Social media impact on household investors and their stock markets participation

**Research Challenge
Technical Report**

**THINK FORWARD
INITIATIVE**

Eric Tham

SOCIAL MEDIA IMPACT ON HOUSEHOLD INVESTORS AND THEIR STOCK MARKETS PARTICIPATION *

TECHNICAL REPORT

Eric Tham
August 2018

Abstract

The growth of social media has transformed how information is transmitted in the financial markets. Through prospect theory and a principal component analysis of market and social media sentiments, the household investors' trust in the social media for investing is extracted. This is found to be pro-cyclical over the last two decades, but has been increasing in the last decade. This increasing trust however does not translate to increased household stock market participation, as evidenced in the triennial Survey of Consumer Finance. On the contrary, a significant monotonic relationship from a ranked correlation test was established between the dynamic correlation of market and social media sentiments, and the household participation rate. The dynamic correlations were obtained from a multi-variate GARCH model after being regressed against the main news sentiment. This suggests that conditional on the news sentiment, household participation in the stock markets depends on their trust on the social media.

Keywords: Household finance, market sentiment, natural language processing, trust, social media, prospect theory, stockholding puzzle

---

1. Introduction

The use of social media since the late 1990s has seen a phenomenal growth worldwide. Social media has changed the mechanism of information transmission. Before the Big Data era, news transmission originated from mass media perceived to be authoritative to the masses. Financial information sharing was primarily by word of mouth. Hong et al. (2017) studied how the word of mouth vis-a-viz a mass media model impacts information transmission and stock trading volume. This has changed with social media which proliferate mobile devices and the masses, promoting peer to peer information transmission. How has this growth in social media impacted household investors' asset allocation who are the primary users of social media and in turn the stock markets?

In the 1970s, beginning from the journal article Noise written by Fisher Black in Black (1986), household investors were perceived as noise traders. These household investors lacked technical sophistication and access to financial information compared to the professional investors. They were deemed to 'provide' liquidity to the market, with welfare benefits transferred from them to the professional traders.

Social media helped in two ways to level the playing field for the household investors. The first is through expedited news transmission with the popularity of Internet platforms like Facebook, Twitter in the western hemisphere and Weixin, Weibo in China. These availed household investors to financial information that was previously available only through dedicated financial channels. The second is through peer to peer social investing platforms that seek to disrupt the investment landscape by improving financial literacy.

Within the social media space, there have been changes in how they operate from the late 1990s to the recent few years. In the late 1990s to the early 2000s, platforms like Motley fool, SiliconInvestor.com and RagingBull.com were popular. These platforms provide chat rooms and internet bulletin boards to discuss stock market movements for the public. Tumarkin and Whitelaw (2001) documented that posts in these early Internet forums correlated with next day stock returns. In the recent few years, there are the popular SeekingAlpha and StockTwits in the USA, Ayondo in Singapore and Xueqiu in China. These platforms adopt typically a pyramidal follower system. At the top of the pyramid are gurus who work their way up the pyramidal system by establishing a reputation amongst its thousands of followers. These followers 'liked' their posts on the platform and subscribed to the gurus' feeds. These gurus also provide stock portfolio recommendations which are replicated by the followers. As the number of followers is sometimes more than a million, it can potentially create a herd mentality in the stock market. In Chen et al. (2014) these posts in SeekingAlpha.com were found to predate stock price movements.

One of the key determinants of household participation in the stock markets is social interaction. Social interaction promotes information exchange which increases stock buying tendency.

Hong et al. (2004) found that households that interact more with their neighbours and attend church are more likely to invest in the stock market, all else being equal than those who do not. This concept of social interaction has changed in meaning with increased use of social media over the last two decades. Social media increased the rate of information transmission and exchange relative to traditional media. Chawla et al. (2017) studied that tweets and re-tweets on the Twitter platform spread stale news in stocks which impacted its trading volume.

Aside from information transmission, financial literacy amongst households has also taken greater attention in recent times. In the latest triennial Survey of Consumer Finance (SCF) by the Federal Reserve Board in 2016, additional questions of self-assessed financial knowledge and financial literacy were posed to the household participants. In Calvet et al. (2009), a study was made of the financial literacy of household investors in Sweden. This financial literacy was measured by their disposition effect (or lack of), under-diversification and risky share inertia. Through an administrative panel study, it is found that wealthier households have higher financial literacy. Whilst social media is unlikely to dispel the psychological biases of the household investors, it may help them in portfolio diversification and increase stock market participation. Calvet et al. (2006) found that households bore welfare costs due to their lack of financial literacy. Social media may thence positively affect households' risk-taking behaviour and the reduction of welfare extraction from them.

Campbell (2006) in the presidential address on household finance cited that an obstacle to stock market participation for households is prevailing fixed costs, especially for lower income households. These fixed costs include time and costs to survey and understand the stock markets. Could social investing platforms by democratising financial information exchange and literacy help lower these fixed costs for household investors?

The SCF reported household participation in the equity markets had fluctuated from 1998[2] to 2016. Household participation here is measured by direct stock holdings and does not include managed funds and retirement account. Retirement account is not included as it is likely to contain bonds and other fixed income. Managed funds do not require financial literacy as much since they are managed by professional managers. A well-known puzzle in finance is the stockholding puzzle. This puzzle points to the low number of household investors not having stocks in their portfolio in spite of a better risk-return payoff. From 1998 to 2004, direct household participation in stocks was relatively higher averaging around 21.2%. In the surveys 2010 to 2016, percentage household participation decreased to 14.2%[3]. Could this be due to earlier Internet platforms waning in popularity post the Internet boom? Notwithstanding the fluctuating percentage participation, the mean value of stock assets had increased over the years from $234k in 1998 to $328k in 2016. The increase in mean value is likely attributed to the wealthier households putting more wealth into the equity markets[4].

---

[2] The start date 1998 is significant as it is about the time when social media starts being popular.
[3] The data is taken from page 291 of the SCF bulletin chart in 2016 showing percent of families with directly held stocks.

[4] The data is taken from page 293 of the same report.

The potential levelling of the investment playing field by social media especially for the less sophisticated household investors has important implications. It may cause lower income households to place a greater allocation to the risky assets, potentially increasing risks to their household wealth. Household investors have been known to be more prone to psychological biases in decision making. Some of these common biases include anchoring or confirmation bias, herding biases and disposition effect. These biases with increased household investor participation may impact the equity risk premium on the macro-economic front. This may cause larger market gyrations which was evident in the China stock market crash in 2015, driven primarily by retail investors as mentioned in the CNBC article and documented in Tham (2016).

Most of the research into these social investing platforms for example in Chen et al.
(2014), Heston and Sinha (2017) and Bollena et al. (2011) have primarily focussed on
the predictability present in these social media posts and how information dissemination through social media has impacted stock prices and volume. The results have been mainly affirmative in that the posts generally predate stock price movements or trading volume. This predictive ability of the social media on stock markets hints at its usefulness for household investors, but how it affects the actual decision making process of the household investors and their overall participation in the equity markets remains to be studied.

This paper attempts to study two questions. First how much do household investors actually trust the social media for investing and how this trust has changed over time? The second is how social media has impacted household participation rate in the equity markets. In the section 2.1, the paper examined the first question from the perspective of Prospect Theory which is a cornerstone for decision making under risk. The section 2.2 studies this in the context of social media sentiment obtained by Natural Language Processing (NLP) techniques and through market sentiment. The section 3 describes the data and methodology used including discussion and robustness tests in section 3.1. The section 3.2 then examines the second question through a dynamic corelation between social media sentiment and the market sentiment with household participation in the stock markets. The final section concludes along with further research that can be done.

2. Background

2.1 Prospect Theory and Decision Making

Prospect Theory is a seminal paper written by Daniel Kahneman and Amos Tversky in Kahneman and Tversky (1979) and is a descriptive model of decision making under risk. The key tenet of Prospect theory is an agent's value is derived from potential loss or gain to a reference point rather than the final outcome. Further, agents tend to be more risk averse with respect to losses than to gains with respect to the reference point. An often assumed tenet of Prospect Theory is decision making is in two phases - editing and valuation phase. The editing phase reflects the investors scanning the information horizon and forming their beliefs on the eventual outcomes. This phase is heavily dependent on psychological biases. The second phase is the valuation phase whence agents value these eventual outcomes based on the beliefs in the first phase. This phase is more dependent on the risk preferences of the investors. These two phases of decision making are distinct albeit closely dependent on each other. Guiso et al. (2008) in a survey of Dutch investors studied that trust (or beliefs) is different from risk preferences, and is important in the selection of stocks for investment.

The rise of social media and Big Data allows to distinguish between these two phases of Prospect Theory. Investors reflect their beliefs on the stock market by writing posts and comments on the social media. These posts and comments in turn may influence the beliefs of other investors. This is the first editing phase. However, the investors do not necessarily act on their beliefs due to the valuation phase. The investor may have just lost his job or just for liquidity needs to send his children to college. The rationale is that even when investors believe in a particular stock and write positive comments on the stock, they do not necessarily trade the stock due to their risk valuation and preferences.

2.2 Natural Language Processing

Natural Language Processing (NLP) is used to extract sentiment of the posts in news and social media. This is done through Thomson Reuters MarketPschy indices (TRMI), which is a global standard in sentiment mining for financial markets. TRMI uses A.I. to obtain different facets of investor sentiments from 2000 sources of news and 800 social media platforms around the world. It include 34 emotional aspects like sentiment, buzz, joy, trust, anger, surprise, fear and other aspects. The AI technique used is supervised learning algorithm, which uses a reference bible of labelled positive or negative statements to classify texts on a sentiment scale. Similarly, NLP has been used for financial studies in Loughran and McDonald (2011), Bollena et al. (2011) and in Chen et al. (2014).

Two different types of TRMI sentiment are used in this study - from the main news media and social media. The main news media considered headlines from the main news media for example, the Wall Street Journal, MarketWatch, and the Financial Times. The posts from these sources tend to be more objective on 'facts' and fundamentals regarding the economy and companies with much less noise than from social media posts. The count of these positive

or negative headlines in the Wall Street Journal has been shown to have an influence on the next day returns, as studied in Tetlock (2007) through a similar textual analysis.

Another widely cited sentiment index, the Baker-Wurgler sentiment index is also used. The index is gleaned from six market activity indicators, including the closed end fund discount, the number and average of first day returns of IPO, NYSE share turnover, equity share in new issues and the dividend premium. Unlike sentiment obtained from NLP, it is gleaned from market activities where investors actually put their 'money where the mouths are'. It thence reflects both phases of investors' decision making - beliefs editing and risk valuation.

3. Data & Methodology

I first look at the static correlation matrix amongst the Baker-Wurgler, social media and news sentiment indices from 1998 Jan to 2015 Sep shown in table 1. The year 1998 represents when TRMI starts collecting data on the social media.

A principal component factor analysis is done on the correlation matrix, and a time series of the principal components extracted. The time series of the first and second principal components is in figure 1. The eignvalues and eigenvectors for the PCAs are shown in tables 2 and 3 respectively.

Figure 1: Time series of PCA components

The PCA2 denoting the household investors' trust in the social media is pro-cyclical. It was initially high at the turn of the century due to the Internet bulletin boards, which subsequently waned in popularity. From 2008 till 2016, trust in the social media slowly increased with the next generation of social media platforms.

Table 1: Correlation matrix of sentiment indices (1998 Jan to 2015 Sep)

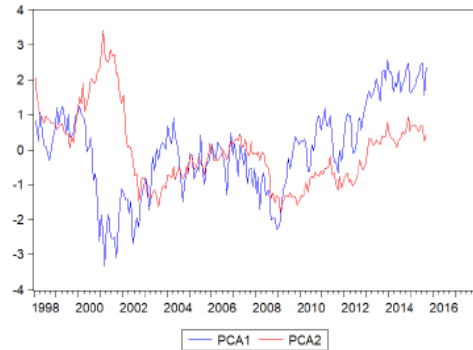|		1	2	3
1.	Social media sentiment	1.0		
2.	News media sentiment	0.54	1.0	
3.	Baker-Wurgler sentiment index	-0.36	0.11	1.0

The correlation between the Baker-Wurgler market sentiment with the sentiment from NLP is actually relatively low over the last 20 years. It is negatively correlated with the social media sentiment and 0.11 with the news sentiment. On the contrary, the social media and news media sentiments from NLP are relatively more correlated.

Table 2: PCA eigen vectors

Variable	PCA 1	PCA 2	PCA 3
1. Social media sentiment	0.59	0.54	−0.59
2. News media sentiment	0.73	−0.06	0.67
3. Baker-Wurgler sentiment Index	−0.33	0.84	0.44

The PCA2 is the extracted beliefs of interest denoting the household investors' trust/ beliefs in social media. It represents the commonality between the social media sentiment and the market sentiment, as it is contributed primarily by the two sentiment indices.

Table 3: PCA eigenvalues

PCA No	Value	Proportion	Cumulative proportion
1	1.60	0.53	0.53
2	1.10	0.37	0.90
3	0.29	0.09	1.0

The PCA2 represents 37% of the total variability compared to the PCA1 of 53%.

4. Results

There is a growing trust of the household investors on social media for investment advice.

The hypothesis is proven in two steps. I first show that this trust can be measured by natural language processing and the commonality between the Baker-Wurgler sentiment and social media sentiment. Then I prove this trust has a positive influence on the stock market.

4.1 Trust as a measure
Beliefs or trust as a measure is hard to quantify. Yet, it is the very fabric that holds the economic system together. From table 3, the second principal component contributes 37% of the total variance and is primarily contributed by the market Baker-Wurgler and social media sentiment as observed in the eigenvectors in table 2. It thence represents the commonality between these indices. Since only the beliefs editing phase is common between the sentiment indices, this second PCA principal is hypothesised to reflect the trust or beliefs in the social media.

4.2 Robustness tests
I further do robustness tests on this extracted second component beliefs using two commonly cited surveys of investors' beliefs. The two surveys are the individual Shiller crash index and the University of Michigan consumers' confidence index. The former is a survey of individual investors on their expectations (or beliefs) of impending market crashes over a 6 month period. The latter survey measures consumer expectations largely with respect to their spending and saving inclinations.

Both surveys are similar to the trust in social media for stock investing in that they reflect expectations or beliefs but do not necessarily reflect actual trading decisions of the investors and their risk preferences. However the surveys are different in that individual Shiller index represents the beliefs of a small fraction of wealthy individuals on the stock market while the University of Michigan confidence index represents the consumers' beliefs on the macro-economy.

The robustness tests are to ascertain that the extracted PCA2 is indeed the beliefs component. This is done in two separate regression tests with the individual Shiller index as the dependent variable as shown in the table 4. The table on the left hand side uses the PCA2 which is the purported beliefs extracted, whilst the right hand side uses the market and social media sentiment. Both indicate a similar R^2 _ 0.06. The likelihood ratio test statistic with LR = 0.999 and $c2$ = 0.002 accepts the null hypothesis that the two models are not distinctly different. This means that from the Baker-Wurgler and social media sentiment indices, the only explanatory component against the Shiller individual index is by the PCA2. Since the Shiller individual index essentially reflects expectations or beliefs, the PCA2 is also a proxy of beliefs.

Next, I use the University of Michigan consumer confidence as the dependent variable in 2 similar regressions. As shown in table 5, on the left hand side the PCA2 is used as a regressor whilst the right hand side uses the market and social media sentiment. The R^2 = 0.35 is even higher than the Shiller individual index. A possible reason is the

University of Michigan consumer confidence index has a larger sample size reflecting the beliefs of the masses. Using the same Michigan index, Li and Li (2011) had showed that its dispersion amongst consumers contributed to greater stock trading volume, which underlined the explanatory power of the index for households.

The likelihood ratio test has LR = 0.984 and $c2 = 0.03$. At degrees of freedom = 1 since the number of parameters lost in the PCA2 is = 1, the null hypothesis that the two models are not distinctly different is not rejected at confidence level = 88%. A lower confidence level is due to University of Michigan index pertains to the consumers' beliefs about the macro-economy, whilst the PCA2 pertains to the household investors' trust in social media. The two models are decidedly similar at 88% confidence level, implying that the PCA2 is the extracted beliefs component from the market and NLP sentiments/

Table 4: Robustness tests against Shiller individual index

Date range	2001 Aug to 2015 Sep		
No of samples	170		
Model with PCA		Original variables model	
Variable	Coefficient	Variable	Coefficient
PCA2	2.40	Baker-Wurgler sentiment	2.17
	(0.72***)		(1.15*)
c	34.3	Social media sentiment	60.6
	(0.61***)		(23.1***)
		c	36.4
			(1.14***)
R^2	0.061	R^2	0.062
Log likelihood	-577.8	Log likelihood	-577.7

The Shiller individual crash index reflects the expectations or beliefs of wealthy individuals. When it is regressed against the market and NLP sentiment, most of the explanatory power should come from its 'extracted beliefs', which is the PCA2 component. This is tested against the model with PCA2. The likelihood test indicates that the two models cannot be rejected for dissimilarity. The bracketed terms below are the standard errors. *** indicates significance at the 1% confidence level. ** indicates significance at the 2.5% confidence level, and * at the 5% confidence level.

Table 5: Robustness tests against University of Michigan consumer confidence index

Date range	1998 Jan to 2015 Sep		
No of samples	213		
Model with PCA		Original variables model	
Variable	Coefficient	Variable	Coefficient
PCA2	7.73	Baker-Wurgler sentiment	6.0
	(0.66***)		(1.04***)
c	85.7	Social media sentiment	274.9
	(0.66***)		(26.5***)
		c	95.4
			(1.2)***
R^2	0.35	R^2	0.43
Log likelihood	-812.8	Log likelihood	-799.8

The University of Michigan consumer confidence index reflects the consumers' beliefs on the general economy, which is slightly different from the PCA2 reflecting the household investors' trust in social media for investing. The bracketed terms below are the standard errors. *** indicates significance at the 1% confidence level. ** indicates significance at the 2.5% confidence level, and * at the 5% confidence level.

3.3 Increased trust in social media for investing

I next try to identify the principal components and their impacts on the stock markets. Individually, they are used as regressors against the excess market return[5]. The PCA1 has a significant coefficient, whilst the PCA2 coefficient has no significance. The PCA1 is contributed by all the three sentiment indices and reflects mainly news headlines sentiment as shown in table 2. Its significant coefficient is consistent with Tetlock (2007) earlier result that WSJ headlines correlates with stock returns.

Although the second PCA belief component does not have a significant impact on the market excess return, it still marginally correlates with stock returns through the HML factor[6]. This is shown in

regression analysis in table 6 with its significant coefficient at 2.5% confidence level.

These two findings - the increasing second PCA belief component since the 2008 and its significant coefficient in the HML factor regression supports the hypothesis there is a growing trust amongst household investors since the 2008 and this trust has a positive impact on the stock markets.

3.4 Social Media impact on household participation in the equity markets

I next postulate how social media impacts household investors and their participation in the stock markets. Does this trust in social media actually motivates them to invest? To examine this question, I use a multi-variate GARCH conditional correlation model as in Tse and Tsui (2002) for the

[5] The S&P 500 index is used to proxy the broad index.

[6] The HML factor in the Fama French 3-factor model is also referred to as the value premium and is the difference in

returns between companies with high book-to-market ratio (value) and low book-to-market ratio (growth) stocks.

Baker-Wurgler and social media sentiment, regressing against the news media sentiment. The model equations are as below.

Hence $i = 1$ denotes for the Baker-Wurgler sentiment index and $i = 2$ for the social media sentiment.

$$y_{i,t} = c_i x_t + \epsilon_{i,t} \quad for\ i \in \{1, 2\} \tag{1}$$

$$\begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix} = H_t^{\frac{1}{2}} v_t \tag{2}$$

$$H_t = D_t^{\frac{1}{2}} R_t D_t^{\frac{1}{2}} \tag{3}$$

$$R_t = (1 - \lambda_1 - \lambda_2)R + \lambda_t \Psi_{t-1} + \lambda_2 R_{t-1} \tag{4}$$

In equation 5, y_t represents a 2 by 1 vector equation for the Baker-Wurgler and social media sentiment which is regressed using the x_t news media sentiment. The dynamic correlation of the residuals $e1$, t and $e2$, t in equation 5 are modelled by the cholesky factor of covariance matrix Ht in equation 2 with i.i.d. $nt \_ N(0, 1)$. Ht itself is modelled by a diagonal matrix Dt of conditional variances given by equation 5.

$$D_t = \begin{bmatrix} \sigma_{1,t}^2 & 0 \\ 0 & \sigma_{2,t}^2 \end{bmatrix} \tag{5}$$

The volatilities of $s2\ 1$, t and $s2\ 2$, t are modelled by ARCH equations after considering AIC criterion.

$$\sigma_{i,t}^2 = s_i + \alpha_i \epsilon_{i,t}^2 \quad for\ i \in \{1, 2\} \tag{6}$$

Rt is the 2 by 2 matrix of conditional correlations updated by an auto-regressive equation 4 with rolling estimator of correlation matrix Yt⊠1 and an unconditional correlation R. $l1$ and $l2$ are parameters that control the dynamics from the rolling estimator and the unconditional matrix. The maximum-likelihood estimates for the parameters are given with the standard error.

The time series plot of the dynamic correlation is observed in figure 3, with a generally positive correlation between the two sentiments.

Table 7: Calibration results for multi-variate GARCH correlation

Baker-Wurgler equations		Social media equations	
c_1	-1.36 (0.29***)	c_1	0.59 (0.02***)
ARCH system of equations			
α_1	0.96 (0.12***)	α_2	0.78 (0.14***)
s_1	0.012 (0.003***)	s_2	0.0001 (0.0000***)
Conditional correlation equation			
R	0.81 (0.12***)		
λ_1	0.15 (0.044***)		
λ_2	0.58 (0.124***)		

The bracketed terms below are the standard errors. *** indicates significance at the 1% confidence level. ** indicates significance at the 2.5% confidence level, and * at the 5% confidence level. All coefficients are significant at the 1% confidence level.

The yearly averages of these dynamic correlations for years when the CSF is published are calculated - 1998, 2001, 2004, 2007, 2010 and 2013[7] against the household direct participation rate in the stock markets is shown in table 8. A null hypothesis that household participation in the stock markets does not increase monotonically with this dynamic correlation between the two sentiment indices is tested using the Spearman's rho correlation. The Spearman rho correlation is a non-parametric test frequently used in Biostatistics and for this case is calculated as 100% in the table 8. This rejects the null hypothesis with degrees of freedom of 6 and one-tailed 0.1% confidence interval having significant critical value at 0.99. See Zar (1984) for the critical values table.

The significant monotonic relationship between the dynamic correlation and the household equity participation rate is telling, especially when the market and social media sentiments were regressed against the news sentiment. In an earlier section, the PCA2 is determined to be the household investors' latent trust in social media for investing. Although this trust is found to increase since 2008, household participation in equity has actually declined. This is counter-intuitive but seen in the light of the two stages in Prospect Theory, trust in the social media for investing need not translate to increased household participation.

Since the news sentiment is a proxy for the main economic news and thence also a proxy for the excess market return, the monotonic relationship suggests that household investors' trust in the social media is conditional on prevailing economic fundamentals. This is consistent with Guiso et al.

(2003) that cited macro-economic factors as an important determinant for the stockholding puzzle.

Guiso et al. (2018) in a study of Italian investors had also found that their risk aversion post 2008 financial crisis had increased and they had divested more stocks. This was attributed to emotional-based changes in the utility functions of the investors. Whilst our data is based on the American households, a similar projection is that American households could have adjusted their risk preferences in the second stage of the Prospect Theory to be more risk averse, and thence decreasing their participation in equities.

---

[7] The Baker-Wurgler sentiment index was updated only till 2015 Sep at the time of writing. Using the year average from the first 9 months of Sep as a proxy for year 2016, the monotonic trend is still observed.

Figure 3: Plot of household participation against multi-variate dynamic correlation
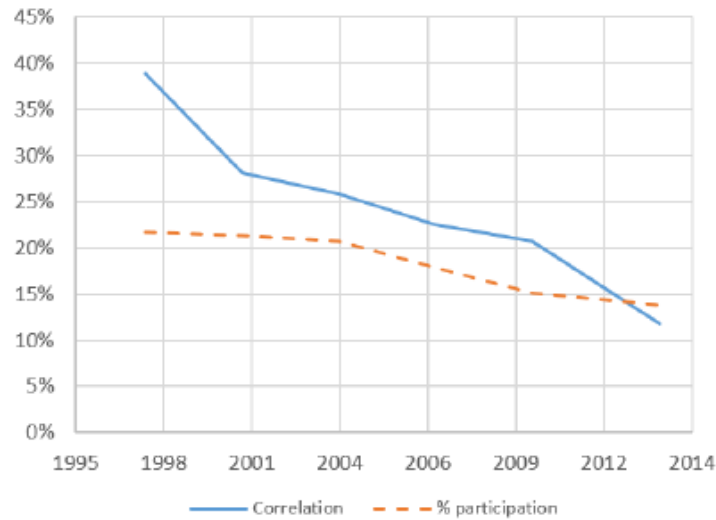
The plot shows the percentage household participation in the stock markets against dynamic correlation between the Baker-Wurgler and social media sentiments in the SCF years 1998 to 2013. A non-parametric ranked correlation test shows a significant monotonic relationship between the two sentiments at confidence interval of 100%.

Table 8: Spearman's rho ranked correlation test

Year	Correlation	Household participation	Correlation rank	Household participation rank
1998	0.39	21.7	1	1
2001	0.28	21.3	2	2
2004	0.26	20.7	3	3
2007	0.23	17.9	4	4
2010	0.21	15.1	5	5
2013	0.12	13.8	6	6

The Spearman's rho is non-parametric ranked correlation test that is specially suited for smaller samples. The variables to be compared are sorted and ranked. The difference in their ranks constitutes a test statistic that is compared to critical values to assess a significant monotoinic relationship. The rank difference in this table adds up to 0, which establishes a significant monotonic relationship between the dynamic correlation and household participation.

13

5. Conclusion

Over the last two decades social media has become increasingly popular. This article has two main results on the impact of social media on household investors. The first is the pro-cyclical trust of household investors of the social media for stocks investing. This trust was high in the late 1990s to early 2000s with formerly popular investment chat boards and Internet bulletins, but gradually declined with the waning of this popularity. In the past decade from 2008 to 2015, this trust has gradually increased again with the recent interest in pyramidal guru-follower system.

This increasing trust does not however translate to greater stock market participation amongst the household investors. In fact from 2006 to 2013, the stock market participation rate has been declining. The paper finding points to that household participation in the stock markets depends on their trust on the social media but this is conditional on the prevailing economic fundamentals.

The paper results are derived using the broad market index. Further research can be extended on a panel study of individual stocks to determine social media impact on their trading pattern. The panel features may include the stock size, beta or even geographical location.

6. References

Black, F. (1986). Noise. *Journal of Finance, 41*.

Bollena, J., Ma, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*, 1–8.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2006). Down or out: assessing the welfare costs of household investment mistakes. *Journal of Political Economy, 115,* 707–747.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Measuring the financial sophistication of households. *American Economic Review*, 99, 393–398.

Campbell, J. (2006). Presidential address: Household finance. *Journal of Finance*, *61*, 1553–1604.

Chawla, N., Da, Z., Xu, J., & Ye, M. (2017). Information diffusion on social media: Does it affect trading, return, and liquidity? SSRN.

Chen, H., De, P., & Hu, Y. (2014). Wisdom of crowds: The value of stock opinionstransmitted through social media. *Review of Financial Studies, 27*.

Guiso, L., Haliassos, M., Jappelli, T., & Claessens, S. (2003). Household stockholding in europe: Where do we stand and where do we go? *Economic Policy, 18*, 123–170.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *Journal of Finance, 63*, 2557–2600.

Guiso, L., Sapienza, P., & Zingales, L. (2018). Time varying risk aversion. *Journal of Financial Economics, 128,* 403–421.

Heston, S. L. & Sinha, N. R. (2017). News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts' Journal, 73*.

Hong, H., Dong, C., Tian, H., & Ungureanu, A. (2017). Mass media versus word of mouth in stock markets. *Journal of Finance*, 137–163.

Hong, H., Kubik, J. D., & Stein, J. C. (2004). Social interaction and stock market participation. *Journal of Finance, 1*, 137–163.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Li, D. & Li, G. (2011). Trusting the stock market. Federal Reserve Board publication.

Loughran, T. & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries and 10-ks. *Journal of Finance, 27*.

Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance, 62*, 1139–1168.

Tham, E. (2016). Unbearable lightness of expectations of the Chinese investor. *Handbook of sentiment analysis in finance*, 400–416.

Tse, Y. K. & Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics, 20*, 351–362.

Tumarkin, R. & Whitelaw, R. F. (2001). News or noise? Internet postings. *Financial Analysts' Journal, 57,* 41–51.

Zar, J. (1984). *Biostatistical analysis.* Prentice-Hall, Inc.

The authors

Eric Tham is a PhD candidate at
Edhec Business School.
E-mail: eric.tham@edhec.com