Journal Pre-proof

When financial literacy meets textual analysis: A conceptual review

Xiao Li

Please cite this article as: X. Li, When financial literacy meets textual analysis: A conceptual review. *Journal of Behavioral and Experimental Finance* (2020), doi: https://doi.org/10.1016/j.jbef.2020.100402.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

When Financial Literacy Meets Textual Analysis: A Conceptual Review

Xiao Li

Email: xiaoli@nankai.edu.cn
School of Finance, Nankai University, Tianjin, 300350, P.R. China

Abstract

Financial Literacy plays a crucial role in individuals' decision-making process. Existing studies have shown that lower level of financial literacy lead to irrational financial decision on investments, pension funds as well as savings and debts. While financial literacy is difficult to measure referring to its qualitative nature, constructing appropriate measurement for financial literacy is vital for both market professionals as well as policy makers. This study focuses on reviewing the measurements of financial literacy commonly used in current literature as well as the corresponding limitations. Then, the study surveys on the recent development and applications of textual analysis in finance domains, through which proposes future directions that textual analysis and financial literacy may merge for the purposes of better understanding about financial literacy as well as further developing effective schemes to improve investors' financial literacy.

Keywords: Financial Literacy; Textual Analysis; Financial Decision-making Process; Machine Learning; Survey-based measurements
JEL classification: G12; G14

When Financial Literacy Meets Textual Analysis: A Conceptual Review

Abstract

Financial Literacy plays a crucial role in individuals' decision-making process. Existing studies have shown that lower level of financial literacy lead to irrational financial decision on investments, pension funds as well as savings and debts. While financial literacy is difficult to measure referring to its qualitative nature, constructing appropriate measurement for financial literacy is vital for both market professionals as well as policy makers. This study focuses on reviewing the measurements of financial literacy commonly used in current literature as well as the corresponding limitations. Then, the study surveys on the recent development and applications of textual analysis in finance domains, through which proposes future directions that textual analysis and financial literacy may merge for the purposes of better understanding about financial literacy as well as further developing effective schemes to improve investors' financial literacy.

1. Introduction

Financial literacy is a concept relating to individuals' understanding of financial knowledge and their ability to use financial skills to improve economic well-beings. It is similar to the concept of financial knowledge in terms of measuring difficulties as a qualitative human capital. However, it also goes beyond financial knowledge referring to its additional practical dimension of measuring individuals' ability to apply financial knowledge to their decision-making process (Huston, 2010). Recent studies have provided ample evidence that financial literate individuals equipped with knowledge and skills are able to better allocate limited resources through making optimal financial decisions in various domains, i.e., retirement planning (e.g., Lusardi and Mitchell, 2007; Alessie, et al., 2011; Almenberg and Save-Soderbergh, 2011), investment decisions (e.g.,van Rooij et al., 2011; Bannier and Neubert, 2016; Zhang et al., 2020), professional advice seeking (e.g., Calcagno and Monticone, 2015; Allgood and Walstad, 2016; Gerrans and Hershey, 2017; Chauhan and Dey, 2020), debt decisions (e.g., Lusardi and Bassa-Scheresberg, 2013; Lusardi and Tufano, 2015), etc. Therefore, improving financial literacy has become a crucial target for public policy makers and regulators for better market efficiency and social welfares.

As a matter of fact, the qualitative nature brings difficulties to developing effective measurement for financial literacy. Generally, survey-based measurement for financial literacy is the norm in current literature. However, two major problems need to be considered. Firstly, the survey-based measurement is inevitably involved with low rate of responds and inaccurate answers. Second, surveys are usually conducted with relatively low frequencies. That is, survey questions only capture the static financial literacy levels and fail to record the development or trend of financial literacy on a timely basis. Textual analysis, also known as text mining, refers to the process of extracting useful information from narrative contents. Through textual analysis, both intended and unintended information can be extracted on a timely basis. In that sense, textual analysis has the potential to resolve the above limitations of survey-based measurements for financial literacy. In fact, an increasing number of current studies has

2

started to use textual analysis to measure the naturally hard-to-value qualitative elements such as article readability and investor sentiment. Thus, the main objective of this review is to discuss future potentials of textual analysis in measuring financial literacy.

As a combination of financial literacy with textual analysis, this review comprises of two phases. Figure 1 shows the detailed logical relationship between financial literacy and textual analysis where this study starts from. As is shown in Figure 1, the first phase surveys current literature on financial literacy. We begin by explaining the concept of financial literacy with both the understanding of financial knowledge and the ability to apply the knowledge for better financial decisions. Second, we discuss the current measurements for financial literacy and the corresponding limitations, with an emphasis on the most commonly used survey-based measurements. We then survey the literature on the effects of financial literacy on individuals' finance decisions, including retirement planning, investment decisions, professional advice seeking and debt decisions. On the other hand, the second phase generates a review on current development of textual analysis. Similar to the first phase, we also start with an explanation of the concept and general process for textual analysis. We then review on the applications of textual analysis in finance domains. Generally, such applications fall into the topics of information contents examination, readability measurement and sentiment analysis. Finally, we envisage future directions through discussing how textual analysis can be used in better capturing financial literacy.

As a review-based qualitative study, we contribute to the literature on measurements of financial literacy by firstly introducing the textual analysis techniques. Also, we provide conceptual interpretation and structures for future empirical studies on the crossing field comprises of financial literacy and textual analysis. Moreover, since improving financial literacy has become one of the major objectives of public policies, our study provides hints to the regulators in terms of constructing effective scheme to reduce and eliminate "financial illiteracy".

Figure 1 Logic diagram of financial literacy and textual analysis
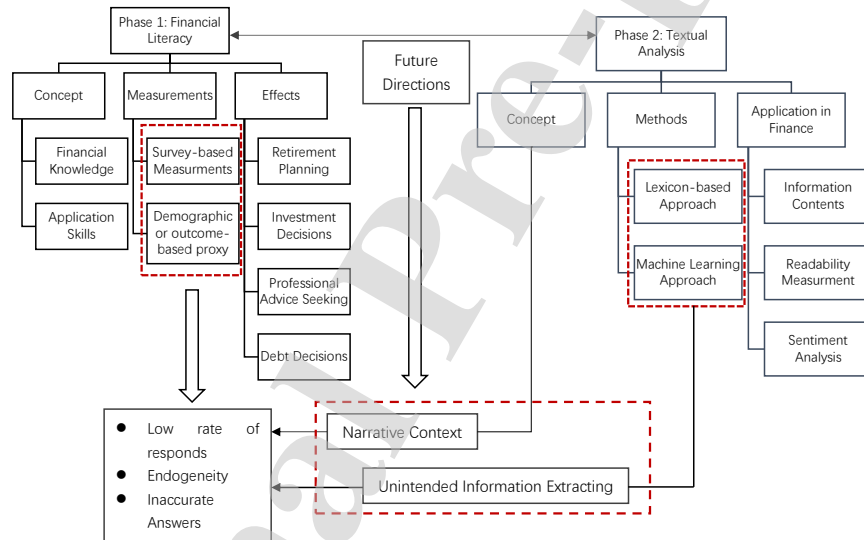
2. What is Financial Literacy

We trace back to the definition of the fundamental (general) literacy for the purpose of a thorough understanding for "financial literacy". According to Zarcadoolas et al. (2006), the fundamental or general literacy refers to an individual's ability to read and write. An illiterate individual does not necessarily mean that he cannot speak, think logically or interpret other forms of nonprint information. Rather, the term illiteracy is actually related to the lack of practical skills to print the information or knowledge. For example, the Literacy Definition Committee and National Adult Literacy Survey define illiteracy as "using printed and written information to function in society, to achieve one's goals, and to develop one's knowledge and potential". In that sense, the concept of fundamental literacy may include three aspects — the ability to prose (print information), document (tabulate information) and quantify (produce arithmetic and numerical information) (Huston, 2010).

Having conducted a detailed review on the concepts of financial literacy, Hung et al. (2009) argue that previous literature has proposed different explanations for financial literacy and there lacks a universally accepted definition for the term. Earlier studies regard financial literacy as an interchangeable term for financial knowledge that refers to the individuals' understanding or familiarity with basic financial principles, instruments and institutions. On the contrary, Hung et al. (2009) describes financial literacy as the "knowledge of basic economic and financial concepts, as well as the ability to use that knowledge and other financial skills to manage financial resources effectively for a lifetime of financial wellbeing" (p.12). Huston (2010) and Atkinson and Messy (2012) hold similar opinion that financial literacy goes beyond financial knowledge and they relate financial literacy to not only individuals' general understanding of financial concepts, but also individuals' practical skills to process, synthesize as well as investigate the implicit contents of financial information. Recall the concept of general literacy, defining financial literacy from both the knowledge and application dimensions coincides with the concept of general literacy through capturing

three basic aspects — ability to prose, document and quantify information. More importantly, such definition of financial literacy is also employed by the Organization for Economic Cooperation and Development (OECD, 2014)[1] in the 2012 Program for International Student Assessment (PISA). Therefore, in this paper, we regard financial literacy as a different term with financial knowledge and we view the concept from two dimensions: one's understanding for financial principles and the ability to apply financial knowledge to making better financial decisions.

---

[1] The detailed definition of financial literacy provided by the OECD(2014) is as follows: "Financial literacy is knowledge and understanding of financial concepts and risks, and the skills, motivation and confidence to apply such knowledge and understanding in order to make effective decisions across a range of financial contexts, to improve the financial well-being of individuals and society, and to enable participation in economic life" (p.33).

3. How to measure financial literacy

The main purpose of this section is to conduct a review on the common measurements for financial literacy, i.e., survey-based measurement, self-assessed measurement and other measurement. Notably, survey-based measurement and self-assessed measurement are closely related to each other and they capture the financial literacy from different dimensions. We then introduce other measurements for financial literacy that has been used in current literature, such as demographic-based proxy and outcomes-based proxy. Lastly, we conduct a discussion on limitations of current measurements.

3.1 Survey-based Measurement

A large number of studies use household surveys to measure individuals' financial literacy. Though the detailed constructions of such surveys vary across different markets, the survey questions often contain three domains. The first domain refers to individuals' knowledge and understanding of financial instruments, i.e., stocks, funds, bonds, or mortgages. The second domain relates to the knowledge of basic financial principles. For example, individuals' understanding about the time value of money and inflations. According to Lusardi (2012), making optimal financial decisions requires the capacity to do calculations, even the complex ones. In that sense, the third domain often focuses on testing individuals' mathematical and numerical skills.

As pioneer scholars in this field, Lusardi and Mitchell (2011a) design a special module for the Health and Retirement Study (HRS)[2] aiming to examine the level of financial literacy of households. The questions they proposed are known as the "Big Three" and have been widely applied in the U.S and other countries in assessing individuals' financial literacy (Stopler and Walter, 2017; Lusardi and Mitchell, 2017). Based on the "Big Three", the OECD develops the INFE (International Network on Financial Education) Core Questionnaire to measure financial literacy internationally.

---

[2] The Health and retirement Study (HRS) is a national survey for Americans over the age of 50. The survey is comprised of questions relating to health conditions, assets, liabilities and patterns of wellbeing.

Since the questionnaire is designed to adjust demographic characteristics of world-wide countries, the concrete questions are displayed differently from the "Big Three" (Atkinson and Messy, 2012).

Referring to Lusardi and Mitchell (2008), the "Big Three" questions are as follows[3]:

*Question 1: Suppose you had $100 in a saving account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow: **more than $102**, exactly $102, less than $102, don't know, refuse to answer.*

*Question 2: Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would be able to buy: more than, exactly the same as, **less than today's money in the account,** don't know, refuse to answer.*

*Question 3: Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock mutual fund." (true, **false**, don't know, refuse to answer)*

The first two questions examine respondents' understanding about inflation and whether they are able to do simple financial calculations. The third question investigates respondents' knowledge for common financial instruments and risk diversification. Notably, Lusardi and Mitchell (2011a) observe that only 34.3% of the respondents answer all the three questions correctly. The accuracy rates for the three questions are 67.1%, 75.2% and 52.3%, respectively. Moreover, 33.7% respondents do not know the answer to the third question, which is the largest percentage compared with the first and second questions. It is not surprising since providing the correct answer to the third question requires professional knowledge about financial instruments. In general, such results indicate low performance on these basic financial literacy questions across the U.S older households in the sample.

---

[3] The correct answers are labeled in bold.

Following Lusardi and Mitchell (2011a), a strand of studies uses the "Big Three" questions as the baseline to construct adjusted surveys to measure financial literacy in different countries. Table 1 displays such studies and surveys. Typical surveys using the "Big Three" questions include the Survey on Household Income and Wealth (SHIW), Sparen und AltersVorsorgE in Deutschland (SAVE), Social Protection Survey (EPS), Russia Financial Literacy Diagnostic Survey (RFLDS), DNB Household Survey, etc. Also, the "Big Three" questions are widely applied in both developed economies (the U.S., Italy, France, Netherland, Germany, New Zealand, Sweden) and less developed economies (Romania, Russia, Chile). Referring to the listed survey scales, Chile has the largest number of respondents whereas New Zealand has the lowest number of respondents. Moreover, respondents in Germany and Netherland are most financial literate since 53% and 45% of the respondents provide correct answers to all the three questions respectively. On the contrary, Russia and Romania have the least financial literate respondents with only 3% and 4% of the respondents have the ability to answer three questions correctly. In general, Table 1 indicates that the more developed of the economy, the more financial literate of the individuals.

Table 1 Surveys and studies based on the "Big Three" questions.

Year of Data	Survey Name	Country	Authors	Survey Questions	Additional Questions	All questions correct	No. of Observations
2006	Survey on Household Income and Wealth (SHIW)	Italy	Forneo and Monticone (2011)	The "Big Three"	No	25%	3992
2009	Sparen und AltersVorsorgE in Deutschland (SAVE)	Germany	Bucher-Koenen and Lusardi (2011)	The "Big Three"	One additional question related to retirement plan	53%	1059
2009	ANZ/Retirement Commission Financial Knowledge Survey	New Zealand	Crossan et al. (2011)	The "Big Three"	No	24%	850
2009	Social Protection Survey (EPS)	Chile	Hastings et al. (2013)	The "Big Three"	No	8%	14243
2009	Russia Financial Literacy Diagnostic Survey (RFLDS)	Russia	Klapper and Panos (2013)	The "Big Three"	No	3%	1366
2010	Survey of Living Preferences and Satisfaction (SLPS)	Japan	Sekita (2011)	The "Big Three"	Self-assessment	27%	5268
2010	DNB Household Survey	Netherlands	Alessie et al (2011)	The "Big Three"	No	45%	1665
2010	Financial Supervisory Authority—Consumer Survey	Sweden	Almenberg and Save-Soderbergh (2011)	The "Big Three"	One additional question related to retirement plan	21%	1302
2011	PATER (PATrimoines et Préférences face au TEmps et au Risque) Household Survey	France	Arrondel et al. (2013)	The "Big Three"	One additional question related to retirement plan	31%	3616
2011	Austrian Central Bank (OeNB) – Euro Survey	Romania	Beckmann (2013)	The "Big Three"	No	4%	1030
2013	Own calculation (FinLit)	USA	Clark et al. (2017)	The "Big Three"	Two additional questions related to tax and 401(k) plan	33.3%	16000

10

To better applicate the "Big Three" questions in formulating surveys, a number of studies extends the domain of "Big Three" to capture other dimensions of financial literacy. The 2009 National Financial Capability Study (NFCS) adds too additional questions regarding to individuals' knowledge about bond pricing and mortgages and form the collectively known "Big Five" questions (Hastings et al., 2013). The additional questions are described as follows:

*Question 4: Do you think that the following statement is true or false: A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest over the life of the loan will be less. (**true**, false, don't know, refuse to answer)*

*Question 5: If interest rates rise, what will typically happen to bond prices: they will rise, **they will fall**, they will stay the same, there is no relationship, don't know, refuse to answer.*

Compared with the "Big Three", the "Big Five" captures individuals' understanding of concrete financial instruments which requires professional financial education in most cases. Therefore, it is reasonable to regard the last two questions in the "Big Five" as advanced questions for measuring financial literacy. In a more recent study, Lusardi and Mitchell (2017) develops a complete set of 13 questions and divide the questions into two main categories, i.e., basic and sophisticated financial literacy questions. The basic financial literacy questions refer to the "Big Three" while the sophisticated questions are an extended version of the two additional questions in the "Big Five". Table 2 contains the detailed construction of sophisticated financial literacy questions and the percentage of respondents providing correct answers to each question respectively. Though the accuracy rate for each question is relatively high[4], Lusardi and Mitchell (2017) report that only 16.5% of the 989 respondents have the ability to answer all the 8 sophisticated questions correctly. Following Lusardi and Mitchell (2017), Deuflhard et al. (2019) combines the sophisticated financial literacy questions with the

---

[4] 7 out of 8 questions have accuracy rate higher than 50%, indicating that for each independent question, more than half of the respondents have the ability to provide correct answers. The only exception is the question relating to the relation between interest rate and bond prices, indicating respondents lack of knowledge for bond pricing.

DNB Household Survey. Another application is the Chinese Survey of Consumer Finance (CSCF) that examines the financial literacy level of Chinese individuals. Since Chinese financial market is generally led by banks—which is slightly different from other markets, the CSCF include questions relating to individuals' understanding for the function of central bank as an adjustment.

Having collected the respondents' answers for survey questions, different methods are then applied to aggregate the answers to obtain the individual level of financial literacy. Such methods include simple indicator variables (Jappelli, 2010; Gathergood, 2012), principal component analysis (Behrmann et al., 2012; Klapper et al., 2013; Lusardi et al., 2014), iterated principle factor analysis (van Rooij et al., 2011) and cluster analysis (Lusardi and Tufano, 2015).

Table 2 Sophisticated Financial Literacy Questions, Answers and Percent Correct

Question domain	Question construction	Answers	Percent Correct
Stock Market Functioning	Which of the following statement describes the main function of the stock market?	The stock market helps to predict stock earnings; The stock market results in an increase in the price of stocks; The stock market brings people who want to buy stocks together with those who want to sell stocks; None of the above; Don't know; Refuse to answer	71.5%
Knowledge of Mutual Funds	Which of the following statement is correct?	Once one invests in a mutual fund, one cannot withdraw the money in the first year; Mutual funds can invest in several assets, for example invest in both stocks and bonds; Mutual funds pay a guaranteed rate of return which depends on their past performance; None of the above; Don't know; Refuse to answer	63%
Interest Rate/Bond Prices Link	If the interest rate falls, what would happen to bond prices?	Rise; Fall; Stay the same; None of the above; Don't know; Refuse to answer	31.6%
Safer: Company Stock or Mutual Fund	True or false? Stocks are normally riskier than bonds.	True; False; Don't know; Refuse to answer.	71.4%
Riskier: Stocks or Bonds	True or false? Stocks are normally riskier than bonds.	True; False; Don't know; Refuse to answer.	80.2%
Long Period Returns	Considering a long period (for example 10 or 20 years), which asset normally gives the highest return?	Savings accounts; Bonds; Stocks; Don't know; Refuse to answer.	62.3
Highest Fluctuation/Volatility	Normally, which asset displays the highest fluctuations over time?	Savings accounts; Bonds; Stocks; Don't know; Refuse to answer.	88.3%
Risk Diversification	When an investor spreads his money among different assets, does the risk of losing money:	Increase; Decrease; Stay the same; Don't know; Refuse to answer.	74.9%

3.2 Self-assessed measurement

Surveys based on the "Big Three" or "Big Five" are indeed test-based measurement to examine individuals' actual financial literacy. A large number of studies has shown that individuals tend to be more optimistic about how much they really know (e.g., Agnew and Szykman, 2005). Such subjective self-assessment of financial literacy is usually labeled as the "perceived" financial literacy which has also been proved to have significant effects on financial behavior separately from the actual financial literacy in various domains: stock market participation (van Rooij et al., 2011), retirement planning and savings (Parker et al., 2012) as well as riskier investment participation such as discount certificates and hedge funds (Bannier and Neubert, 2016). On the other hand, perceived financial literacy measures individuals' financial confidence. Individuals with high perceived financial literacy and low actual financial literacy are regarded as overconfident. Referring to the literature on individuals' overconfidence and financial behaviors, overconfidence induces individuals to take more risks and conduct biased financial decisions since they believe they have advantages in predicting asset prices (e.g., Odean, 1998; Barber and Odean, 2001). Taken together, self-assessed financial literacy is as important as the test-based financial literacy in explaining financial consequences (Allgood and Walstad, 2016).

In that sense, it is prevalent among current studies to combine the test-based financial literacy with the self-assessed financial literacy when constructing survey questions. For example, Hung et al. (2009) evaluates 18 studies and report that 6 (one-third) of them use self-assessed measurement in addition to the test-based measurement to better capture financial literacy. The description of a typical self-assessed question is usually worded as follows (Lusardi and Mitchell, 2014):

On a scale from 1 to 7, where 1 means very low and 7 means very high, how would you assess your overall financial knowledge?

Notably, studies may use adjusted self-assessment questions to examine the perceived financial literacy. For example, Hastings et al. (2013) develop three-

dimensional self-assessment questions, i.e., self-assessed overall financial knowledge, self-assessed numerical knowledge and self-assessed skills of dealing with financial matters. The results are still comparable since they observe that such different constructions of self-assessments are highly correlated with each other. Another interesting question is to explore how self-assessed (perceived) financial literacy relates to the test-based (actual) financial literacy. As a subjective financial literacy measurement, the self-assessed financial literacy is found to be positively correlated to the more objective test-based measurement (Lusardi and Mitchell, 2009; Parker et al., 2012). To be specific, Agnew and Szykman (2005) document that the correlations between actual and perceived financial literacy vary from 0.10 to 0.78 across demographic groups, and the median correlation is 0.49. Therefore, it is better to combine the actual financial literacy with perceived financial literacy and cautions should be taken when using the latter to replace the former.

The self-assessed financial literacy is subjective to individuals' demographics. In general, older respondents tend to be more confident. Compared with younger respondents, the older respondents generate poorer performance in answering the test questions. However, they give themselves higher scores in the self-assessment question (Lusardi and Tufano, 2015; Gamble et al., 2015). Bucher-Koenen et al. (2014) find strikingly similar gender differences in self-assessed financial literacy across countries including the U.S., Germany and Netherlands. Specifically, they observe that women have lower self-assessed financial literacy compared with men, indicating that men are more likely to be overconfident. Through analyzing the data from SAVE survey, Bannier and Neubert (2016) observe contradicting result for gender difference in self-assessed financial literacy in Germany. Interestingly, they observe that women are more numerous than men in the group with high perceived whereas low actual financial literacy, indicating that women are more likely to rate their financial literacy as above average though the actual level is below average.

3.3 Other measurements

Survey-based measurement (test-based and self-assessed) for financial literacy is the norm in literature, while other approaches including demographic-based proxy and outcomes-based proxy may also be worth noting.

Since the robust correlation between demographic characteristics and financial literacy has been proved, several studies turn to use respondents' demographic characteristics to proxy for financial literacy instead of surveys. Different demographic characteristics are usually used as a combined dataset to measure financial literacy, including age (Calvet et al., 2007; Calvet et al., 2009; Georgarakos and Pasini, 2011), disposable income and wealth conditions (Dhar and Zhu, 2006; Calvet et al., 2007; Calvet et al., 2009), educational backgrounds (Calvet et al., 2007; Christiansen et al., 2008; Calvet et al., 2009) as well as professional status (Calvet et al., 2009). Notably, a two-step methodologically advanced approach is conducted by Stolper (2018) to examine the predicted financial literacy through combining the direct survey-based measurement with demographic-based proxies. Firstly, the author estimates an empirical model for financial literacy using a supplementary data source containing both a direct measurement of financial literacy and several demographic characteristics that have been proved as explanatory factors for cross-sectional variation in individuals' financial literacy levels[5]. Secondly, the author applies this model to predict the financial literacy levels of the households in the primary sample, where only the demographic characteristics associated with financial literacy are available.

Alternative to the demographic-based proxy, the quality and outcomes of individuals' financial decisions are used by several studies to measure financial literacy. For example, individuals that have the ability to conduct financially sophisticated behavior such as consistently refinancing a mortgage when interest rates fall, are regarded as financial literate and vice versa. Other financially sophisticated behavior

---

[5] The demographic characteristics used by the author include age, gender, education, employment status, income and wealth. Such factors are consistent with the key demographic characteristics relating to financial literacy which has been documented by Lusardi and Mitchell (2014).

may include managing risk diversification in portfolios (Grinblatt and Keloharju, 2001; Goetzmann and Kumar, 2008), successfully investing in complex financial instruments (Goetzmann and Kumar, 2008) as well as prior investment performance (Goetzmann and Kumar, 2008; Nicolosi et al., 2009), etc. Taking Sweden as an example, Calvet et al. (2009) formulate a financial sophistication index based on the examination of sampled households' ability to avoid three typical financial mistakes: under-diversification in managing portfolios, inertia in risk taking and involving disposition effect when holding stocks. Compared with standard survey-based measurement, such outcomes-based measurement has the advantage to directly unfold individuals' real financial literacy level through tracking their financial behaviors, rather than simply depending on limited scope of survey questions. However, the outcomes-based measurement hinges more on the availability of valid dataset and thus, it may be used as a fruitful way to predict future behavior and performance (Hastings et al., 2013; Stolper and Walter, 2017).

3.4 Limitations of current measurements

Although the survey-based measurement has become the international benchmark for assessment of financial literacy, there are several limitations. As a rule of thumb, surveys have the natural shortages of low rate of responds and biased or inaccurate answers. Specifically, individuals are sensitive to the construction of questions. Lusardi and Mitchell (2011b) and van Rooij et al. (2011) find that wording has a significant impact on respondents' answers. Such phenomenon is more likely to happen when respondents answer questions based on their intuitive guesses. For example, individuals may treat "buying company stock usually provides a safer return than a stock mutual fund" and "buying a stock mutual fund usually provides a safer return than a company stock" differently though the latter is simply a reverse version of the former. In that sense, answers classified as "correct" do not necessarily mean that individuals are financial literate (Lusardi and Mitchell, 2014).

Another limitation of survey-based measurement refers to its low frequency, i.e., surveys are conducted yearly in most cases. In that sense, the survey data cannot capture the changes or development of individuals' financial literacy within a shorter period of time on a continuous basis. What captures by the survey is the static documentation of financial literacy at a specific time point when the survey is conducted. Such limitation also leads to the situation where majority of the current studies mainly focuses on analyzing financial literacy on a cross-sectional basis, rather than on a time-series basis. Compared with the static conclusion, it is more important to disclose the dynamic changes of financial literacy and explore its effects on individuals' financial decisions. Moreover, understanding the trend of individuals' financial literacy helps the government or policy makers construct more effective schemes on financial education to further eliminate financial illiteracy as well as improve the market efficiency.

As for the alternative measurement such as the demographic-based proxy and outcome-based proxy, exogenous problems are inevitable which needs cautious treatments. Specifically, Hastings et al. (2013) argue that the outcome-based proxy may

be more appropriate in forecasting future trading behavior, rather than being treated as

an alternative measurement for financial literacy.

4. Why Financial Literacy Matters?

A large number of studies has provided ample evidence showing that financial literacy plays a crucial role in affecting individuals' financial decisions. As we will review in this section, financial literate individuals equipped with knowledge and skills are able to better allocate limited resources through making optimal financial decisions in various domains, i.e., retirement planning, investment decisions, professional advice seeking, debt decisions, etc. Through developing an augmented stochastic life cycle model, Lusardi et al. (2017) further prove that financial literacy can explain 30-40 percent of wealth inequality in the U.S market, indicating the importance of financial literacy in improving individuals' lifetime economic welfare. Therefore, improving individuals' financial literacy is of great importance for governments and policy makers.

4.1 Financial Literacy and Retirement Planning

Essential financial knowledge and skills are necessary for individuals to understand the pension system and conduct retirement plans. Generally, studies have shown that individuals' financial literacy has positive effect on retirement planning (Lusardi and Mitchell, 2007; Alessie, et al., 2011; Almenberg and Save-Soderbergh, 2011; Bucher-Koenen and Lusardi, 2011; Lusardi and Mitchell, 2011a; Lusardi and Mitchell, 2011b; Sekita, 2011; Fornero and Monticone, 2011; Lusardi and Mitchell, 2017).

As one of the pioneer studies in this field, Lusardi and Mitchell (2007) focus on the U.S market and document that individuals' financial literacy level is positively related to retirement planning and further affect the retirement wealth after controlling for many socio-demographic factors. Also, Lusardi and Mitchell (2011b) observe similar phenomenon that respondents answer the "Big Three" questions correctly are much more likely to calculate how much they need to save for retirement. However, some may argue about possible reverse causality that financial literacy may be resulting from individuals' choice for retirement plans. That is, individuals who plan for retirement may be more willing to invest in financial education which further improves their financial literacy. At this point, Lusardi and Mitchell (2011b) re-estimate the impact of

financial literacy on retirement plans using the instrumental variables approach and find a consistently positive relationship between financial literacy and retirement planning. That is, they do not find the opposite reverse causality. Given the fact that there are many differences in pension systems across countries, the positive effect of financial literacy on retirement planning remains remarkably consistent in countries including Germany (Bucher-Koenen and Lusardi, 2011), Sweden (Almenberg and Save-Soderbergh, 2011), Netherland (Alessie, et al., 2011), Japan (Sekita, 2011) and Italy (Fornero and Monticone, 2011). Either based on the "Big Three" or "Big Five" questions, such studies generally show that whether individuals have the ability to provide correct answers to questions relating to interest rate and risk diversification matters most for retirement planning. This is quite reasonable since understanding of interest rate characterizes individuals' basic numerical skills to figure out how much to save for retirement and the knowledge of risk diversification indicates individuals' ability to effectively manage the retirement plans. Having conducted a review on the worldwide evidence for financial literacy, Lusardi and Mitchell (2011c) conclude that in Netherlands, a correct answer to an additional financial literacy question generates 10 percentage point higher probability of retirement planning. As for the U.S., Germany, Japan and Sweden, individuals answering one additional question correctly only associates a 3-4 percentage point higher possibility of conducting retirement plans. Therefore, in general, the effect of financial literacy is the highest in the Netherlands among the above countries.

In a more recent study, Lusardi and Mitchell (2017) not only provide an additional evidence for the positive effect of financial literacy on retirement plans, but also prove a positive effect of self-assessed financial literacy (perceived financial literacy) on individuals' tendency to conduct retirement plans. In addition, Clark et al. (2017) take the 401(k) plan as an example and show that the higher the individuals' financial literacy levels, the more likely they participate in the plan and obtain investment profits.

4.2   Financial Literacy and Investment Decisions

Understanding of financial principles and products determines individuals' investment decisions on choosing different assets, managing risks, diversifying investment portfolios, etc. Specifically, financial literate individuals tend to have higher risk tolerance and are more likely to take risks. Bannier and Neubert (2016) show that men are more financial literate than women and they are more willing to invest in sophisticated financial assets with higher risks. In a similar vein, Zhang et al. (2020) provide additional empirical evidence for the positive relationship between financial literacy and risk-taking behavior.

In regarding to how financial literacy affect individuals' investment decision on risky assets, a robust finding is that financial literacy is positively related to stock market participation. Having considered various individuals' socio-demographic characteristics (e.g., gender, income, wealth, education backgrounds), van Rooij et al. (2011) conclude that lack of financial knowledge is the key reason for "stock participation puzzle" in Netherland, i.e., households do not hold stocks. According to their empirical results, a one standard deviation increase in advanced financial literacy improves the stock market participation by roughly 9 percentage points. Based on the U.S dataset, Yoong (2011) also finds that financial illiteracy acts as an impediment to stock participation. More importantly, consistent results emerge when comparing stocks with other assets. For example, Almenberg and Widmark (2011) relate financial literacy to households' investment in two assets — stocks and real estates. Typically, they observe that when taking into consideration of demographic characteristics, the positive relationship between financial literacy and stock market participation still remains whereas housing market participation is no longer significantly related to financial literacy. Similarly, Clark et al. (2017) investigates how financial literacy affects individuals' construction of 401 (k) retirement investments. The study compares stocks with various assets such as bonds, mutual fund, REITs fund, etc. The results show that as the financial literacy level increases, the percentage of stocks held in investment portfolios also increases. Above evidence of financial literate individuals' preference

22

for stocks rather than other relatively safer assets also confirms the role of financial literacy in determining risk tolerance. In addition, Almenberg and Dreber (2015) observe that women are more financial illiterate than men and are less likely to participate in stock market, indicating that financial literacy also helps explain the gender difference in stock market participation.

Similar to financial literacy, sociability has also been regarded as an important factor affecting stock market participation since it reduces the fixed cost of information through information sharing (Bonte and Filipiak, 2012). To this point, Balloch et al. (2014) construct a stock-based financial literacy index and further compare the effects of financial literacy as well as sociability on stock market participation. Their empirical results show that once the financial literacy is taken into consideration, sociability can no longer explain stock market participation. In the additional test, the authors construct two separate groups—highly sociable households that have low stock market literacy and lower sociable households with high stock market literacy. Specifically, they find no association between sociability and participation among highly sociable households with low stock market literacy. Also, households having low sociability and high financial literacy generate significant association with stock market participation. Such results indicate that financial literacy is more important than sociability in determining stock market participation.

Another strand of studies relating to financial literacy and investment decisions mainly focus on investigating how financial literacy affect the portfolio diversifications. For instance, Abreu and Mendes (2010) hypothesize that diversification behavior is determined by investors' financial literacy of digesting all the pertinent financial information. Notably, the empirical findings validate such hypothesis by showing that financial literacy significantly contributes to an increase in the number of assets included in investment portfolios. In a similar vein, Guiso and Jappelli (2009) prove that poor financial literacy is the main variable explaining the lack of portfolio diversification. To be specific, Clark et al. (2017) document that investors with highest financial literacy hold 18 percentage points more stocks than the counterparts with

23

lowest financial literacy. On the other hand, von Gaudecker (2015) uses the return loss[6] to measure the cost of under-diversification and observes that for individuals with no professional advice provided, one standard deviation decrease in financial literacy leads to an increase in return loss by 0.7 percentage points on average, indicating a positive relationship between financial literacy and portfolio diversification.

---

[6] Von Gaudecker (2015) employs the return loss as the diversification measure by following Calvet et al. (2007). The return loss is the difference between the maximum expected return attainable through conducting a fully-diversified portfolio at a given standard deviation and the actual expected return for a particular portfolio.

4.3   Financial Literacy and Advice Seeking

As is explained previously, lack of financial literacy impedes individuals from making sound decisions, i.e., constructing retirement plans and diversify investment portfolios. In that sense, professional advice from financial advisers is quite important and beneficial for financial illiterate individuals (Bhattacharya et al., 2012). In that sense, financial advice acts as a substitute to financial literacy and we may observe a negative relationship between financial literacy and advise seeking behavior. However, current studies conclude the opposite — financial literacy is positively related to advice seeking. The more financial literate the individuals are, the more likely they seek for professional advice.

Colllins (2012) concludes that financial advice should be regarded as a complement to financial literacy rather than a substitute since his empirical findings show that as income, education and financial literacy increases, advice seeking behavior also increases. In regard to the determinants for advice seeking behavior, Gerrans and Hershey (2017) introduce financial adviser anxiety (e.g., individuals' concerns about embarrassment or other forms of apprehension while meeting with a financial adviser) as another factor that may affect the association between financial literacy and advice seeking. Notably, they still find consistent results that financial literacy retains a positive relationship with the financial advice seeking behavior in the future. In fact, individuals' seeking for advice includes various domains relating to different financial matters. To this point, Allgood and Walstad (2016) examine how financial literacy positively relates to a set of advice including savings or investments, taking mortgages or loans, insurance, tax planning and debt counseling. As a result, individuals with high financial literacy are more likely to seek for investment advice, whereas individuals with low financial literacy tend to seek more tax planning advice. An exception to the positive relationships is the debt counseling advice, demonstrating that as financial literacy improves, the probability of seeking such advice reduces.

While above studies mainly focus on the U.S market, Calcagno and Monticone (2015) investigate such issues in Italian market. The study not only provides additional

evidence for the positive relationship between financial literacy and advice seeking, but also explores the underlying reasons for the phenomenon. Typically, they construct a stylized model comprised of two agents: a financial intermediary who sells a risky asset and provide advice to clients at the same time, and an investor. Through solving the model, the authors observe that the financial advisors only provide useful information to financial literate individuals who are anticipated to be more likely to consult them. On the other hand, financial advisors do not provide useful information to financial illiterate individuals. The empirical findings coincide with the model implications by showing a positive relationship between financial literacy and the probability of advice seeking though a number of socio-demographic factors have been taken into consideration. Therefore, the study concludes that the conflicts of interests and advisors' incentive to take advantage of clients lead to the positive relationship between financial literacy and advice seeking.

In summary, the robust positive relationship between financial literacy and advice seeking provide insights for policy makers. Improving financial literacy through education programs and reducing conflicts of interests are possible ways to motivate effective advice seeking. In addition to this topic, a recent study by Chauhan and Dey (2020) shows that individuals with high financial literacy weight more for financial advice and have high willingness to pay (WTP) for suitable advice, which further emphasizes on the importance of how financial literacy affects advice seeking behavior.

4.4   Financial Literacy and Debt decisions

Financial knowledge and skills are essential for individuals to interpret as well as understand credit-related financial products. Lack of financial literacy may exhibit confusion over different financial products, resulting in individuals' unawareness of their debt positions. In regard to this topic, a large number of studies has shown that individuals with low financial literacy are more likely to make poor decisions and fail to conduct effective debt management. For example, Lusardi and Bassa-Scheresberg (2013) find that low financial literacy is closely related to the high-cost borrowings in the U.S., i.e., financial literacy accounts for about 20 percentage of the reduction in the use of high-cost borrowing. Similarly, Lusardi and Tufano (2015) document a negative relationship between financial literacy and over-indebtedness, indicating that individuals with low financial literacy are unable to judge their debt position and may involve in excessive debt loads. Also, Disney and Gathergood (2013) provide an additional evidence for the negative relationship between financial literacy and participation in credit market from the U.K market. Specifically, they present that individuals with low financial literacy typically hold portfolios with larger shares of high cost credit such as home credit and payday loans.

Apart from high-cost borrowing, low financial literacy has also been proved to be associated with inefficient or inappropriate credit use. Referring to Allgood and Walstad (2013), compared with individuals with low financial literacy, those with high financial literacy are more likely to pay their credit card balance in full each month and are less likely to be charged additional fees for late payment or exceeding credit limit. Moreover, Klapper et al. (2013) observe that individuals with low financial literacy tend to engage more in informal financial markets and involve in informal sources of borrowings. In addition, Gerardi et al. (2014) even find that financial illiterate individuals, especially those with low numerical skills, are more likely to default on mortgages when there is a market depression.

27

5. Textual analysis meets financial literacy

5.1   What is Textual Analysis?

Textual analysis, also known as text mining, refers to the process of excavate useful information from the text. Generally, textual analysis may include computational linguistics, natural (or statistical) language processing, information retrieval, content analysis, and stylometrics (Loughran and Mcdonald, 2016). With the recent development of computing power, textual analysis has been permeated to the field of accounting and finance to quantify various qualitative information that are hard to value using traditional measurements. According to Guo et al. (2016), textual analysis includes three steps — text harvesting, text cleaning and text parsing or analyzing.

Text harvesting refers to the process of collecting multi-dimensional datasets. Traditionally, researchers may obtain data from financial database, i.e., Compustat, Bloomberg, Thompson Reuters, etc. Recent development of the Internet search engines and social networks allow researchers to obtain both structured and unstructured data from the Internet using web-crawling techniques. Such data used in accounting and finance discipline includes filings and disclosures (e.g., Securities and Exchange Commission filings, 10-K or 8-K annual reports, etc.), analyst reports, earnings conference calls, news or media articles, internet message boards (e.g., Yahoo! Finance, Raging Bull and Seeking Alpha) and posts on online social networks (e.g., Tweets, Facebook status).

Having finished the text harvesting, we need to clean the data and transform the unstructured data into the format that computers can process. Compared with the structured data, unstructured data does not have pre-defined data model and is not organized in a pre-defined manner. Typical unstructured data includes voice, images, text contents, etc. In that sense, unstructured data is usually text heavy and difficult to analyze. Through text cleaning, the qualitative contents are transformed into word vectors and the unstructured data is then ready to be used.

The last process of text parsing and analyzing aims to extract valuable information using different techniques. Among different approaches, the Lexicon-based approach or namely bag-of-words approach is a simple but powerful way of quantifying textual contents. Technically, the key is to set up a group of targeted phrases relating to the contents under investigation. For instance, Loughran et al. (2009) concern about the corporate governance and lawsuits of "sin stocks", so they set up a group of phrases indicating the business ethics (e.g., "ethic", "corporate responsibility", "social responsibility", "socially responsible", etc.) and calculate the words frequencies or attributes to the whole text. A more evolutionary approach beyond the targeted phrases refers to compiling a word list that share common characteristics, i.e., positive sentiment, negative sentiment, risk aversions, risk appetites and even financial literacy. Equipped with such "dictionary", we may abstract the emotions and ideologies underlying individuals' behavior by matching the words in textual contents with the word lists. Therefore, the construction of dictionary is quite crucial in the Lexicon-based approach. Commonly used dictionaries by studies in finance include The Henry Word List (Henry, 2008)[7], Harvard General Inquirer (GI) Word Lists[8], Diction Optimism and Pessimism Word Lists[9] and Loughran and McDonald Word Lists (Loughran and McDonald, 2011)[10]. Among the four dictionaries, Harvard GI word lists and Diction Word Lists are not created by rooting specifically in finance context, but researchers use the lists to conduct semantic analysis in accounting and finance field.

Alternatively, some researchers use machine learning techniques to classify and recognize textual patterns. Machine learning is closely related to computational statistics and comprises of two major steps. The first step refers to giving a sample dataset and certain mathematical algorithms for computers to learn from "training". As for the second step, the computer with "learning experience" needs to make predictions

---

[7] This is the first dictionary for financial textual analysis, including 85 negative words created by examining earnings press releases for the telecommunications and computer services industries.
[8] Harvard General Inquirer Word Lists consist of several categories relating to humans' semantic contents, i.e., the Harvard IV-4 dictionary, the Lasswell value dictionary, etc.
[9] Similar to Harvard General Inquirer Word Lists, Diction Optimism and Pessimism Word Lists have 35 subcategories covering 686 optimism words and 920 pessimism words.

[10] Loughran and McDonald [2011] Word Lists are quite extensive with 354 positive and 2329 negative words.

or decisions without being programmed to do so. In that sense, high-quality dataset and effective algorithms are two key issues in machine learning since they may affect the predictability. In the field of finance, commonly used machine training algorithms include the Naïve Bayes, Support Vector Machine and Neural Network. In reality, machine learning techniques may not necessarily generate better performance than the Lexicon-based approach in financial textual analysis. For example, Guo et al. (2016) test different methods in textual analysis and show that the Harvard General Inquirer (GI) Word Lists works as well as the Naïve Bayes technique, whereas Loughran and McDonald Word Lists performs even better. To sum up, the Lexicon-based approach is highly depending on the word lists while machine learning techniques are determined by the "training data" and training algorithms. Both approaches may generate great performance when the key factors and issues are effectively considered. Therefore, we may not conclude which one is better, one may outperform the other as long as it suits the problem being investigated.

5.2   Textual analysis in finance

During the past few years, the fast development of computing power and internet technologies has brought textual analysis into the world of finance for the purpose of constructing measurements for factors that are difficult to quantify. Compared with traditional measurements of investors' behavior, i.e., market variables or survey-based methods, textual analysis techniques provide a more direct and accurate approach to quantify individual behavior through revealing both intended as well as unintended messages hidden under the surface. Generally, current studies applying textual analysis to finance and accounting fall into the topics of information contents examination, readability measurement and sentiment analysis.

5.2.1   Information Content

With textual analysis, researchers are able to measure the actual information content conveyed by various sources and further quantify the impacts on financial markets. Therefore, studies in this field can be grouped into different categories according to the information sources.

Generally, most studies focus on investigating the information content of the traditional information dissemination channels, such as the official filings and disclosures (e.g., Securities and Exchange Commission filings, 10-K or 8-K annual reports, etc.), analyst reports and news articles. Based on Diction Optimism and Pessimism Word Lists, Davis et al. (2012) examine full texts of 23000 published quarterly earnings press releases and observe that the net optimistic language in earnings press releases is positively related to future Return on Assets (ROA) as well as the market returns around the earnings announcement dates. They further conclude that the language in earnings press releases conveys credible information about firm's future performance. Similarly, Hanley and Hoberg (2012) provide additional evidence for the information contents of official fillings from the IPO market. They decompose the initial IPO prospectus into standard content and informative content components and observe that offers with more informative contents have smaller absolute changes

in offer prices and lower underpricing, indicating that informative contents contribute to improving pricing accuracy. Apart from the official fillings and disclosures, Huang et al. (2014) examine the information contents of analyst reports and observe the predictability for future earnings growth in the subsequent five years. Notably, they further use the Naïve Bayes algorithm to extract positive and negative components from analyst report text. Cross-sectionally, they find that the negative component exerts stronger impact on investors compared with the positive component. As for the research on news articles, a more recent study by Boudoukh and Feldman (2019) uses machine learning techniques to group the massive news from Dow Jones Newswire into different categories, i.e., "identified" value-relevant events, "identified" value-relevant events with different complexity and "unidentified" news. In particular, they examine the stock market reaction during trading hours and overnight around different types of news and conclude that the intensity of information contents differs across news categories.

Alternatively, a strand of studies refers to the more burgeoning internet-based information dissemination channels. Among such studies, Internet stock message board is one of the most common information dissemination channels being investigated. For instance, Wysocki (1999) conducts the first study to explore the information content of Internet stock message board. Focusing on the message postings on Yahoo! Finance, the author reports that the volume of postings has predictability for the next-day trading volume and abnormal stock returns. In a similar vein, Bagnoli et al., (1999) provide the first evidence that the online unofficial forecasts from netters (commonly referred as "whispers") perform even better in predicting corporate earnings compared with analyst official forecasts. Though textual analysis techniques are not involved in the above two pioneer studies, the novel trial on exploring information content of new-media has inspired voluminous studies on this cutting-edge topic. For example, Antweiler and Frank (2004) use two types of machine learning techniques, i.e., Naïve Bayes and Support Vector Machine to study the effect of more than 1.5 million message boards posted on Yahoo! Finance and Raging Bull. They find that such message postings help to predict the stock returns as well as the trading volume on the next day, indicating that

-financial relevant information content is included in online "talks".

5.2.2   Readability Measurements

Readability measures the ease that readers can understand and interpret textual material. In that sense, readability is concerned with the problem of matching the reader and text, which depends on both effective written communication as well as readers' comprehension. Though as a general concept, readability has been widely examined in the field of finance referring to its importance in effectively disseminating information contained in corporate disclosures such as the annual reports and analyst reports.

Due to the limitation of computing power, earlier studies in this field have to suffer from small sample sizes and flawed methods (e.g., Lewis et al., 1986; Tennyson et al., 1990). Development of textual analysis allow researchers to construct index-based measurements for readability by decomposing long texts into targeted phrases and comparing with the pre-defined word lists. As one of the pioneer studies in this field, Li (2008) uses the Fog Index and the number of words to measure the readability for 10K annual reports. As is shown in the following equation, Fog Index is a function containing variables representing the length of sentences and complex words (complex words refer to words with more than two syllables):

$$Fog\ Index = 0.4\ (Average\ number\ of\ words\ per\ sentence + \\ percentage\ of\ complex\ words)$$

As a result, Li (2008) finds that firms with annual reports that are easier to read and interpret tend to have more persistent positive earnings, whereas the annual reports of firms with lower earnings are usually harder to read. Following Li (2008), a strand of studies starts to use the Fog Index to measure readability. For example, Miller (2010) finds that small investors tend to trade fewer shares of firms whose annual reports have lower readability. Small investors tend to be less-sophisticated with lower level of financial literacy, thus less readable annual reports are harder for them to process that further affect their trading behavior. Based on a sample comprised of 78000 U.S households, Lawrence (2013) observe a consistent positive association between annual

report readability and trading behavior for the retail investors. Apart from investor behavior, Lehavy et al. (2011) use the Fog Index to relate annual report readability to analyst coverings by providing a positive relationship between the two. On top of that, they even observe that more readable annual reports are associated with lower analyst dispersion and greater earnings forecast accuracy. Notably, Guay et al. (2016) construct a composite readability index based on six measurements used in previous studies. Table 3 shows the detailed components of the six readability measurements. Unlike above studies, Guay et al. (2016) focus on exploring the relationship between the financial statement readability and information environment. On the one hand, the within-firm evidence shows a positive relationship between readability and voluntary disclosure. On the other hand, the cross-sectional evidence shows that the relationship between readability and voluntary disclosure is weaker for firms with poor performance and greater earnings management and vice versa.

Table 3 Readability Measurements

Measurements	Calculations
Flesch Kincaid	0.39 * (number of words/number of sentences) þ 11.8 * (number of syllables/number of words) – 15.59
LIX	(number of words / number of sentences) þ (number of words over 6 letters * 100/ number of words)
RIX	(number of words with 7 characters or more)/(number of sentences)
Fog	0.4 * (number of words/number of sentences) þ 40 * (number of words with more than two syllables / number of words)
ARI	4.71 * (number of characters / number of words) þ 0.5 * (number of words / number of sentences) 21.43
SMOG	1.043 * sqrt (30 * number of words with more than two syllables / number of sentences) + 3.1291

5.2.3 Sentiment Analysis

Traditional research on analyzing sentiment effect on stock market movements usually rely on market fundamental variables. Specifically, Baker and Wurgler (2006) in their highly-cited study construct a composite sentiment index using six variables: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium. Following Baker and Wurgler (2006), a strand of empirical studies applies the market-based sentiment index to further investigate the sentiment effects on trading behavior, asset prices and other market dynamics (e.g., Baker and Wurgler, 2007; Karlsson et al., 2009; Stambaugh et al., 2012; Oliveira et al., 2013; Sayim and Rahman, 2015).

Recent studies in this field have shifted the focus to using textual analysis and machine learning techniques to detect the implicit semantics and emotions from the narrative texts. Generally, there are two sources that researchers extract sentiment from — traditional media platform (news and articles) and the so-called new media platform (internet stock message boards and social media networks). As the first study on predictability news content for stock market movements, Tetlock (2007) focus on analyzing the textual contents of the daily "Abreast of the Market" column in the Wall Street Journal and construct a news-based pessimism sentiment factor by referring to the Harvard General Inquirer (GI) Word Lists. Furthermore, Tetlock et al. (2008) extend the news sample to the whole Wall Street Journal and Dow Jones News Service (DJNS). In addition to stock returns, they further observe that the news contents with negative sentiment even forecasts low future earnings for individual S&P 500 firms. Through enlarging the time period to 100 years (1905 to 2005), Garcia (2013) revisits the Tetlock (2007) and Tetlock (2008)'s evidence while considering the effects of business cycles. Typically, Garcia (2013) find that the predictability of news sentiment on stock returns is concentrated in recessions, i.e., a one standard deviation shock to the sentiment measure during recessions forecasts a 12 basis points change in Dow Jones Industrial Average returns over one single day. Apart from the stock market, similar evidence for the predictability of news sentiment on asset prices has also been found in the real estate

market. Having analyzed 125000 U.S newspaper article headlines from Bloomberg, the Financial Times, Fobes and The Wall Street Jounals, Ruscheinsky et al. (2018) document a positive relationship between sentiment measure and REIT returns. Precisely, positive (negative) changes in the news-based sentiment induce upward (downward) pressure on REIT returns three to four months later.

Rather than focusing on linguistic news contents, another strand of studies aims to extract the investors' sentiment that are usually hidden in narrative texts on the new media platforms. One of the commonly used new media platforms refer to the internet stock message boards, i.e., Yahoo! Finance, Raging Bull and SeekingAlpha, through which investors are able to disseminate information and communicate ideas. Based on the stock message postings on Yahoo! Finance and Raging Bull, Antweiler and Frank (2004) use the Naïve Bayes method and construct a bullishness index to measure sentiment of retail investors. Following Antweiler and Frank (2004), Leung and Ton (2015) employ similar method to extract the bullish sentiment of Australian investors by focusing on 2.5 million messages from HotCopper — the largest online stock message board in Australia. Typically, they find that small cap stocks are more easily to be influenced by investors' sentiment. Moreover, Das and Chen (2007) even apply five machine-learning algorithms and classify the investors' sentiment extracted from online message boards into three categories: bullish (optimistic), bearish (pessimistic) and neutral (either bullish nor bearish). Focusing on stocks in the tech sector, they find that the aggregated sentiment tracks the index returns and such effect is quite weak for individual stocks. On the contrary, Kim and Kim (2014) extract investor sentiment from Yahoo! Finance and find no significant evidence that sentiment has predictive power for future stock returns, volatility as well as trading volume. Rather, they observe that prior stock price performance is positively related to future investor sentiment.

Another type of way that investors communicate with each other is through the social media networks such as Facebook and Twitter. For example, Siganos et al. (2014) study 20 international markets and document that sentiment on Facebook is positively related to stock returns, but is negatively related to trading volume and return volatility.

Through measuring Twitter mood into six-dimensional sentiment (Calm, Alert, Sure, Vital, Kind and Happy), Bollen et al. (2011) firstly observe that sentiment on Twitter helps to predict future Dow Jones Industrial Average index returns. Following Bollen et al. (2011), a large number of studies has started to investigate the causal relationship between sentiment on twitter and stock market variables across different countries including the U.S., European countries and Asia-pacific countries (Sprenger et al., 2014; Ranco et al., 2015; Zhang et al., 2016; Li et al., 2017; Li et al., 2018; Zhang et al., 2018; Li, 2019; Saurabh and Dey, 2020). Interestingly, Li et al. (2018) provide an empirical evidence for foreign sentiment contagion from the U.S to Chinese stock market. Generally, it is notable that such studies reach consistent results that Twitter sentiment actually tracks stock market. Moreover, a more recent study by Behrendt and Schmidt (2018) takes an intraday perspective and investigate how sentiment on Twitter affect the absolute 5-minute stock returns. As a result, Behrendt and Schmidt (2018) document that the high-frequency sentiment is not that useful in predicting intraday stock returns or volatility. That is, the causal relationship between investor sentiment extracted from social media platforms and stock market variables tend to be more rendered for lower frequencies.

5.3   Textual Analysis and Financial Literacy

Referring to the previous review on financial literacy, we may propose two future directions for this field relating to two major issues that need to be addressed urgently.

On the one hand, more effective measurements for financial literacy are crucially needed in order to better understand how financial literate are individuals. In general, current studies have shown that individuals tend to be relatively financial illiterate worldwide. Also, financial literacy levels are proved to be closely related to individuals' economic welfare, i.e., investment decisions, debt decisions, retirement planning, etc. In that sense, a more accurate measurement for current financial literacy levels as well as developing more efficient approaches to improve financial literacy is of great importance for policy makers. As we have explained, current measurements for financial literacy have several limitations: survey-based measurements have problems of low rate of responds, biased or inaccurate answers and low frequencies; whereas demographic-based proxy and outcome-based proxy have inevitable exogenous problems. In fact, current studies have shown an increasing trend of using textual analysis to quantify the hard-to-measure qualitative variables such as information contents, financial context readability and investor sentiment, indicating textual analysis as a potentially more accurate approach for measuring financial literacy. Specifically, textual analysis techniques provide the opportunity to detect individuals' actual financial literacy levels hidden behind their day-to-day behavior on the social media networks, e.g., contents of online postings, search frequencies about specific words, dissemination action for official announcements. Unlike survey-based measurements that only captures the intended information, individuals' online behavior discloses both their intended as well as unintended characteristics that contribute to a more accurate measurement for financial literacy.

On the other hand, current measurements fail to capture the changing pattern of financial literacy on a time-series basis. Due to the limitation, current studies on investigating the effects of financial literacy are generally focus on individuals' one-off financial decisions such as the retirement planning, investment decisions, debt

decisions, etc. In fact, a current research blank lies in the field of examining how financial literacy relates to individuals' more frequent day-to-day trading behavior. Such topic is quite crucial for evaluating asset prices, investigating market dynamics as well as improving market efficiency as a whole. Also, capturing the changing pattern of financial literacy lays a solid foundation for developing efficient programs to improve financial literacy. Notably, a key advantage of textual analysis is the more frequent time-series based narrative texts which underlay the dataset. Therefore, researchers may refer to textual analysis techniques to capture the changing pattern of financial literacy on a time-series basis and fills the current gap through relating financial literacy to frequent trading behaviors.

In summary, textual analysis can be regarded as an alternative approach for financial literacy that helps to resolve the limitations of current measurements. Additionally, through combing narrative dataset with psychological theories, the machine-learning techniques may also be used to formulate more effective questions that better discloses financial literacy. In that sense, textual analysis may improve the effectiveness of traditional survey-based approach through contributing to better-established questionnaires. Moreover, the recent development for textual analysis techniques (commonly referred to Layered Voice Analysis, LVA) even allow researchers to analyze non-verbal information, i.e., vocal cues, tones, gestures. The textual analysis provides researchers an opportunity to collect vocal materials when conducting surveys to obtain a better understanding about financial literacy.

6 Conclusions

Financial literacy has been regarded as one of the important characteristics that lead to material impacts on individuals' decision-making process as well as their behaviors. Studies have shown that financial illiteracy actually relates to irrational financial decisions that may reduce the efficiency of financial markets as well as impose negative effects on social economic welfares. Specifically, financial literacy levels affect individuals' financial decisions including retirement planning, investment decisions, professional advice seeking and debt decisions. In that sense, measuring the level of individuals' financial literacy is not only important for professionals and researchers in the field of finance, but also vital for policy makers in terms of constructing effective financial education scheme to improve financial literacy so as to increase social economic welfare as a whole. However, similar to investors' sentiment, financial literacy is difficult to measure referring to its qualitative nature. Current literature mainly uses survey-based measurement for financial literacy. Despite the low rate of responds and biased questionnaires generated by surveys, we argue that survey-based measurement fail to capture the time-series changing patterns of financial literacy. Referring to the rapid development in textual analysis techniques, we propose that such methods may bring more opportunities for constructing more appropriate ways of measuring financial literacy. Specifically, textual analysis allows us to extract both intended and unintended information through investors' behavior on online social networks. Such information not only contributes to developing new measurements for financial literacy, but also make it possible to obtain financial literacy levels on a time-series basis. Also, this study concludes that future research may combine textual analysis techniques with survey-based methods to measure financial literacy more accurately. Generally, our study sheds lights on future studies on both the field of financial literacy as well as textual analysis.

References

Abreu, M., & Mendes, V. (2010). Financial literacy and portfolio diversification. Quantitative Finance, 10(5), 515-528.

Antweiler, W., & Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. The Journal of Finance, 59(3), 1259-1294. doi: 10.1111/j.1540-6261.2004.00662.x

Agnew, J. R., & Szykman, L. R. (2005). Asset Allocation and Information Overload: The Influence of Information Display, Asset Choice, and Investor Experience. Journal of Behavioral Finance, 6(2), 57-70. doi: 10.1207/s15427579jpfm0602_2

Alessie, R., Van Rooij, M., & Lusardi, A. (2011). Financial literacy and retirement preparation in the Netherlands. Journal of pension economics & finance, 10(4), 527-545.

Allgood, S., & Walstad, W. B. (2016). The Effects of Perceived and Actual Financial Literacy on Financial Behaviors. Economic Inquiry, 54(1), 675-697.

Almenberg, J., & Dreber, A. (2015). Gender, Stock Market Participation and Financial Literacy. Economics Letters, 137(137), 140-142.

Almenberg, J., & Save-soderbergh, J. (2011). Financial Literacy and Retirement Planning in Sweden. Journal of pension economics & finance, 10(4), 585-598.

Arrondel, L., Debbich, M., & Savignac, F. (2013). Financial Literacy and Financial Planning in France. Numeracy, 6(2), 8.

Atkinson, A., Messy, F.-A., 2012. Measuring financial literacy: results of the OECD/ International Network on Financial Education (INFE) pilot study. OECD Working Papers on Finance, Insurance and Private Pensions 15, OECD.

Bagnoli, M., Beneish, M. D., & Watts, S. G. (1999). Whisper forecasts of quarterly earnings per share. Journal of Accounting and Economics, 28(1), 27-50. doi: https://doi.org/10.1016/S0165-4101(99)00018-X

Baker, M. P., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. Journal of Finance, 61(4), 1645-1680.

Baker, M. P., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. Journal of Economic Perspectives, 21(2), 129-152.

Balloch, A. G., Nicolae, A., & Philip, D. (2014). Stock Market Literacy, Trust and Participation. Review of Finance, 19(5), 1925-1963.

42

Barber, B. M., & Odean, T. (2001). Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment. Quarterly Journal of Economics, 116(1), 261-292.

Bannier, C. E., & Neubert, M. (2016). Gender differences in financial risk taking: The role of financial literacy and risk tolerance. Economics Letters, 145, 130-135. doi: https://doi.org/10.1016/j.econlet.2016.05.033

Beckmann, E. (2013). Financial Literacy and Household Savings in Romania. Numeracy, 6(2),9.

Behrendt, S., & Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. Journal of Banking & Finance, 96, 355-367. doi: https://doi.org/10.1016/j.jbankfin.2018.09.016

Behrman, J. R., Mitchell, O. S., Soo, C. K., & Bravo, D. (2012). How Financial Literacy Affects Household Wealth Accumulation. The American Economic Review, 102(3), 300-304.

Bhattacharya, U., Hackethal, A., Kaesler, S., Loos, B., & Meyer, S. (2012). Is Unbiased Financial Advice To Retail Investors Sufficient? Answers from a Large Field Study. Review of Financial Studies, 25(4), 975-1032.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2(1), 1-8. doi: https://doi.org/10.1016/j.jocs.2010.12.007

Bonte, W., & Filipiak, U. (2012). Financial literacy, information flows, and caste affiliation: Empirical evidence from India. Journal of Banking and Finance, 36(12), 3399-3414.

Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2019). Information, Trading, and Volatility: Evidence from Firm-Specific News. Review of Financial Studies, 32(3), 992-1033.

Bucher-Koenen, T., & Lusardi, A. (2011). Financial Literacy and Retirement Planning in Germany. Journal of pension economics & finance, 10(4), 565-584.

Bucher-Koenen, T., Lusardi, A., Alessie, R., & van Rooij, M. (2014). How financially literate are women? An overview and new insights. National Bureau of Economic Research Working Paper Series, No. 20793. doi: 10.3386/w20793

Calcagno, R., & Monticone, C. (2015). Financial Literacy and the Demand for Financial Advice. Journal of Banking and Finance, 50(1), 363-380.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. Journal of Political Economy, 115(5), 707-747.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Measuring the Financial Sophistication of

Households. The American Economic Review, 99(2), 393-398.

Chauhan, Y., & Dey, D. K. (2020). Does financial literacy affect the value of financial advice? A contingent valuation approach. Journal of Behavioral and Experimental Finance, 25, 100268.

Christiansen, C., Joensen, J. S., & Rangvid, J. (2008). Are Economists More Likely to Hold Stocks. Review of Finance, 12(3), 465-496.

Clark, R. L., Lusardi, A., & Mitchell, O. S. (2017). FINANCIAL KNOWLEDGE AND 401(K) INVESTMENT PERFORMANCE: A CASE STUDY. Journal of pension economics & finance, 16(03), 324-347.

Collins, J. M. (2012). Financial Advice: A Substitute for Financial Literacy?. Financial Services Review, 21(4).

Crossan, D., Feslier, D., & Hurnard, R. (2011). Financial literacy and retirement planning in New Zealand. Journal of pension economics & finance, 10(04), 619-635.

Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. Management Science, 53(9), 1375-1388. doi: 10.1287/mnsc.1070.0704

Davis, A. K., Piger, J. M., & Sedor, L. M. (2012). Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language*. Contemporary Accounting Research, 29(3), 845-868. doi: 10.1111/j.1911-3846.2011.01130.x

Deuflhard, F., Georgarakos, D., & Inderst, R. (2019). Financial Literacy and Savings Account Returns. Journal of the European Economic Association, 17(1), 131-164.

Dhar, R., & Zhu, N. (2006). Up Close and Personal: Investor Sophistication and the Disposition Effect. Management Science, 52(5), 726-740

Disney, R., & Gathergood, J. (2013). Financial literacy and consumer credit portfolios. Journal of Banking & Finance, 37(7), 2246-2254. doi: https://doi.org/10.1016/j.jbankfin.2013.01.013

Fornero, E., & Monticone, C. (2011). Financial Literacy and Pension Plan Participation in Italy. Journal of pension economics & finance, 10(04), 547-564.

Garcia, D. (2013). Sentiment during Recessions. The Journal of Finance, 68(3), 1267-1300. doi: 10.1111/jofi.12027

Gathergood, J. (2012). Self-control, financial literacy and consumer over-indebtedness. Journal of Economic Psychology, 33(3), 590–602

Gamble, K. J., Boyle, P. A., Yu, L., & Bennett, D. A. (2015). Aging and Financial Decision Making.

44

Management Science, 61(11), 2603-2610.

Gerardi, K. S., Goette, L., & Meier, S. (2013). Numerical ability predicts mortgage default. Proceedings of the National Academy of Sciences of the United States of America, 110(28), 11267-11271.

Gerrans, P., & Hershey, D. A. (2017). Financial Adviser Anxiety, Financial Literacy, and Financial Advice Seeking. Journal of Consumer Affairs, 51(1), 54-90.

Georgarakos, D., & Pasini, G. (2011). Trust, Sociability and Stock Market Participation. Review of Finance, 15(4), 693-725.

Goetzmann, W. N., & Kumar, A. (2008). Equity Portfolio Diversification. Review of Finance, 12(3), 433-463.

Grinblatt, M., & Keloharju, M. (2001). How Distance, Language, and Culture Influence Stockholdings and Trades. Journal of Finance, 56(3), 1053-1073.

Guay, W. R., Samuels, D., & Taylor, D. J. (2016). Guiding Through the Fog: Financial Statement Complexity and Voluntary Disclosure. Journal of Accounting and Economics, 62(2), 234-269.

Guiso, Luigi, and Tullio Jappelli, 2009, Financial literacy and portfolio diversification, Working paper 212, CSEF, University of Naples.

Guo, L., Shi, F., & Tu, J. (2016). Textual analysis and machine leaning: Crack unstructured data in finance and accounting. The Journal of Finance and Data Science, 2(3), 153-170. doi: https://doi.org/10.1016/j.jfds.2017.02.001

Hanley, K. W., & Hoberg, G. (2010). The Information Content of IPO Prospectuses. The Review of Financial Studies, 23(7), 2821-2864. doi: 10.1093/rfs/hhq024

Hastings, J. S., Madrian, B. C., & Skimmyhorn, W. L. (2013). Financial Literacy, Financial Education, and Economic Outcomes. Annual Review of Economics, 5(1), 347-373. doi: 10.1146/annurev-economics-082312-125807

Henry, E. (2008). Are Investors Influenced by How Earnings Press Releases are Written. Journal of Business Communication, 45(4), 363-407.

Huang, A. H., Zang, A. Y., & Zheng, R. (2014). Evidence on the Information Content of Text in Analyst Reports. The Accounting Review, 89(6), 2151-2180.

Hung, A. a. P., Andrew M. and Yoong, Joanne. (2009). Defining and Measuring Financial Literacy. RAND Working Paper Series WR-708. doi: https://ssrn.com/abstract=1498674

Jappelli, T. (2010). Economic literacy: an international comparison. Economic Journal, 120: F429–F451. doi:10.1111/j. 1468-0297.2010.02397.x

Karlsson, N., Loewenstein, G., & Seppi, D. J. (2009). The ostrich effect: Selective attention to information. Journal of Risk and Uncertainty, 38(2), 95-115.

Klapper, L., Lusardi, A., & Panos, G. A. (2013). Financial literacy and its consequences: Evidence from Russia during the financial crisis. Journal of Banking and Finance, 37(10), 3904-3923.

Kim, S.-H., & Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. Journal of Economic Behavior & Organization, 107, 708-729. doi: https://doi.org/10.1016/j.jebo.2014.04.015

Lawrence, A. (2013). Individual investors and financial disclosure. Journal of Accounting and Economics, 56(1), 130-147.

Lehavy, R., Li, F., & Merkley, K. J. (2011). The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. The Accounting Review, 86(3), 1087-1115.

Lewis, N. R., Parker, L. D., Pound, G. D., & Sutcliffe, P. (1986). Accounting Report Readability: The Use of Readability Techniques. Accounting and Business Research, 16(63), 199-213.

Loughran, T., Mcdonald, B., & Yun, H. (2009). A Wolf in Sheep's Clothing: The Use of Ethics-Related Terms in 10-K Reports. Journal of Business Ethics, 89(1), 39-49.

Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. Journal of Accounting Research, 54(4), 1187-1230.

Lusardi, A., & Mitchell, O. S. (2007). Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth. Journal of Monetary Economics, 54(1), 205-224. doi: https://doi.org/10.1016/j.jmoneco.2006.12.001

Lusardi, A. (2008). Household saving behavior: The role of financial literacy, information, and financial education programs: National Bureau of Economic Research.

Lusardi, A., & Mitchell, O. S. (2009). How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness. National Bureau of Economic Research Working Paper Series, No. 15350. doi: 10.3386/w15350

Lusardi, A., & Mitchell, O. S. (2011a). Financial literacy and planning: Implications for retirement wellbeing: National Bureau of Economic Research

Lusardi, A., & Mitchell, O. S. (2011b). Financial literacy and retirement planning in the United States. Journal of pension economics & finance, 10(4), 509-525.

Lusardi, A., & Mitchell, O. S. (2011c). Financial literacy around the world: an overview. Journal of pension economics & finance, 10(4), 497-508.

Lusardi, A. (2012). Numeracy, financial literacy, and financial decision-making. National Bureau of Economic Research Working Paper Series, No. 17821. doi: 10.3386/w17821

Lusardi, A., & Scheresberg, C. d. B. (2013). Financial Literacy and High-Cost Borrowing in the United States. National Bureau of Economic Research Working Paper Series, No. 18969. doi: 10.3386/w18969

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. Journal of economic literature, 52(1), 5-44.

Lusardi, A., & Tufano, P. (2015) Debt literacy, financial experiences, and overindebtedness. Journal of Pension Economic Finance 14(4):332–368. doi:10.1017/S1474747215000232

Lusardi, A., & Mitchell, O. S. (2017). How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness. Quarterly Journal of Finance, 07(03), 1750008. doi: 10.1142/s2010139217500082

Lusardi, A., Michaud, P., & Mitchell, O. S. (2017). Optimal Financial Knowledge and Wealth Inequality. Journal of Political Economy, 125(2), 431-477.

Leung, H., & Ton, T. (2015). The impact of internet stock message boards on cross-sectional returns of small-capitalization stocks. Journal of Banking & Finance, 55, 37-55. doi: https://doi.org/10.1016/j.jbankfin.2015.01.009

Li, F. (2008). Annual Report Readability, Current Earnings, and Earnings Persistence. Journal of Accounting and Economics, 45(45), 221-247.

Li, X., Shen, D., Xue, M., & Zhang, W. (2017). Daily happiness and stock returns: The case of Chinese company listed in the United States. Economic Modelling, 64, 496-501. doi: https://doi.org/10.1016/j.econmod.2017.03.002

Li, X., Shen, D., & Zhang, W. (2018). How Does Foreign Sentiment Affect the Chinese Stock Markets?-Some Empirical Evidence. China Accounting and Finance Review, 20, 1-25.

Li, X. (2019). Does Chinese investor sentiment predict Asia-pacific stock markets? Evidence from

a nonparametric causality-in-quantiles test. Finance Research Letters, 101395. doi: https://doi.org/10.1016/j.frl.2019.101395

Miller, B. P. (2010). The Effects of Reporting Complexity on Small and Large Investor Trading. The Accounting Review, 85(6), 2107-2143.

Nicolosi, G., Peng, L., & Zhu, N. (2009). Do Individual Investors Learn from Their Trading Experience. Journal of Financial Markets, 12(2), 317-336.

Odean, T. (1998). Are Investors Reluctant to Realize Their Losses. Journal of Finance, 53(5), 1775-1798.

Oliveira N., Cortez P., Areal N. (2013). On the Predictability of Stock Market Behavior Using StockTwits Sentiment and Posting Volume. Progress in Artificial Intelligence (pp.355-365). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40669-0_31

Parker, A. M., De Bruin, W. B., Yoong, J. K., & Willis, R. J. (2012). Inappropriate Confidence and Retirement Planning: Four Studies with a National Sample. Journal of Behavioral Decision Making, 25(4), 382-389.

Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The Effects of Twitter Sentiment on Stock Price Returns. PLOS ONE, 10(9), e0138441. doi: 10.1371/journal.pone.0138441

Ruscheinsky Jessica, R., Lang, M., & Schäfers, W. (2018). Real estate media sentiment through textual analysis. Journal of Property Investment & Finance, 36(5), 410-428. doi: 10.1108/JPIF-07-2017-0050

Sayim, M., & Rahman, H. (2015). The relationship between individual investor sentiment, stock return and volatility. International Journal of Emerging Markets, 10(3), 504-520.

Sekita, S. (2011). Financial literacy and retirement planning in Japan. Journal of pension economics & finance, 10(04), 637-656.

Sprenger, T. O., Sandner, P. G., Tumasjan, A., & Welpe, I. M. (2014). News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. Journal of Business Finance & Accounting, 41(7-8), 791-830. doi: 10.1111/jbfa.12086

Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. Journal of Economic Behavior & Organization, 107, 730-743. doi: https://doi.org/10.1016/j.jebo.2014.06.004

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. Journal of Financial Economics, 104(2), 288-302. doi: https://doi.org/10.1016/j.jfineco.2011.12.001

Stolper, O. A., & Walter, A. (2017). Financial literacy, financial advice, and financial behavior. Journal of Business Economics, 87(5), 581-643.

Stolper, O. (2018). It takes two to Tango: Households' response to financial advice and the role of financial literacy. Journal of Banking & Finance, 92, 295-310. doi: https://doi.org/10.1016/j.jbankfin.2017.04.014

Saurabh, S., & Dey, K. (2020). Unraveling the relationship between social moods and the stock market: Evidence from the United Kingdom. Journal of Behavioral and Experimental Finance, 26, 100300. doi: https://doi.org/10.1016/j.jbef.2020.100300

Tennyson, B. M., Ingram, R. W., & Dugan, M. T. (2008). ASSESSING THE INFORMATION CONTENT OF NARRATIVE DISCLOSURES IN EXPLAINING BANKRUPTCY. Journal of Business Finance & Accounting, 17(3), 391-410.

Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance, 62(3), 1139-1168. doi: 10.1111/j.1540-6261.2007.01232.x

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More Than Words: Quantifying Language to Measure Firms' Fundamentals. The Journal of Finance, 63(3), 1437-1467. doi: 10.1111/j.1540-6261.2008.01362.x

Tumarkin, R., & Whitelaw, R. F. (2001). News or Noise? Internet Postings and Stock Prices. Financial Analysts Journal, 57(3), 41-51. doi: 10.2469/faj.v57.n3.2449

Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. Journal of Financial Economics, 101(2), 449-472. doi: https://doi.org/10.1016/j.jfineco.2011.03.006

Von Gaudecker, H. (2015). How does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice. Journal of Finance, 70(2), 489-507.

Yoong, J. (2010). Financial Illiteracy and Stock Market Participation: Evidence from the RAND American Life Panel. Pension Research Council Working Paper, No. 2010-29. doi: http://dx.doi.org/10.2139/ssrn.1707523

Wysocki, P., 1998. Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards. In: University of Michigan Business School Working Paper. 98025,

Zarcadoolas, C., Pleasant, A., & Greer, D. S. (2009). Advancing health literacy: A framework for understanding and action (Vol. 45). John Wiley & Sons.

Zhang, W., Li, X., Shen, D., & Teglio, A. (2016). Daily happiness and stock returns: Some

international evidence. Physica A: Statistical Mechanics and its Applications, 460, 201-209. doi: https://doi.org/10.1016/j.physa.2016.05.026

Zhang, W., Wang, P., Li, X., & Shen, D. (2018). Twitter's daily happiness sentiment and international stock returns: Evidence from linear and nonlinear causality tests. Journal of Behavioral and Experimental Finance, 18, 50-53.

Zhang, Y., Jia, Q., & Chen, C. (2020). Risk attitude, financial literacy and household consumption: Evidence from stock market crash in China. Economic Modelling. doi: https://doi.org/10.1016/j.econmod.2020.02.040

Xiao Li: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Roles/Writing – original draft; Writing – review & editing.