

Article

Interactions and Sentiment in Personal Finance Forums: An Exploratory Analysis

Maurizio Naldi ^{1,2} 

¹ Department of Civil Engineering and Computer Science, University of Rome Tor Vergata, Via del Politecnico 1, 00133 Rome, Italy; maurizio.naldi@uniroma2.it

² Department of Law, Economics, Politics and Modern languages, LUMSA University, Via Marcantonio Colonna 19, 00192 Rome, Italy; m.naldi@lumsa.it

Received: 18 June 2019; Accepted: 9 July 2019; Published: 10 July 2019

Abstract: The kinds of interactions taking place in an online personal finance forum and the sentiments expressed in its posts may influence the diffusion and usefulness of those forums. We explore a set of major threads on a personal finance forum to assess the degree of participation of posters and the prevailing sentiments. The participation appears to be dominated by a small number of posters, with the most frequent poster contributing even more than a third of all posts. Just a small fraction of all possible direct interactions actually take place. Dominance is also confirmed by the large presence of self-replies (i.e., a poster submitting several posts in succession) and rejoinders (i.e., a poster counter-replying to another poster). Though trust is the prevailing sentiment, anger and fear appear to be present as well, though at a lower level, revealing that posts exhibit both aggressive and defensive tones.

Keywords: online communities; forums; personal finance; concentration

1. Introduction

Personal finance forums represent a major example of online communities [1]. The influence of online communities on the formation of opinions cannot be overstated; the electronic word-of-mouth (eWOM) phenomenon has been studied in several contexts. In [2], the factors that drive consumers to adopt and use the messages from other consumers were investigated, with information relevance and information comprehensiveness resulting as the most vital elements for influencing information adoption within an online consumer community information. It has been shown that both eWOM content and observing other consumers' purchases significantly affect consumers' intentions to buy a product [3]. However, although more positive comments lead to more purchases in many cases, involvement and prior knowledge partially moderate the relationship between the ratio of messages and the eWOM effect, since the credibility of web sites and eWOM messages can be damaged in the long run if all the eWOM messages are positive [4]. Potential customers (i.e., those who would like to purchase a product in the near future and are currently reading reviews with the intention to decide whether or not to buy that product) find the negative reviews containing service failure information and the positive reviews containing information on core functionalities, technical aspects, and aesthetics to be more helpful [5]. The impact of eWOM is particularly relevant for the financial industry. It has been shown that banks' higher profitability is related to more verbalised positive feelings [6], and internet postings significantly affect stock prices [7]. Approaches to forecasting the prices of financial assets through the analysis of texts from different sources (financial reports, news, message boards, social media, etc.) are spurring an increasing interest in the literature, as shown in the survey contained in [8].

The behaviour of participants in online communities is a major issue in any case, since their comments may be driven by several factors. Though posters are often led to join the discussion by their search for social support and social identification [9], members of brand communities may develop “oppositional brand loyalty” towards other rival brands [10], so their opinions are often biased. In fact, a basic analysis of the major personal finance forums reveals that almost 95 percent of conversations taking place are either neutral (mostly statement-based) or negative (with 30% directed specifically at brands) [1].

In general, we can therefore assume that personal finance forums may be a significant source of influence on personal finance choices and can lead to opinions directed by personal biases rather than objective analyses. It has been shown in [11,12] that the collection of sentiments on finance-related forums may provide a useful input to machine learning tools to manage one’s own portfolio and outperform traditional asset allocation methods. Personal finance forums should therefore be monitored by companies acting in the sector (e.g., financial companies, insurance companies, banks, etc.) in order to understand which forces drive them and how the interests of companies may be affected.

In this paper, we analyse personal finance forums, exploring their interaction dynamics and the sentiments conveyed by posts. In particular we wish to address the following research questions (RQs):

1. Is participation in online personal finance forums uniform?
2. Do some participants take a lead and possibly exert a higher influence on others?
3. Are major participants inclusive? Or are other participants put off by their dominance?
4. What sentiments are exhibited on posts?

We employ a mixture of approaches, adopting tools from industrial economics, social network analysis, and sentiment analysis. We apply those tools to the top threads of a major personal finance forum to get an exploratory view of the behaviour of posters on personal finance forums.

We provide the following contributions and findings:

- We propose some measures of dominance phenomena, borrowing them from the field of industrial economics.
- We find that dominance phenomena are present, which answers RQ1 negatively: the participation is quite far from being uniform, with a small fraction of posters submitting most papers, and the most frequent poster contributing even more than a third of the total number of posts, allowing us to answer positively to RQ2 (see Section 3).
- We find that direct interaction (in the form of a reply to the latest post) involves a small minority of posters, which adds to the negative answer to RQ1 (Section 4).
- We find that self-replies, consisting in a poster submitting a sequence of posts and therefore unduly reducing the possibilities for others to interact, make up a significant portion of the thread, which allows us to answer RQ3 (Section 4).
- We find that rejoinders, which may represent aggressive behaviour, are themselves a significant portion of the overall number of posts, providing us with further support to answer RQ3 (Section 4).
- We find that trust is the most present emotion in the language of the posts, but anticipation and anger are strongly present as well, allowing us to answer RQ4 (Section 5).

2. Datasets

Our analysis of behaviours in personal finance forums relies on datasets extracted from the SavingAdvice website [13]. In their own words, Saving Advice is a community built around the notion that saving money is good, releasing free articles and information on all aspects of personal finance including debt reduction, how to save money, how to make money, and how to invest. In this section, we describe the datasets that have been employed in this paper.

We considered the top 5 threads (by number of posts) in the personal finance section of the SavingAdvice website. We could have considered a larger number of threads; however, we randomly selected other threads below the top 5 and observed that the results do not change significantly. In addition, our purpose is rather to show the applicability of the method rather than perform an exhaustive survey of all personal finance forums. For that reason, we stick to the top 5 threads as being a representative sample for our purposes. These threads are:

- *Should US wealth be redistributed?*, whose aim is to examine if redistributing the wealth of the richest people in the US to the poorest people would make any difference at all [14].
- *The problem with being budget minded is other people*, which examines the strain on the relations within a family, in the work environment, or with friends, when you are budget-minded and they are not [15].
- *Health Care and the alternative point of view*, where a lot of topics are discussed concerning the adoption of health insurance [16].
- *Indexed Universal Life Insurance*, where the value of life insurance policies is debated [17].
- *First Real Job...No Idea How to Save*, where advice is given and sought regarding saving money and plan one's financial future at the beginning of one's working life [18].

The size of these threads, as measured by the number of posts, is shown in Table 1. While the number of posts is rather homogeneous, around 200, since they are the top 5 threads on the website, the number of contributing posters varies quite widely among them. As a consequence, the average number of posts submitted by a poster varies between 4 and slightly more than 10.

Table 1. Dataset size.

Thread	No. of Posts	No. of Posters	Avg No. of Posts
Wealth Redistribution	248	31	8.00
Budget-Minded	259	64	4.05
Healthcare	217	22	9.86
Life Insurance	192	19	10.11
First Real Job	190	24	7.92

3. Dominance in a Thread

A major aspect we wish to examine was the presence of dominance phenomena, i.e., of posters dominating the interaction that takes place within the thread. For the purpose of this analysis, a thread is seen as a sequence of posts and the associated sequence of posters. In this section, we analyse dominance phenomena by employing three tools: rank-size plots, the Hirschman–Herfindahl Index (HHI), and the concentration ratio of top 4 (CR_4). After describing these tools, we examine the results of their application to our datasets. The choice of adopting those three indices rather than relying on established centrality indices in social networks, such as the degree centrality index, is that these do not consider the presence of self-loops and cannot, therefore, properly represent the phenomena we are interested in.

The notion of dominance is considered here with reference to the number of submitted posts: individuals (or a group of individuals) are dominant if they submit a large fraction of the posts.

We first examine how posts are distributed by looking at the rank-size plot: after ranking posters by the number of posts they submit, the frequency of posts is plotted against the rank of the poster. It is to be noted that the rank-size distribution of the number of posts coincides with the degree distribution of the associated social network, examined, e.g., on Twitter [19]. We observe, of course, a decreasing curve, but the rate of decrease provides us with relevant information. First, the higher the rate of decrease, the higher the concentration of posts in the hands of a restricted number of posters (hence, their dominance). Second,

we may observe a power law (a.k.a. a generalised Zipf law), where the number of posts submitted by the poster with rank i decreases as the following expression (see [20] for a description of its characteristics):

$$N_{(i)} = \frac{k}{i^\alpha}, \tag{1}$$

where k is a normalising constant and α is the Zipf exponent. On a doubly logarithmic scale, the power law relationship appears as a linear one, with the Zipf exponent measuring the slope of the log-linear curve. Hence, the higher that slope α , the higher the dominance. If we focus on the frequency ratio between the poster of rank i and its runner up in the ranking list, we easily obtain

$$R_i = \frac{N_{(i)}}{N_{(i+1)}} = \left(1 + \frac{1}{i}\right)^\alpha. \tag{2}$$

The sequence of R_i values gives us a more intuitive idea of how fast the frequency of posters decreases with the rank and how prevailing a poster is with respect to its less frequent colleagues. The Zipf law, first introduced by Zipf himself in [21] to describe the distribution of words, has been observed in a number of contexts, from the distribution of populations over cities [22] to the distribution of links on the web [23], the size of companies [24], and the distribution of traffic over a telecommunications network [25]. Assessing whether the distribution of posts follows Zipf’s law is relevant since it allows us to apply the body of knowledge developed for such a distribution.

In Figure 1, we have plotted the observed rank–size relationship for our 5 threads. We see a roughly linear trend, indicating an adherence to Zipf’s law for 4 out of 5 cases, leaving out the life insurance dataset, where the frequency of posts falls at a lower rate than what we would expect from a power law.

By performing a best quadratic fit, we get a rough estimate of α , reported in Table 2. We see that the Zipf exponent is around 1.5 for 4 out of 5 cases and is much lower for the budget-minded dataset. In order to get a more intuitive feeling of how dominant some posters are, we can recall the R_1 and R_{10} ratio, as defined in Equation (2). For the slowest decrease (lowest α), exhibited by the budget-minded thread, we see that the top poster submits nearly twice as many posts as its runner-up ($R_1 \simeq 1.97$), while the tenth most frequent poster still submits 10% more posts than its runner-up ($R_{10} \simeq 1.1$). For the fastest decrease (the healthcare thread, since the life insurance thread deviates significantly from a power law, so that we cannot apply Equation (2)), we see that the first poster submits nearly three times as many posts as the second poster $R_1 = 2.81$, and the tenth poster is 15% higher than its runner-up ($R_{10} \simeq 1.15$).

Table 2. Concentration measures. HHI—Hirschman–Herfindahl Index.

Thread	Zipf Exponent	HHI	CR ₄
Wealth Redistribution	1.336	0.1277	60.08%
Budget-Minded	0.981	0.0592	37.07%
Healthcare	1.490	0.1556	68.20%
Life Insurance	1.504	0.1446	70.31%
First Real Job	1.444	0.2008	71.58%

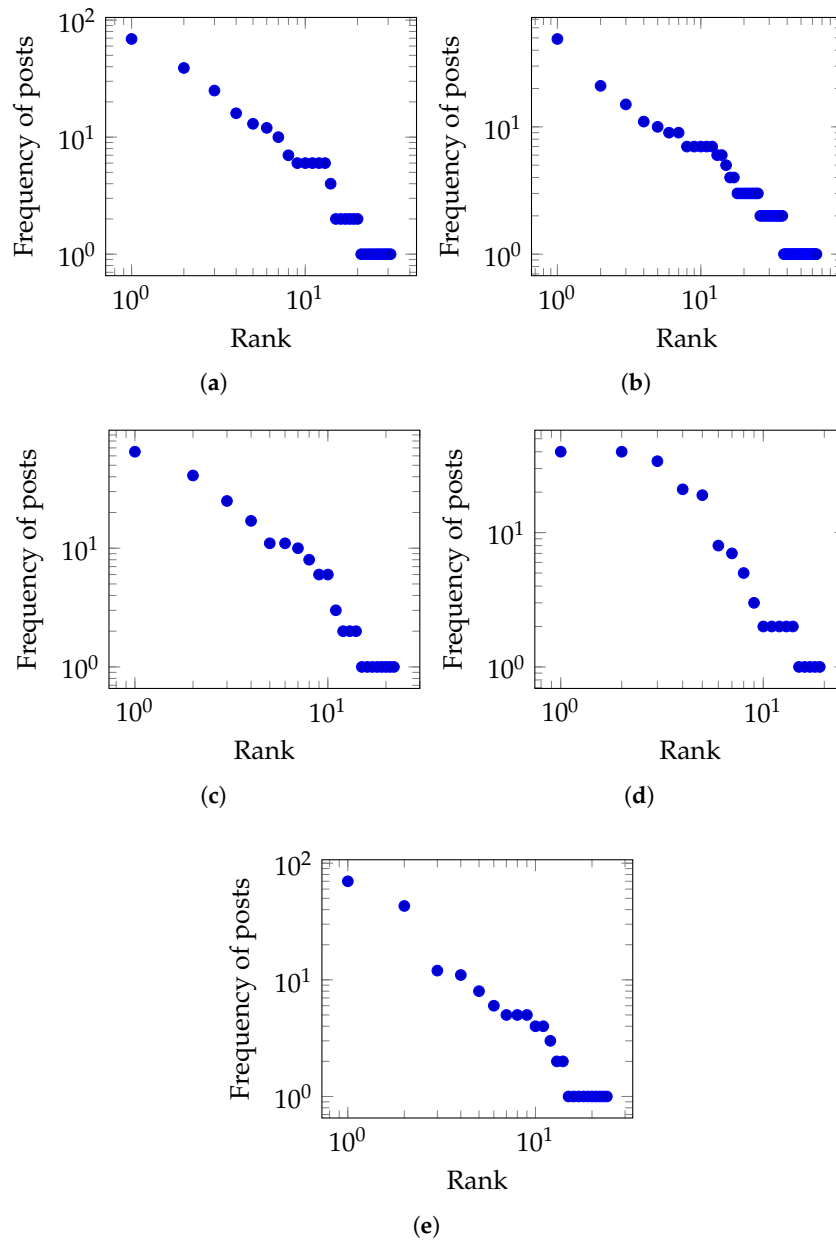


Figure 1. Rank–Size plot. (a) Wealth Redistribution dataset; (b) Budget-Minded dataset; (c) Healthcare dataset; (d) Life Insurance dataset; (e) First Real Job dataset.

As more general indices to assess dominance position, we borrow two from industrial economics: the Hirschman–Herfindahl Index (HHI) [26–28] and the CR₄ [29,30]. For a market where n companies operate, whose market shares (expressed as a fraction of the total market) are v_1, v_2, \dots, v_n , the HHI is

$$HHI = \sum_{i=1}^n v_i^2. \tag{3}$$

The HHI satisfies the inequality $1/n \leq HHI \leq 1$, where the lowest value corresponds to the case of no concentration (perfect uniform distribution of the market) and the highest value represents the case of

monopoly. Therefore, the larger the HHI, the larger the concentration (hence, the dominance). Instead, the CR₄ measures the percentage of the whole market owned by the top four companies:

$$CR_4 = \sum_{i=1}^4 v_i. \tag{4}$$

Similarly, the higher the CR₄, the heavier the concentration.

In our case, the fraction of posts submitted by a poster can be considered as its market share, so that the HHI can be redefined as

$$HHI = \sum_{i=1}^n \left(\frac{N_i}{m}\right)^2 = \frac{\sum_{i=1}^n N_i^2}{m^2}, \tag{5}$$

where *m* is the number of posters and *N_i* is the number of posts submitted by the generic *i*-th poster.

Instead, the CR₄ can be defined in terms of poster frequency as

$$CR_4 = \frac{\sum_{i=1}^4 N_{(i)}}{\sum_{i=1}^n N_i}. \tag{6}$$

For our datasets, we get the results reported in Table 2. According to the guidelines provided by the U.S. Department of Justice, the point of demarcation between unconcentrated and moderately concentrated markets is set as HHI = 0.15 [31]. In our datasets, we find that the HHI is higher than that threshold in the Healthcare and First Real Job cases, very close in the Life Insurance case, and rather close in the Wealth Redistribution case. We can therefore conclude that a significant concentration is present in 2 out of 5 datasets.

If we turn to the CR₄, we can adopt the correspondence between market structure and CR₄ values reported in [32]. According to that classification, if CR₄ < 0.4 we have an effective competition, while we have a loose oligopoly if 0.4 < CR₄ < 0.6 and a tight oligopoly or a dominant firm if CR₄ > 0.6. In Table 2 we see, by looking at the CR₄ indicator, that the top 4 posters submit more than 60% of all the posts in 4 out of 5 cases, leaving out just the Budget-Minded dataset. Delving deeper into the top 4, as reported in Table 3, we also see that the most frequent poster typically contributes a large portion of the posts, over 20% in 4 out of 5 cases (again with the exception of the Budget-Minded dataset, where the percentage is, however, a close 18.92%), with a peak in the First Real Job dataset, where the most frequent poster contributes a staggering 36.84% of the total number of posts. According to the CR₄, we have a tight oligopoly in 4 out of 5 threads and a loose oligopoly in the fifth one.

Summing up, all three indicators suggest that a significant level of concentration is present in those forums.

Table 3. Contribution of most frequent poster.

Thread	Most Frequent Poster
Wealth Redistribution	27.82%
Budget-Minded	18.92%
Healthcare	29.95%
Life Insurance	20.83%
First Real Job	36.84%

4. Interaction Dynamics

In addition to detecting dominance phenomena through measures that encompass the whole sequence of posts, we wish to understand interaction phenomena that take place at a smaller scale, i.e., over a

very smaller subsequence of posts. In this section, we build the social network that is embedded in each thread and define measures that help us understand how posters interact. In particular, we look at two indicators—the number of self-replies and the number of rejoinders—that may reveal obnoxious behaviour.

We first build the social network embedded in each thread. The network graph is built by employing the following rules:

- the nodes of the network represent the posters;
- an edge is drawn between node i and node j if poster i replies to a post submitted by poster j (i.e., poster i comes immediately after poster j in the sequence of posts);
- the weight of the edge between node i and node j is the number of times poster i replies to a post submitted by poster j .

We end up with a weighted directed network. The resulting network for the sample case of the Healthcare thread is shown in Figure 2, where we have adopted a degree centrality-based visualisation (actually, we show here the core of the network, removing all nodes whose degree is just 1). In that network, we do not show self-loops (representing self-replies) in order to not clutter the graph. Even if we do not consider self-replies, that degree-based representation shows a dominance of a restricted number of posters (shown in the core of the graph). Similar graphs are obtained for the other threads.

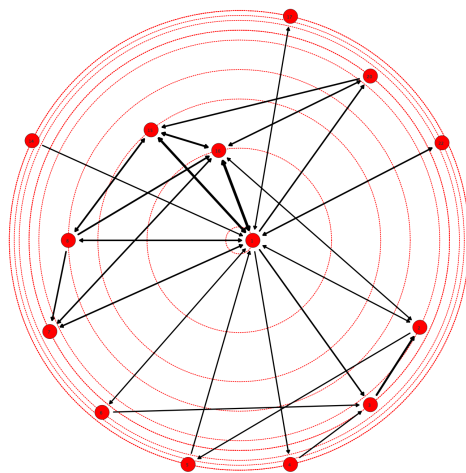


Figure 2. Social network for the Healthcare thread.

As we can see, just a small subset of the posters interact directly. We would have a complete direct interaction if any poster would reply to the post of any other poster at least once. In that case we would get a fully meshed network, which is not the case of Figure 2. Instead, we see that the number of edges is much lower than its potential maximum (i.e., $n(n - 1)$, which is roughly the square of the number of posters). The lack of direct interaction can be quantified by computing the sparsity of the associated weighted adjacency matrix, i.e., the percentage of zero entries. Its complement to 100% is the percentage of direct interactions. As shown in Table 4, the percentage of direct interactions can be lower than 5% (i.e., a sparsity larger than 95%, as in the Budget-Minded thread) and is anyway not larger than 20%.

Table 4. Sparsity in weighted adjacency matrices.

Thread	Sparsity [%]
Wealth Redistribution	89.49
Budget-Minded	95.31
Healthcare	83.68
Life Insurance	81.16
First Real Job	87.84

We now turn to the nature of those direct interactions by examining self-replies and rejoinders.

A poster may exert an excessive influence over the others not just by submitting more posts (which we detect by measuring the degree of concentration through one of the indices described in Section 3), but also by submitting posts in a rapid sequence, one after the other like a machine gun. We would therefore observe a sequence of posts submitted by the same poster; we call those posts self-replies because it is like the posters replying to themselves. Self replies may represent an obnoxious behaviour because they occupy the virtual space and reduce the possibilities for other posters to state their opinion. It is a behaviour counter to normal turn-taking. A related issue has been studied in verbal communications, where turn-taking has been shown to have a significant impact on the dynamics of a conversation [33,34].

Another relevant phenomenon is that of rejoinders. In fact, we may have the case of a poster A replying to a poster B who had just themselves replied to the poster A. This behaviour, which could be the manifestation of a flame or a sort of tit-for-tat, can be obnoxious as well, though the real nature should be better assessed by checking for the presence of aggressive or offensive language. We call the corresponding posts rejoinders. We classify as rejoinders not just the quick ABA patterns, but also the more complex ABBA, ABBBA, and so on, where the rejoinder comes after a sequence of self-replies.

In Table 5, we report the number of self-replies and rejoinders detected in our 5 threads. We can compare those figures with the overall number of posts shown in Table 1. Self-replies represent a large portion of posts for the Life Insurance thread (20.3%), while they are just a minor portion (4.2%) for the Budget-Minded thread. Rejoinders are instead a widespread phenomenon for all threads, representing more than a quarter of all posts for the Wealth Redistribution, Healthcare, and First Real Job threads.

Table 5. Self-replies and rejoinders.

Thread	No. of Self-Replies	No. of Rejoinders
Wealth Redistribution	38 (15.3%)	64 (25.8%)
Budget-Minded	11 (4.2%)	35 (13.5%)
Healthcare	32 (14.7%)	55 (25.3%)
Life Insurance	39 (20.3%)	41 (21.4%)
First Real Job	23 (12.1%)	51 (26.8%)

Overall, the dynamics show a heavy presence of direct interactions between a small subset of specific pairs of posters and a significant presence of self-replies.

5. Sentiment Analysis

After describing the social network structure embedded in forums, in this section we delve into the text making up the posts and investigate the sentiments expressed by posters. Actually, the dialogues among the posters may convey emotions rather than neutral statements of facts in support of their opinions concerning finance decisions. Sentiment analysis can help us uncover those emotions, though expressed in natural language. According to [35], sentiment analysis tasks can be classified into two categories: emotion detection and polarity analysis. For our purposes, the first category seems more appropriate since most forum threads concern a topic more complex than just the binary approach associated with polarity analysis, where there is only either favour or opposition towards a specific item could be observed.

Here we have opted for a simple bag-of-words approach (see Chapter 2.2 of [36]). Though more advanced models may be employed, such as those based on neural networks [37], the bag-of-words approach retains the major advantages of being very intuitive and directly related to the text, as well as not requiring any training phase, the relevant information being contained in the lexicon employed. Though several tools exist for detecting emotions, here we have opted for open source tools. In particular, we have chosen the NRC Word–Emotion Association Lexicon (a.k.a. EmoLex). This lexicon, as described in [38–40], associates words to emotions, adopting the classification proposed in [41], which considers eight emotions. These emotions can be naturally paired into four pairs of opposites: joy–sadness, anger–fear, trust–disgust, and anticipation–surprise. This lexicon comprises 13,889 words. The sentiment computation is rather straightforward. The occurrence in the text of a word appearing in one of the categories (say, the disgust one) counts as 1 in the sentiment score for that category. So, if a sentence contains 3 words listed in the list of words for disgust, the score for that sentence in the disgust category will be 3. When using the NRC lexicon, rather than receiving the algebraic score due to positive and negative words, as in a basic polarity analysis, each sentence gets a score for each sentiment category. The NRC lexicon was applied through the *syuzhet* (see <http://cran.r-project.org/web/packages/syuzhet/>) package in R.

In Figure 3a through Figure 3e, a spider web diagram is used to illustrate the degree to which those emotions are present in our threads. We can now examine the four pairs.

We see that in all threads, trust largely prevails over its opposite, disgust. Anger is present roughly as strongly as fear, which may envisage a roughly equal distribution of posts between those exhibiting an aggressive behaviour and a defensive one. Anticipation has a slight prevalence over its counterpart surprise, meaning that contents are somewhat expected. Joy and sadness are instead present roughly at an equal level.

This view is confirmed in Figure 4, where the mean scores are reported for all emotions across the threads; each symbol (dot, square, etc.) represents a single thread, and the bar width represents the range of scores obtained for a specific sentiment over the set of 5 threads. Trust confirms itself as the emotion present at the largest extent, while its opposite, disgust, is that showing the lowest score. This is a positive note for this kind of forum, showing that a high degree of common ground is present. The second-placed emotion is anticipation, while the second worst-placed is its opposite, surprise, again showing that posters' expectations about the content of other posts are somewhat fulfilled. Anger and fear are unfortunately a close third and fourth position respectively, revealing that both aggressive and defensive texts are present.

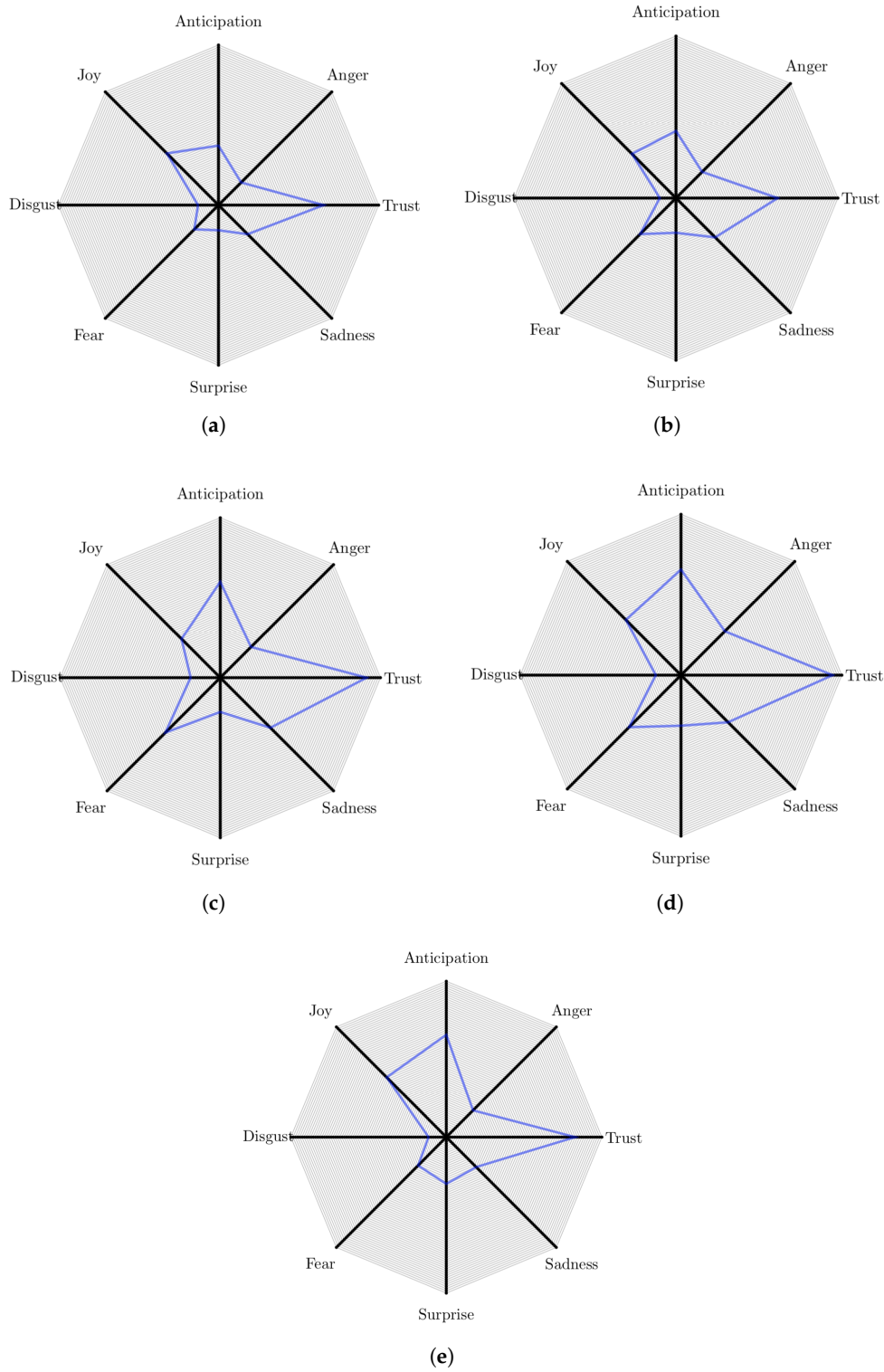


Figure 3. Mean sentiment scores. (a) Wealth Redistribution dataset; (b) Budget-Minded dataset; (c) Healthcare dataset; (d) Life Insurance dataset; (e) First Real Job dataset.

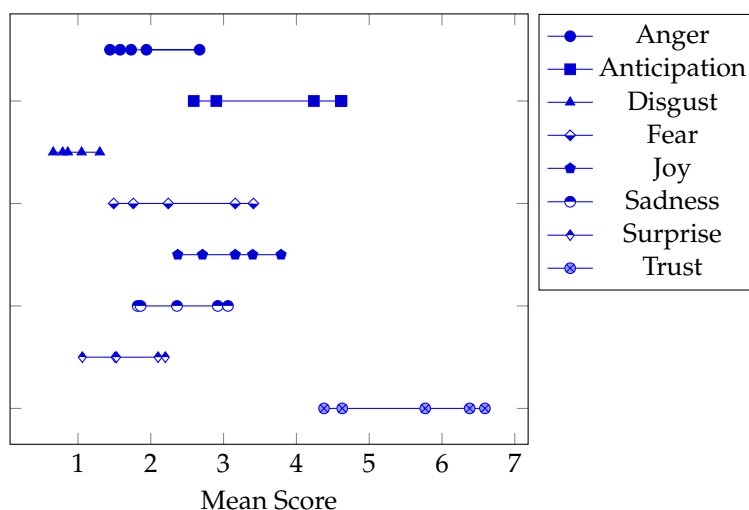


Figure 4. Mean sentiment scores for the five threads.

6. Discussion

6.1. Theoretical Contributions

The main contribution we provide is the introduction of dominance indicators, self-replies, and rejoinder patterns as tools to analyse the nature of interactions in an online forum.

Dominance indicators, in the form of the HHI, CR_4 , and Zipf exponents, allow us to detect the presence of dominant players. Though centrality indices have been defined for a long time for social networks and could be used to identify whether some posters are more central than others (a concept related to dominance), the measures we introduce, being borrowed from industrial economics, are more apt for describing concentration and dominance phenomena than degree centrality, which does not consider self-replies.

Similarly, self-replies and rejoinders can be computed by going through the sequence of posts and reveal other facets of dominance phenomena. In fact, a poster submitting posts in succession reveals their intent to occupy the virtual space and monopolise the discussion. Similarly, a poster counter-replying to another poster may be led by the intention to demolish other posters’ arguments and have their opinion prevail.

6.2. Implications for Practice

Dominance in an online forum may lead to an excessive influence in forming forum observers’ opinions. All indicators point out that dominance phenomena exist to a large extent in personal finance forums. This is to be avoided both for the forum’s moderators and the companies involved in the discussion. These companies may be, e.g., banks providing financial services (e.g., loans and mortgages), insurance companies, or companies listed on the stock exchange. Whatever their field of operation, those companies must be aware that the virtual space provided by online forums may be poisoned and invaded by biased opinions. A continuous monitoring of such online discussions is therefore advisable for those companies, and may allow them to intervene promptly and correct biased opinions that may unduly damage them. Their monitoring may employ the techniques that have been described, as well as a more specific text-mining approach tailored to their interest.

7. Conclusions

We examined five threads in a personal finance forum to investigate the characteristics of interactions and the sentiments conveyed by posts. Strong dominance phenomena appeared in most cases, with a few posters contributing most posts, even by delivering posts in succession (self-replies). Just a small fraction of posters interacted directly. Rejoinders, where a poster counter-replies to another poster, represented more than a quarter of all posts. Though trust emotions are the most conveyed by the texts, anger and fear were, respectively, third and fourth, betraying the presence of aggressive behaviour, which could also be the reason for such a large presence of rejoinders.

All these indicators point to a high degree of aggressiveness in personal finance forums, which may be a hindrance to their diffusion and usefulness as an unbiased source of financial information.

Funding: This research received no external funding

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

HHI Hirschman–Herfindahl Index

CR₄ Concentration ratio

References

- Hallam, J. *The Social Media Manifesto*; Springer: New York, NY, USA, 2012.
- Cheung, C.M.; Lee, M.K.; Rabjohn, N. The impact of electronic word-of-mouth: The adoption of online opinions in online customer communities. *Internet Res.* **2008**, *18*, 229–247. [[CrossRef](#)]
- Wang, Y.; Yu, C. Social interaction-based consumer decision-making model in social commerce: The role of word of mouth and observational learning. *Int. J. Inf. Manag.* **2017**, *37*, 179–189. [[CrossRef](#)]
- Doh, S.J.; Hwang, J.S. How consumers evaluate eWOM (electronic word-of-mouth) messages. *Cyberpsychol. Behav.* **2009**, *12*, 193–197. [[CrossRef](#)] [[PubMed](#)]
- Ahmad, S.N.; Laroche, M. Analyzing electronic word of mouth: A social commerce construct. *Int. J. Inf. Manag.* **2017**, *37*, 202–213. [[CrossRef](#)]
- Tang, C.; Mehl, M.R.; Eastlick, M.A.; He, W.; Card, N.A. A longitudinal exploration of the relations between electronic word-of-mouth indicators and firms' profitability: Findings from the banking industry. *Int. J. Inf. Manag.* **2016**, *36*, 1124–1132. [[CrossRef](#)]
- Tumarkin, R.; Whitelaw, R.F. News or Noise? Internet Postings and Stock Prices. *Financ. Anal. J.* **2001**, *57*, 41–51. [[CrossRef](#)]
- Xing, F.Z.; Cambria, E.; Welsch, R.E. Natural language based financial forecasting: A survey. *Artif. Intell. Rev.* **2018**, *50*, 49–73. [[CrossRef](#)]
- Chiu, C.M.; Huang, H.Y.; Cheng, H.L.; Sun, P.C. Understanding online community citizenship behaviors through social support and social identity. *Int. J. Inf. Manag.* **2015**, *35*, 504–519. [[CrossRef](#)]
- Kuo, Y.F.; Feng, L.H. Relationships among community interaction characteristics, perceived benefits, community commitment, and oppositional brand loyalty in online brand communities. *Int. J. Inf. Manag.* **2013**, *33*, 948–962. [[CrossRef](#)]
- Xing, F.Z.; Cambria, E.; Welsch, R.E. Intelligent asset allocation via market sentiment views. *IEEE Comput. Intell. Mag.* **2018**, *13*, 25–34. [[CrossRef](#)]
- Malandri, L.; Xing, F.Z.; Orsenigo, C.; Vercellis, C.; Cambria, E. Public mood-driven asset allocation: The importance of financial sentiment in portfolio management. *Cogn. Comput.* **2018**, *10*, 1167–1176. [[CrossRef](#)]
- SavingAdvice Website. Available online: <https://www.savingadvice.com/forums/> (accessed on 28 December 2018).

14. Should US Wealth Be Redistributed? Available online: <https://www.savingadvice.com/forums/forum/financial-chit-chat/personal-finance/4905-should-us-wealth-be-redistributed> (accessed on 28 December 2018).
15. The Problem With Being Budget Minded Is Other People. Available online: <https://www.savingadvice.com/forums/forum/financial-chit-chat/personal-finance/8056-> (accessed on 28 December 2018).
16. Health Care and the Alternative Point of View. Available online: <https://www.savingadvice.com/forums/forum/financial-chit-chat/personal-finance/42558-> (accessed on 28 December 2018).
17. Indexed Universal Life Insurance. Available online: <https://www.savingadvice.com/forums/forum/financial-chit-chat/personal-finance/55109-> (accessed on 28 December 2018).
18. First Real Job...No Idea How to Save. Available online: <https://www.savingadvice.com/forums/forum/financial-chit-chat/personal-finance/53153-> (accessed on 28 December 2018).
19. Bodrunova, S.S.; Blekanov, I.S. Power laws in ad hoc conflictual discussions on Twitter. In Proceedings of the International Conference on Digital Transformation and Global Society, St. Petersburg, Russia, 30 May—2 June 2018; pp. 67–82.
20. Read, C.B. Zipf's Law. In *Encyclopedia of Statistical Sciences*; Wiley: New York, NY, USA, 2004.
21. Zipf, G.K. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*; Ravenio Books: Cambridge, MA, USA, 2016.
22. Blank, A.; Solomon, S. Power laws in cities population, financial markets and internet sites (scaling in systems with a variable number of components). *Phys. A Stat. Mech. Appl.* **2000**, *287*, 279–288. [[CrossRef](#)]
23. Adamic, L.A.; Huberman, B.A. Power-law distribution of the world wide web. *Science* **2000**, *287*, 2115–2115. [[CrossRef](#)]
24. Axtell, R.L. Zipf distribution of US firm sizes. *Science* **2001**, *293*, 1818–1820. [[CrossRef](#)] [[PubMed](#)]
25. Naldi, M.; Salaris, C. Rank-size distribution of teletraffic and customers over a wide area network. *Trans. Emerg. Telecommun. Technol.* **2006**, *17*, 415–421. [[CrossRef](#)]
26. Rhoades, S.A. The Herfindahl-Hirschman Index. *Fed. Res. Bull.* **1993**, *79*, 188.
27. Naldi, M. Concentration indices and Zipf's law. *Econ. Lett.* **2003**, *78*, 329–334. [[CrossRef](#)]
28. Naldi, M.; Flamini, M. Censoring and Distortion in the Hirschman–Herfindahl Index Computation. *Econ. Papers J. Appl. Econ. Policy* **2017**, *36*, 401–415. [[CrossRef](#)]
29. Pavic, I.; Galetic, F.; Piplica, D. Similarities and Differences between the CR and HHI as an Indicator of Market Concentration and Market Power. *Br. J. Econ. Manag. Trade* **2016**, *13*, 1–8. [[CrossRef](#)]
30. Naldi, M.; Flamini, M. Correlation and Concordance between the CR₄ Index and the Herfindahl-Hirschman Index. SSRN Working Paper Series 2014. Available online: <https://ssrn.com/abstract=2502764> (accessed on 9 July 2019).
31. The U.S. Department of Justice and the Federal Trade Commission. Horizontal Merger Guidelines. 19 August 2010. Available online: <https://www.justice.gov/atr/horizontal-merger-guidelines-08192010> (accessed on 9 July 2019).
32. Naldi, M.; Flamini, M. The CR₄ Index and the Interval Estimation of the Herfindahl-Hirschman Index: An Empirical Comparison. 2014. Available online: <https://ssrn.com/abstract=2448656> (accessed on 9 July 2019).
33. Cowley, S.J. Of timing, turn-taking, and conversations. *J. Psycholinguist. Res.* **1998**, *27*, 541–571. [[CrossRef](#)]
34. Levinson, S.C.; Torreira, F. Timing in turn-taking and its implications for processing models of language. *Front. Psychol.* **2015**, *6*, 731. [[CrossRef](#)] [[PubMed](#)]
35. Cambria, E. Affective computing and sentiment analysis. *IEEE Intell. Syst.* **2016**, *31*, 102–107. [[CrossRef](#)]
36. Kwartler, T. *Text Mining in Practice with R*; John Wiley & Sons: New York, NY, USA, 2017.
37. Majumder, N.; Poria, S.; Hazarika, D.; Mihalcea, R.; Gelbukh, A.; Cambria, E. Dialoguernn: An attentive rnn for emotion detection in conversations. *arXiv* **2018**, arXiv:1811.00405.
38. Mohammad, S.M.; Turney, P.D. Crowdsourcing a word–emotion association lexicon. *Comput. Intell.* **2013**, *29*, 436–465. [[CrossRef](#)]
39. Mohammad, S.M.; Turney, P.D. Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, CA, USA, 5 June 2010; pp. 26–34.

40. Naldi, M. A review of sentiment computation methods with R packages. *arXiv* **2019**, arXiv:1901.08319.
41. Plutchik, R. Emotions: A general psychoevolutionary theory. *Approaches Emot.* **1984**, *1984*, 197–219.



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).