

# Media sentiment e mercati finanziari: analisi delle relazioni

## Abstract

Questo lavoro indaga l'esistenza di una relazione tra il *media sentiment* e l'andamento del mercato azionario. Su un campione di 7132 notizie, ottenute tra novembre 2011 e febbraio 2013, sono stati stimati per mezzo dell'analisi testuale quattro indicatori dello stato d'animo del Wall Street Journal: due di pessimismo e due di sentiment, utilizzando le categorie di due tra i più diffusi dizionari in ambito di *text analysis* (H4 e Loughran&McDonald).

I risultati delle stime di regressione, anche per mezzo dei modelli VAR, mostrano che il modello cattura la relazione tra media sentiment e cinque tra i principali indici azionari americani. I risultati, però, differiscono a seconda dell'elenco di termini utilizzato, e ciò è attribuibile a due ragioni: il Loughran&McDonald è più orientato ai termini finanziari e, soprattutto, ha un numero di elementi della categoria *positive* decisamente ridotto rispetto all'H4. Per questo motivo il primo si rivela più efficace sia in termini di evidenziazione della relazione, che di capacità predittiva (con un nesso di causalità anche bilaterale tra il mercato ed il sentiment).

**Keywords:** Media sentiment, Text analysis, Investor sentiment, Financial markets

**Jel classification:** G2, G17

**Parole chiave ADEIMF:** C107: Other Financial Markets, E102: Behavioural Finance

## Introduzione

In questo paper è descritta una analisi empirica delle relazioni esistenti tra il sentiment desumibile dai media ed i mercati azionari, al fine di valutare il nesso di causalità tra i due e la possibilità di anticiparne i movimenti.

Seguendo le constatazioni fornite già da Herbert Simon nel 1955, gli studi più recenti in ambito finanziario hanno declinato le ipotesi insite nei modelli tradizionali, in cui gli individui perfettamente razionali operano in modo da massimizzare il profitto. La *behavioral finance* pone l'enfasi su fattori che in precedenza non erano considerati, tra cui le caratteristiche del mercato che non permettono l'esistenza della concorrenza perfetta, gli elementi inerenti la psicologia degli investitori, e l'assunto che le scelte di questi ultimi non seguano sempre la perfetta razionalità.

In questo paper viene proposto un metodo di determinazione del media sentiment che prevede l'utilizzo e l'analisi dei *feed* RSS del Wall Street Journal, ovvero flussi di notizie ottenibili anche per ampi intervalli di tempo. L'assunto principale, come proposto da Baker e Wurgler (2007) e Bollen et.al (2010), è che gli investitori siano soggetti al sentiment e siano in grado di condizionare il mercato. Dopo aver ottenuto uno schema analizzabile con un foglio di calcolo, questo metodo prevede una analisi testuale, che consiste in una media di parole positive e/o negative di due dei più utilizzati dizionari psicologici.

La questione cui si cerca di dare risposta è, dunque, se esistono delle relazioni tra il sentiment così determinato da uno dei più popolari quotidiani americani e l'andamento del mercato azionario USA, descritto da cinque differenti indici: il Nasdaq 100, il Nasdaq Composite, il Dow Jones Industrial Average, il New York Stock Exchange e lo Standard & Poor's 500.

La ricerca è strutturata in quattro parti: la prima introduce le differenze tra gli approcci più obsoleti e la finanza comportamentale, ponendo l'accento sugli elementi di debolezza del modello classico come l'efficienza dei mercati e l'assenza di limiti all'arbitraggio. Inoltre vengono descritti il concetto di sentiment e le metodologie per determinarlo. La seconda parte descrive i passaggi necessari per convertire le informazioni qualitative contenute nei *feed* in informazioni quantitative, con le caratteristiche principali dei due dizionari utilizzati per la text analysis. La terza parte descrive nel dettaglio la metodologia seguita per la creazione del database, partendo dalla descrizione dei *feed* e della scelta in merito a tale strumento, cui segue l'analisi testuale e la determinazione di due tipologie di indici: uno che indica la media di termini negativi all'interno dei *feed* di ciascuna giornata, ed uno che include la differenza tra i termini positivi e termini negativi.

Nella quarta ed ultima parte, infine, sono presentati i risultati delle stime del modello, mostrando come tra i dizionari vi siano delle differenze in termini sia di efficacia che di capacità predittiva.

## **Dall'efficienza alla finanza comportamentale**

Secondo l'approccio classico alla finanza, un mercato in cui i prezzi rispecchiano pienamente tutte le informazioni disponibili è detto efficiente. In tale contesto il mercato svolge alcune funzioni aventi valenza sufficientemente generale come il trasferimento dei fondi dai settori in surplus a quelli in deficit, la formazione dei prezzi delle attività finanziarie in condizioni di concorrenza tendenzialmente perfetta, e la generazione dei tassi attesi di rendimento delle attività coerenti con il rischio.

L'idea di fondo che ha fornito le linee essenziali per lo sviluppo dei contributi più vicini a questo modello è che gli individui operino razionalmente e, di conseguenza, che i mercati siano caratterizzati da gradi più o meno ampi di efficienza.

L'impostazione tradizionale circa la razionalità degli individui nell'operare scelte di investimento è stata introdotta da Harry Markowitz (1959), e in letteratura ci si riferisce più in generale all'EMH (*Efficient Market Hypothesis*), ma tra i contributi più forti alla affermazione della ipotesi di efficienza dei mercati vi sono quelli di Eugene Fama (1970). Secondo le impostazioni dell'autore americano si evidenziano tre elementi che costituiscono il fondamento del modello di efficienza. Innanzitutto i mercati seguirebbero degli andamenti casuali (*Random Walks*) perché qualora insieme ad operatori di mercato perfettamente razionali operassero anche soggetti in modo non razionale, il comportamento di questi ultimi seguirebbe un percorso del tutto casuale. In secondo luogo, i movimenti dei prezzi si verificano soltanto in risposta a nuove informazioni e notizie, per definizione imprevedibili. Terzo aspetto rilevante è legato al concetto di arbitraggio: qualora il comportamento irrazionale di alcuni operatori non dovesse avere un andamento casuale, ma seguire un'unica direzione, si aprirebbero ampi spazi di intervento per gli arbitraggisti che riporterebbero i prezzi verso livelli coerenti con i valori fondamentali.

Nonostante siano numerosi gli elementi a favore della teoria dei mercati finanziari efficienti, la storia recente del mercato finanziario, caratterizzata da eventi come crash finanziari, scoppio di bolle speculative, evidenzia una certa difficoltà nell'adeguare alla realtà il modello di finanza tradizionale, in cui investitori perfettamente razionali "forzano" i prezzi del mercato al valore atteso dei flussi di cassa. Secondo Shiller (2000), se valesse la teoria dei mercati efficienti ad abilità diverse degli investitori non corrisponderebbero diverse prestazioni di investimento. Le persone più razionali non sarebbero in grado di operare meglio di altre in termini di attività di investimento, e ciò a causa del fatto che la loro maggiore comprensione sarebbe già del tutto incorporata nel prezzo delle azioni. Pertanto, nella descrizione che Shiller fornisce del modello tradizionale, sforzo ed intelligenza non hanno alcun valore negli investimenti e, in termini di rendimenti attesi, si potrebbero benissimo scegliere dei titoli a caso.

Quando, tra gli anni '80 e '90, cominciarono ad apparire le prime critiche all'approccio classico alla finanza, fino ad allora considerato il miglior impianto teorico, uno dei primi aspetti messi in discussione fu il presupposto secondo cui gli individui agiscono razionalmente e che, qualora fossero presenti nel mercato operatori non perfettamente razionali, questi agirebbero in maniera del tutto casuale.

Parlando di allontanamento dalla razionalità teorica, secondo gli studi di Kahneman e Tversky (1973; 1974; 1979), si possono individuare una serie di illusioni cognitive ricorrenti, tra cui l'ancoraggio, la rappresentatività, l'*overconfidence* o la *regression to the mean*. Inoltre, nei postulati della *Prospect Theory* dei due autori israeliani, si descrive il comportamento degli individui come se la funzione che descrive il valore attribuito alle perdite sia generalmente più ripida di quella che descrive l'atteggiamento rispetto alle vincite. In altre parole, gli individui assegnano un valore più elevato ad un guadagno certo, rispetto al valore che assegnano ad un guadagno atteso, benché sensibilmente più elevato. Ne deriva che il disagio determinato da una perdita risulta maggiore della soddisfazione legata ad una vincita analoga. Le evidenze psicologiche denotano, dunque, che gli investitori difficilmente operano nel mercato perseguendo le strategie passive indicate nella teoria tradizionale. Per tali ragioni assume rilevanza l'assunto della finanza comportamentale secondo il quale quando in un mercato individui razionali ed irrazionali operano congiuntamente, l'irrazionalità può determinare significativi e duraturi effetti sui prezzi.

Sono numerose le ricerche che puntano a determinare questi effetti, e sono state svolte già dalla metà degli anni '90. A titolo di esempio, Saunders (1993) e Hirshleifer e Shumway (2003) hanno studiato gli effetti dei cambiamenti climatici sugli andamenti dei titoli azionari, Kamstra, Kramer e Levi (2000), hanno trovato nel passaggio dall'ora solare a quella legale una causa dei ribassi azionari, o ancora Edmans, Garcia e Norli (2007) hanno evidenziato su un campione di 39 paesi come gli indici azionari locali siano influenzati dai risultati calcistici internazionali.

Anche i mezzi di diffusione delle informazioni di tipo digitale come i *social media* si sono rivelati un utile strumento. Antweiler e Frank (2004) hanno utilizzato i messaggi nelle chat di internet, filtrando quelli orientati ai mercati grazie al contenuto dei messaggi "buy", "sell" e "hold", riscontrando delle evidenze circa la relazione tra tali messaggi e il volume dell'attività di investimento, e tra gli stessi messaggi e la volatilità dei prezzi. Inoltre Bollen, Mao e Pepe (2009) e Bollen, Mao e Zeng (2010), hanno studiato grazie all'analisi testuale gli effetti che il sentiment, determinabile dalla attività di microblogging degli utenti di Twitter, ha sui principali indicatori economici e sociali e sul mercato finanziario americano, determinando con una accuratezza prossima al 90% la possibilità di predire le variazioni nel giorno successivo del Dow Jones Industrial Average.

Secondo gli studi sulla psicologia cognitiva degli investitori, dunque, i prezzi dei titoli potrebbero essere, entro certi limiti, prevedibili. Come mostrato anche da De Bondt (1993) infatti, gli investitori risultano ottimisti in situazioni di mercato al rialzo, e pessimisti in condizioni contrarie.

Il punto è dunque non solo valutare gli effetti dello stato d'animo degli investitori sul mercato, ma soprattutto misurarlo con la massima accuratezza (ad esempio Schmeling nel 2009 utilizza la fiducia dei consumatori come indicatore del sentiment degli investitori). È molto diffusa a tal proposito la determinazione del *media sentiment*, ovvero lo stato d'animo e le sensazioni desumibili dalla analisi di riviste, quotidiani e social media in generale. Nello specifico, in questo lavoro il *media sentiment* è definito dalla quota di termini positivi e negativi all'interno dei *feed* emessi giornalmente dal Wall Street Journal, pertanto l'ipotesi di partenza è che il sentiment all'interno del quotidiano finanziario sia in grado di anticipare i movimenti del mercato. Lo stesso Tetlock (2007), partendo da una ipotesi equivalente, ottiene dei risultati compatibili con il modello dei *noise traders* e definisce il contenuto dei media come una buona proxy del sentiment di una parte rilevante di investitori, che pertanto non agisce secondo le ipotesi tradizionali di perfetta razionalità.

Ciò che rimane ora da chiarire è come ottenere una misura quantitativa del *media sentiment* e, a tal proposito, si rivela efficace l'analisi testuale.

### **Due dizionari per la *text analysis***

Passando alla descrizione del metodo, l'analisi dei *feed* RSS consiste nel conteggio della media di parole positive e negative tratte dalle categorie di due dei più importanti dizionari in ambito di analisi testuale: l'H4 dell'Università di Harvard e il dizionario per il business di Tim Loughran e Bill McDonald.

Il primo contiene 182 categorie di tag, tra cui ad esempio 2291 parole negative, 1902 parole che implicano forza, 911 parole definite passive e 168 che denotano piacere e benessere.

Come notato da Tetlock (2005), i software che effettuano questo tipo di indagini sono semplicemente dei programmi di conteggio delle parole, non essendo pertanto in grado di categorizzare certe combinazioni di termini che spesso hanno differenti significati. Proprio l'ambiguità del linguaggio umano è il maggior rischio di un approccio come quello del presente lavoro, in cui viene calcolata una "quota" di termini appartenenti ad una specifica categoria.

Nonostante ciò, i risultati ottenuti con tale metodologia sono stati nel tempo confortanti, e i ricercatori in ambito economico e finanziario si focalizzano generalmente sulle categorie positive e negative del dizionario Harvard IV-4, anche se tanti hanno posto delle questioni sull'elenco di quelle positive. Il numero di termini contenuti in quest'ultima categoria (1915, minori rispetto alle negative che sono 2291) è attribuibile alla loro frequente negazione. Secondo Loughran e

McDonald la giustificazione di ciò può essere riscontrata nel fatto che è più comune vedere degli annunci di notizie negative che utilizzano parole positive, piuttosto che comunicazioni che utilizzano parole negative nell'annuncio di situazioni positive.

Uno dei più conosciuti studi in cui viene utilizzato il dizionario di Harvard è stato svolto da Tetlock in "Giving the content to investor sentiment: the role of media in stock market". Secondo l'autore, gli andamenti casuali che si osservano sulle news finanziarie riguardanti i mercati devono essere associate a fattori inerenti la psicologia e la sociologia. È chiaro, dunque, che le notizie di finanza condizionano, amplificano, o semplicemente riflettono le idee che hanno gli investitori del mercato. L'autore analizza la sezione "Abreast of the market" del Wall Street Journal e confronta i risultati della analisi testuale con i rendimenti azionari ed i volumi di transazioni. In tal modo mira a evidenziare le relazioni tra i report dei media e l'attività giornaliera dei mercati azionari. I commenti utilizzati sono quelli nel periodo tra il 1984 ed il 1999, e la considerazione di questa sezione da parte di Tetlock è stata una "scelta naturale", in quanto essa riflette ed influenza il sentiment degli investitori per tre ragioni: il Wall Street Journal, tra le pubblicazioni finanziarie è quella con la maggior diffusione; ha un'ottima reputazione tra gli investitori; e la sezione utilizzata è facilmente accessibile anche per periodi passati.

Nonostante l'elenco di termini appena descritto sia tra i più utilizzati, Loughran e McDonald (2011) hanno evidenziato che l'H4 può generare degli errori di classificazione delle parole. In tal caso gli effetti sulla analisi del sentiment derivanti da un termine possono essere del tutto fuorvianti.

Per tale ragione hanno sviluppato un dizionario, partendo da quello di Harvard, in cui le parole sono modificate per comprenderne differenti forme. Il risultato è un elenco, aggiornato nel 2011, contenente 3912 termini, suddivisi in 6 categorie: *negative*, *positive*, *uncertainty*, *litigious*, *strong modal words* e *weak modal words*.

Una delle peculiarità di questo dizionario è l'orientamento prettamente finanziario, dato che le categorie sono state create esaminando le parole ricorrenti almeno nel 5% di report aziendali alla SEC (10-Ks) tra il 1994 e il 2008. Un ulteriore aspetto da sottolineare è la lista di parole positive, nettamente inferiore rispetto all'H4, con solo 354 termini rispetto ai 1915 di Harvard. Ciò deriva dal fatto che gli autori hanno voluto includere esclusivamente i termini che denotano un tono unilateralmente positivo, al fine di limitare al massimo il rischio di ambiguità.

### **Creazione del database**

Secondo la classificazione di Baker e Wurgler (2007), l'approccio per la determinazione del sentiment qui seguito è un approccio "bottom-up" ed ha l'obiettivo di determinare le relazioni tra

l'andamento del mercato azionario americano e le notizie del *Wall Street Journal*, sotto l'assunto che i *noise traders* siano in grado di condizionare il mercato.

Come punto di partenza è stato necessario creare, per mezzo della analisi testuale, un indice che racchiuda il pensiero prevalente all'interno dei *feed* RSS. Questi ultimi sono flussi di informazioni formattate secondo specifiche prestabilite in modo da facilitare lo scambio dei contenuti, raccolte generalmente tramite un aggregatore.

La scelta in merito a questo strumento è motivata da una serie di vantaggi che vanno dalla facilità di reperimento (dopo aver effettuato la sottoscrizione, con l'utilizzo dell'aggregatore Google Reader è stato possibile ottenere i *feed* meno recenti), alla semplicità di analisi (in seguito alla conversione in un formato analizzabile, si prestano all'analisi testuale in quanto costituiscono una sorta di sintesi degli articoli di giornale, privi certamente di termini ridondanti).

Il campione è costituito da 7132 *feed* RSS, pubblicati nella sezione "US News" del *Wall Street Journal* tra l'8 novembre 2011 ed il 15 febbraio 2013. Nel periodo considerato ogni giornata contiene in media 15 *feed*, ciascuno dei quali contiene mediamente 28 parole.

Per la conversione di informazioni qualitative in quantitative l'analisi testuale è stata svolta secondo due tipologie, che hanno portato a due differenti categorie di indici. Un primo approccio è stato quello di utilizzare esclusivamente i termini negativi dei dizionari, in modo da ottenere due distinti indicatori di pessimismo dei media. I termini delle categorie *negative* sono stati ricercati all'interno di ciascun *feed* e poi se ne è calcolata la media per giornata, in modo da ottenere un valore (che nel campione oscilla tra 0 e +5) che indichi quanti termini inequivocabilmente negativi sono presenti in quella data. La scelta sugli indicatori del solo pessimismo deriva da studi di scienza cognitiva come Baumeister, Bratslavsky, Finkenauer e Vohs (2001), Rozin e Royzman (2001), Taylor (1991), secondo cui le notizie positive e negative hanno un impatto differente sulla mente degli individui. Generalmente, le informazioni negative esercitano un impatto più forte rispetto a quelle positive e questa è la ragione per cui, secondo Shoemaker e Reese (1996), i giornalisti finanziari tendano solitamente a drammatizzare gli eventi, in modo da rendere i loro articoli più interessanti per gli investitori.

Il secondo approccio, che ha portato alla determinazione di altri due indici, contempla invece la possibilità che dal *Wall Street Journal* sia desumibile anche un sentiment positivo o ottimista. In questo caso è calcolata la differenza tra la media di parole della categoria *positive* e quelle della categoria *negative*. Un primo indice chiamato H4, è creato con due delle categorie dell'omonimo dizionario di Harvard, mentre il secondo (Fin) nasce dalla differenza tra la media di parole positive e negative del dizionario di Loughran e McDonald. Qui le differenze tra i dizionari sono evidenti e

mostrano come, effettivamente, i termini positivi del dizionario di Harvard rendano meno attendibili i risultati.

L'ultima considerazione preliminare sul database così determinato riguarda l'andamento delle serie. Come spesso accade operando con dati in serie storica, i prezzi degli indici azionari presentano un andamento tendenziale, mentre ciò non si verifica per il sentiment. Grazie alla analisi della funzione di autocorrelazione ed il test di Dickey-Fuller sulle serie è possibile constatare con certezza la presenza di radice unitaria per tutti e cinque gli indici azionari considerati.

I problemi che possono derivare dalla non stazionarietà della serie considerata sono legati alla cosiddetta regressione spuria, che si ha quando le variabili hanno radice unitaria e non sono cointegrate. In questo caso le stime OLS possono fornire risultati completamente fuorvianti. Potendo escludere, tenuto conto che tutti gli indici di sentiment sono stazionari, l'ipotesi di cointegrazione tra le variabili, le serie storiche sono utilizzate nelle differenze piuttosto che nei livelli, pertanto sono stati stimati i parametri delle regressioni che vedono come variabile dipendente la variazione di valore dell'indice azionario, e come variabile esplicativa la variazione dell'indice di media sentiment.

Per determinare l'esistenza di una relazione tra il media sentiment calcolato in precedenza e cinque tra i principali indici azionari americani, risultano esplicativi i risultati delle stime OLS.

Ponendo l'attenzione inizialmente sugli indici di pessimismo, questi indagano la presenza di parole negative e secondo gran parte della letteratura sono più efficaci nella determinazione del sentiment proprio per l'enfasi posta dagli stessi autori dei giornali finanziari. Secondo le ipotesi del presente lavoro, ad una variazione dell'indicatore di pessimismo corrisponderà una variazione nel senso opposto del valore dell'indice azionario, per lo studio di tale relazione la formula è la seguente:

$$\Delta close_t = \beta_0 + \beta_1 \Delta close_{t-1} + \beta_2 \Delta msci_t + \beta_3 \Delta sent_{t-1} + u$$

Dove tra i regressori sono state inserite due variabili di controllo come la variazione della dipendente nel periodo precedente, quella dell'indice azionario globale MSCI (Morgan Stanley Capital International), e la variazione dell'indice di sentiment/pessimismo.

### **Risultati statistici**

Indicato con il nome H4-Neg (così denominato per l'utilizzo dei soli termini negativi del dizionario di Harvard) il primo indice fornisce risultati che sembrano confermare l'ipotesi.

Come mostrato nella tabella 1, le stime OLS che vedono come variabile dipendente la variazione del pessimismo nel giorno precedente hanno un buon livello di significatività con i Nasdaq 100 e Composite, ed uno più basso con gli altri. In particolare, il segno negativo del coefficiente  $\Delta H4-$



Neg(-1) sembra già mostrare l'esistenza di una relazione inversa tra le variazioni di pessimismo e le variazioni del mercato azionario con un livello di significatività prossimo al 97%.

La tabella 2, invece, mostra i risultati ottenuti con l'altro indice di pessimismo, Fin-Neg, il cui nome deriva dalla tipologia di dizionario utilizzato (con termini finanziari).

Anche in questo secondo caso il segno dei coefficienti è negativo per le cinque regressioni, ma un risultato ancora più interessante viene fornito dai p-value indicati tra parentesi. Tutte le stime OLS risultano significative con un ottimo livello di confidenza, confermando quindi l'esistenza di una relazione inversa tra il pessimismo del WSJ e l'andamento azionario. A titolo di esempio, un aumento di un punto della negatività insita negli articoli del quotidiano finanziario genera, il giorno seguente, una riduzione di 23,42 punti dell'indice Dow Jones Industrial Average, termine in valore assoluto superiore rispetto a tutti gli altri, come a mostrare che il pessimismo desumibile dal Wall Street Journal sia più influente su questo che su altri indici azionari. L'affidabilità del modello, confermata dalla assenza di autocorrelazione tra i residui, avvalora dunque l'ipotesi che il sentiment dei media, estratto dai *feed*, influenzi il mercato azionario.

**Tabella 1: stime OLS con la variabile H4-Neg**

<b>Variabile esplicativa</b>	<b><math>\Delta</math>close_n100</b>	<b><math>\Delta</math>close_nc</b>	<b><math>\Delta</math>close_dj</b>	<b><math>\Delta</math>close_sp500</b>	<b><math>\Delta</math>close_nyse</b>
<b>C</b>	0,109 (0,93)	0,563 (0,70)	3,14 (0,55)	0,51 (0,38)	2,69 (0,47)
<b><math>\Delta</math>close(-1)</b>	-0,122 (0,01)	-0,12 (0,01)	-0,12 (0,01)	-0,12 (0,01)	-0,10 (0,03)
<b><math>\Delta</math>msci</b>	69,62 (0,00)	79,72 (0,00)	295,5 (0,00)	33,6 (0,00)	202,1 (0,00)
<b><math>\Delta</math>H4-Neg(-1)</b>	-4,90 (0,03)	-5,38 (0,03)	-16,56 (0,07)	-1,77 (0,08)	-11,24 (0,08)
<b>R<sup>2</sup></b>	0,21	0,23	0,24	0,24	0,22

p-value indicati tra parentesi

Fonte: elaborazione eviews

**Tabella 2: stime OLS con la variabile Fin-Neg**

Variabile esplicativa	$\Delta\text{close\_n100}$	$\Delta\text{close\_nc}$	$\Delta\text{close\_dj}$	$\Delta\text{close\_sp500}$	$\Delta\text{close\_nyse}$
<b>C</b>	0,09 (0,94)	0,55 (0,70)	3,11 (0,55)	0,51 (0,38)	2,67 (0,48)
<b><math>\Delta\text{close}(-1)</math></b>	-0,11 (0,02)	-0,12 (0,01)	-0,12 (0,01)	-0,11 (0,02)	-0,10 (0,04)
<b><math>\Delta\text{msci}</math></b>	70,71 (0,00)	80,97 (0,00)	299,8 (0,00)	34,12 (0,00)	204,91 (0,00)
<b><math>\Delta\text{Fin-Neg}(-1)</math></b>	-6,26 (0,004)	-7,07 (0,003)	-23,42 (0,005)	-2,61 (0,006)	-15,13 (0,01)
<b>R<sup>2</sup></b>	0,22	0,24	0,25	0,25	0,23

p-value indicati tra parentesi

Fonte: elaborazione eviews

I commenti circa i valori e la significatività dei coefficienti portano ad identificare l'esistenza della relazione ipotizzata nel presente lavoro, ma approfondendo l'analisi verso lo studio della capacità predittiva del pessimismo, è necessario introdurre le considerazioni sui modelli VAR e sul test di Causalità di Granger.

**Tabella 3: test causalità di Granger**

	H <sub>0</sub>	$\Delta\text{close\_n100}$	$\Delta\text{close\_nc}$	$\Delta\text{close\_dj}$	$\Delta\text{close\_sp500}$	$\Delta\text{close\_nyse}$
<b><math>\Delta\text{H4-Neg}</math></b>	market→sent	(0,534)	(0,714)	(0,92)	(0,95)	(0,96)
	sent→market	(0,068)	(0,056)	(0,026)	(0,062)	(0,03)
<b><math>\Delta\text{Fin-Neg}</math></b>	market→sent	(0,542)	(0,41)	(0,16)	(0,395)	(0,27)
	sent→market	(0,09)	(0,066)	(0,027)	(0,06)	(0,046)

Fonte: elaborazione eviews

Considerando gli elementi della Tabella 3, i valori della prima riga mostrano tra parentesi i p-value associati alle ipotesi che il valore azionario non causi il pessimismo (in alto) e che il pessimismo non causi il valore azionario (nella parte bassa delle celle). Come è evidente, è possibile accettare con sufficienti livelli di significatività l'ipotesi che sia il pessimismo a causare nel senso di Granger l'andamento del mercato finanziario, e non viceversa.

Anche i test su  $\Delta\text{Fin-Neg}$  forniscono i medesimi risultati, potendosi constatare con un certo livello di significatività che vi è una relazione di causalità unidirezionale del pessimismo (misurato in questo caso per mezzo del dizionario di Loughran e McDonald) sul mercato azionario americano.

Le conferme derivanti dallo studio della causalità di Granger portano, infine, alla stima dei modelli VAR. Si tratta di modelli in forma vettoriale in cui ciascuna equazione utilizza come variabili esplicative i valori ritardati di tutte le variabili da studiare. Questi permettono una visione più completa e mostrano la eventuale capacità predittiva della variabile in oggetto: il pessimismo dei media.

Nella Tabella 4 sono indicati i coefficienti (con le relative statistiche  $t$ ) delle stime con il modello vettoriale autoregressivo per il primo indice di pessimismo, l'H4-Neg.

Oggetto di quest'ultima parte del lavoro è la determinazione di una eventuale capacità del sentiment di anticipare i movimenti di mercato. Ciò, con lo sviluppo di un modello più completo, che include anche altre variabili esplicative e perfeziona l'analisi del sentiment dei media (non includendo ad esempio soltanto i *feed* RSS), può portare ad una vera e propria strategia di previsione dei movimenti finanziari.

Le considerazioni in termini di predittività del pessimismo sono visibili nei coefficienti in cui la variabile dipendente è l'indice azionario. In questi casi, per il Nasdaq 100 e Composite l'intervallo di confidenza è superiore al 95%, mentre per i restanti tre è tra il 90 ed il 95%, potendo pertanto concludere che l'enfasi posta dagli autori del Wall Street Journal in una giornata particolarmente pessimista genera incertezza, causando dei ribassi del valore dei titoli (più influente sui primi due che sugli altri).

Un ultimo commento è necessario per la prima equazione di questo VAR. Qui i coefficienti associati a  $\Delta\text{close\_dj}(-1)$  ed a  $\Delta\text{close\_sp500}(-1)$  sono significativi, a dimostrazione del fatto che è anche il pessimismo a muoversi in risposta al mercato. In particolare giornate di ribassi nei due indici appena menzionati hanno fatto registrare, nel periodo considerato, incrementi del pessimismo all'interno degli articoli del Wall Street Journal.

**Tabella 4: VAR(1); variabili dipendenti sulle colonne, coefficienti e t-statistics in parentesi**

	$\Delta H4\text{-Neg}$	$\Delta \text{close\_n100}$	$\Delta \text{close\_nc}$	$\Delta \text{close\_dj}$	$\Delta \text{close\_sp500}$	$\Delta \text{close\_nyse}$
<b><math>\Delta H4\text{-Neg}(-1)</math></b>	-0,3636 (-6,924)	-4,9158 (-1,866)	-5,4797 (-1,892)	-16,4213 (-1,584)	-1,740 (-1,474)	-10,9213 (-1,481)
<b><math>\Delta \text{close\_n100}(-1)</math></b>	0,0088 (1,0138)	-0,7380 (-1,683)	-0,7105 (-1,474)	-3,4923 (-2,025)	-0,3643 (-1,855)	-2,113 (-1,722)
<b><math>\Delta \text{close\_nc}(-1)</math></b>	-0,0121 (-1,293)	0,500 (1,058)	0,4686 (0,902)	3,002 (1,615)	0,293 (1,385)	1,713 (1,295)
<b><math>\Delta \text{close\_dj}(-1)</math></b>	-0,0026 (-1,922)	-0,0388 (-0,564)	-0,0388 (0,513)	-0,0345 (-0,127)	-0,0011 (-0,0382)	0,036 (0,188)
<b><math>\Delta \text{close\_sp500}(-1)</math></b>	0,0547 (2,007)	1,014 (0,74)	1,0171 (0,676)	2,9362 (0,545)	0,353 (0,577)	1,527 (0,398)
<b><math>\Delta \text{close\_nyse}(-1)</math></b>	-0,0036 (-1,282)	-0,058 (-0,411)	-0,0641 (-0,411)	-0,4840 (-0,866)	-0,052 (-0,823)	-0,288 (-0,726)
<b>C</b>	-0,0143 (-0,469)	0,6108 (0,399)	1,155 (0,686)	4,831 (0,802)	0,7201 (1,050)	4,0192 (0,937)
<b>R<sup>2</sup></b>	0,17	0,031	0,028	0,027	0,024	0,023

Fonte: elaborazione eviews

**Tabella 5: VAR(1); variabili dipendenti sulle colonne, coefficienti e t-statistics in parentesi**

	$\Delta\text{Fin-Neg}$	$\Delta\text{close\_n100}$	$\Delta\text{close\_nc}$	$\Delta\text{close\_dj}$	$\Delta\text{close\_sp500}$	$\Delta\text{close\_nyse}$
<b><math>\Delta\text{Fin-Neg}(-1)</math></b>	-0,3798 (-7,444)	-4,975 (-2,044)	-5,673 (-2,122)	-18,092 (-1,891)	-1,998 (-1,835)	-11,146 (-1,635)
<b><math>\Delta\text{close\_n100}(-1)</math></b>	0,0080 (0,875)	-0,7432 (-1,697)	-0,716 (-1,489)	-3,509 (-2,039)	-0,366 (-1,868)	-2,125 (-1,733)
<b><math>\Delta\text{close\_nc}(-1)</math></b>	-0,0125 (-1,272)	0,4908 (1,040)	0,458 (0,885)	2,978 (1,606)	0,291 (1,379)	1,693 (1,282)
<b><math>\Delta\text{close\_dj}(-1)</math></b>	-0,003 (-2,232)	-0,0436 (-0,635)	-0,044 (-0,585)	-0,051 (-0,190)	-0,003 (-0,098)	0,025 (0,133)
<b><math>\Delta\text{close\_sp500}(-1)</math></b>	0,0611 (2,139)	1,1869 (0,870)	1,209 (0,807)	3,508 (0,654)	0,414 (0,678)	1,911 (0,500)
<b><math>\Delta\text{close\_nyse}(-1)</math></b>	-0,0041 (-1,393)	-0,0724 (-0,511)	-0,079 (-0,513)	-0,531 (-0,955)	-0,057 (-0,906)	-0,319 (-0,806)
<b>C</b>	-0,0059 (-0,184)	0,6011 (0,393)	1,145 (0,681)	4,811 (0,800)	0,718 (1,050)	3,998 (0,933)
<b>R<sup>2</sup></b>	0,19	0,034	0,031	0,031	0,028	0,025

Fonte: elaborazione eviews

Considerazioni analoghe possono essere espresse con i dati contenuti nella Tabella 5. In particolare, dai valori della statistica  $t$  relativa ai coefficienti di  $\Delta\text{Fin-Neg}(-1)$  nelle cinque equazioni esplicative del mercato azionario si può dedurre la capacità predittiva del *media sentiment* con livelli di significatività elevati. Il dizionario di Loughran e McDonald, considerato fin qui solo nei termini negativi, proprio grazie al suo orientamento finanziario, di rivela più efficace nella possibilità di effettuare previsioni sull'andamento del mercato. Risultano significativi con un livello di confidenza superiore al 95% i coefficienti associati a quattro dei cinque indici azionari.

Dalle considerazioni appena svolte, dunque, emerge come vi sia una relazione tra il pessimismo dei *media* ed il mercato azionario. Avvalorando la tesi secondo cui i quotidiani, specie quelli finanziari, puntano molto sui termini negativi per dare "più effetto" ai propri contenuti. Gli investitori meno informati, poi, recepiscono questo sentiment e lo trasferiscono al mercato, esercitando delle pressioni al ribasso sui cinque indici del mercato americano considerati.

Passando alla analisi degli indicatori di sentiment, quelli che contemplano anche la possibilità che sia desumibile dell'ottimismo dalla analisi testuale dei *feed*, si evidenziano sorprendenti differenze sulla base dell'elenco di parole utilizzato. Il primo indice è denominato H4, e prende il nome dall'omonimo dizionario dell'Università di Harvard in quanto sono stati utilizzati sia i termini positivi che quelli negativi di questo.

Come evidente dalla Tabella 6, le stime dei coefficienti della variabile  $\Delta H4$  non presentano alcun risultato significativo. Vengono dunque confermate le ipotesi di Tim Loughran e Bill McDonald che hanno portato alla nascita del loro nuovo dizionario: i termini positivi di Harvard, molto più numerosi dell'altro elenco, fanno cambiare notevolmente i risultati e si adattano male a questo tipo di analisi, generando spesso ambiguità, soprattutto in un caso come quello presentato in questo lavoro in cui sono oggetto di studio i *feed*, per natura molto sintetici. Ciò avvalora inoltre la tesi secondo cui i giornalisti utilizzano volutamente una ampia gamma di termini negativi, per porre maggior enfasi al loro contributo, generando effetti sulla incertezza dei *noise traders*, come descritto in precedenza.

Si rivelano più interessanti, invece, i risultati con il dizionario per il business. Gli output delle stime di regressione in cui la variabile esplicativa è  $\Delta\text{Fin}$  presentano coefficienti significativi con tutti gli indici azionari con un livello di confidenza pari al 99%. Bisogna qui ricordare un importante punto di divergenza tra i due elenchi di termini utilizzati: il numero di parole positive.

I due autori americani, infatti, hanno ridotto drasticamente il numero di quelli contenuti in questa categoria in modo da limitare al massimo il rischio di ambiguità, e i risultati sembrano dare ragione a loro. I coefficienti positivi associati al sentiment mostrano che esiste una relazione diretta tra i due. Un aumento dell'ottimismo (o, specularmente, una riduzione del pessimismo) genera nel

giorno seguente una tendenza al rialzo dei prezzi azionari. I *noise* traders, seguendo l'entusiasmo crescente all'interno del quotidiano finanziario di riferimento americano, acquisiscono fiducia e la trasmettono per mezzo del loro modo di operare al mercato.

**Tabella 6: stime OLS con la variabile H4**

<b>Variabile esplicativa</b>	$\Delta\text{close\_n100}$	$\Delta\text{close\_nc}$	$\Delta\text{close\_dj}$	$\Delta\text{close\_sp500}$	$\Delta\text{close\_nyse}$
<b>C</b>	0,05 (0,96)	0,50 (0,73)	2,93 (0,58)	0,49 (0,41)	2,55 (0,50)
<b><math>\Delta\text{close}(-1)</math></b>	-0,11 (0,02)	-0,12 (0,02)	-0,11 (0,02)	-0,11 (0,02)	-0,09 (0,05)
<b><math>\Delta\text{msci}</math></b>	69,21 (0,00)	79,42 (0,00)	291,26 (0,00)	33,38 (0,00)	201,1 (0,00)
<b><math>\Delta\text{H4}(-1)</math></b>	2,25 (0,40)	1,89 (0,52)	15,37 (0,14)	1,07 (0,36)	4,40 (0,55)
<b>R<sup>2</sup></b>	0,20	0,22	0,23	0,23	0,21

Fonte: elaborazione eviews

**Tabella 7: stime OLS con la variabile Fin**

<b>Variabile esplicativa</b>	$\Delta\text{close\_n100}$	$\Delta\text{close\_nc}$	$\Delta\text{close\_dj}$	$\Delta\text{close\_sp500}$	$\Delta\text{close\_nyse}$
<b>C</b>	0,07 (0,95)	0,52 (0,72)	3,04 (0,56)	0,50 (0,39)	2,62 (0,48)
<b><math>\Delta\text{close}(-1)</math></b>	-0,12 (0,02)	-0,12 (0,01)	-0,12 (0,01)	-0,11 (0,02)	-0,10 (0,04)
<b><math>\Delta\text{msci}</math></b>	71,78 (0,00)	81,48 (0,00)	301,57 (0,00)	34,33 (0,00)	205,9 (0,00)
<b><math>\Delta\text{Fin}(-1)</math></b>	6,48 (0,005)	7,39 (0,003)	26,18 (0,003)	2,98 (0,003)	16,78 (0,009)
<b>R<sup>2</sup></b>	0,22	0,24	0,25	0,25	0,23

Fonte: elaborazione eviews

Le differenze tra i due dizionari sono evidenti, chiaramente, anche in termini di capacità predittiva e di causalità di Granger. La Tabella 8 mostra i p-value associati alle ipotesi che l'indice azionario non causi il sentiment (parte alta) e quella che l'indice di sentiment non causi l'azionario (in basso). Usando l'elenco Harvard non è possibile rifiutare con un livello di confidenza adeguato alcuna

ipotesi nulla, pertanto è impossibile, nel caso della variabile H4, determinare l'esistenza di un nesso di causalità tra le variabili.

**Tabella 8: test causalità di Granger**

	$H_0$	$\Delta\text{close\_n100}$	$\Delta\text{close\_nc}$	$\Delta\text{close\_dj}$	$\Delta\text{close\_sp500}$	$\Delta\text{close\_nyse}$
<b><math>\Delta H4</math></b>	market→sent	(0,27)	(0,33)	(0,33)	(0,30)	(0,32)
	sent→market	(0,33)	(0,39)	(0,10)	(0,27)	(0,43)
<b><math>\Delta Fin^*</math></b>	market→sent	(0,32)	(0,16)	(0,03)	(0,10)	(0,04)
	sent→market	(0,07)	(0,05)	(0,06)	(0,06)	(0,10)

\*Test di causalità con un solo ritardo

Fonte: elaborazione eviews

Per quel che riguarda invece l'indice di media sentiment ottenuto col dizionario di Loughran e McDonald, è interessante notare come vi sia un nesso di causalità per tutti gli indici azionari (fatta eccezione, per il NYSE), ma soprattutto come questa causalità sia bidirezionale. Ponendo attenzione al caso del Dow Jones Industrial Average, infatti, è possibile accettare l'ipotesi che nel periodo considerato sia anche il mercato azionario a causare l'ottimismo/pessimismo del Wall Street Journal, ed è evidente che questa relazione tra il quotidiano e l'indice Dow Jones è molto forte, considerati anche gli altri casi esaminati in precedenza. Questo fattore si rivela particolarmente interessante in quanto il quotidiano reagisce in risposta al mercato, soprattutto in casi di crescita di valore e conseguente crescita dell'ottimismo (la causalità non era bidirezionale in sede di analisi del solo pessimismo).

Il risultato cui si perviene è dunque ragionevole e corrispondente con l'ipotesi, la causalità bidirezionale conferma l'esistenza di una certa relazione e il dizionario di Loughran e McDonald sembra catturarla.

Le considerazioni sulla capacità predittiva infine, rispecchiano in parte quanto appena sottolineato. Nella Tabella 9 sono mostrati i coefficienti e le statistiche t delle 6 equazioni che compongono il VAR. Solo il coefficiente associato al sentiment nella equazione dell'indice Dow Jones risulta significativo con un intervallo di confidenza superiore al 95%, portando ad affermare con un buon livello di significatività che variazioni del sentiment nel WSJ producono effetti sull'indice. È necessario ridurre l'intervallo di confidenza a poco più del 90% per poter accertare anche la capacità predittiva sul Nasdaq 100 e sullo S&P 500.



Il media sentiment determinato con il dizionario di Loughran e McDonald invece, già rivelatosi più efficace in sede di analisi delle stime OLS dei coefficienti delle regressioni, mostra una migliore capacità predittiva su tutti gli indici di mercato azionario considerati. Con un livello di confidenza superiore al 95% risultano significativi i coefficienti associati alla variabile  $\Delta\text{Fin}(-1)$  in tutte le equazioni del VAR della Tabella 10, ad esclusione del solo New York Stock Exchange dove il coefficiente è significativo tra il 90 ed il 95%.

L'elenco di termini proposto dai due autori americani, dunque, è anche nel presente lavoro più efficace del dizionario definito generalmente "psicologico" dalla illustre università americana. Il primo non soltanto evidenzia relazioni tra il sentiment ed il mercato finanziario, ma è addirittura in grado di anticiparne i movimenti con una certa affidabilità sia utilizzando solo le parole negative, che facendo ricorso ad entrambe le categorie. Sembra quindi più affidabile in ambito finanziario un vocabolario studiato apposta per includere i termini orientati al business.

Una considerazione interessante, infine, deriva dai coefficienti della prima equazione del VAR, quelli che spiegano la variabile  $\Delta\text{Fin}(-1)$ . Il  $\beta$  associato alle variabili  $\Delta\text{close\_dj}(-1)$  e  $\Delta\text{close\_sp500}(-1)$  sono significativi con livelli di confidenza superiori al 99% e l' $R^2$  di questa è circa 0,30. Tutto ciò porta a concludere che, come notato nella tabella con i test di causalità di Granger, la relazione tra il sentiment ed il mercato è bilaterale, per cui la sintesi dello stato d'animo del quotidiano finanziario si muove anche in risposta all'azionario.

**Tabella 9: VAR(1); variabili dipendenti sulle colonne, coefficienti e t-statistics in parentesi**

	$\Delta H4$	$\Delta \text{close\_n100}$	$\Delta \text{close\_nc}$	$\Delta \text{close\_dj}$	$\Delta \text{close\_sp500}$	$\Delta \text{close\_nyse}$
$\Delta H4(-1)$	-0,508 (-9,909)	3,927 (1,307)	3,872 (1,172)	22,357 (1,90)	1,842 (1,373)	9,244 (1,10)
$\Delta \text{close\_n100}(-1)$	-0,006 (-0,822)	-0,715 (-1,624)	-0,688 (-1,422)	-3,349 (-1,944)	-0,352 (-1,794)	-2,058 (-1,673)
$\Delta \text{close\_nc}(-1)$	0,007 (0,953)	0,449 (0,949)	0,414 (0,795)	2,798 (1,50)	0,273 (1,292)	1,599 (1,207)
$\Delta \text{close\_dj}(-1)$	-1.47E-05 (-0,012)	-0,035 (-0,519)	-0,036 (-0,475)	-0,014 (-0,052)	0,0003 (0,012)	0,043 (0,225)
$\Delta \text{close\_sp500}(-1)$	-0,017 (-0,748)	1,060 (0,771)	1,087 (0,719)	2,753 (0,512)	0,353 (0,576)	1,611 (0,419)
$\Delta \text{close\_nyse}(-1)$	0,001 (0,593)	-0,055 (-0,388)	-0,062 (-0,398)	-0,441 (-0,79)	-0,049 (-0,779)	-0,28 (-0,702)
<b>C</b>	0,006 (0,235)	0,578 (0,376)	1,116 (0,661)	4,771 (0,793)	0,711 (1,036)	3,95 (0,92)
<b>R<sup>2</sup></b>	0,27	0,025	0,02	0,031	0,023	0,02

Fonte: elaborazione eviews

**Tabella 10: VAR(1); variabili dipendenti sulle colonne, coefficienti e t-statistics in parentesi**

	$\Delta Fin$	$\Delta close\_n100$	$\Delta close\_nc$	$\Delta close\_dj$	$\Delta close\_sp500$	$\Delta close\_nyse$
<b><math>\Delta Fin (-1)</math></b>	-0,37 (-7,14)	4,675 (1,80)	5,415 (1,90)	18,853 (1,851)	2,134 (1,841)	11,50 (1,58)
<b><math>\Delta close\_n100(-1)</math></b>	-0,007 (-0,805)	-0,749 (-1,70)	-0,724 (-1,50)	-3,536 (-2,05)	-0,369 (-1,88)	-2,141 (-1,74)
<b><math>\Delta close\_nc(-1)</math></b>	0,013 (1,438)	0,473 (1,002)	0,439 (0,846)	2,919 (1,57)	0,284 (1,349)	1,656 (1,254)
<b><math>\Delta close\_dj(-1)</math></b>	0,003 (2,402)	-0,047 (-0,69)	-0,048 (-0,64)	-0,069 (-0,255)	-0,005 (-0,164)	0,014 (0,077)
<b><math>\Delta close\_sp500(-1)</math></b>	-0,072 (-2,65)	1,268 (0,928)	1,303 (0,868)	3,83 (0,714)	0,45 (0,738)	2,108 (0,55)
<b><math>\Delta close\_nyse(-1)</math></b>	0,005 (1,77)	-0,071 (-0,50)	-0,078 (-0,50)	-0,529 (-0,949)	-0,057 (-0,902)	-0,318 (-0,801)
<b>C</b>	0,005 (0,18)	0,578 (0,377)	1,11 (0,66)	4,73 (0,787)	0,71 (1,038)	3,95 (0,922)
<b>R<sup>2</sup></b>	0,18	0,03	0,028	0,03	0,028	0,024

Fonte: elaborazione eviews

## Conclusioni

I risultati appena elencati permettono, dunque, di giungere a tre importanti conclusioni.

Innanzitutto è stata dimostrata empiricamente l'esistenza di una relazione tra il media sentiment ed i mercati azionari USA. Questo deriva dal fatto che le regressioni in cui è posto un indice azionario come variabile dipendente ed uno di pessimismo/sentiment come indipendente presentano coefficienti OLS significativi. Vale, inoltre, la considerazione del segno del coefficiente: infatti in caso di relazione con un indice di pessimismo questo è negativo, mentre accade l'inverso con il sentiment (possibile ottimismo).

Un secondo risultato rilevante riguarda ancora i test di causalità e la stima con i modelli Vector Autoregressive. Posto che la relazione tra sentiment e mercato esiste, è possibile concludere che in gran parte dei casi è il primo a causare il secondo (ciò pone degli ottimi spunti circa la possibilità di predire gli andamenti azionari, come molti lavori di finanza comportamentale puntano a fare).

La terza ed ultima considerazione riguarda le differenze tra i due dizionari utilizzati. L'H4, sviluppato dall'Università di Harvard, nasce come dizionario psicologico, pertanto meno orientato ai termini finanziari rispetto a quello di Loughran e McDonald. Questa differenza si nota anche all'interno della presente analisi poiché i coefficienti sono meno significativi per gli indici di pessimismo e addirittura non significativi per quelli di sentiment. Quest'ultimo risultato deriva in particolar modo dalla mole di termini positivi. Sono numerosi, infatti, gli studi che utilizzano esclusivamente i quelli negativi per due ragioni: in primis sono più utilizzati dai giornalisti che hanno l'obiettivo di colpire i lettori, inoltre sono meno soggetti a negazioni nel linguaggio comune. La grande pecca del dizionario di Harvard, stando anche ai risultati di questo lavoro, è proprio questa: l'eccessiva inclusione di termini positivi.

Tali risultati aprono ad ulteriori scenari di ricerca. Per prima cosa è stato possibile testare in un primo lavoro l'efficacia dei *feed* RSS degli articoli del Wall Street Journal. Per questa ragione si potrebbe affinare la ricerca tramite la creazione di un nuovo dizionario, ex novo, che sia specifico per questo sempre più diffuso strumento di scambio di informazioni, oltre chiaramente allo sviluppo di un modello più completo, che limiti al massimo l'omissione di variabili significative.

## **Riferimenti bibliografici**

Antweiler W., Frank M.Z. (2004) "Is all that talk just noise? The information content of internet stock message boards", *The Journal of Finance* vol. 59 (num. 3): pp. 1259-1293.

Baker M., Wurgler J. (2007) "Investor sentiment in the stock market", *Journal of economic perspectives*, vol. 21 (num. 2), pp. 129-151.

Barberis N., Shleifer A., Vishny R. (1998) "A model of investor sentiment", *Journal of financial economics*, vol. 49 (num. 3), pp. 307-343.

Baumeister R. F., Bratslavsky E., Finkenauer C., Kathleen D. (2001) "Bad is Stronger than Good", *Review of General Psychology*, vol. 5 (num. 4), pp. 323-370.

Bollen J., Mao H., Pepe A. (2009) "Modeling public mood and emotion: twitter sentiment and socio-economic Phenomena", *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*.

Bollen J., Mao H., Zeng X. (2010) "Twitter mood predicts the stock market", *Journal of Computational Science*, vol. 2 (num.1), pp. 1-8.

De Bondt W., Thaler R. (1985) "Does the stock market overreact?", *Journal of finance*, vol. 40 (num. 3), pp. 793-805.

De Bondt W. (1993), "Betting on trends: Intuitive forecasts of financial risk and return", *International Journal of Forecasting*, vol. 9, pp. 355-371.

Carretta A., Farina V., Graziano E., Reale M. (2011) "Does investor attention influence stock market activity? The case of spin-off deals" *MPRA Paper*.

De Long J., Shleifer A., Summers L., Waldmann R. (1990), "Noise Trader Risk in Financial Markets", *Journal of Political Economy*, Vol. 98 (num. 4), pp. 703-738.

Edmans A., Garcia D., Norli O. (2007), "Sports Sentiment and Stock Returns", *The Journal of Finance*, vol. 62 (num. 4), pp. 1967-1998.

Fama E. (1970) "Efficient capital markets: A review of theory and empirical work", *Journal of finance*, vol. 25 (num. 2), pp. 338-417.

Fama E. (1998) "Market efficiency, long term returns and behavioral finance", *Journal of financial economics*, vol. 49 (sept.), pp. 283-306.

- Fehle F., Tsyplakov S., Zdorovtsov V. (2005) “Can Companies Influence Investor Behavior through Advertising? Super Bowl Commercials and Stock Returns”, *European financial management*, vol. 11 (num. 5), pp. 625-647.
- Ferguson N., Guo J., Lam H., Philip D. (2011), Department of Economics and Finance, Durham University.
- Finter P., Niessen-Ruenzi A., Ruenzi S. (2011), cfr-working paper n. 10-03.
- Hirshleifer D., Shumway T. (2003) “Good Day Sunshine: Stock Returns and the Weather”, *The Journal of Finance*, vol. 58 (num. 3), pp. 1009-1032.
- Kahneman D., Tversky A. (1979) “Prospect Theory: An Analysis of Decision under Risk”, *Econometrica*, vol. 47 (num. 2), pp. 263-292.
- Kamstra M.J., Kramer L.A., Levi M.D. (2000), “Losing Sleep at the Market: The Daylight Saving Anomaly”, *The American Economic Review*, vol. 90 (num. 4), pp. 1005-1011.
- Koop G. (2000), *Analysis of economic data*, John Wiley & Sons, Ltd. (trad. It: *Logica statistica dei dati economici*, UTET, Torino, 2001).
- Li N., Dash Wu D. (2010) “Using text mining and sentiment analysis for online forums hotspot detection and forecast”, *Decision support systems*, vol. 48 (num. 2), pp. 354-368.
- Loughran T., McDonald B. (2011) “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks”, *Journal of finance*, vol. 66 (num. 1), pp. 35-65.
- Ozik G., Sadka R. (2011) “Media and Investment Management”, EDHEC Working paper.
- Piras L. (2005) *Dalla finanza classica a quella comportamentale*, Giuffrè, Milano.
- Rigoni U. (2006) “Finanza comportamentale e gestione del risparmio”, Giappichelli, Torino.
- Rozin P., Royzman E. B. (2001) “Negativity Bias, Negativity Dominance, and Contagion” *Personality and Social Psychology Review*, vol. 5, pp. 296–320.
- Saunders E.M.J. (1993) “Stock prices and wall street weather”, *American Economic Review* Vol. 83 (num. 5), 1337–1345.
- Schmeling M. (2009) “Investor sentiment and stock returns: Some international evidence”, *Journal of empirical finance*, vol. 16 (num. 3), pp. 394-408.
- Shiller R. (2000) *Irrational exuberance*, Princeton University Press.

- Shleifer A., Vishny R. (1997) "The limits of arbitrage", *Journal of finance*, vol. 52 (num.1), pp. 35-55.
- Shleifer A. (2000) *Inefficient markets. An introduction to behavioral finance*, Oxford University Press.
- Shoemaker P. J., Reese S. D. (1996) *Mediating the Message: Theories of Influences on Mass Media Content*, Longman Publishers USA.
- Subrahmanyam A. (2007) "Behavioural finance: a review and synthesis", *European Financial Management*, vol. 14 (num. 1), pp. 12-29.
- Taylor S. E. (1991) "Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis", *Psychological Bulletin*, vol. 110 (num. 1), pp. 67-85.
- Tetlock P. (2007) "Giving content to investor sentiment: the role of media in the stock market", *Journal of finance*, vol. 62 (num. 3), 1139-1168.
- Tetlock P., Saar-Tsechansky M., Macskassy S. (2007), "More than words: quantifying language to measure firms' fundamentals", *Journal of finance*, vol. 63 (num. 3), pp. 1437-1467.