sciendo

Financial Internet Quarterly „e-Finanse" 2017, vol.13/ nr 4, s. 1-13

DOI: 10.1515/fiqf-2016-0031

INFORMATION NEEDS OF FINANCIAL MARKET PROFESSIONALS IN THE BIG DATA AND SOCIAL MEDIA ERA. THE EMPIRICAL EVIDENCE FROM POLAND

Andrzej Cwynar[1], Wiktor Cwynar[2], Robert Pater[3], Piotr Kaźmierkiewicz[4]

Abstract

To meet general objectives of the article, i.e. to check the extent to which the information needs of financial market institutions are satisfied, and to learn about whether there is a transition in this realm triggered by the advent of social media and big data, we surveyed a sample of 415 financial market professionals working in Poland. We also used logit regression models, through which we processed the survey results, to identify which factors are responsible for meeting the needs. We showed that although the information needs of financial market professionals are met to a large degree, still some potential for improvement remains in this regard. We found also that respondent-specific traits are insignificant in explaining the degree of satisfaction with data and information that is used by financial market professionals. Out of firm-specific characteristic and, the value of assets under the institution's management turned out to be the key factors explaining the distribution of responses concerning satisfaction.

JEL classification: G14, G23, G24
Keywords: financial market professionals, information needs, big data, social media, financial institutions

Received: 16.08.2016 Accepted: 20.09.2017

1 Corresponding author. University of Economics and Innovation, Lublin, Poland, andrzej.cwynar@wsei.lublin.pl.
2 University of Economics and Innovation, Lublin, Poland.
3 University of Information Technology and Management, Rzeszów, Poland.
4 CDM Pekao SA Securities.

Introduction

The conditions in which financial markets function are constantly being rearranged. This also regards the way financial market professionals provide themselves with data and information. Recently, we have observed two phenomena that can potentially alter the *status quo* in this field. The first of them is rapid development of big data technologies that refer to computer science achievements, while the other is the realm of social media that has emerged as a new source of informal data and information concerning billions of human individuals – their thoughts, habits, preferences, moods, opinions, etc. – available at low cost and in real-time. Since information is key for financial market participants and its overall efficiency, discussion on the consequences of both big data and social media advent for the usage of data and information by financial market professionals is certainly required. In Poland, only the information needs of retail investors had been investigated (SII, 2015). What we lack is a comprehensive empirical study of such needs reported by professionals working at financial institutions more or less closely related to financial markets (for the convenience of further discussion we named them "financial market professionals", in short – FMP, hereafter). We decided to bridge the gap. To do that, we designed a questionnaire through which we surveyed the sample of 415 FMP representing various financial institutions that conduct their operations in Poland, to determine their perception of the information needs, the extent to which they are satisfied, and sources used to feed them. We also processed the responses gathered through the survey within logit regression models to get better insight into the framework of financial market professionals' information needs.

Specifically, we formulated the following research questions that reflect the objectives of the article: (1) what are current information needs of FMP in Poland, (2) how are they met, (3) to what degree are they fulfilled, (4) how do FMP perceive the present state in the market for data and information underlying decision-making processes in financial markets, (5) how do big data and social media affect FMP in their efforts to get relevant and timely data and information.

The remainder of the paper is organized as follows: the next section includes a synthetic review of the respective literature used to form research hypotheses. This is followed by the methodological part in which we demonstrate data collection procedure and specification of the utilized models. After that we provide the results and discussion on them. Finally, the article ends with conclusions summing up key points.

Literature review and hypothesis development

One of the essential functions of financial markets is the production of information (Levine, 2004). It facilitates price discovery, allows for the appraisal of investment opportunities, and, ultimately, enhances the optimum allocation of resources on the entire economy level. Lee and So (2014) claim that the market for information feeding financial markets deserves at least the same attention of researchers as the financial market itself. Such an assertion is plausible as the understanding of the forces underlying the market for information is vital for exploring the mechanisms of financial asset valuation. Currently, two schools of thinking compete in explaining the mechanisms – modern portfolio theory based on the efficient market hypothesis (EMH, hereafter) and behavioural finance theory built on the noise trader models (NTM, henceforth).

EMH underpinnings were constructed by Fama (1970) and are commonly recognized both by academia and in practice. It assumes the rationality of investors who make their decisions exclusively on the basis of the information they are provided with. As a result, the market prices – set by these decisions – reflect all publicly available information and can't be predicted. Yet the theory has been challenged by repeatedly observed market anomalies – events questioning the rational actions of the investors. According to NTM, introduced by Kahneman and Tversky (1979) and developed by Shiller (1984), the anomalies occur since the choices of market participants are driven by irrational factors – i.e. noise, as Black (1986) termed it. Noise-driven investors do not trade on information. Instead, sentiment can be responsible for a disproportionately large fraction of their buy or sell decisions. This creates the opportunity for predicting securities' returns.

Both the financial market and the market for financial information have changed significantly since the publication of the ground-breaking concepts of Fama and financial behaviourists. The most material changes are

twofold: (i) the availability of information has increased sharply which has resulted in the decrease of the cost of information acquisition and, hence, trading costs, and (ii) the costs of information processing has dropped substantially due to rapid technological improvements. Bai, Philippon and Savov (2013) showed on US data that financial market prices have become more informative since the 1960s, confirming that the market evolution has led to better informed market participants. Thereby, we hypothesized that in the world of easily available and relatively cheap information the information needs of FMP in Poland are satisfied to a large degree:

Hypothesis 1 – The information needs of FMP are satisfied to a high extent.

The key sources of data and information needed by FMP are well-recognized. They include all kinds of reports and announcements emitted by the security issuers, also through online channels, professional industry and macroeconomic analyses of all sorts, various media feeds, etc. However, the spectrum of data and information sources is permanently changing, as still newer opportunities emerge while others fade away. Recently, we witnessed growing interest of the financial community in social media which can herald a shift in the set of data and information sources used by FMP towards more informal outlets in the nearest future (Greenwich Associates, 2015). Additionally, as KPMG (2008, p. 5) rightfully asserts, "technology has been the single biggest change agent in the transformation of business and the finance function", including the utilization of the technology in exchanging data and information and processing it. Hence, we hypothesized that the surveyed FMP would report data and information needs different from what was needed a couple of years ago:

Hypothesis 2 – The information needs of FMP are different nowadays compared to what was needed a couple of years ago.

As mentioned, the evolution of the market for financial information largely accelerated in the past decade due to the advent and proliferation of social media and big data. The phenomena, being two sides of the same coin, put the competing theories – EMH and NTM – in a new perspective. Social media can be considered an emergent source of informal information, potentially useful for market participants. The media create unprecedented opportunities of market sentiment monitoring and trading on it. Promising results of sentiment-based analysis literature, showing that market sentiment can be a good predictor of future stock returns (e.g. Sprenger & Welpe, 2010; Chen, De, Hu & Hwang, 2013; Heston & Sinha, 2014), prompted us to assume that social media content monitoring is prevalent in the finance community:

Hypothesis 3 – Majority of FMP monitor social media content for professional purposes on at least weekly basis.

Statistics show that the amount of data and information at the disposition of FMP is massive and rapidly growing: 2,5 quintillion bytes of data is produced per day. Moreover, 90% of the data we deal with today was produced solely within a couple of the past years (PwC, 2013). A comprehensive discussion on the current state of, and prospects before the "big data" phenomenon can be found in Einav and Levin (2014), while an elaborate and exhaustive review of diversified solutions in big data analytics was made by Elgendy and Elragal (2014) as well as by Chong and Shi (2015). In turn, Chen et al. (2016) reviewed 186 journal publications on big data issued in the years 2011-2015 to integrate their findings within one, all-inclusive report. Such huge volumes of data and information can be both an opportunity and a threat, or at least a challenge. We were curious about the perception of the phenomenon by FMP, therefore we hypothesized that they would have ambivalent attitudes to the phenomena. However, we supposed that the emergence and fast-moving development of big data technologies should result in the prevalence of optimistic views over the negative. Financial industry reports reinforce such belief. For example, according to CGMA (2013), 87% of surveyed CFOs and finance professionals (more than 2,000) thinks that big data is going to alter the way business is done over the next ten years. Turner, Shroeck and Shockley (2013) report that the share of the banking and financial market firms that agree with the assertion that big data creates a competitive advantage for their organizations has doubled in recent years. Thus, we also hypothesized that the majority of FMP would strive for taking the advantage of the plethora of data and information through adoption of specialized analytical tools based on big data technology:

Hypothesis 4 – The attitude of the FMP community towards big data is ambivalent, yet with a prevalence of positive opinion.

Following KPMG (2008) predictions, we assumed also that in the pursuit of automated decision-making

dashboards, the majority of FMP would be convinced of the increasing importance of visualising data and information. The hypothesis can be also justified by the empirical findings of CGMA (2013) according to which more than 1/3 of investigated finance professionals admitted that what still fails in their organizations is the intelligent visualisation and reporting of data. The bulk of literature devoted to the adoption of visualisation tools in business, including in the financial domain, is largely diversified and rapidly growing. Valuable reviews of various visualisation techniques and approaches for the users of multidimensional financial data can be found in Marghescu (2007) and Ko et al. (2016). Rigorous demonstrations of novel data visualisation applications and tools are delivered by, for example, Khalil, Reza, Junaedi and Kanigoro (2015) and Schaefer et al. (2011). External and internal conditions of financial data visualisation development are comprehensively discussed by Flood, Lemieux, Varga and Wong (2015) on the example of monitoring systemic financial stability and Kandel, Paepcke, Hellerstein and Heer (2012) through interviews with data analysts. To recap, we formulated the following hypothesis:

Hypothesis 5 – The majority of FMP recognize the increasing relevance of data and information visualisation.

Finally, we hypothesized that despite large volumes of data and information to be potentially used, FMP lack one, all-inclusive database equipped with analytical tools interfacing with them. This seems to be plausible as according to a CGMA study (2013) 62% of surveyed finance professionals pointed to the fact that bringing data together from different sources poses a great challenge in capturing and extracting valuable insights. Importantly, this was the most frequently indicated weak point in the CGMA research. Therefore, we proposed as follows:

Hypothesis 6 – The majority of FMP need an all-encompassing database that integrates data and information from various sources.

Methodology

In the first step of the research we designed a questionnaire to learn the opinions of FMP on their professional information needs and the way they are fulfilled. In the second step we carried out the pilot study to check whether the research tool was properly designed. Since we found no reasons for revising the tool, we conducted the survey between April 25, 2016 and June 6, 2016. We used the CATI technique with the partnership of market research firm. 415 financial market professionals working at financial institutions active in Poland participated in the survey. In the sampling process we strived to trace the population of financial institutions operating in Poland that offer investment services. This includes various types of funds (mutual funds, pension funds, PE/ VE funds, etc.), asset management firms, brokerage houses, and also banks specialized in investment services as well as universal banks having capital market departments in their internal structures.

In another step we estimated models to explain the extent to which information needs of respondents are satisfied by data and information sources used in their institutions (Y_i), that take the following values for an *i*-th respondent:

$$Y_i = \begin{cases} 1 & \text{very low} \\ 2 & \text{low} \\ 3 & \text{average} \\ 4 & \text{high} \\ 5 & \text{very high} \end{cases}$$

Since the variable is ordinal we used an ordered logit model. It takes the following form (Greene 2002, p. 736):

$$y^* = \boldsymbol{m}'\boldsymbol{\alpha} + \varepsilon \quad (1)$$

where y^* is a latent variable,

y is the observable variable,

$\boldsymbol{m}$ is a vector of explanatory variables,

$\boldsymbol{\alpha}$ is a vector of parameters.

There are *N* surveyed individuals. Vector $\boldsymbol{m}$ will contain basic characteristics of surveyed professionals and institutions represented by them, as well as additional variables that will describe possible spatial patterns and how respondents perceive 'information'.$y=J$ if $\mu_{J-1}<y^*\leq\mu_J$, where $J=1,\ldots,5$ and μ's are parameters to be estimated with $\boldsymbol{\alpha}$. The probability to select an alternative J is $p_J=p(y=J|\boldsymbol{m})=p(\mu_{J-1}<y^*\leq\mu_J)=\Lambda(\mu_J-\boldsymbol{m}'\boldsymbol{\alpha})-\Lambda(\mu_{J-1}-\boldsymbol{m}'\boldsymbol{\alpha})$. $\Lambda(\boldsymbol{m}'\boldsymbol{\alpha})=exp(\boldsymbol{m}'\boldsymbol{\alpha})/(1+exp(\boldsymbol{m}'\boldsymbol{\alpha}))$ is a logistic cumulative distribution function.

To test for statistical significance of the whole set of variables, we used a likelihood ratio (LR) test, which has a χ^2 distribution with *k* degrees of freedom, where k is the number of explanatory variables. We calculated a variance inflation factor (VIF) for each variable. We eliminated (imperfectly) collinear variables, most often

taking them as the benchmark cases. Results confirm that explanatory variables were not significantly correlated among each other.

In each of the models we applied different sets of explanatory variables (basically, the models are incrementally capacious in terms of the number of applied independent variables; in this sense, model 4, that includes the largest set of such variables, should be considered the most informative and, hence, the most valuable).

To realize the article's goals, we asked survey participants four questions. The first question (Q1) was designed to check the extent to which the information needs of FMP are satisfied. In the question we utilized a Likert-like scale from "very low" through "very high". However, we also allowed them to respond: "difficult to say" (see Table 1 for details). The replies to the second question (Q2) allowed us to get insight into the reasons behind the particular levels of satisfaction reported by surveyed FMP in Q1, alongside the variables recorded through the respondents' particular section of the questionnaire. The replies to Q2 were used as five independent variables in Model 4 (see Table 1 and Table 2 for details).The third question (Q3) responses revealed the values placed by FMP on various sources of data and information. We were also curious about the way usage of social media influences the degree to which information needs of FMP are met, therefore in the fourth question (Q4) we asked whether they use such media. Various social media were used as independent variables in Model 4 (see Table 1 and Table 2 for details).To recap, the questions from 2 through 4 served as a vehicle enabling us to better understand the reasons behind perceptions reported in the responses to the first question.

Results and discussion

Distribution of survey responses

Our survey showed that almost ¾ of the sample believe that the information needs in their work are fulfilled to at least a high degree (46% – high; 27% – very high), positively verifying Hypothesis 1. Those who perceive the degree as low or very low comprised merely 5% of the surveyed sample. Some explanations of such results can be inferred from the responses to the third question through which we were examining key information sources of the studied sample. On its supply side, the market for financial data and information in Poland seems to be quite well-developed. There are many diversified sources of such content, which was reflected in the cafeteria-style option list for selection by respondents featured in our questionnaire. This means that financial market professionals have got many possibilities they can choose among and this can partly explain their satisfaction in the field of providing them with information relevant to their work. A considerable percent (more than 20%) uses Thomson Reuters / Bloomberg or similar terminals that are commonly acknowledged as well-fitted to typical needs of financial institutions and thus their users' disappointment is rather unlikely.

However, despite the plethora of financial information sources, the majority of respondents (56%) agreed with the statement that what we lack is one, all-encompassing database and information platform for financial market professionals (which can be considered a positive verification of Hypothesis 6). This suggests that there is room for a data and information provider that would integrate the resources dispersed among many individual platforms.

Our study showed that the share of social media users in the entire sample is sizable (almost 64%), however they refer to them for informational purposes mostly occasionally. Hence, Hypothesis 3 turned out to be the only one that was negatively verified. Seemingly, the distribution of responses to the first question might challenge the prospects before social media as the source of information content having an incremental value. Yet, one must remember that what is of a pivotal meaning in the social media content is far beyond information – it is its tone that reflects emotions and mood states of those who wrote the text. Therefore, we think that large satisfaction concerning the feed of information needs doesn't mean neglect of social media value.

Thought provoking is what showed up as a consequence of confronting the distribution of responses to the two initial questions (Q1 vs. Q2). Apparently, there is a dissonance between large satisfaction of respondents in covering their informational needs reported in the first question, and the percentage of those who agreed with the statement that the needs are sufficiently well-met in the succeeding question (only 23%). We considered the option "f" in second question ("The information needs

are sufficiently well-met, hence there is no necessity for a change in this field") a kind of a "control statement" when designing the questionnaire, and hence we expected that the percent of individuals who agree with the statement should roughly confirm what was revealed by the responses to the first question. To our surprise it turned out that it didn't. We can't exclude that this may denote an inconsistency in responding to the survey questions, obscuring superficiality of the respondents' attitude towards the research. There is an avenue for further investigation in this respect.

In contrast, we were not surprised that almost 85% of all respondents believe that we need different data and information today compared to what was needed several years ago, which supports Hypothesis 2. In conjunction with the large satisfaction concerning the information needs fulfilment, this finding can suggest that the financial data and information market evolves rapidly, yet on its supply side it seems to be able to adapt equally fast and effectively resulting in demand-side contentment. Alternatively, in the light of the striking percent of respondents who indicated the own resources of the institution wherein they work as key information source utilized by them (72%), such an observation can indicate well-developed abilities of financial institutions to gather useful data and information internally, or at least well-developed data mining competencies, with little innovative supply-side market participants. The latter seems to be more plausible since the scale of start-up activity – in terms of Western, particularly American standards – is at best moderate in Poland. The US experiences show that this start-up community is what threw down a challenge to such giants as Thomson Reuters or Bloomberg which are behind the data and information revolution in the finance industry.

Not surprisingly, the majority of respondents confirm that data and information visualisation plays an increasingly significant role (69%), positively verifying Hypothesis 5. We perceive such an observation as being consistent with broader phenomenon – what we call "image civilization", i.e. civilization of icons, pictograms, emoticons, etc. – and reflecting the abundance and massive complexity of data we deal with nowadays and, finally, the need to also synthetize it in a graphic or visual way. With reference to this phenomenon, we expected a greater percentage of respondents sharing the opinion that the magnitude of data and information is so big that it is hard to work through it (a bit less than 52% agreed with such a statement). Possibly, this can imply that a remarkable share of investigated institutions, particularly the biggest ones, have implemented some big data infrastructure solutions that alleviate the problems resulting from massive amounts of unstructured and fast moving data.

What is known for sure is that 62% of financial market professionals surveyed in this study think that big data technology can give them an edge on the market. Such a finding confirms our expectation that the availability of big data solutions should result in more FMP perceiving the high volume of data an opportunity as compared to the number of FMP perceiving it as a threat (i.e. positive verification of Hypothesis 4).

What is of key concern in the distribution of responses to the question regarding the most important information sources (Q3) is the astonishingly low share of those who indicated the data and information supplied by the security issuer (26%). Not only the existing literature (e.g. EFRAG, 2009), but also common sense suggests that the outlet is crucial. Some light on the puzzling observation can be shed by rare empirical studies devoted to corporate web sites as the tool of communication between companies and their stakeholders carried out so far in Poland. For example, Sajnóg (2015) found in a comparative study that German companies were more effective in communicating through the web site as compared to Polish firms. However, we noticed also that our sample was substantially diversified across the types of institutions in terms of their replies to Q3. For those that are directly engaged in investing in the capital market, data and information emitted by the security issuer turned out to be clearly more important as compared to the entire sample (it was indicated by 56% of brokerage houses, as well as by 43% of investment funds, asset management, private equity and venture capital firms). By contrast, only 16% of universal banks pointed out this source.

The most frequently indicated sources perceived as key in providing the respondents with data and information in our study included: (1) own resources of the institution (72%), (2) institutions supplying basic statistics concerning the economy and financial market – Central Statistical Office, National Polish Bank, and Warsaw Stock Exchange (56%), (3) news platforms (newswires) specialized in financial topics and issues (41%). Interestingly, firms specialized in delivering financial data – and not news –

were indicated by merely 12% of surveyed individuals.

To summarize, the views expressed in our study by financial market professionals confirm rapid mutation of the market for data and information supporting the decision-making process in financial institutions. They also show that the supply side of the market keeps pace with the changes, although from our study it is not clear if it is due to self-provision with data and information by financial institutions, or due to innovative transformation of external data providers (the first explanation seems to be more plausible, particularly in the light of the tremendous popularity of own informational resources of the investigated entities). Such a conclusion can be supported not only by the small percentage of respondents reporting a low degree to which their information needs are satisfied, but also by the small share of those who reported an adverse effect of massive amounts of data and information.

Table 1: Distribution of survey responses

	Entire sample, N=415	Investment bank. N=44	Universal bank. N=105	Brokerage house. N=27	Other brokerage activity, N=49	Investment funds, asset management, private equity, venture capital. N=28	Insurance company. N=54	Financial advisory company. N=43	Other financial institutions. N=17	Other. N=48
Q1. To what extent are your information needs satisfied by the sources you use in your work?										
Very high	**27,00%**	38.6%	34.3%	29.6%	8.2%	35.7%	18.5%	27.9%	29.4%	20.8%
High	**46.3%**	47.7%	44.8%	51.9%	46.9%	28.6%	48.1%	48.8%	52.9%	47.9%
Moderate	**20.2%**	9.1%	12.4%	14.8%	38.8%	25,00%	24.1%	20.9%	17.6%	25.0%
Low	**2.4%**	2.3%	1.9%	0,00%	2,00%	7.1%	3.7%	2.3%	0.0%	2.1%
Very low	**2.4%**	0,00%	4.8%	3.7%	0.0%	0.0%	5.6%	0.0%	0.0%	2.1%
Difficult to say	**1.7%**	2.3%	1.9%	0.0%	4.1%	3.6%	0.0%	0,00%	0.0%	2.1%
Q2. Indicate the statements you identify with in your work.										
We need different data and information from what we needed a couple of years ago (info_change)	**84.8%**	86.4%	93.3%	70.4%	85.7%	71.4%	87,00%	81.4%	76.5%	83.3%
The magnitude of data and information is so big that it is hard to work through it (info_much)	**51.6%**	52.3%	56.2%	48.1%	49,00%	53.6%	55.6%	44.2%	47.1%	47.9%
The advantage is hidden in the analyses of massive amounts of non-homogenous data and information (info_large)	**61.9%**	75%	65.7%	59.3%	61.2%	50%	59.3%	58.1%	64.7%	56.3%
Visualization of data and information gains in value (info_visual)	**69.2%**	81.8%	81.9%	51.9%	73.5%	60.7%	66.7%	39.5%	58.8%	72.9%
Despite progress, we lack an all-encompassing database for financial market professionals (info_one)	**56.4%**	56.8%	58.1%	48.1%	67.3%	50%	59.3%	51.2%	52.9%	52.1%
Information needs are sufficiently well-met, hence there is no necessity for a change in this field	**23.1%**	20.5%	26.7%	37%	20.4%	7.1%	27.8%	18.6%	11.8%	25.0%
None	**0.5%**	0.0%	0.0%	7.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Q3. Indicate three key sources of data and information used in your work.										
Security issuer (periodic reports, website, meetings with representatives)	**26,00%**	34.1%	16.2%	55.6%	20.4%	42.9%	18.5%	41.9%	5.9%	20.8%

Resources of the institution where I work	**72,00%**	79.5%	87.6%	55.6%	73.5%	53.6%	68.5%	62.8%	64.7%	64.6%
Bloomberg. Reuters. SP Capital IQ. Factset. HIS etc.	**21.4%**	29.5%	18.1%	66.7%	10.2%	32.1%	7.4%	16.3%	17.6%	22.9%
Central Statistical Office, National Polish Bank, Warsaw Stock Exchange	**55.9%**	56.8%	66.7%	40.7%	59.2%	50%	46.3%	48.8%	58.8%	56.3%
General information agencies (e.g. Polish Press Agency, Interactive Poland, Polish Information Agency)	**21,00%**	15.9%	20.0%	29.6%	18.4%	28.6%	22.2%	20.9%	41.2%	12.5%
Specialized information (news) services (e.g. Bankier.pl, Money.pl. Stooq.pl)	**41.2%**	40.9%	42.9%	18.5%	46.9%	46.4%	44.4%	46.5%	41.2%	33.3%
Specialized providers of financial data and information (e.g. Notoria, Quant Research, Financial Data Box)	**16.9%**	6.8%	18.1%	14.8%	20.4%	14.3%	25.9%	11.6%	17.6%	16.7%
Other (which?)	**11.6%**	4.5%	8.6%	7.4%	12.2%	7.1%	24.1%	9.3%	11.8%	16.7%
Q4. Indicate the key social media you use for professional purposes										
Discussion forums / QandA portals (media_forum)	**37.6%**	47.7%	33.3%	33.3%	28.6%	39.3%	27.8%	44.2%	52.9%	47.9%
Blogs (media_blog)	**19.5%**	25.0%	9.5%	22.2%	18.4%	35.7%	9.3%	27.9%	35.3%	25.0%
Microblogs (including Twitter) (media_mblog)	**13.7%**	13.6%	7.6%	25.9%	8.2%	17.9%	9.3%	13.6%	23.5%	22.9%
Social networks (e.g. Facebook or Linkedin) (media_network)	**39.8%**	38.6%	31.4%	29.6%	49.0%	42.9%	37.0%	48.8%	58.8%	41.7%
Social news platforms (media_news)	**14.9%**	20.5%	12.4%	3.7%	6.1%	14.3%	11.1%	25.6%	17.6%	25.0%
Other (media_other)	**11.1%**	4.5%	10.5%	14.8%	8.2%	17.9%	13.0%	11.6%	5.9%	14.6%
None (media_none)	**36.1%**	34.1%	42.9%	37.0%	38.8%	28.6%	46.3%	30.2%	23.5%	22.9%

Source: Own study

Logit regression results

Table 2 contains the results of logit regression analysis. We used it to model the extent to which information needs are satisfied by data sources used in the surveyed institutions. Essentially, respondent-specific characteristics (experience and job position) turned out to be insignificant in all studied models. Although insignificant, the relation between experience in financial asset management (exp_general) and the information requirements fulfilment was positive, reflecting slightly higher information need satisfaction reported by more experienced respondents. Possibly, it may be due to more experience in managing information. Years of practice on the job (exp_occup) also did not have significant meaning for information need satisfaction. However, managers who spent more time in their current positions tended to be slightly less satisfied with the information they had access to. Both exp_occup and exp_general variables signs are robust to sample length and model specification. Interestingly, out of all the studied job positions, CFOs / accountants (occup_findir) belonged to the least satisfied respondents, while the executives were the only job position class outperforming the benchmark (directors / managers) in terms of satisfaction.

We observed some differences among the types of economic entities respondents work at and the extent to which their information needs were satisfied, as reported by the surveyed individuals. In comparison to investment banks (selected as the benchmark), all other institutions were less satisfied with the information they possessed. Significantly less satisfied were, among others, investment (mutual) funds, asset management, private equity, and venture capital firms (firm_fund). As mentioned before, only about half of surveyed managers reported the value of assets under management of his / her institution (asset_value). Interestingly, the variable turned out to be important in explaining the degree of information

requirements fulfilment. We found that its relation to the dependent variable was statistically significant at p=0.05 (Model 1). The higher the value of managed assets, the more satisfied respondents were as to their information needs coverage. We believe that this may be explained, at least partly, by the funds they have access to, as greater resources mean access to more timely and accurate information. Alternatively, assuming that the asset value is a proxy of the institution size, the relationship we detected seems to be rational as bigger institutions can afford more, including information, and have the means to better serve their information needs. We noticed also that the type of assets managed by surveyed institutions matters in explaining the dependent variable, yet it is not crucial (statistically insignificant). Firstly, institutions that managed more types of assets were slightly less satisfied with the available information (asset_number in Model 1). Secondly, more risky assets came with more satisfaction of managers with information they had at hand, but only to a certain point. Beyond that point, greater risk was accompanied by smaller satisfaction (a non-linear, parabolic relationship). This result may indicate that more dedicated information is needed by the institutions having different risk profiles of managed assets.

We found no visible spatial pattern of data and information needs fulfilment, despite some differences between NUTS2 regions in Poland (with Mazowieckie as the benchmark region, where the capital city is located). Podkarpackie (south-eastern region) and Pomorskie (northern region) exhibited the lowest, while Warminsko-Mazurskie (north-eastern region) – the highest level of satisfaction.

When designing the questionnaire, we assumed that Q2 replies would serve as the means of shedding additional light on the reasons behind the respective levels of satisfaction with data and information sources mined by the surveyed institutions. We found two of them especially important in terms of the association with the explained variable: the need for one, complete and comprehensive source of data and information (info_one), and that financial advantage may be gained by analysing large collections of heterogeneous data and information (info_large). In both cases we obtained negative signs concerning the relationship of the responses with the explained variable. It means that data and information provision dissatisfaction is more likely among those who agree with the statement that what we lack is one, all-encompassing data delivery platform for financial market professionals, as hypothesized by us. The increased likelihood of dissatisfaction is shared with those who agreed that market advantage could be achieved through mining diversified, non-uniform data sets which implicitly requires the combination of various data and information needs, and, possibly, integrating them into a comprehensive system. In this sense the responses of the individuals surveyed by us can be perceived as coherent.

We evidenced also that not using social media for professional (i.e. data and information delivery) purposes (media_none) decreases the extent to which the needs in this field are met, though insignificantly in a statistical sense. Additionally, the regression results indicate that the usage of social news platforms (media_news) improves the satisfaction and that the relationship is statistically significant.

Table 2: Logit regression results

Variable	Model 1		Model 2		Model 3		Model 4	
exp_occup	−0.056	(0.096)	−0.004	(0.075)	−0.005	(0.076)	−0.025	(0.081)
exp_general	0,038	(0.077)	0,033	(0.063)	0,039	(0.066)	0,034	(0.068)
asset_value	0,182	(0.088)**						
asset_number	-0,107	(0.141)						
firm_fund			−0.821	(0.526)	−1.291	(0.556)**	−1.535	(0.572)***
firm_house			−0.571	(0.492)	−0.817	(0.514)	−0.970	(0.538)*
firm_insurer			−1.049	(0.428)**	−1.189	(0.445)***	−1.345	(0.465)***
firm_advisory			−0.531	(0.428)	−0.702	(0.440)	−1.098	(0.466)**
firm_otherfin			−0.204	(0.573)	−0.227	(0.590)	−0.550	(0.610)
firm_ubank			−0.302	(0.366)	−0.547	(0.383)	−0.593	(0.398)
firm_broker			−1.569	(0.423)***	−1.926	(0.449)***	−1.928	(0.467)***

firm_other			−0.750	(0.441)*	−0.906	(0.462)*	−1.176	(0.478)**
occup_analyst			−0.077	(0.499)	−0.267	(0.517)	−0.107	(0.535)
occup_executive			0,219	(0.326)	0,104	(0.339)	0,161	(0.350)
occup_specialist			−0.189	(0.440)	−0.398	(0.447)	−0.272	(0.460)
occup_broker			−0.054	(0.332)	−0.469	(0.352)	−0.438	(0.367)
occup_findir			−0.682	(0.451)	−0.853	(0.465)*	−0.620	(0.506)
occup_owner			−0.181	(0.310)	−0.271	(0.322)	−0.173	(0.331)
occup_other			−1.114	(0.641)*	−1.493	(0.657)**	−1.327	(0.702)*
asset_pfixed			−0.236	(0.331)	−0.140	(0.345)	−0.040	(0.355)
asset_ffixed			−0.300	(0.472)	−0.304	(0.494)	−0.297	(0.510)
asset_pstock			0,14	(0.297)	0,072	(0.307)	0,061	(0.315)
asset_fstock			0,343	(0.387)	0,543	(0.404)	0,593	(0.415)
asset_estate			0,108	(0.317)	0,264	(0.332)	0,379	(0.354)
asset_hrisk			−0.328	(0.279)	−0.284	(0.288)	−0.185	(0.305)
doln					−0.384	(0.435)	−0.435	(0.445)
kuja					−0.028	(0.390)	0,146	(0.398)
lube					−0.846	(0.441)*	−0.692	(0.452)
lubu					−0.220	(0.740)	0,031	(0.770)
lodz					0,588	(0.600)	0,6	(0.607)
malo					0,297	(0.420)	0,483	(0.434)
opol					0,875	(1.375)	1,225	(1.431)
podk					−1.009	(0.435)**	−0.915	(0.462)**
podl					0,201	(0.548)	0,345	(0.570)
pomo					−0.852	(0.413)**	−0.717	(0.430)*
slask					0,601	(0.429)	0,457	(0.458)
warm					1,431	(0.620)**	1,644	(0.642)**
wiel					−0.124	(0.396)	−0.172	(0.406)
zach					0,249	(0.696)	0,3	(0.705)
info_change							0,065	(0.327)
info_much							−0.103	(0.216)
info_large							−0.391	(0.221)*
info_visual							−0.380	(0.249)
info_one							−0.668	(0.216)***
media_forum							−0.300	(0.271)
media_blog							0,03	(0.294)
media_mblog							−0.122	(0.347)
media_network							0,063	(0.276)
media_news							0,677	(0.323)**
media_other							0,111	(0.350)
media_none							−0.453	(0.341)

N	207		400		400		400	
Log-likelihood	−255.6		−471.1		−457.3		−440.9	
AIC	527,3		996,1		996,6		987,9	
BIC	554		1103,9		1160,3		1199,4	
HQ	538,1		1038,8		1061,5		1171,7	
% predicted	47,1		47,2		51,2		52,8	
LR χ^2	52,2	[<0.01]***	123,2	[<0.01]***	150,7	[<0.01]***	183,4	[<0.01]***

Standard errors in parentheses, p-values in square brackets. * means significant at 0.1, ** means significant at 0.05, *** means significant at 0.01. Constants were included in the models, but omitted in the table. In the case of each variable *VIF*<3.

Source: Own study

Conclusions

We have evidenced that the information needs of financial market professionals at institutions operating in Poland are well-satisfied, despite some differences across sample sections distinguished on the basis of firm-specific characteristics. Adapted regression models showed that the differences can be primarily explained by the value of assets under the management of surveyed institutions, with institutions greater in terms of the asset value reporting larger satisfaction. Unfortunately, the variable was reported by a mere half of the survey participants which limits the inference potential possibly attributable to the factor. Respondent-specific characteristics turned out to be statistically insignificant as factors responsible for the extent to which the information needs of studied entities are satisfied. Despite such positive assessment of the degree to which information needs are met, the majority of respondents share the view that what we lack today is an integrated, all-inclusive database for financial market professionals. We proved that affirming such an opinion is statistically significant in explaining the differences as to reported extents to which information needs are fulfilled. Those who opted for the statement "Despite progress, we lack an all-encompassing database for financial market professionals" were less likely to report high level of satisfaction with information sources. We also showed that the majority of FMP acknowledge the potential created by the big data revolution, and that more FMP perceive it as an opportunity rather than a threat. Social media are used by the majority however only occasionally, and still the media do not add significantly to informing decisions or recommendations made by FMP. Intriguingly, a small percentage of the entire sample (about ¼) indicated data and information coming from the security issuer as key in their work. This finding requires further examination as it contradicts our expectations based on both theory and earlier empirical findings.

References

Bai, J., Philippon, T., Savov, A. (2013). Have Financial Markets Become More Informative? NBER Working Paper 19728. Retrieved from http://www.nber.org/papers/w19728.pdf (referred on 21/10/2016).

Black, F. (1986). Noise. *Journal of Finance, 41*(3), 529-543.

CGMA (Chartered Global Management Accountant). (2013). From Insight to Impact. Unlocking Opportunities in Big Data. Retrieved from http://www.cgma.org/Resources/Reports/DownloadableDocuments/From_insight_to_impact-unlocking_the_opportunities_in_big_data.pdf (referred on 10/09/2016).

Chen, H., De, P., Hu, Y., Hwang, B. (2013). Customers As Advisors: The Role of Social Media in Financial Markets. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1807265.

Chen, Y., Chen, H., Gorkhali, A., Lu, Y., Ma, Y., Li, L. (2016). Big Data Analytics and Big Data Science: A Survey. *Journal of Management Analytics, 3*(1), 1-42.

Chong, D., Shi, H. (2015). Big Data Analytics: A Literature Review. *Journal of Management Analytics, 2*(33), 175-201.

EFRAG (European Financial Reporting Advisory Group). (2009). The Needs of Users of Financial Information. A User Survey. Retrieved from http://www.efrag.org/Assets/Download?assetUrl=%2Fsites%2Fwebpublishing%2FSiteAssets%2FPAAinE%2520Paper%2520-%2520The%2520Needs%2520of%2520Users%2520of%2520Financial%2520Information%2520-%2520May%25202009.pdfandAspxAutoDetectCookieSupport=1 (referred on 10/09/2016).

Einav, L., Levin, J. (2014). The Data Revolution and Economic Analysis, *NBER paper.* Retrieved from http://web.stanford.edu/~jdlevin/Papers/BigData.pdf (referred on 21/09/2016).

Elgendy, N., Elragal, A. (2014). Big Data Analytics: A Literature Review Paper. *Lecture Notes in Computer Science*, 8557, 214-227.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, *Journal of Finance, 25*(2), 383-417.

Flood, M.D., Lemieux, V.L., Varga, M., Wong, B.L.W. (2015). The Application of Visual Analytics to Financial Stability Monitoring, Office for Financial Research Working Paper 14-02c. Retrieved from https://financialresearch.gov/working-papers/files/OFRwp2014-02c_Application-Visual-of-Analytics-to-Financial-Stability-Monitoring_revised.pdf (referred on 21/09/2016).

Greenwich Associates. (2015). Institutional Investing in the Digital Age: How Social Media Informs and Shapes the Investing Process. Retrieved from https://www.greenwich.com/sites/default/files/reports_pdf/IS-Digital_Media-2015-GR.pdf (referred to 22/09/2016).

Heston, S.L., Sinha, N.R. (2014). News versus Sentiment: Comparing Textual Processing Approaches for Predicting Stock Returns. *SSRN Electronic Journal*, DOI: 10.2139/ssrn.2311310. Retrieved from http://finpko.faculty.ku.edu/myssi/FIN938/Heston%20%26%20Sinha_News%20vs%20Sentiment_WP_2014.pdf.

Kahneman, D., Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263–91.

Kandel, S., Paepcke, A., Hellerstein, J.M., Heer, J. (2012). Enterprise Data Analysis and Visualization: An Interview Study. *IEEE Transactions on Visualization and Computer Graphics, 18*(2), 2917-2926.

Ko, S., Cho, I., Afzal, S., Yau, C., Chae, J., Malik, A., Beck, K., Jang, Y., Ribarsky, W., Ebert, D.S. (2016). A Survey On Visual Analysis Approaches for Financial Data. *Computer Graphics Forum, 35*(3), 599-617.

Khalil, A.-A., Reza, A., Junaedi, P.A., Kanigoro, B. (2015). Data Visualization Application for Analyzing Public Company Financial Statement. *Procedia Computer Science*, 59, 45-53.

KPMG. (2008). Finance of the Future – Looking Forward to 2020. Retrieved from http://www.kpmg.com/CN/en/IssuesAndInsights/ArticlesPublications/Documents/looking-forward-2020-O-0812.pdf (referred on 10/09/2016).

Lee, C.M.C., So, E.C. (2014). Aplhanomics: The Informational Underpinnings of Market Efficiency, *Foundations and Trends® in Accounting, 9*(2-3). 59-258. Retrieved from https://www.gsb.stanford.edu/sites/gsb/files/leeso2015-alphanomics.pdf (referred on 21/10/2016).

Levine, R. (2004). Finance and Growth: Theory and Evidence, *NBER Working Paper*, 10766. Retrieved from http://www.nber.org/papers/w10766.pdf (referred on 21/20/2016).

Marghescu, D. (2007). Multidimensional Data Visualization Techniques for Financial Performance data: A Review (Turku Centre for Computer Science Technical Report No. 810). Retrieved from https://www.researchgate.net/publication/31597178_Multidimensional_Data_Visualization_Techniques_for_Financial_Performance_Data_A_Review (referred on 21/09/2016).

PwC. (2013). Where Have You Been All My Life? How the Financial Services Industry Can Unlock the Value in Big Data. Retrieved from https://www.pwc.com/us/en/financial-services/publications/viewpoints/assets/pwc-unlocking-big-data-value.pdf (referred on 10/09/2016).

Schaefer, M., Zhang, L., Wanner, F., Schreck, T., Kahl, R., Keim, D.A. (2011, September). *A Novel Explorative Visualization Tool for Financial Time Series Data Analysis*. Proceedings from Third International UKVAC Workshop on Visual

Analytics, University College London, UK. Retrieved from https://bib.dbvis.de/uploadedFiles/345.pdf (referred on 21/09/2016).

Sajnóg, A. (2015). Relacje inwestorskie w polskich i niemieckich spółkach giełdowych. *Zarządzanie i Finanse. Journal of Management and Finance, 13*(1), 221-240.

Shiller, R. (1984). Stock Prices and Social Dynamics. *The Brookings Papers on Economic Activity*, 2, 457-510.

SII (Stowarzyszenie Inwestorów Indywidualnych). (2015). Ogólnopolskie badanie inwestorów–OBI 2015. Retrieved from http://www.sii.org.pl/8867/edukacja-i-analizy/badania-i-rankingi/ogolnopolskie-badanie-inwestorow-obi-2015.html#ak8867 (referred on 10/09/2016).

Sprenger, T.O., Welpe, I.M. (2010). Tweets and Trades: The Information Content of Stock Microblogs. *Working paper.* Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1702854 (referred on 12/01/2017).

Turner, D., Shroeck, M., Shockley, R. (2013). Analytics: The Real-World Use of Big Data in Financial Services. Retrieved from http://www-935.ibm.com/services/multimedia/Analytics_The_real_world_use_of_big_data_in_Financial_services_Mai_2013.pdf (referred on 10/09/2016).

Acknowledgements

The authors acknowledge financial support from the European Regional Development Fund, Priority Axis 1: "Competitive and Innovative Economy" (grant no. RPPK.01.02.00-18-0045/15).