

Is Big Brother Watching Us?

Google, Investor Sentiment and the Stock Market

深澤 風土

総合政策学部 4 年

平成 27 年 10 月 30 日作成

読んだ論文

タイトル: Is Big Brother Watching Us? Google, Investor Sentiment and the Stock Market

著者: ■ Francisca Beer
(California State University, San Bernardino)
■ Fabrice Herve
(Universite de Bourgogne - CREGO - UBFC)
■ Mohamed Zouaoui
(University of Burgundy)

日付: December 6, 2012

ジャーナル: Economics Bulletin

要約

- 新しい投資家心理の指標として、Google Trends Negative Sentiment Index (GTNS) を作った。
- フランスのデータを使っている。
- この指標は、既存の投資家心理の指標と相関がある。
- この指標は、短期的にはリターンを予測し得る可能性を秘めている。

既存の投資家心理の指標

大きく分けて、サーベイによる指標とマーケットデータによる指標の二つがある。

サーベイデータ: American Association of Individual Investor Survey など

マーケットデータ: mutual fund flows, VIX など

mutual fund flows とは、投資信託のキャッシュフローのことである。投資信託に多くの人がお金を預けるほど、キャッシュフローはプラスになり、お金を引き上げるとマイナスになる。このことから、投資家が投資したいか否かの指標として使われる。

新しい指標 GTNS

まず、経済に関するネガティブな検索ワードを選定する。

- 1 Harbard IV-4 dictionary から、negative な時に使われる単語を持ってくる。(1160 個)

<http://www.wjh.harvard.edu/inquirer/homecat.htm>

- 2 google で economy と一緒に検索されるものに絞る。(63 個)
- 3 google で検索数が少ないものを消す (8 個)

"bankruptcy", "debtor", "deficit", "inflation", "liquidation",
"poverty", "recession", "crisis"

新しい指標:GTNS

これら 8 つの単語の検索数を主成分分析し、その第 1 主成分を GTNS とする。

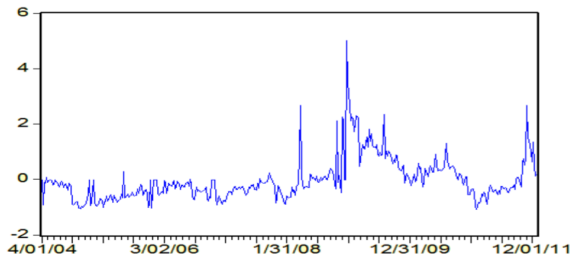
$$\begin{aligned} GTNS_t = & 0.745bankruptcy_t + 0.406debtor_t + 0.615deficit_t \\ & + 0.417inflation_t + 0.467liquidation_t + 0.402poverty_t \\ & + 0.661recession_t + 0.776crisis_t \end{aligned}$$

第 1 主成分の寄与度は 36%である。ちなみに第 3 主成分までとると 62%までいく。

新しい指標:GTNS

Figure 1: Google trends negative sentiment index

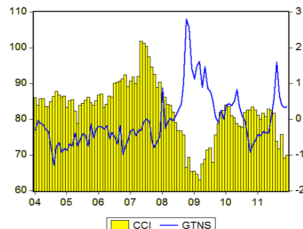
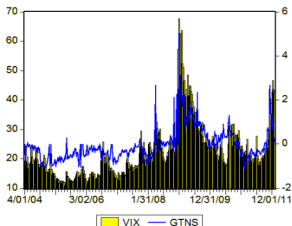
This figure displays the development of the Google trends negative sentiment index (GTNS) from January 04, 2004 to December 25, 2011.



GTNS を既存指標と比較

Figure 2: Google trends negative sentiment vs. alternative sentiment measures

These graphics display the development of Google Trends Negative Sentiment index (GTNS) against two alternative sentiment measures: the implied volatility index (VIX) derived from CAC 40 options and the consumer confidence index (CCI) published by the French National Institute of Statistics and Economic Studies. The weekly sample (left graph) contains 417 observations from January 04, 2004 to December 25, 2011. The monthly sample (right graph) contains 96 observations from January 2004 to December 2011.



GTNS を既存指標と比較

指標名1	指標名2	相関係数
GTNS	VIX	0.74
GTNS	CCI	-0.663

Figure: 相関係数

グレンジャーの因果性もみて、GTNS \rightarrow CCI は存在したがその逆は存在しなかったと言っているが、VIX については触れていない。多分なかったのだろう。

用語説明

- VIX Implied Volatility Index の略。日本語では恐怖指数。マーケットデータの代表である。オプション取引のボラティリティを元に算出される。ここでは、CAC 40 option(フランスの日経平均株価的存在のオプション) を元に算出されている。
- CCI Consumer Confidence Index の略。日本語では消費者信頼指数。サーベイデータの代表である。消費者に対して景気に関するアンケートを実施して、指数を割り出す。

mutual fund flows と GTNS の関係

$$Flow_{i,t+k} = \alpha_0 + \alpha_1 GTNS_t + \sum_m \beta_m Control_{i,t}^m + \epsilon_{i,t+k}$$

というモデルで回帰分析をする。

- Flow : mutual fund flows
- GTNS : Google Trends Negative Sentiment Index
- Control : コントロール変数。ここでは、VIX, CCI, Market return(5 lags) の3つ。

Table 1: Relationship between GTNS and mutual fund flows

This table presents the results of estimating the model (2). Fund flow represents the ratio inflow minus outflow scaled by fund total net asset for each category of funds. GTNS denotes the coefficient estimated on the sentiment variable. The coefficients on the control variables are not reported here to conserve space. The sample period includes monthly data from January 2008 to December 2011. ***, **, * indicate statistical significance at the 0.01, 0.05 and 0.10 level, respectively.

	Flow _{t+1}	Flow _{t+2}	Flow _{t+3}	Flow _{t+4}	Flow _{t+5}
Panel A: Equity fund flow					
GTNS _t	-0.003*	-0.009**	-0.001	-0.001	-0.000
Control variables	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.115	0.114	0.112	0.113	0.112
Panel B: Bond fund flow					
GTNS _t	0.005	0.004**	0.001	0.000	0.000
Control variables	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.057	0.092	0.006	0.007	0.006

結果の解釈

GTNS が上がると、Equity fund flow のキャッシュフローがマイナスになり、Bond fund flow がプラスになる。つまり、株ではなく国債を持とうとする。

このことから、GTNS が悲観的な投資家心理の度合いを示すものということが示唆される。

return と GTNS の関係

return に対して VAR モデルで推定する。

$$Y_t = \mu + \sum_{i=1}^p \phi_i Y_{t-i} + \epsilon_t$$

- Y_t : $GTNS$, R_{small} , R_{large} を含む行列。
- R_{small} : 時価総額の下位 30% の企業で組んだポートフォリオのリターン。
- R_{large} : 時価総額の上位 30% の企業で組んだポートフォリオのリターン。
- ϕ : 係数行列
- p : ラグの次数
- weekly データで、417 個ある

R_{small} と R_{large} に分ける理由

既存の研究 (Baker and Wurgler 2006,2007) によると投資家心理は全てのストックに対して同じように反応するわけではない。そこで、時価総額で大型株と小型株に分けた。

リターンの計算方法

- 時価総額の下位 30%と上位 30%のそれぞれでポートフォリオを組む。
- 毎年 6 月にリバランスを行う。
- リバランスの基準、いつを基準にしたリターンなのかは明記されていないが、リターンはおそらく前の週を基準にしたものと思われる。

単位根検定の結果

	Obs	Mean	Std.dev	ADF	PP
GTNS	417	0	1	-6.056***	-6.112***
R_{small}	417	0.004	0.021	-9.953***	-15.946***
R_{large}	417	0.001	0.029	-23.090***	-23.042***
VIX	417	22.689	9.428	-2.894**	-2.949**
CCI	96	82.552	7.956	-1.315	-1.189

- ADF : The augmented Dickey-Fuller unit root test
- PP : Phillips-Perron unit root test

VARの結果

Panel A: VAR estimation results				
	Lag	GTNS	R_{small}	R_{large}
GTNS	1	0.806*** (15.950)	-0.014*** (-5.307)	-0.005*** (-2.878)
	2	0.032 (0.642)	0.009*** (3.667)	0.002 (1.571)
R_{small}	1	0.124 (0.100)	-0.128** (-1.968)	-0.115 (-1.220)
	2	1.747 (1.133)	0.037 (0.612)	0.013 (0.320)
R_{large}	1	0.196 (0.108)	0.086*** (2.893)	0.142*** (2.165)
	2	-3.490*** (-2.063)	0.034 (0.395)	0.183*** (2.969)
Constant		0.014 (0.514)	0.001 (1.173)	0.002 (2.665)
Adj. R-squared		0.705	0.071	0.150

VAR の解釈

- GTNS に対する VAR モデルでは、lag1 では負の係数なのに lag2 では正の係数になっている。GTNS は悲観度合いを示す指標なので、lag1 ではリターンが下がると、悲観度が増すことを示している。明言していなかったが、lag2 はそれに対するカウンターと思われる。
- 各リターンに対する GTNS の係数では、 R_{large} に対する、lag2 の GTNS が有意となっている。

グレンジャーの因果性

Panel B: Granger Causality results

Null Hypothesis	Obs	lag	F-statistics	P-value
GTNS does not Granger Cause R_{Small}	415	2	14.209	0.000
GTNS does not Granger Cause R_{large}	415	2	5.737	0.003
R_{Small} does not Granger Cause GTNS	415	2	0.693	0.500
R_{large} does not Granger Cause GTNS	415	2	0.645	0.525

グレンジャーの因果性の解釈

- GTNS は R_{small} , R_{large} に対するグレンジャー因果性が存在しないという帰無仮説が棄却されたので、GTNS で R_{small} , R_{large} を説明することには正当性があると思われる。

分散分解

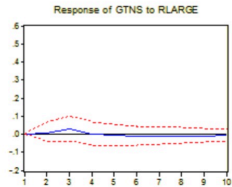
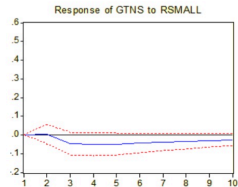
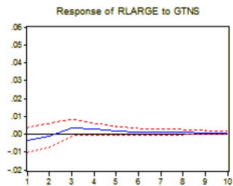
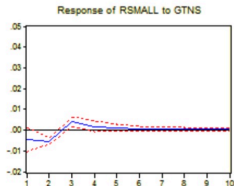
Panel C: Variance decomposition results			
	GTNS	R_{small}	R_{large}
GTNS	91.05 (0.000)	3.111 (0.112)	5.839 (0.087)
R_{small}	8.930 (0.047)	46.282 (0.000)	44.788 (0.000)
R_{large}	3.801 (0.092)	3.389 (0.104)	92.81 (0.000)

- GTNS に対するショックは、 R_{small} に対して少し働く。(約 9%)

インパルス応答関数

Figure 3: Impulse response functions

The graphs represent the impulse response functions. The dashed lines represent the upper and lower bands of the confidence interval obtained using Monte Carlo simulation. The responses are statistically significant at the 5% level when the upper band and lower band have the same sign. The horizon is on the horizontal axis and the percentage returns are on the vertical.



- 順番は $GTNS, R_{small}, R_{large}$ の順にしてある。どの順番にしても結果に大きな影響を与えなかったと書いてある。
- $GTNS$ の値が高くなると、投資家心理による価格のミスプライシングが起こる。つまり、悲観的になるとストックを買わなくなり、価格が下がり、リターンも下がる。というのを説明していると言っている。

結論

新しい指標 GTNS は既存の指標とも相間があるし、リターンの予測についても有用かもしれない、と言ってる。

おまけ:google trend の仕様について

Google Trends は 2006 年にリリースされ、Google insight は 2008 年にリリースされた。

その後、2012 年 9 月に両者が統合され、今の Google Trends になっている。

かつての google insight では、検索数の指標に fixed scaling が使われ、Google Trends では relative scaling が使われていた。

fixed scaling その検索ワードの max で割ったもの。

relative scaling 要求された期間の平均検索数で割ったもの。

しかし、統合時に fixed Trends のみになった。著者は relative scaling を使ったのは trivial に見えるが重要な新規性なんだみたいなことを言っていたので、著者には残念な感じ。

おまけ:検索数の解釈について

検索数の増加が、投資家の悲観具合か楽観具合を示すかは、あらかじめ決めるのではなく、分析してみて、価格等と負の相間があれば悲観具合であるし、正の相間があれば楽観具合を示すものとした方がいいみたいなことをおそらくこの著者は考えている。確かにという感じ。

おまけ:Google Trends データの利点

著者は、Google Trends のデータを使う利点についていくつか挙げていたが、要はこの辺が重要らしい。

- 1 データ数が多く、高頻度。
- 2 アンケートと異なり、自発的な行動に関するデータを取っているので、アンケート調査のように受け身な調査であることによる弊害などがない。(適当に答えるとか、選ばれた一部分の人の回答であることとかだろう)

おまけ:GoogleTrends のデータを使う時に暗に定まる仮定

Google Trends のデータを経済的な事象の説明に使う時、分析者は、「人は経済的な事象の情報を集める時に、サーチエンジンを使う」という仮定を置いていることになる、と言っていて、Google で調べることが当たり前になってしまった僕にとっては、なるほど、という感じで新鮮だった。ちなみにこの仮定は、Holscher and Strube(2000) という論文で、経済に関して詳しくない人は、情報を集める時にサーチエンジンを使うという実証研究があるらしく、それによって正当化されているらしい。