

文章编号:1007-6735(2016)05-0457-08

DOI:10.13255/j.cnki.jusst.2016.05.008

物流人员舆情指标与生产企业股票波动 趋势预测实证研究

赵敬华¹, 程婉芸²

(1. 上海理工大学 管理学院, 上海 200093; 2. 上海证券交易所 资本市场研究所, 上海 200120)

摘要: 以物流专业论坛的交流信息内容作为研究对象, 利用语义处理技术构建了反映某类产品的物流人员的工作繁忙程度和情绪状态高低的物流人员情绪综合指标; 选取该行业产品的多个生产企业的股票走势合成一个股票指数, 通过实证方法检验了同一产品行业的物流人员情绪指标对该类产品生产企业的股票走势具有预测能力. 研究表明, 情绪指数的波动趋势常常比相应股票综合指数的波动趋势提前 40 周呈现一致.

关键词: 物流论坛; 证券市场; 预测能力; 语义分析

中图分类号: F 832 **文献标志码:** A

Empirical Study on the Logistics Staff Public Opinion Index and the Trend Prediction of Stock Movements of Production Enterprises

ZHAO Jinghua¹, CHENG Wanyun²(1. Business School, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. Capital Market Research Institute, Shanghai Stock Exchange, Shanghai 200120, China)

Abstract: Taking the exchange of information contents on the professional forum for logistics as the research object, the semantic processing technology was used to construct a comprehensive index for the logistics personnel's emotion, which reflects the busy degree of logistics personnel and the level of emotional state of a product. Meanwhile, the stock movements of a number of manufacturing enterprises of the products of a certain kind of industry were selected to synthesize a stock index, and the ability of the logistics personnel's emotion index to predict the stock movements trend of manufacturing enterprises was inspected by the empirical method. The results show that the fluctuation of sentiment index in some direction often comes out earlier in advance of 40 weeks than the fluctuation of the corresponding stock price index in the same direction.

Keywords: logistics message board; stock market; prediction capability; text analysis

收稿日期: 2016-06-22

基金项目: 国家自然科学基金资助项目(61403255); 教育部人文社科项目(14YJCZH218); 上海市高校青年教师培养资助计划(1015303801)

第一作者: 赵敬华(1984-), 女, 讲师. 研究方向: 管理信息系统. E-mail: sunnyjhlove@163.com

近10年来,文本挖掘技术在金融市场上的应用越来越广泛.使用文本挖掘技术对社交媒体内容作分析处理,可以获得语义层面上的信息,这些语义信息为分析上市公司的基本面信息、评估公司未来的盈利水平以及预测其在证券市场表现提供了有效信息.这些研究所使用的数据来源大致可以分为3类:上市公司的信息披露和第三方的专业分析报告、普通投资者的交流与讨论以及供应链其他参与者的交流与讨论.

第一类数据来源主要包括上市公司的定期与非定期报告、证券分析师的分析报告以及新闻媒体报告等.文献[1-6]研究发现,从这类数据中抽取出的语义信息(如褒(贬)义词语的使用量、语气及态度等)是辅助投资者作出及时、正确的投资决策的有效信息.然而,Kothari等^[7]认为,上市公司的定期与非定期报告的可靠性和及时性并不高,证券分析师的分析报告会存在个人偏见问题.文献[8-12]认为新闻媒体报道虽然具有较高的可靠性、及时性和客观性,但它所具备的报道内容和角度多元化的特点,使其不能成为长期稳定的投资决策信息来源.

第二类数据来源主要指普通投资者在财经类社交媒体上发布和交流的信息等.最近的研究结果表明,利用文本挖掘技术可以从不同的财经类社交媒体中获得市场层面、行业层面和个股层面的度量指标,如Bollen等^[12],Sprenger等^[13],Zhang等^[14]分别从美国社交网络及微博客服务网站Twitter中获取Twitter用户情绪和用户活跃量,Gilbert等^[15]从博客中获得用户焦虑指数,张旭^[16]从财经论坛获得发帖量的异常变动量,以及刘颖等^[17]、张永杰等^[18]、方勇等^[19]分别获取搜索引擎的搜索量和信息量等.这些度量指标与证券市场中市场指数、行业指数或个股股价的表现有着紧密关系,此类数据往往可以获得投资者的短期关注程度和情绪波动等信息,因此,其对分析市场短期走势具有借鉴性,而对于中长期的股票走势波动不具有很好的相关性.

第三类数据来源主要指供应链中各个参与者的反馈信息.Luo^[20]对美国运输部公开的公众对各个航空公司的投诉意见进行文本挖掘,发现负面意见对航空公司个股的收益和波动率均存在短期和长期影响.Tirunillai等^[21]对3个电子商务网站(Amazon.com, Epinions.com和Yahoo! Shopping)的产品评论进行分析,发现总评论量和负面情绪可以有效地预测个股的短期走势.

本文基于供应链各个参与者的反馈信息来预测

股票波动趋势.供应链中供应商、制造商、分销商以及最终客户等的产品传递靠物流人员,供应链上的物流人员是最快、最准确地感受到供应链物流流动的快慢,也是最先、最全面地察觉到行业变化的群体.因此,他们的舆情对供应链的生产企业的效益有准确的预测.

本文以中国的一个专业物流论坛为数据来源,使用文本分析技术对论坛中的用户活动情况及其发表帖子的内容进行分析,抽取出某一产品行业的物流人员的情绪高低、工作忙碌状况和意见分歧程度等信息合成物流人员情绪指数,用来预测同一产品行业的生产企业的股票走势的中期趋势.

1 机理分析

供应链结构视图如图1所示.物流人员是完成原材料、中间品及最终产品在供应链间(供应商、制造商、分销商、零售商以及最终用户)实物产品转移的执行人.他们忙碌,买卖双方必然交易频繁;他们空闲,整个供应链上的生产停滞.因此,从该类产品行业总体来说,物流人员忙碌程度与核心生产企业生产忙碌具有一致性.从某类产品的物流角度来洞察其生产情况,更具有全面性、及时性和客观性.

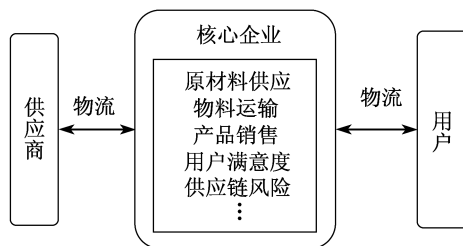


图1 供应链结构视图

Fig.1 Structural view of the supply chain

物流人员的工作忙碌程度必然跟他们的收入、论坛中交流的内容直接相关,在交流用语方面也有显著区别.例如,工作繁忙时,必然有类似“…不能参加,最近我工作实在太忙了”等用语出现且出现频率增加;同时,所遇到的问题也会相应增加,论坛中就实际问题提问、讨论以及请求帮助的帖子会增加,供应链领域的专业术语、发问和感谢等用语的出现频率都会增加.此外,物流人员的收入增加和减少又影响了他们在发帖时的情绪和用词,如“…不敢买,等过几个月再说”等表达消费用词的出现和使用频率也有显著不同.

产品的主要生产企业如果不断有产品销售出

去,那这个产品生产行业的利润必然在增加,而利润增加就有其股票走上趋势的动力。

本文以服装行业为研究对象,通过分析服装物流行业人员在物流专业论坛上的交流记录,构建服装行业物流人员的情绪指数,然后基于上海深圳股票交易所主要从事服装生产企业的股票交易构建股票走势指数,从而分析其相关性。

2 数据与实证模型

截至2007年底,中国证券市场中97%的上市公司完成了股权分置改革。大量的研究发现,在股权分置时期,中国证券市场中操纵信息披露等做法切实存在,股价未能及时准确地反映公司绩效。本文选择的研究样本周期从2008年1月1日到2013年2月8日,共261周。

2.1 物流论坛指数构建

本文的物流论坛样本数据来自物流沙龙论坛(www.logclub.com),物流沙龙论坛是目前国内最大的专业物流与供应链论坛。自2004年运营至2013年初,注册会员已达15万人,帖子总量达40万条。该论坛根据不同的行业设立了行业物流子论坛,本文从论坛发帖指数 $AVLM_t$,帖子篇幅指数 ML_t ,情绪指数 SI_t ,意见分歧指数 $DISAG_t$ 和服装生产企业股票指数 ECI 等指标入手作研究。本研究使用物流子论坛在样本区间的所有帖子作为数据样本。在此期间,平均每周的发帖数量约为77条,平均每周的活跃人数约为47人。

现介绍物流子论坛的度量指标计算步骤。

a. 整理物流论坛帖子。

与中国证券市场的工作时间对应,第 t 周论坛帖子的时间区间为:第 $t-1$ 周的周五15:00至第 t 周的周五15:00。每个帖子包含编号、作者、标题、发帖时间、内容以及所回复帖子的编号等信息。

b. 计算发帖指数 $AVLM_t$ 。

$AVLM_t$ 表示服装物流人员的活跃程度。

$$AVLM_t = H(avlm_t) \quad (1)$$

式中: $avlm_t$ 是第 t 周发帖量; $H(\cdot)$ 表示使用Holt-Winters指数平衡法进行平滑处理。

c. 计算帖子篇幅指数 ML_t 。

ML_t 表示物流人员的活跃程度。

$$ML_t = H\left(\frac{1}{AVLM_t} \sum_{a=1}^{AVLM_t} ml_a\right) \quad (2)$$

式中: $ml_a(a=1,2,\dots,AVLM_t)$ 为每个帖子包含的

字数。

d. 计算情绪指数 SI_t 。

SI_t 表示物流人员的情绪变化。该指数使用ROST EA工具^[22]计算获得。在ROST EA中,否定词用作改变情绪倾向,程度词用作加强(减弱)情绪程度。态度词是表示人们观点态度和情感倾向的词语,每个态度词根据其表达观点、情绪倾向的正负极性和程度人为给定权值,权值范围为 $[-20,20]$ 。-20表示消极情绪最大,20表示积极情绪最大。针对服装物流行业,从物流子论坛的高频词中挑选并增加到工具的自定义态度词,包括表示供应链问题的态度词(短缺、召回、闲置等),表示工作繁忙的态度词(忙、没空、加班、停工等),表示收入变动的态度词(加工资、年终奖、加班费等),表示心情好坏的态度词(开心、满足、担忧、失望等)。

若帖子 x 中有 n 个包含情感的句子,第 $k(k=1,2,\dots,n)$ 个句子中有 $Nneg_k$ 个否定词, $Ndeg_k$ 个程度词以及 $Nview_k$ 个态度词,第 j 个态度词的权重为 $w_{k,j}(j=1,2,\dots,Nview_k)$,帖子 x 的情绪值 SS_x 计算式为

$$SS_x = \frac{1}{n} \sum_{k=1}^n \left((-1)^{Nneg_k} (Ndeg_k + 1) \sum_{j=1}^{Nview_k} w_{k,j} \right) \quad (3)$$

第 t 周的情绪指数 SI_t 的计算式为

$$SI_t = H\left(\frac{1}{AVLM_t} \sum_{x=1}^{AVLM_t} SS_x\right) \quad (4)$$

$SI_t > 0$,表示服装物流人员的情绪乐观; $SI_t < 0$,表示情绪悲观; $SI_t = 0$,表示无明显情绪倾向。

e. 计算意见分歧指数 $DISAG_t$ 。

$DISAG_t$ 表示物流人员的意见分歧程度。

$$DISAG_t = H\left(\left|1 - \frac{N_{pos} - N_{neg}}{N_{pos} + N_{neg}}\right|\right) \quad (5)$$

式中: N_{pos} 表示包含积极情绪的帖子数量; N_{neg} 表示包含消极情绪的帖子数量。 $DISAG_t$ 的取值在 $[0,1]$ 之间,数值越大,表示意见分歧程度越大。

2.2 服装生产行业股票综合指数构建

本研究从中国证券市场中选取从事服装生产的55支股票作为样本股,按照以下规则,使用上海大智慧股份有限公司的大智慧传统行情软件构造服装生产企业股票指数 ECI 。

a. 以2000年1月1日为基准日期,基准日收盘价计算出的总市值为基准。

b. 市值计算式为

$$(P_1 W_1 + P_2 W_2 + \dots + P_n W_n) / m$$

式中: m 为股票总数; $P_b(b=1,2,\dots,n)$ 为第 b 支

股票的价格, $W_b (b = 1, 2, \dots, n)$ 为第 b 支股票的流通股权重.

c. $ECI = \frac{NS}{BS} \times 1\ 000$. NS 为当前市值, BS 为基准日市值.

d. 新股上市当天开始参与计算, 市值需要乘上当日原股票市值与当日股票总市值的比值.

e. 除权的计算方法. 送红股, 不影响指数; 分红, 总市值需要减去分红总额; 配股, 总市值需要加上配股总市值.

f. 停牌按上一交易日收盘价计算.

服装生产企业股票指数 ECI 的收益率 R_t 、波动率 SIG_t 和成交量 VLM_t 计算式为

$$R_t = \ln P_t - \ln P_{t-1} \quad (6)$$

$$SIG_t = \sqrt{\frac{(R_t - \bar{R}_t)^2}{\sigma_t^2}} \quad (7)$$

$$VLM_t = \ln OVLM_t \quad (8)$$

式中: P_t 是第 t 周服装生产企业股票指数的收盘价格; $\bar{R}_t$ 是样本区间的平均收益; σ_t^2 是样本区间收益的方差; $OVLM_t$ 为原始成交量.

2.3 实证模型设定

向量自回归 (VAR) 模型^[23] 将需要分析的变量都视为内生变量纳入一个系统进行分析, 不仅刻画了变量之间的动态关系, 还提供了 Granger 因果检验和脉冲响应分析等分析工具. 一个 p 阶 VAR 模型可以表示为

$$Y_t = \mu + \sum_{i=1}^p \varphi_i Y_{t-i} + e_t \quad (9)$$

式中: $Y_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{k,t})$ 是内生变量构成的 k 维向量; μ 是 k 维常量向量; φ_i 是待估计参数; p 是滞后阶数; e_t 是 k 维误差向量.

a. Granger 因果检验.

Granger 因果检验可以检验系统内某一个变量预测其他变量的能力. 以两元 p 阶 VAR 模型为例, 式(9)可以改写为

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} C_{1t} \\ C_{2t} \end{pmatrix} + \begin{pmatrix} \tau_{11}^1 & \tau_{12}^1 \\ \tau_{21}^1 & \tau_{22}^1 \end{pmatrix} \begin{pmatrix} Y_{1t-1} \\ Y_{2t-1} \end{pmatrix} + \dots + \begin{pmatrix} \tau_{11}^p & \tau_{12}^p \\ \tau_{21}^p & \tau_{22}^p \end{pmatrix} \begin{pmatrix} Y_{1t-p} \\ Y_{2t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (10)$$

式中: Y_{1t}, Y_{2t} 是内生变量; C_{1t}, C_{2t} 是常量; $\tau_{11}^i, \tau_{12}^i, \tau_{21}^i, \tau_{22}^i (i = 1, 2, \dots, p)$ 是待估计参数; $\varepsilon_{1t}, \varepsilon_{2t}$ 是误差项.

检验 Y_{1t} 和 Y_{2t} 之间的因果关系. 如果 $\tau_{12}^i \neq 0, \tau_{21}^i \neq 0$, 则存在双向因果关系; 如果 $\tau_{12}^i = 0, \tau_{21}^i = 0$, 则相互独立; 如果 $\tau_{12}^i \neq 0, \tau_{21}^i = 0$, 则 Y_{1t} Granger 引起 Y_{2t} ; 如果 $\tau_{12}^i = 0, \tau_{21}^i \neq 0$, 则 Y_{2t} Granger 引起 Y_{1t} .

b. 脉冲响应函数.

脉冲相应函数可以描述 VAR 系统内变量之间的动态影响和时滞关系. 一般地, 当式(9)可逆, VAR 模型可以作为向量移动平均模型 (vector moving average model), 用 Y_t 表示.

$$Y_t = \mu + e_t + \Phi_1 e_{t-1} + \Phi_2 e_{t-2} + \dots \quad (11)$$

式中: $\varphi_{ij} = \frac{\partial y_{i,t+s}}{\partial \varepsilon_{j,t}}$ 是系数矩阵 Φ_s 中第 i 行、第 j 列元素, 它表示在保持其他变量在所有期间不变的情况下, y_i 对由 y_j 在 t 期产生的单位冲击的 $t+s$ 期反应, 即 VAR 系统中 y_i 对 y_j 的 s 期脉冲响应. $s = 1, 2, \dots, t$.

3 实证分析结果及讨论

3.1 统计描述

表 1 给出了服装生产企业股票指数 ECI 数据和服务物流论坛数据的基本统计特征描述. 样本数据的峰度均在 2~10 之间, 大多呈现出较明显的尖峰特征. 样本数据的偏度在 -1~3 之间波动, 呈现大小不等的左(右)偏特征, 表明全部样本序列都表现出非正态性.

表 1 数据变量的统计特征描述

Tab. 1 Statistical features description of data variables

取值	P_t	R_t	SIG_t	VLM_t	SI_t	$AVLM_t$	$DISAG_t$	ML_t
平均值	3241.051	-0.002	0.785	16.720	4.408	86.090	0.363	99.854
最大值	5051.760	0.166	4.011	17.741	8.106	411.756	0.568	259.023
最小值	1571.450	-0.167	0.006	15.220	0.000	0.000	0.000	0.000
标准差	707.411	0.043	0.680	0.466	1.512	96.197	0.155	49.515
偏度	-0.062	-0.301	2.116	-0.349	-0.976	1.134	-0.888	0.860
峰度	2.674	5.212	9.014	3.186	4.217	3.565	2.641	4.698

3.2 变量相关性分析

考察样本数据之间相关性,如表2所示.在5%显著性水平下,服装生产企业股票指数 ECI 的上涨率与服装物流人员在论坛上的情绪指数、意见分歧指数和帖子篇幅指数均呈正向相关性; ECI 的波动率与物流论坛的4个指数都呈负向相关;而 ECI 的成交量和除意见分歧指数以外的3个物流论坛指数都表现出正向相关性.可见,物流论坛内容确实与服装生产企业股票指数存在紧密的相关性.

图2刻画了服装生产企业股票指数与服装物流论坛情绪指数在2008年1月1日到2013年6月28

日的走势.由图2中可见,情绪指数 SI_t 的走势与生产企业股票指数的走势具有相似性,而且先行于生产企业股票指数 ECI . SI_t 从A点开始上涨(ECI 从A'开始上涨),经历了B-C区间的震荡后(ECI 经历了B'-C'的震荡区间)在D点达到最高值(ECI 在D'达到峰值),随后一路下跌到E点(ECI 一路下跌到E'),在E-F区间和F-G区间是两波上涨、下跌趋势(ECI 在E'-F'区间、F'-G'区间小幅度的上涨及下跌).图2再一次证明了服装物流人员的情绪指标与服装生产企业的股票指数之间的相关性.

表2 数据序列的相关系数矩阵

Tab.2 Correlation coefficient matrix of data series

指数	P_t	R_t	SIG_t	VLM_t	$AVLM_t$	SI_t	$DISAG_t$	ML_t
P_t	1							
R_t	0.052	1						
SIG_t	-0.200**	-0.153**	1					
VLM_t	-0.041	0.259**	0.003	1				
$AVLM_t$	0.063	0.022	-0.269**	-0.240**	1			
SI_t	-0.172**	0.213**	-0.222**	0.363**	0.113*	1		
$DISAG_t$	0.301**	0.125**	-0.450**	0.080	0.519**	0.527**	1	
ML_t	-0.033	0.168**	-0.123**	0.439**	-0.113*	0.858**	0.384**	1

注:*,**分别表示显著性水平10%和5%.

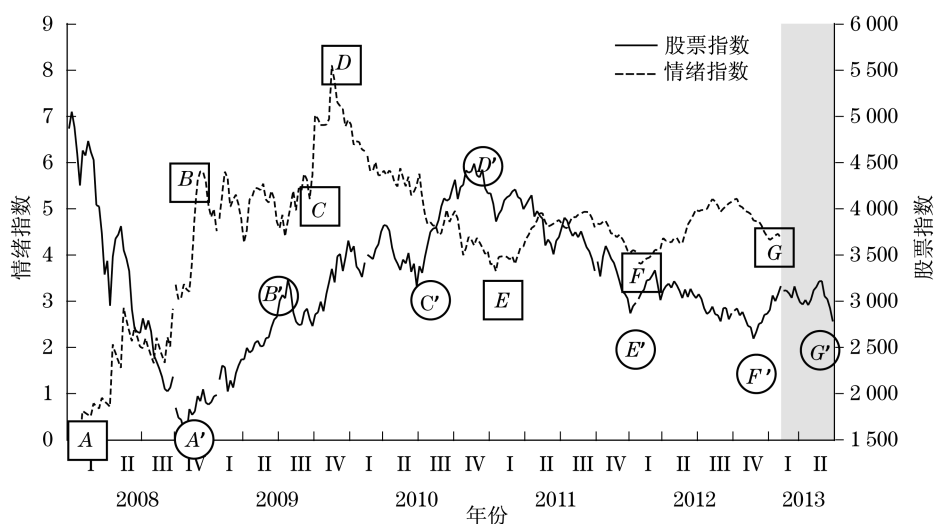


图2 服装物流子论坛情绪指数与服装生产企业股票指数的相关性

Fig.2 Relationship between the clothing logistics sub forum sentiment index and clothing production enterprise stock index

图2阴影区间为2013年2月22日至2013年6月28日,不属于样本区间,仅刻画 ECI 的走势.

3.3 VAR 模型分析

首先对全部变量采用带有截距项(趋势项)的Phillips Perron(PP)单位根检验方法进行平稳性检

验.表3(见下页)中的PP单位根检验结果表明,在1%的显著性水平下, ECI 的收益率、波动率和成交量均为平稳时间序列,物流论坛的4个指数不是平稳的时间序列,经过一阶差分后则满足时间序列的平稳性要求.

表3 单位根检验
Tab.3 Unit root test

变量	$AVLM_t$	SI_t	$DISAG_t$	ML_t	R_t	SIG_t	VLM_t	$\Delta AVLM_t$	ΔSI_t	$\Delta DISAG_t$	ΔML_t
PP 统计量	-1.057	-2.964	-2.464	-2.210	-16.570	-14.253	-8.161	-16.593	-15.162	-14.638	-16.728
结论	非平稳	非平稳	非平稳	非平稳	平稳	平稳	平稳	平稳	平稳	平稳	平稳

注:PP 统计值使用的 1%,5%,10% 显著性水平的临界值分别为 -3.455, -2.872, -2.573.

根据平稳性检验结果,建立 4 个四元向量(R_t , SIG_t , VLM_t , $\Delta AVLM_t$)', (R_t , SIG_t , VLM_t , ΔSI_t)', (R_t , SIG_t , VLM_t , $\Delta DISAG_t$)' 和 (R_t , SIG_t , VLM_t , ΔML_t)',再按照式(9)建立 4 个 VAR 模型用以研究相应变量间的动态关系. $\Delta AVLM_t$, ΔSI_t , $\Delta DISAG_t$, ΔML_t 分别为发帖指数、情绪指数、意见分歧指数、帖子篇幅指数的增量. 每个 VAR 模型的最优滞后阶数采用赤池信息量准则(AIC)进行选取,依次分别为滞后 1 阶、滞后 8 阶、滞后 3 阶和滞后 2 阶.

3.3.1 Granger 因果分析

采用 Granger 因果检验考察物流论坛内容与证券市场之间的关系的具体互动方向. 表 4 的 Granger 因果检验结果表明:a. 在 1% 的显著性水平下,情绪指数的变动与收益率存在双向 Granger 关系;情绪指数的变动与成交量之间在 10% 的显著性水平下也互为 Granger 原因. b. 在 5% 的显著性水平下,发帖指数的变动是成交量的 Granger 原因. c. 在 5% 的显著性水平下,意见分歧指数的变动是成交量的 Granger 原因,而收益率能 Granger 引起意见分歧指数的变化. d. 在 10% 的显著性水平下,帖子篇幅指数的变动与收益率存在双向 Granger 关系. 成交量在 5% 显著性水平下能 Granger 引起帖子篇幅指数的变动.

由此可见,物流论坛内容对证券市场活动也具有较强的预测能力,其中,物流子论坛的情绪指数和帖子篇幅指数可以帮助预测生产企业股票综合指数的上涨率;情绪指数、发帖指数和意见分歧指数可以帮助预测成交量.

3.3.2 脉冲响应函数分析

分别对 4 个 VAR 模型构建脉冲响应函数. 由于本文主要关注物流论坛内容能否预测证券市场相应股票波动趋势,因此,仅列出证券市场的上涨率、波动率和成交量分别对物流论坛 4 个度量指标的脉冲响应. 此外,情绪指数会在 0 期产生一个标准差大小的脉冲冲击,而这种脉冲冲击在随后的 30 期内对股票走势产生的影响围绕 0 逐步减弱,除情绪指数外的 3 个物流论坛度量指标与股票指数走势的相关影响则在随后的第 5 期后基本趋近于 0,而股票指数走势与情绪指数的响应在第 40 期后基本稳定. 因

表4 Granger 因果关系检验结果

Tab.4 Results of the Granger causality test

指数对比	原假设	卡方检验
面板 A: 情绪指数的变动对服装生产企业股票指数	R_t 不是 ΔSI_t 的 Granger 原因	26.861***
	ΔSI_t 不是 R_t 的 Granger 原因	22.751***
	SIG_t 不是 ΔSI_t 的 Granger 原因	5.852
	ΔSI_t 不是 SIG_t 的 Granger 原因	11.211
	VLM_t 不是 ΔSI_t 的 Granger 原因	17.879**
	ΔSI_t 不是 VLM_t 的 Granger 原因	15.250*
面板 B: 发帖指数的变动对服装生产企业股票指数	R_t 不是 $\Delta AVLM_t$ 的 Granger 原因	1.813
	$\Delta AVLM_t$ 不是 R_t 的 Granger 原因	0.007
	SIG_t 不是 $\Delta AVLM_t$ 的 Granger 原因	0.056
	$\Delta AVLM_t$ 不是 SIG_t 的 Granger 原因	0.347
	VLM_t 不是 $\Delta AVLM_t$ 的 Granger 原因	0.634
	$\Delta AVLM_t$ 不是 VLM_t 的 Granger 原因	3.876**
面板 C: 意见分歧指数的变动对服装生产企业股票指数	R_t 不是 $\Delta DISAG_t$ 的 Granger 原因	7.535*
	$\Delta DISAG_t$ 不是 R_t 的 Granger 原因	0.645
	SIG_t 不是 $\Delta DISAG_t$ 的 Granger 原因	4.060
	$\Delta DISAG_t$ 不是 SIG_t 的 Granger 原因	0.231
	VLM_t 不是 $\Delta DISAG_t$ 的 Granger 原因	2.273
	$\Delta DISAG_t$ 不是 VLM_t 的 Granger 原因	10.421**
指数: 帖子篇幅指数的变动对服装生产企业股票指数	R_t 不是 ΔML_t 的 Granger 原因	4.819*
	ΔML_t 不是 R_t 的 Granger 原因	6.947**
	SIG_t 不是 ΔML_t 的 Granger 原因	2.846
	ΔML_t 不是 SIG_t 的 Granger 原因	1.682
	VLM_t 不是 ΔML_t 的 Granger 原因	6.469**
	ΔML_t 不是 VLM_t 的 Granger 原因	1.806

注:*, **, *** 分别为表示显著性水平 10%, 5%, 1%.

此, 本文将短期影响定义为 5 个周期(即 5 周)内各期股票指数走势的脉冲响应, 将中长期影响定义为 10~40 个周期(10~40 周)股票指数走势的累计脉冲响应.

从表 5 可以看出:a. 情绪指数的变动对收益率存在短期和中长期影响. 短期内造成 ECI 震荡走势, 并在随后 40 周左右内都造成反向影响. b. 发帖指数的变动对 ECI 无影响. c. 帖子篇幅指数的变动对上涨率存在短期和中期影响. 短期内也造成 ECI 震荡走势, 反向影响在随后 10 周内逐渐稳定. d. 意见分歧指数的变动对上涨率存在短期和中期影响. 在第 4 周对上涨率有正向影响, 正向影响在随后 20

周内逐渐稳定。

表6的脉冲响应函数结果表明:a.物流论坛4个指数的变动对 ECI 的波动率都有短期影响,在5期内都会造成波动率的变化.b.情绪指数的变动对 ECI 的波动率有长期正向影响,影响时间将持续40周左右.帖子篇幅指数的变动和意见分歧指数的变动对 ECI 的波动率存在中期影响,持续时间分别为10周左右和20周左右.发帖指数的变动对波动率造成的中期反向影响将持续10周左右。

物流论坛4个指数对 ECI 的成交量的脉冲响应函数结果如表7所示,可以发现:a.物流论坛4个指数的变动对 ECI 的成交量均存在短期影响,在5期内都会影响成交量的变化.b.情绪指数的变动对 ECI 的成交量的长期反向影响将持续40周左右.发

帖指数的变动和意见分歧指数的变动对 ECI 的成交量存在中期影响,持续时间分别为10周左右和20周左右.意见分歧指数的变动对成交量造成的中期正向影响将持续30周左右。

综上所述,物流子论坛内容对股票指数走势的上涨率、波动率和成交量都存在显著的相关性.短期(1~5周)内,情绪指数、帖子篇幅指数和意见分歧指数对上涨率、波动率和成交量的变动都有相关性,发帖指数仅对波动率和成交量有相关性.在中期(10~30周)相关程度方面,帖子篇幅指数和意见分歧指数会对 ECI 的上涨率、波动率和成交量有相关影响.发帖指数仅与波动率和成交量的变化相关,但对上涨率不存在相关性.在长期(40周)相关性方面,情绪指数对波动率有正向相关性,且相关周期在40周左右。

表5 服装物流论坛内容与服装生产企业股票指数上涨率的影响

Tab.5 Influences of the clothing logistics forum contents and the rising rate of clothing production enterprise stock index

指数	短期					长期			
	1周	2周	3周	4周	5周	10周	20周	30周	40周
ΔSI_t	0.000	-0.004	0.001	-0.005	-0.008	-0.006	-0.008	-0.011	-0.012
$\Delta AVL M_t$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ΔML_t	0.000	-0.006	0.003	-0.001	0.001	-0.004	-0.004	-0.004	-0.004
$\Delta DISAG_t$	0.000	0.000	0.000	0.002	0.000	0.003	0.004	0.004	0.004

注:1.短期影响均为0期的单位冲击所引起的 t 期响应,长期影响则均为累计脉冲响应;

2.所有数据均处在正负两倍的标准差的可置信区域之间。

表6 服装物流论坛内容与服装生产企业股票指数波动率的相关性分析

Tab.6 Correlation analysis between the contents of clothing logistics forum and the stock index fluctuation

指数	短期					长期			
	1周	2周	3周	4周	5周	10周	20周	30周	40周
SI_t	0.000	0.032	-0.029	0.087	0.051	0.242	0.384	0.445	0.487
$AVLM_t$	0.000	-0.025	-0.005	-0.001	0.000	-0.032	-0.032	-0.032	-0.032
ML_t	0.000	0.037	-0.019	-0.002	-0.002	0.014	0.014	0.014	0.014
$DISAG_t$	0.000	-0.005	-0.004	0.014	0.001	0.012	0.008	0.008	0.008

注:1.短期影响均为0期的单位冲击所引起的 t 期响应,长期影响则均为累计脉冲响应;

2.所有数据均处在正负两倍的标准差的可置信区域之间。

表7 服装物流论坛内容与服装生产企业股票指数成交量的相关性分析

Tab.7 Correlation analysis between the logistics forum contents and the stock index volume

指数	短期					长期			
	1周	2周	3周	4周	5周	10周	20周	30周	40周
SI_t	0.000	-0.044	0.009	-0.045	-0.075	-0.228	-0.217	-0.232	-0.244
$AVLM_t$	0.000	-0.045	-0.023	-0.013	-0.007	-0.098	-0.098	-0.098	-0.098
ML_t	0.000	-0.022	-0.002	-0.011	-0.001	-0.043	-0.044	-0.044	-0.044
$DISAG_t$	0.000	0.009	-0.035	0.045	0.028	0.101	0.121	0.122	0.122

注:1.短期影响均为0期的单位冲击所引起的 t 期响应,长期影响则均为累计脉冲响应;

2.所有数据均处在正负两倍的标准差的可置信区域之间。

4 结 论

在对物流子论坛中物流人员日常讨论内容进行抽取的基础上,通过数据挖掘构建了反映服装物流人员工作繁忙程度和情绪变动的4个指数:情绪指数、发帖指数、帖子篇幅指数和意见分歧指数,分析了4个指数与上市服装生产企业股票指数的波动方向的相互关系.实证研究表明,同一产品的物流人员的综合情绪指标(由情绪指数、发帖指数、帖子篇幅指数和意见分歧指数4个指数合成)的波动方向与该类产品的上市生产企业公司股票指数的波动方向具有一致性,并且提前期约为40周.因此,通过分析某类产品的物流人员的综合情绪指标,对该类生产企业的上市公司股票指数的波动方向具有一定的预测性.

参考文献:

- [1] TETLOCK P C. Giving content to investor sentiment: the role of media in the stock market[J]. The Journal of Finance, 2007, 62(3): 1139 - 1168.
- [2] TETLOCK P C, SAAR-TSECHANSKY M, MACSKASSY S. More than words: quantifying language to measure firms' fundamentals[J]. The Journal of Finance, 2008, 63(3): 1437 - 1467.
- [3] FELDMAN R, GOVINDARAJ S, LIVNAT J, et al. Management's tone change, post earnings announcement drift and accruals[J]. Review of Accounting Studies, 2010, 15(4): 915 - 953.
- [4] 杨庆国, 陈敬良, 徐君兰. 我国新闻出版企业上市融资模式研究[J]. 上海理工大学学报(社会科学版), 2014, 36(1): 79 - 84.
- [5] DAVIS A K, PIGER J M, SEDOR L M. Beyond the numbers: measuring the information content of earnings press release language[J]. Contemporary Accounting Research, 2012, 29(3): 845 - 868.
- [6] DEMERS E, VEGA C. Soft information in earnings announcements: news or noise? [C]// Proceedings of the 2009 Meeting Papers. Marina Azzimonti: Society for Economic Dynamics, 2008, 80: 1 - 56.
- [7] KOTHARI S P, LI X, SHORT J E. The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: a study using content analysis[J]. The Accounting Review, 2009, 84(5): 1639 - 1670.
- [8] IRVINE P J. Analysts' forecasts and brokerage-firm trading[J]. The Accounting Review, 2004, 79(1): 125 - 149.
- [9] CHEN Q, JIANG W. Analysts' weighting of private and public information[J]. The Review of Financial Studies, 2006, 19(1): 319 - 355.
- [10] AGRAWAL A, CHEN M A. Do analyst conflicts matter? Evidence from stock recommendations[J]. The Journal of Law and Economics, 2008, 51(3): 503 - 537.
- [11] FRANCIS J, PHILBRICK D. Analysts' decisions as products of a multi-task environment[J]. Journal of Accounting Research, 1993, 31(2): 216 - 230.
- [12] BOLLEN J, MAO H N, ZENG X J. Twitter mood predicts the stock market[J]. Journal of Computational Science, 2010, 2(1): 1 - 8.
- [13] SPRENGER T O, TUMASJAN A, SANDNER P G, et al. Tweets and trades: the information content of stock microblogs[J]. European Financial Management, 2014, 20(5): 926 - 957.
- [14] ZHANG X, FUEHRES H, GLOOR P A. Predicting stock market indicators through twitter "I hope it is not as bad as I fear"[J]. Procedia-Social and Behavioral Sciences, 2011, 26: 55 - 62.
- [15] GILBERT E, KARAHALIOS K. Widespread worry and the stock market[C]// Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 2010: 59 - 65.
- [16] 张旭. 中国网络字频波动与股票市场关系研究[J]. 统计与决策, 2011(7): 145 - 147.
- [17] 刘颖, 吕本富, 彭康. 网络搜索对股票市场的预测能力: 理论分析与实证检验[J]. 经济管理, 2011(1): 172 - 180.
- [18] 张永杰, 张维, 金曦, 等. 互联网知道的更多么? ——网络开源信息对资产定价的影响[J]. 系统工程理论与实践, 2011, 31(4): 577 - 586.
- [19] 方勇, 孙绍荣. 影响证券投资者情绪的因素分析[J]. 上海理工大学学报, 2008, 30(2): 157 - 161.
- [20] LUO X M. Quantifying the long-term impact of negative word of mouth on cash flows and stock prices[J]. Marketing Science, 2009, 28(1): 148 - 165.
- [21] TIRUNILLAI S, TELLIS G J. Does chatter really matter? Dynamics of user-generated content and stock performance[J]. Marketing Science, 2012, 31(2): 198 - 215.
- [22] SHEN Y, LI S C, ZHENG L, et al. Emotion mining research on micro-blog[C]// Proceedings of the 1st IEEE Symposium on Web Society. Lanzhou: IEEE, 2009: 71 - 75.
- [23] SIMS C A. Macroeconomics and reality[J]. Econometrica, 1980, 48(1): 1 - 48.

(编辑: 石 瑛)