0)	(Constant)	23.469	0.000	11.292	0.000
	ExpDum1	0.998	0.229	−0.266	0.609
	ExpDum2	1.211	0.208	−0.692	0.252
Diversification (*Less Diversification* = 0)	(Constant)	22.273	0.000	11.582	0.000
	DiversDum1	1.355	0.168	0.000	0.999
	DiversDum2	2.831*	0.003	−1.116**	0.058

Note: * Significant at 5 the percent level. **Significant at the 10 percent level.

4. DISCUSSION

The ANOVA and the multiple regression analysis yielded similar results. The previous literature has studied the relationship between different aspects of what is characterized here as investor sophistication and come to different conclusions.

The frequency of trading is found not to be significantly associated with herding and overconfidence. This is in contrary to the findings of Kawshala et al. (2020), Prosad et al. (2015), Odean (2001), which reported significant associations. The finding of the current study is not against the tide of the literature, however, as the directions found in the papers contradict each other. There is no consensus in the literature regarding the relationship between frequency of trading and herding or overconfidence.

No significant relationship between occupation and herding or overconfidence are reported by this study's results. This again goes contrary to Kawshala et al. (2020), Prosad et al. (2015), Tekçe et al. (2016) and Sarkar and Sahu (2018), which report significant relationships for herding. But again, these studies do not agree on the direction, even in this case. The finding is consistent with Lin (2011), Kumar and Goyal (2016) and Elizabeth

et al. (2020), who also reported insignificant results for overconfidence.

The results do not support the existence of a significant relationship between experience and herding or overconfidence. This again contradicts Kawshala et al. (2020), Prosad et al. (2015), Menkhoff et al. (2006), and Bodnaruk and Simonov (2015) who find a negative relationship with herding, at least according to the interpretation as discussed in the literature review. Glaser et al. (2004), Baker et al. (2018), and Barber and Odean (2001) also find significant relationships with overconfidence, but this time again, the directions in these papers contradict. Therefore, the findings are not against any general trend in the literature.

Finally, the study found significant relationships of diversification with both herding and overconfidence. Investors with a higher level of diversification engage less in herding but are found to be more overconfident than those with less diversification. Kim and Pantzalis (2003) also find that more experienced investors herd less. The result of the current study regarding overconfidence contradicts Fuertes et al. (2012), Merkle (2017), and Cervellati et al. (2013).

The results obtained, either in consonance with or in opposition to those of other studies, differ

significantly in their interpretation. Most other studies have explored the associations in contexts other than investor sophistication, and therefore all the variables that have been used as controls here have not been used in previous studies. Even studies that explicitly target investor sophistication, such as Prosad et al. (2015), use definitions of sophistication that differ from those presented in this study. This makes this study novel and distinct from other studies done in the past. Thus, the results obtained in this study are not expected to necessarily match those of earlier studies, as they have not controlled for all the variables that are relevant in the conceptual model.

CONCLUSION

This paper aims to propose a new conceptual characterization of investor sophistication, where the level of sophistication has been defined as the distance from unsophistication (random trading) or the closeness to sophistication (rational trading). This is determined by learning and competition, although the causal relationships are complicated. The paper considers time-experience, frequency of trade, profession and diversification as proxies for investor sophistication that affect and are affected by learning and competition proposed in the model.

The study also relates these variables of investor sophistication to overconfidence and biases. The study establishes significant relationships of diversification with overconfidence and herding bias using ANOVA and post-hoc analysis. However, the other findings of the study are not found to be significant. These results are also supported by multiple regression using the identity function. The findings of the study imply that investor sophistication can potentially help reduce susceptibility towards behavioral biases while making investment decisions.

The conceptual framework proposed by the study has important implications too. While the findings of the study, that only diversification is found to be significantly related to the biases, may reflect a flaw in the conceptual model, note that diversification is only a crude measure of a strategy. An alternative interpretation is that perhaps diversification already reflects the effects of experience and knowledge. The model is a qualitative conceptual one that can be used to generate a multitude of quantitative models. This can help researchers identify and choose variables to study investor sophistication.

AUTHOR CONTRIBUTIONS

Conceptualization: Ashutosh Yadav, Deepshikha Yadav, Ishan Kashyap Hazarika.
Data curation: Ashutosh Yadav, Deepshikha Yadav.
Formal analysis: Ashutosh Yadav, Ishan Kashyap Hazarika.
Investigation: Ashutosh Yadav, Deepshikha Yadav.
Methodology: Ashutosh Yadav, Deepshikha Yadav, Ishan Kashyap Hazarika.
Project administration: Ishan Kashyap Hazarika.
Resources: Ashutosh Yadav, Deepshikha Yadav, Ishan Kashyap Hazarika.
Software: Ishan Kashyap Hazarika.
Supervision: Ashutosh Yadav, Deepshikha Yadav, Ishan Kashyap Hazarika.
Validation: Ishan Kashyap Hazarika.
Visualization: Ashutosh Yadav, Deepshikha Yadav.
Writing – original draft: Ishan Kashyap Hazarika.
Writing – review & editing: Ashutosh Yadav, Deepshikha Yadav.

REFERENCES

1. Adadan, E., & Savasci, F. (2011). An Analysis of 16-17-Year-Old Students' Understanding of Solution Chemistry Concepts Using a Two-Tier Diagnostic Instrument. *International Journal of Science Education, 34*(4), 513-544. https://doi.org/10.1080/09500693.2011.636084

2. Allen, W. D., & Evans, D. A. (2005). Bidding and Overconfidence in Experimental Financial Markets. *The Journal of Behavioral Finance, 6*(3), 108-120. https://doi.org/10.1207/s15427579jpfm0603_1

3. Arrow, K. J. (1987). The Demand for Information and the Distribution of Income. *Probability in the Engineering and Informational Sciences, 1*(1), 3-13. https://doi.org/10.1017/S0269964800000243

4. Asch, S. E. (1952). *Social Psychology*. Englewood Cliffs, NJ, US: Prentice Hall, Inc. http://dx.doi.org/10.1037/10025-000.

5. Baddeley, M. (2010). Herding, Social Influence and Economic Decision-Making: Socio-Psychological and Neuroscientific Analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1538), 281-290. https://doi.org/10.1098/rstb.2009.0169

6. Baddeley, M., Burke, C., Schutz, W., & Tobler, T. (2010). *Impacts of Personality on Herding in Financial Decision-Making* (CWPE Series 1006). Retrieved from https://www.econ.cam.ac.uk/research-files/repec/cam/pdf/cwpe1006.pdf

7. Baker, H. K., Kumar, S., Goyal, N., & Gaur, V. (2019). How Financial Literacy and Demographic Variables Relate to Behavioral Biases. *Managerial Finance, 45*(1), 146-168. Retrieved from https://ideas.repec.org/a/eme/mfipps/mf-01-2018-0003.html

8. Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics, 116*(1), 261-

292. http://dx.doi.org/10.2139/ssrn.139415

9. Bartov, E., Radhakrishnan, S., & Krinsky, I. (2000). Investor Sophistication and Patterns in Stock Returns after Earnings Announcements. *The Accounting Review, 75*(1), 43-63. Retrieved from https://www.jstor.org/stable/248632

10. Beatrice, V., Murhadi, W. R., & Herlambang, A. (2021). The Effect of Demographic Factors On Behavioral Biases. *Jurnal Siasat Bisnis, 25*(1), 17-29. https://doi.org/10.20885/jsb.vol25.iss1.art2

11. Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy, 100*(5), 992-1026. Retrieved from https://www.jstor.org/stable/2138632

12. Block, R. A., & Harper, D. R. (1991). Overconfidence in Estimation: Testing the Anchoring-and-Adjustment Hypothesis. *Organizational Behavior and Human Decision Processes, 49*(2), 188-207. http://dx.doi.org/10.1016/0749-5978(91)90048-X

13. Bodnaruk, A., & Simonov, A. (2015). Do Financial Experts Make Better Investment Decisions? *Journal of Financial Intermediation, 24*(4), 514-536. https://doi.org/10.1016/j.jfi.2014.09.001

14. Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Measuring the Financial Sophistication of Households. *American Economics Review, 99*(2), 393-98. https://doi.org/10.1257/aer.99.2.393

15. Carlin, B. I., & Manso, G. (2011). Obfuscation, Learning, and the Evolution of Investor Sophistication. *The Review of Financial Studies, 24*(3), 754-785. Retrieved from http://hdl.handle.net/1721.1/64773

16. Cervellati, E. M., Pattitoni, P., Savioli, M., & Yazdipour, R. (2013). *Entrepreneurial Under-*

diversification: Over Optimism and Overconfidence (Working Paper Series 09-13). The Rimini Centre for Economic Analysis. Retrieved from https://ideas.repec.org/p/rim/rimwps/09_13.html

17. Chandra, A., Sanningammanavara, K., & Nandini, A. S. (2017). Does Individual Heterogeneity Shape Retail Investor Behaviour? *International Journal of Social Economics, 44*(5), 578-593. https://doi.org/10.1108/IJSE-04-2015-0097

18. Chen, G. M., Kim, K. A., Nofsinger, J. R., & Rui, O. M. (2004). *Behavior and Performance of Emerging Market Investors: Evidence from China* (Unpublished Working paper). Washington State University. Retrieved from http://cicfconf.org/past/cicf2006/cicf-2006paper/20060104133427.pdf

19. Christie, W. G., & Huang, R. D. (1995). Following the Pied Piper: Do Individual Returns Herd Around the Market? *Financial Analysts Journal, 51*(4), 31-37. https://doi.org/10.2469/faj.v51.n4.1918

20. Collins, D. W., Gong, G., & Hribar, P. (2003). Investor Sophistication and the Mispricing of Accruals. *Review of Accounting Studies, 8*(2), 251-276. https://doi.org/10.1023/A:1024417513085

21. Corporate Finance Institute (CFI). (n.d.) *What is a Sophisticated Investor?* Retrieved from https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/sophisticated-investor/

22. Davis, A. M. (2019). Biases in individual decision-making. In K. Donohue, E. Katok, & S. Leider (Eds.), *The handbook of behavioral operations* (pp. 151-198). Wiley Blackwell. Retrieved from https://psycnet.apa.org/record/2018-21882-005

23. Dhar, R., & Zhu, N. (2006). Up close and Personal: Investor Sophistication and the Disposition Effect. *Management Science, 52*(5), 726-740. https://doi.org/10.1287/mnsc.1040.0473

24. Elizabeth, J., Murhadi, W. R., & Sutejo, B. S. (2020). Investor Behavioral Bias Based on Demographic Characteristics. *Proceedings of the 17th International Symposium on Management (INSYMA 2020).* https://dx.doi.org/10.2991/aebmr.k.200127.002

25. Fehr, E., & Tyran, J. R. (2005). Individual Irrationality and Aggregate Outcomes. *Journal of Economic Perspectives, 19*(4), 43-66. https://doi.org/10.1257/089533005775196651

26. Fellner, G., & Krügel, S. (2012). Judgmental Overconfidence: Three Measures, One Bias? *Journal of Economic Psychology, 33*(1), 142-154. https://doi.org/10.1016/j.joep.2011.07.008

27. Feng, L., & Seasholes, M. S. (2005). Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets? *Review of Finance, 9*(3), 305-351. https://doi.org/10.1007/s10679-005-2262-0

28. Filiz, I., Nahmer, T., Spiwoks, M., & Bizer, K. (2018). Portfolio Diversification: the Influence of Herding, Status-Quo Bias, and the Gambler's Fallacy. *Financial Markets and Portfolio Management, 32*(2), 167-205. https://doi.org/10.1007/s11408-018-0311-x

29. Fuertes, A. M., Muradoglu, G., & Ozturkkal, B. (2014). A Behavioral Analysis of Investor Diversification. *The European Journal of Finance, 20*(6), 499-523. https://doi.org/10.1080/1351847X.2012.719829

30. Gervais, S., & Odean, T. (2001). Learning to be Overconfident. *The Review of Financial Studies, 14*(1), 1-27. Retrieved from http://faculty.haas.berkeley.edu/odean/papers%20current%20versions/learningrfs.pdf

31. Gigerenzer, G. U., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic Mental Models: A Brunswikian Theory of Confidence. *Psychological Review, 98*(4), 506-528. https://doi.org/10.1037/0033-295x.98.4.506

32. Glaser, M., Nöth, M., & Weber, M. (2004). Behavioral Finance. In D. J. Koehler and N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision-Making* (pp. 527-546). Blackwell, Hoboken, NJ. Retrieved from http://www.communicationcache.com/uploads/1/0/8/8/10887248/blackwell_handbook_of_judge-ment_and_decision_making_-_framing_loss_aversion_and_men-tal_accounting.pdf

33. Gode, D. K., & Sunder, S. (1993). Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. *Journal of Political Economy, 101*(1), 119-137. http://www.jstor.org/stable/2138676

34. Goo, Y. J., Chen, D. H., Chang, S. H. S., & Yeh, C. F. (2010). A Study of the Disposition Effect for Individual Investors in the Taiwan Stock Market. *Emerging Markets Finance and Trade, 46*(1), 108-119. https://doi.org/10.2753/REE1540-496X460110

35. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010) Multivariate Data Analysis (7th ed.). Pearson, New York.

36. Jamal, K., & Sunder, S. (1996). Bayesian Equilibrium in Double Auctions Populated by Biased Heuristic Traders. *Journal of Economic Behavior & Organization, 31*(2), 273-291. Retrieved from https://ideas.repec.org/a/eee/jeborg/v31y-1996i2p273-291.html

37. Kacperczyk, M., Nosal, J., & Stevens, L. (2019). Investor Sophistication and Capital Income Inequality. *Journal of Monetary Economics, 107,* 18-31. https://doi.org/10.1016/j.jmone-co.2018.11.002

38. Kawshala, B. A. H., Anuradha, P. S., & Shamil, M. M. (2020). Socio-Economic, Trading Sophistication and Self-Reflection on Investors' herding Bias: Evidence from Colombo Stock Exchange. *International Journal of Management, Innovation & Entrepreneurial Research, 6*(2), 128-138. https://doi.org/10.18510/ijmier.2020.6212

39. Kim, C., & Pantzalis, C. (2003). Global/Industrial Diversification and Analyst Herding. *Financial Analysts Journal, 59*(2), 69-79. https://doi.org/10.2469/faj.v59.n2.2515

40. Kumar, S., & Goyal, N. (2016). Evidence on Rationality and Behavioural Biases in Investment Decision Making. *Qualitative Research in Financial Markets, 8*(4), 270-287. https://doi.org/10.1108/QRFM-05-2016-0016

41. Lin, H. W. (2011). Elucidating Rational Investment Decisions and Behavioral Biases: Evidence from the Taiwanese Stock Market. *African Journal of Business Management, 5*(5), 1630-1641. Retrieved from https://academicjournals.org/article/ar-ticle1380559829_Lin.pdf

42. Menkhoff, L., Schmidt, U., & Brozynski, T. (2006). The Impact of Experience on Risk Taking, Overconfidence, and Herding of Fund Managers: Complementary Survey Evidence. *European Economic Review, 50*(7), 1753-1766. https://doi.org/10.1016/j.euroecorev.2005.08.001

43. Merkle, C. (2017). Financial Overconfidence over Time: Foresight, Hindsight, and Insight of Investors. *Journal of Banking & Finance, 84,* 68-87. https://doi.org/10.1016/j.jbank-fin.2017.07.009

44. Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin, 105*(1), 156-166. https://doi.org/10.1037/0033-2909.105.1.156

45. Mushinada, V. N. C., & Veluri, V. S. S. (2020). Self-attribution, Overconfidence and Dynamic Market Volatility in Indian Stock Market. *Global Business Review, 21*(4), 970-989. https://doi.org/10.1177/0972150918779288

46. Odean, T. (1998). Volume, Volatility, Price and Profit When All Traders Are Above Average. *The Journal of Finance, 53*(6), 1887-1934. https://doi.org/10.1111/0022-1082.00078

47. Pope, D. G., & Sydnor, J. R. (2015). Behavioral Economics.

In G. Keren and G. Wu (Eds.), *The Wiley Blackwell Handbook of Judgment and Decision Making.* John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118468333.ch28

48. Prosad, J. M., Kapoor, S., & Sengupta, J. (2015). Behavioral biases of Indian investors: a survey of Delhi-NCR region. *Qualitative Research in Financial Markets, 7*(3), 230-263. https://doi.org/10.1108/QRFM-04-2014-0012

49. Prosad, J. M., Kapoor, S., Sengupta, J., & Roychoudhary, S. (2018). Overconfidence and Disposition Effect in Indian Equity Market: An Empirical Evidence. *Global Business Review, 19*(5), 1303-1321. https://doi.org/10.1177/0972150917726660

50. Sarkar, A. K., & Sahu, T. N. (2018). *Investment Behaviour: Towards an Individual-Centred Financial Policy in Developing Economies.* Emerald Group Publishing.

51. Shanmugham, R., & Ramya, K. (2012). Impact of Social Factors on Individual Investors' Trading Behaviour. *Procedia Economics and Finance, 2,* 237-246. https://doi.org/10.1016/S2212-5671(12)00084-6

52. Shusha, A. A., & Touny, M. A. (2016). The Attitudinal Determinants of Adopting the Herd Behavior: An Applied Study on the Egyptian Exchange. *Journal of Finance and Investment Analysis, 5*(1), 55-69. Retrieved from http://www.scienpress.com/Upload/JFIA%2fVol%205_1_3.pdf

53. Smith, V. L. (2009). *Rationality in Economics.* Cambridge Books.

54. Tekçe, B., Yılmaz, N., & Bildik, R. (2016). What Factors Affect Behavioral Biases? Evidence from Turkish Individual Stock Investors. *Research in International Business and Finance, 37,* 515-526. https://doi.org/10.1016/j.ribaf.2015.11.017

55. Thaler, R. H. (1999). Mental Accounting Matters. *Journal of Behavioral Decision Making, 12*(3), 183-206. https://doi.org/10.1002/(SICI)1099-0771(199909)12:3%3C183::AID-BDM318%3E3.0.CO;2-F

56. Toyokawa, W., Whalen, A., & Laland, K. N. (2019). Social Learning Strategies Regulate the Wisdom and Madness of Interactive Crowds. *Nature Human Behaviour, 3*(2), 183-193. https://doi.org/10.1038/s41562-018-0518-x

57. Vo, X. V., & Phan, D. B. A. (2017). Further evidence on the herd behavior in Vietnam stock market. *Journal of Behavioral and Experimental Finance, 13,* 33-41. https://doi.org/10.1016/j.jbef.2017.02.003

58. Weber, M., & Camerer, C. F. (1998). The Disposition Effect in Securities Trading: An Experimental Analysis. *Journal of Economic Behavior & Organization, 33*(2), 167-184. https://doi.org/10.1016/S0167-2681(97)00089-9

59. Wubie, A. W., Dibabe, T. M., & Wondmagegn, G. A. (2015). The Influence of Demographic Factors on Saving and Investment Decision of High School Teachers in Ethiopia: A Case Study on Dangila Woreda. *Research Journal of Finance and Accounting, 6*(9), 64-68. Retrieved from https://core.ac.uk/download/pdf/234630741.pdf

60. Yadav, A., & Narayanan, G. B. (2021). Do Personality Traits Predict Biasedness While Making Investment Decisions? *International Journal of Accounting & Finance Review, 6*(1), 19-33. https://doi.org/10.46281/ijafr.v6i1.939

61. Zhang, J., & Liu, P. (2012). Rational Herding in Microloan Markets. *Management Science, 58*(5), 892-912. https://doi.org/10.1287/mnsc.1110.1459