



THE ROLE OF TWITTER IN THE U.S. FINANCIAL MARKET

FERNANDO MARTINELLI RAMACCIOTTI

10 de Agosto de 2018

Apresentação da Dissertação de Mestrado Profissional em Economia

Banca Examinadora:

Bruno Giovannetti (Orientador)

Fernando Chague

Rodrigo de Losso

(EESP-FGV)

(EESP-FGV)

(FEA-USP)

AGENDA

Introdução / motivação

Revisão de literatura

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

AGENDA

Introdução / motivação

Revisão de literatura

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

MOTIVAÇÃO

“Although the news media — newspapers, magazines, and broadcast media, along with their new outlets on the Internet — present themselves as detached observers of market events, they are themselves an integral part of these events. Significant market events generally occur only if there is similar thinking among large groups of people, and the news media are essential vehicles for the spread of ideas.”

Shiller, Robert. “*Irrational Exuberance*”. Ch. Four.
Princeton University Press (2000)

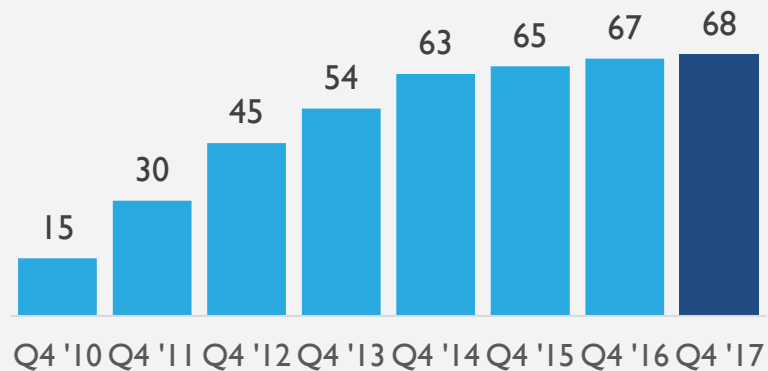
MOTIVAÇÃO

74%

dos americanos obtém
notícias de mídias sociais

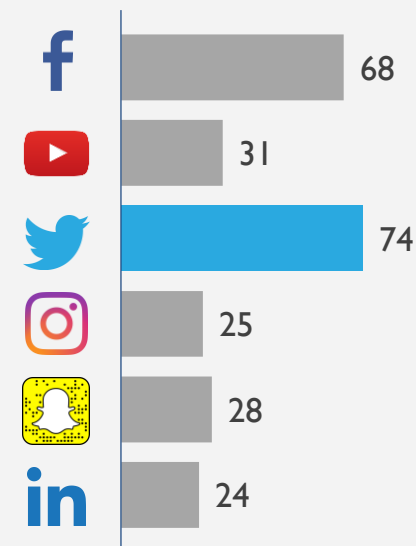
Fonte: PEW Research Center

Usuários Ativos no Twitter nos EUA (em milhões)



Fonte: Statista

% dos usuários que obtém notícias de
cada mídia social



Fonte: PEW Research Center

AGENDA

Introdução / motivação

Revisão de literatura

Psicologia

Market/Investor Sentiment

Mídia

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

REVISÃO DE LITERATURA

Tversky e Kahneman

Lei dos Pequenos Números
(TVERSKY e KAHNEMAN, 1971)

Psicologia por trás das previsões
(KAHNEMAN e TVERSKY, 1973)

Julgamentos sob incertezas
(TVERSKY e KAHNEMAN, 1974)

Inferência Bayesiana

Modelo formal da lei dos pequenos números usando inferência bayesiana
(RABIN 2002)

Psicologia do investidor

Conservadorismo e representatividade influenciam como o investidor atualiza suas crenças originais
(BARBERIS, SHLEIFER e VISHNY, 1998)

Como o excesso de confiança, informações públicas e privadas são incorporadas nos preços
(DANIEL, HIRSHLEIFER e SUBRAHMANYAM, 1998)

AGENDA

Introdução / motivação

Revisão de literatura

Psicologia

Market/Investor Sentiment

Mídia

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

REVISÃO DE LITERATURA

Market/investor sentiment

Market/investor sentiment index é uma medida de como está a atitude do mercado ou do investidor em geral (i.e. mais ou menos propensos ao risco). Isso pode revelar se há alguma relação entre tal medida e o mercado de ações

Taxa de desconto de *closed-end funds* como medida de sentimento (LEE, SCHLEIFER e THALER 1991; CHOPRA, 1993)

NEAL e WHEATLEY (1998) encontraram que para companhias pequenas a taxa de desconto pode prever retornos

OTOO *et al.* (1999) encontrou que o sentimento dos investidores são influenciados pelos preços das ações e não o contrário

AGENDA

Introdução / motivação

Revisão de literatura

Psicologia

Market/Investor Sentiment

Mídia

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

REVISÃO DE LITERATURA

SHILLER (2000) descreve vários impactos que a mídia pode trazer para o mercado financeiro. Ela ainda pode ajudar a espalhar especulações que podem até ter papel relevante em bolhas

Jornais



Wall Street Journal, Fator de Pessimismo (via PCA).

↑pessimismo - ↓preços ↑ volume;

↓retornos - ↑ pessimismo (TETLOCK, 2007)

GARCÍA (2013) encontrou que o poder preditivo das notícias é muito maior durante períodos de recessão

Twitter



Twitter para medir o sentimento do público frente a eventos sociais, culturais, políticos e econômicos (BOLLEN, MAO e PEPE, 2011)

Calma e Felicidade ajudam a prever o valor do índice DJIA (BOLLEN, MAO e ZENG, 2011)

Geolocalização



Período pré-Internet: coberturas de mídia locais influenciam negociações em corretoras locais. Mais especificamente, a cobertura local tem impacto somente quando publicada no primeiro dia possível (ENGELBERG e PARSONS, 2011)

Política Monetária



Atas do COPOM (CHAGUE *et al.*, 2015)

Efeito das políticas monetárias dependem do sentimento do mercado (KUROV, 2010)

AGENDA

Introdução / motivação

Revisão de literatura

Dados e metodologia

Resultados e conclusões

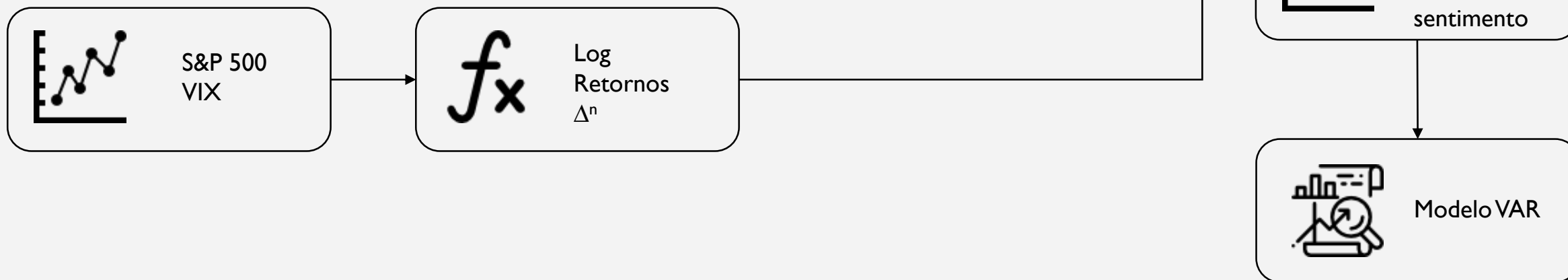
Trabalhos futuros

DADOS E METODOLOGIA

Dados não-estruturados



Dados estruturados



DADOS E METODOLOGIA

O período analisado foi de **2007-03-21** ate **2017-10-31**:

- **Tweets**

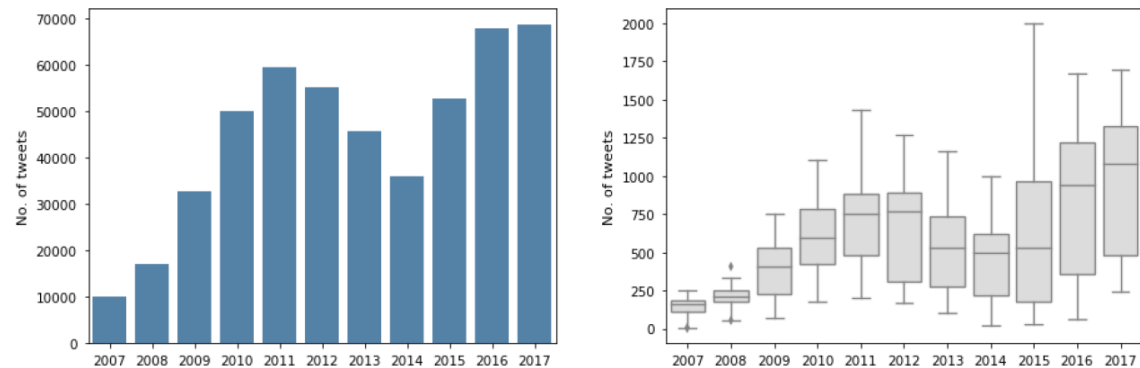
- N = 495.222
- Fontes: termos de pesquisa (NYSE, NASDAQ, DJIA, Financial Market) e contas de noticiários relacionados ao mercado financeiro (Wall Street Journal, Financial Times Finance News, CNBC, NY Times Business)

- **Financeiros**

- N = 2.651
- S&P500 (volume e valor)
- VIX (*Volatility Index*)

DADOS E METODOLOGIA

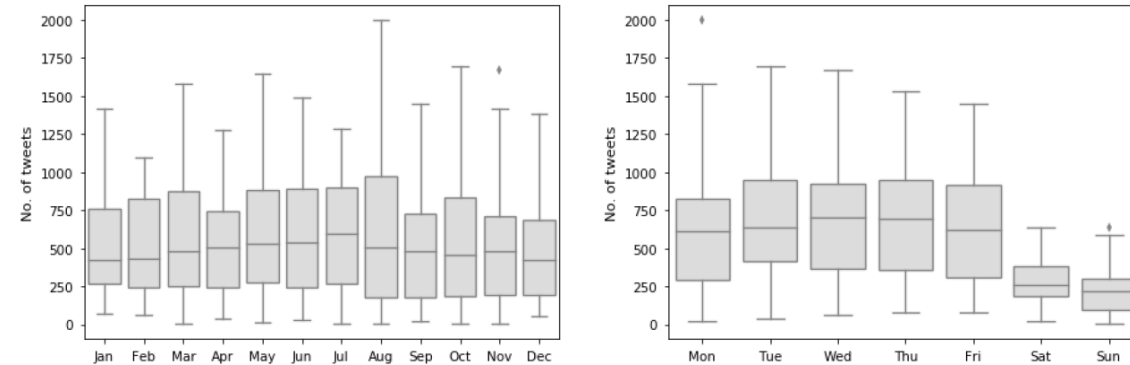
Figure 3.1 – Distribution of tweets over the years



(a) Number of tweets over the years

(b) Tweets distribution over the years

Figure 3.3 – Intra-year seasonality patterns

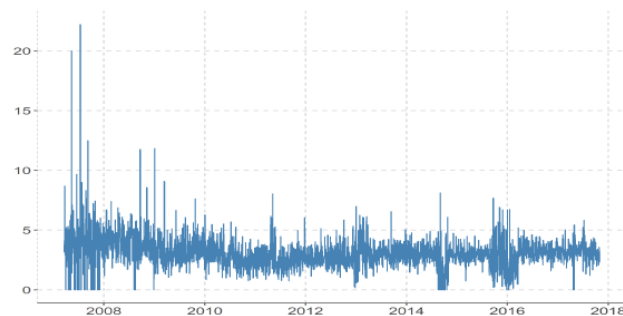


(a) Tweets distribution by month

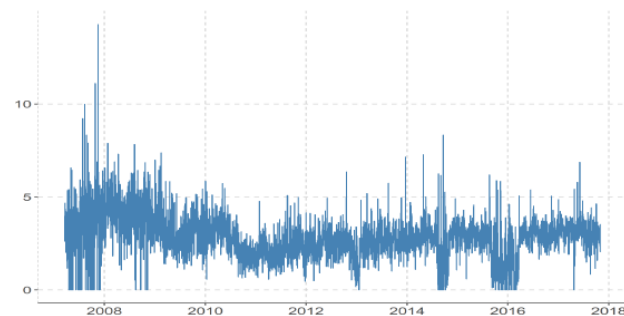
(b) Tweets distribution by weekday

DADOS E METODOLOGIA

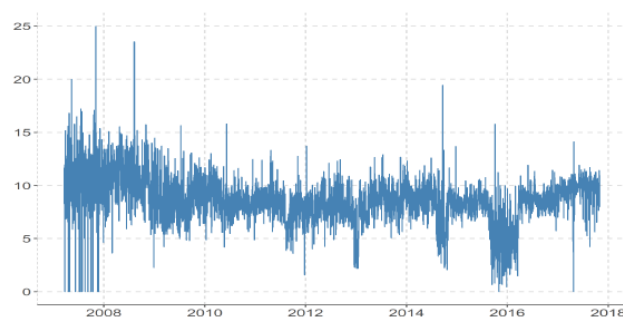
Figure 4.1 – GI categories relative frequency over time



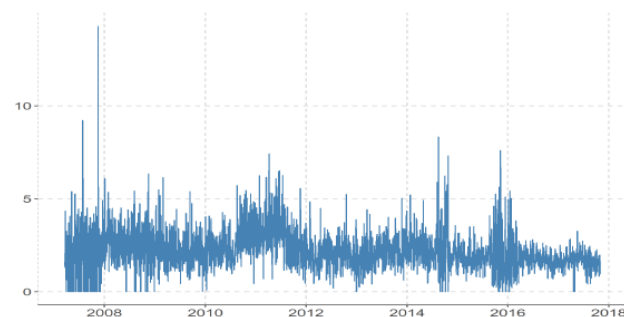
(a) Pstv



(b) Ngvt



(c) Strong



(d) Weak

	PC1	PC2	PC3	PC4
Pstv	.4599	.5440	.6723	.2012
Ngvt	.5804	-.2239	-.4147	.6641
Strong	.5777	.2490	-.3966	-.6686
Weak	.3433	-.7694	.4677	-.2674

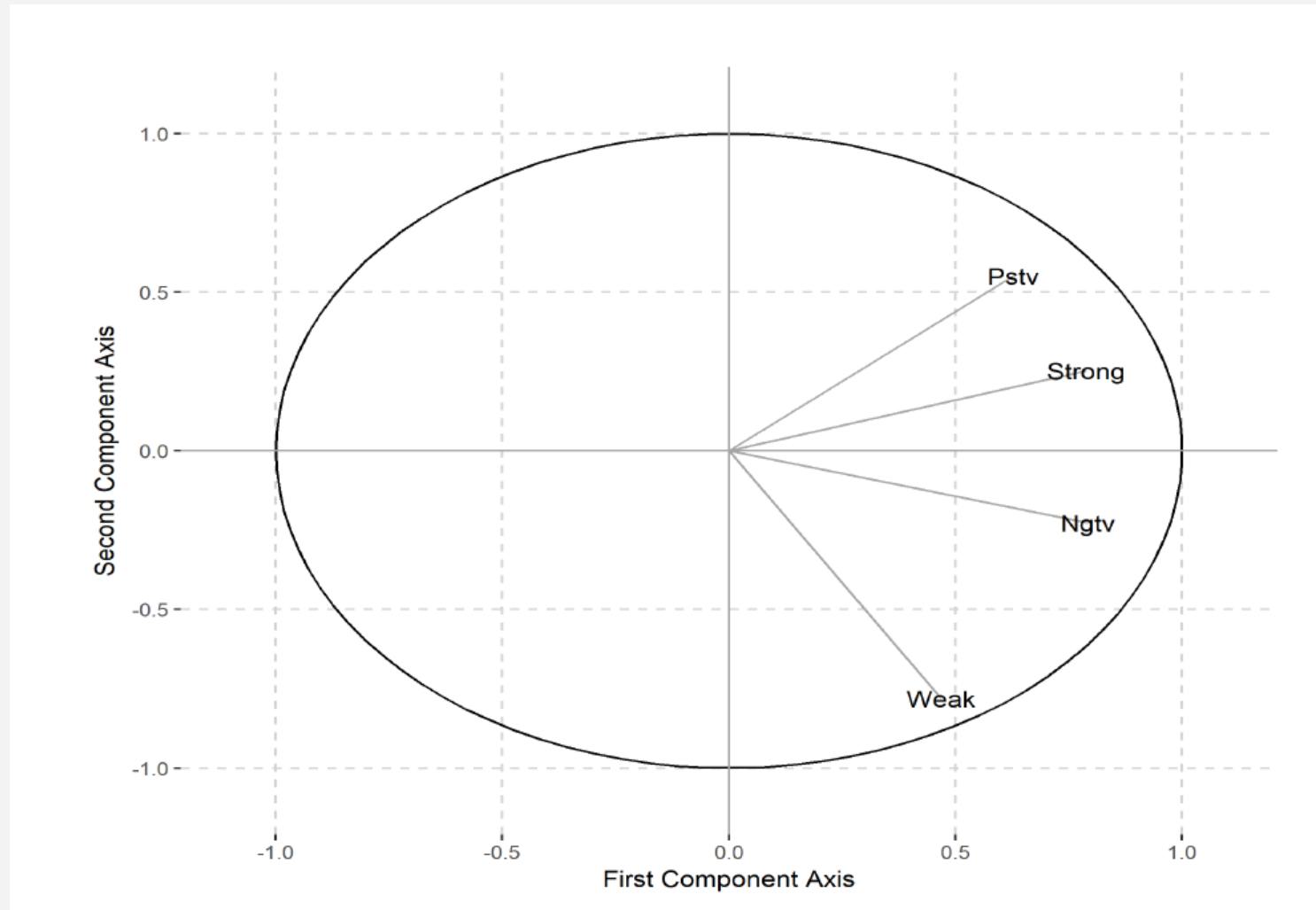
Variância explicada por cada um dos 4 componentes

	PC1	PC2	PC3	PC4
Standard deviation	1.3655	1.0149	0.7887	.6953
Proportion of Variance	.4661	.2575	.1555	.1209
Cumulative Proportion	.4661	.7236	.8791	1.0000

DADOS E METODOLOGIA

- Usaremos os dois primeiros componentes: **72% da variância** das categorias
- Os coeficientes do primeiro componente ("PC1") são todos positivos, com valores entre .34 e .58 - podemos interpretá-lo como **Fator de Engajamento** (*Engagement Factor* ou **EG**), ou seja, o quanto as pessoas estão ativas nas redes sociais.
- Vários sentimentos diferentes estão sendo expressados, diferentemente de uma coluna de jornal diária que vai, seguramente, seguir um ponto de vista e, portanto, dificilmente englobaria todas as categorias de sentimentos do dicionário
- Os coeficientes do segundo componente ("PC2") obviamente terão os sinais alternados (uma vez que são os componentes são ortogonais entre si) e nota-se que é positivamente correlacionado com categorias que indicam positivas e fortes. Podemos interpretá-lo como **Fator de Otimismo** (*Optimism Factor* ou **OF**)

DADOS E METODOLOGIA

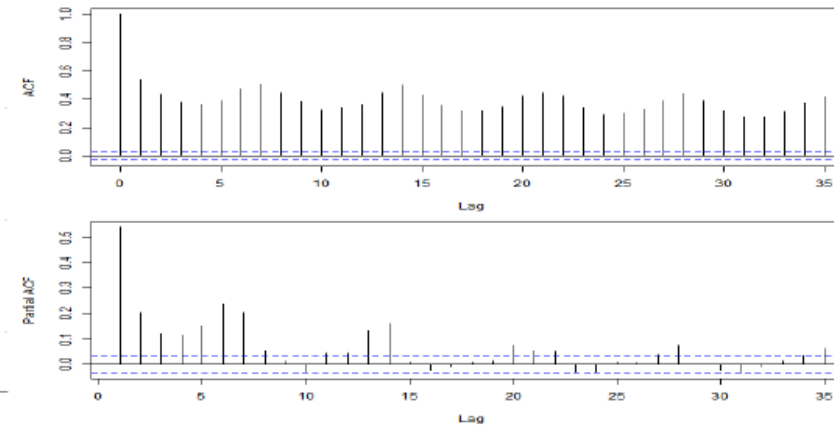


DADOS E METODOLOGIA

Engagement Factor



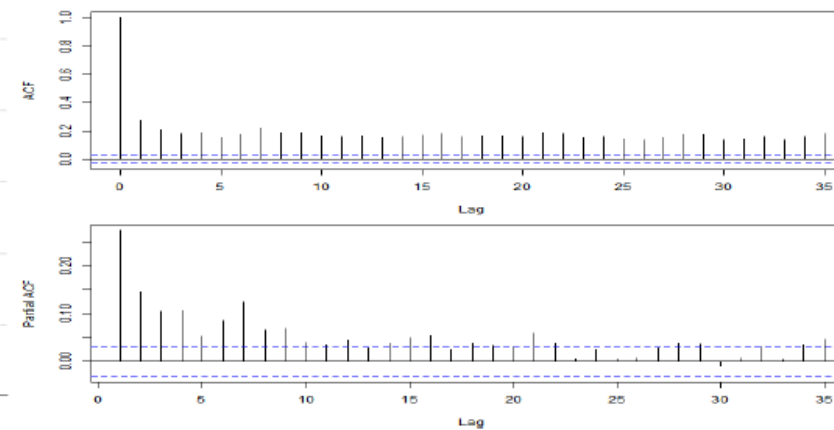
EF ACF e PACF



Optimism Factor



OF ACF e PACF



DADOS E METODOLOGIA

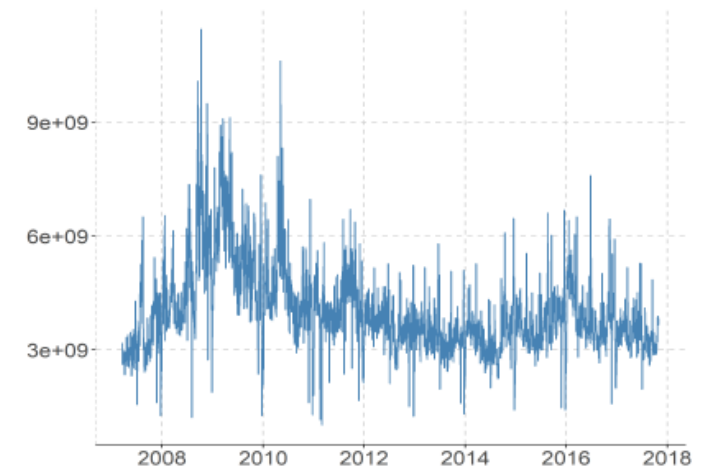
Figure 3.4 – Level value of S&P500, VIX and S&P500 trading volume



(a) S&P500 index, level



(b) VIX index, level



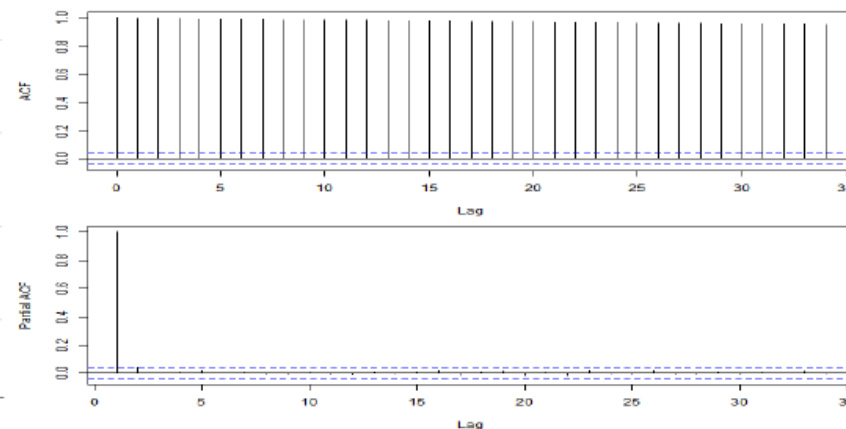
(c) Volume of S&P500, level

DADOS E METODOLOGIA

S&P 500 nível



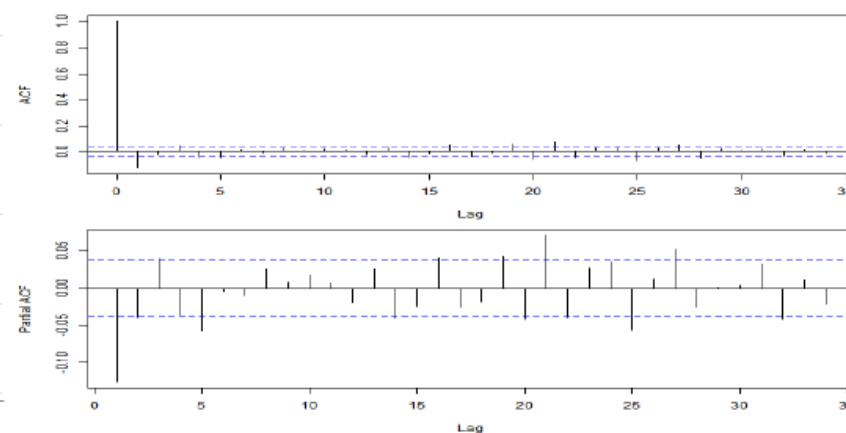
Nível ACF e PACF



S&P 500 retorno diário



Retorno ACF e PACF

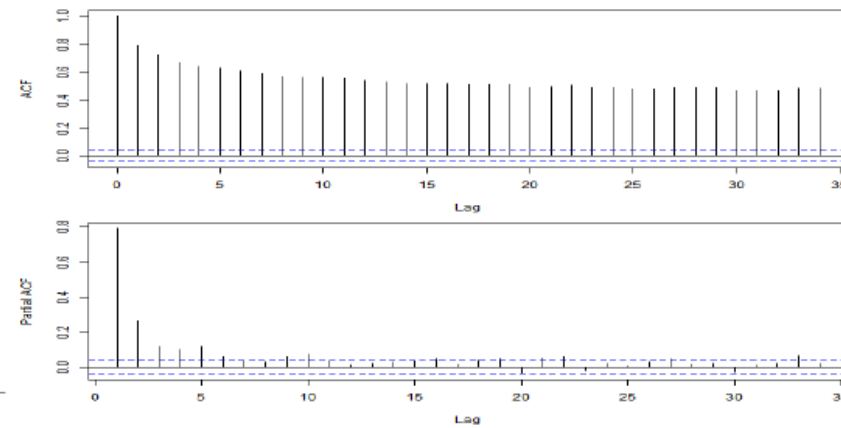


DADOS E METODOLOGIA

S&P 500 Volume nível



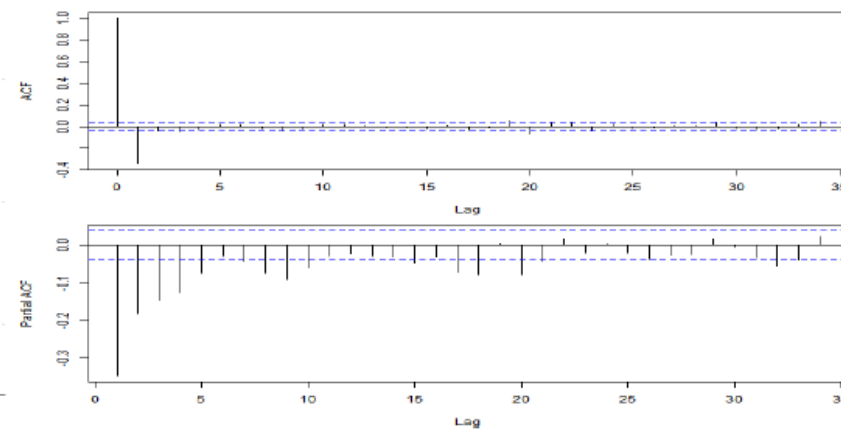
Nível ACF e PACF



S&P 500 Volume primeira dif.



Primeira dif. ACF e PACF

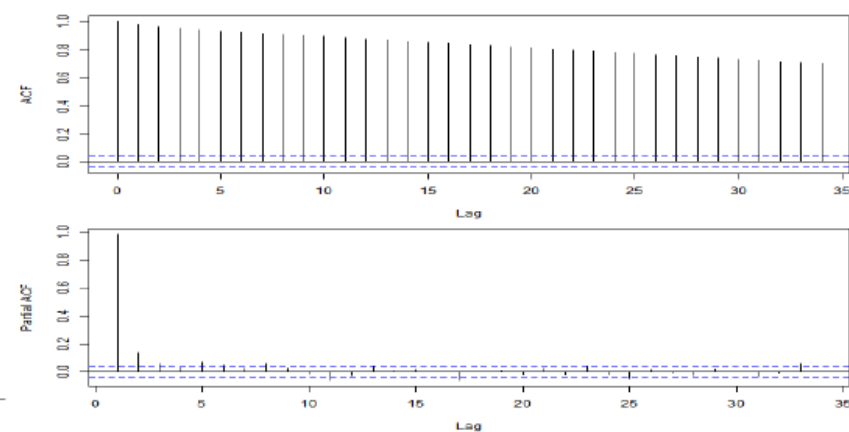


DADOS E METODOLOGIA

VIX nível



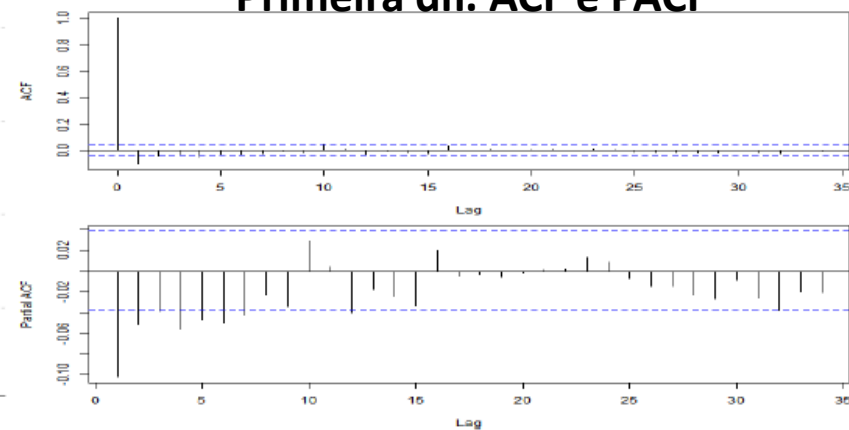
Nível ACF e PACF



VIX primeira dif.



Primeira dif. ACF e PACF



DADOS E METODOLOGIA

Table 4.4 – Augmented Dickey-Fuller test statistics

Series	Level	log	Δ	$\Delta \log$
<i>EF</i>	<u>-6.16***</u>	n/a	-23.58***	n/a
<i>OF</i>	<u>-8.23***</u>	n/a	-25.83***	n/a
S&P500	-2.12	-2.39	-14.27***	-14.27***
VIX	-4.21***	-15.46***	<u>-15.46***</u>	-16.13***
S&P500 Volume	-6.26***	-7.16***	<u>-19.57***</u>	-18.90***
Excess of log return S&P500	<u>-14.30***</u>	n/a	n/a	n/a

Note: *p<0.1; **p<0.05; ***p<0.01

log is the the log transformed series and Δ means the first difference

$$\mathbf{X}_t = \begin{bmatrix} EF_t \\ OF_t \\ Ret_t \\ \Delta Vlm_t \\ \Delta Vix_t \end{bmatrix}$$

DADOS E METODOLOGIA

- Modelo **VAR(5)**:

$$\Phi(L)X_t = c + \varepsilon_t,$$

$$EF_t = c_1 + \alpha_1 L_{1-5}(EF_t) + \beta_1 L_{1-5}(OF_t) + \gamma_1 L_{1-5}(Ret_t) + \eta_1 L_{1-5}(\Delta Vlm_t) + \theta_1 L_{1-5}(\Delta Vix_t) + \varepsilon_{1,t} \quad (4.7)$$

$$OF_t = c_2 + \alpha_2 L_{1-5}(EF_t) + \beta_2 L_{1-5}(OF_t) + \gamma_2 L_{1-5}(Ret_t) + \eta_2 L_{1-5}(\Delta Vlm_t) + \theta_2 L_{1-5}(\Delta Vix_t) + \varepsilon_{2,t} \quad (4.8)$$

$$Ret_t = c_3 + \alpha_3 L_{1-5}(EF_t) + \beta_3 L_{1-5}(OF_t) + \gamma_3 L_{1-5}(Ret_t) + \eta_3 L_{1-5}(\Delta Vlm_t) + \theta_3 L_{1-5}(\Delta Vix_t) + \varepsilon_{3,t} \quad (4.9)$$

$$\Delta Vlm_t = c_4 + \alpha_4 L_{1-5}(EF_t) + \beta_4 L_{1-5}(OF_t) + \gamma_4 L_{1-5}(Ret_t) + \eta_4 L_{1-5}(\Delta Vlm_t) + \theta_4 L_{1-5}(\Delta Vix_t) + \varepsilon_{4,t} \quad (4.10)$$

$$\Delta Vix_t = c_5 + \alpha_5 L_{1-5}(EF_t) + \beta_5 L_{1-5}(OF_t) + \gamma_5 L_{1-5}(Ret_t) + \eta_5 L_{1-5}(\Delta Vlm_t) + \theta_5 L_{1-5}(\Delta Vix_t) + \varepsilon_{5,t} \quad (4.11)$$

AGENDA

Introdução / motivação

Revisão de literatura

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

RESULTADOS

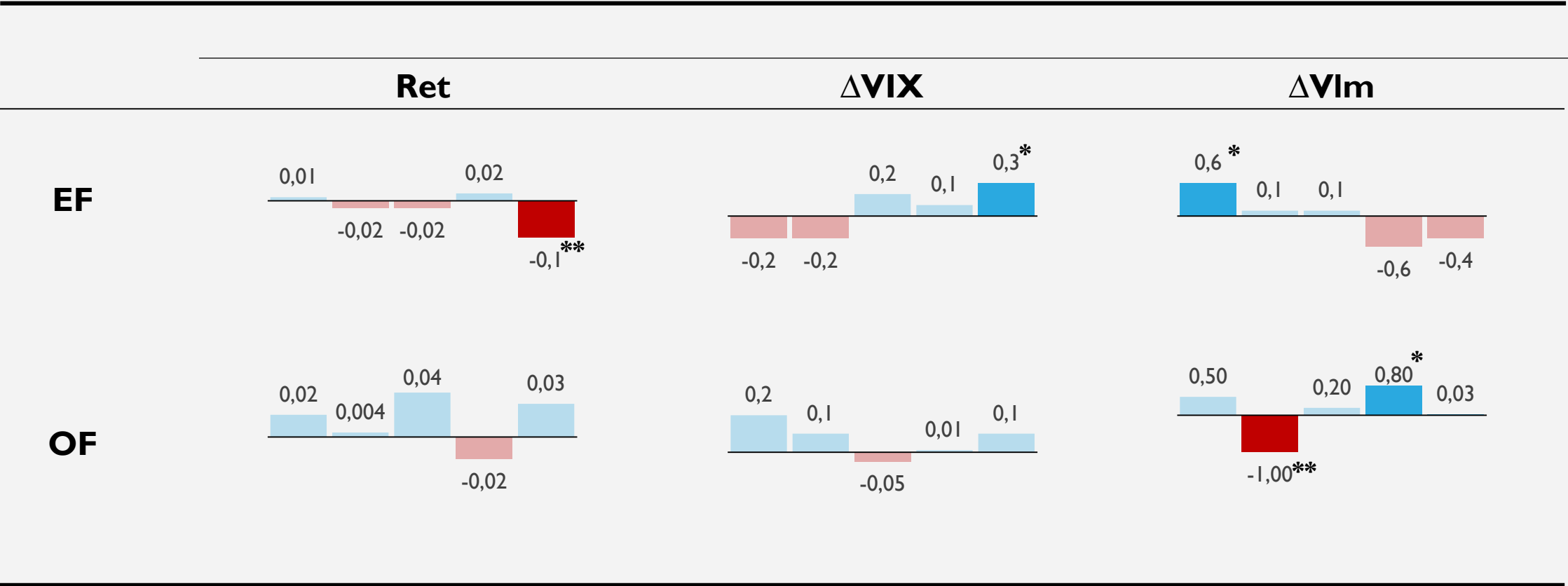
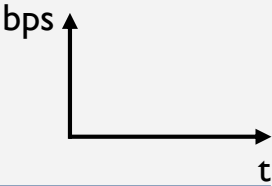
Table 5.1 – Social Media factors estimates

	<i>Dependent variable:</i>		
	<i>Ret</i>	ΔVix	ΔVlm
EF_{t-1}	-.001** (.0003)	.003* (.002)	-.004 (.004)
EF_{t-2}	.0002 (.0003)	.001 (.002)	-.006 (.004)
EF_{t-3}	-.0002 (.0003)	.002 (.002)	.001 (.004)
EF_{t-4}	-.0002 (.0003)	-.002 (.002)	.001 (.004)
EF_{t-5}	.0001 (.0003)	-.002 (.002)	.006* (.004)
OF_{t-1}	.0003 (.0003)	.001 (.002)	.0003 (.005)
OF_{t-2}	-.0002 (.0004)	.0001 (.002)	.008* (.005)
OF_{t-3}	.0004 (.0004)	-.0005 (.002)	.002 (.005)
OF_{t-4}	.00004 (.0004)	.001 (.002)	-.010** (.005)
OF_{t-5}	.0002 (.0003)	.002 (.002)	.005 (.005)
Observations	2,646	2,646	2,646
R ²	.033	.028	.206
Adjusted R ²	.023	.019	.198
Residual Std. Error (df = 2620)	.013	.073	.170
F Statistic (df = 25; 2620)	3.544***	3.023***	27.172***

Note:

*p<0.1; **p<0.05; ***p<0.01
Standard deviations in parenthesis

RESULTADOS



*p<0.1; **p<0.05; ***p<0.01

RESULTADOS

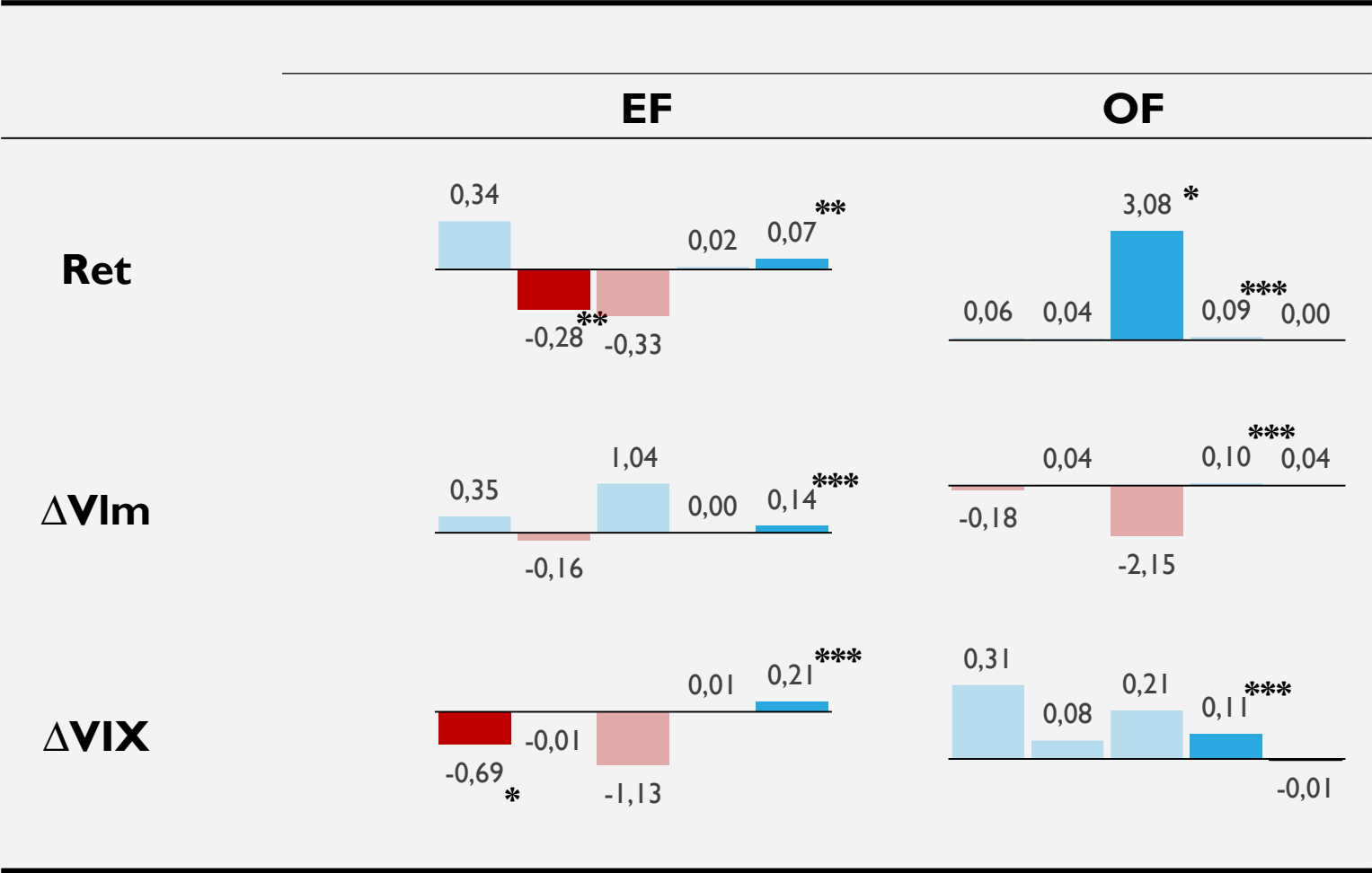
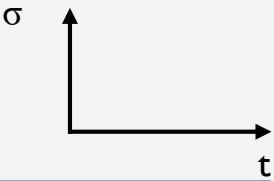
Table 5.2 – Market effects on social media sentiment

	<i>Dependent variable:</i>	
	<i>EF</i>	<i>OF</i>
Ret_{t-1}	.073*** (.020)	.001 (.018)
Ret_{t-2}	.017 (.024)	.088*** (.021)
Ret_{t-3}	-.328 (2.052)	-3.081* (1.764)
Ret_{t-4}	-.280** (.113)	.041 (.097)
Ret_{t-5}	.341 (.357)	-.059 (.307)
Vlm_{t-1}	.142*** (.020)	.037** (.017)
Vlm_{t-2}	.004 (.024)	.102*** (.021)
Vlm_{t-3}	1.039 (2.054)	-2.153 (1.766)
Vlm_{t-4}	-.158 (.112)	.040 (.096)
Vlm_{t-5}	.347 (.357)	-.181 (.307)
VIX_{t-1}	.212*** (.019)	-.005 (.017)
VIX_{t-2}	.012 (.023)	.106*** (.020)
VIX_{t-3}	-1.127 (2.039)	.205 (1.753)
VIX_{t-4}	-.012 (.101)	.078 (.087)
VIX_{t-5}	-.691* (.353)	.313 (.304)
Observations	2,646	2,646
R ²	.574	.212
Adjusted R ²	.570	.204
Residual Std. Error (df = 2620)	.881	.758
F Statistic (df = 25; 2620)	141.344***	28.159***

Note:

*p<0.1; **p<0.05; ***p<0.01
Standard deviations in parenthesis

RESULTADOS



*p<0.1; **p<0.05; ***p<0.01

CONCLUSÕES

- De maneira geral, parece que o sentimento público é mais responsivo aos eventos de mercado
- Euforias momentâneas (respostas rápidas a eventos de retorno e incerteza ~dias)
- O efeito das euforias parece se reverter (em 4 dias a relação é negativa)
- Volume parece dar otimismo no curto prazo (1-2 dias)
- EF e OF são muito correlacionado com os valores passados (até o 5o lag são todos significativos);
- O EF é influenciado positivamente pelo otimismo (OF) no primeiro e segundo lag;

AGENDA

Introdução / motivação

Revisão de literatura

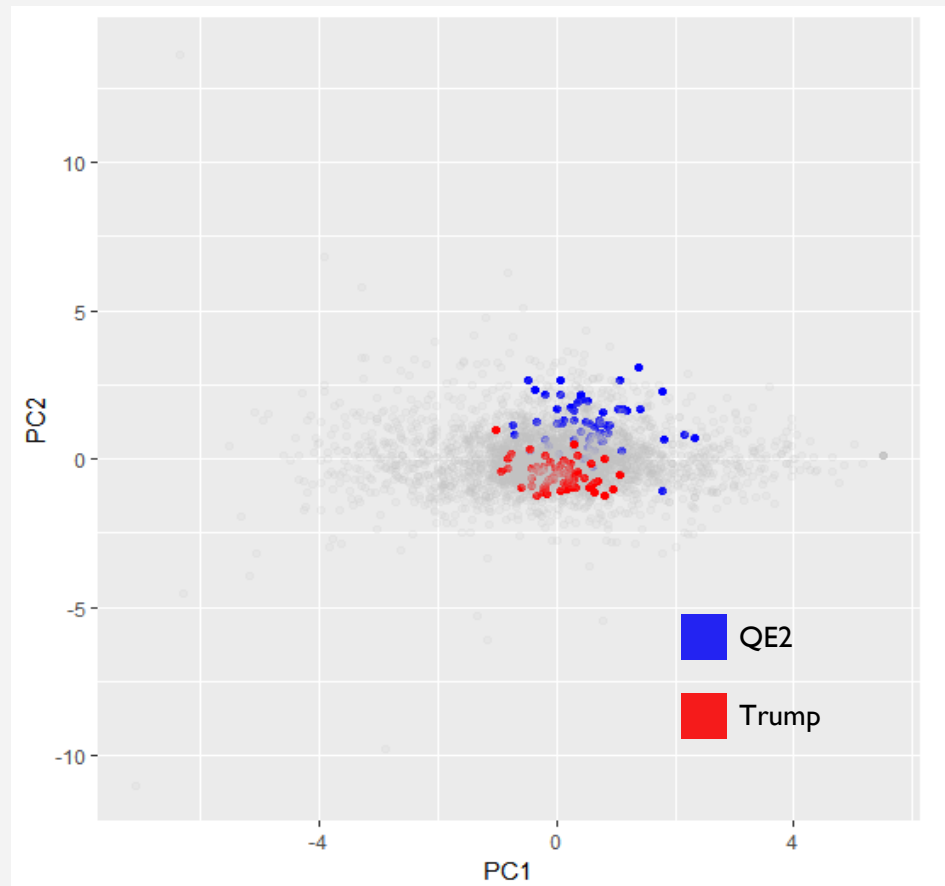
Dados e metodologia

Resultados e conclusões

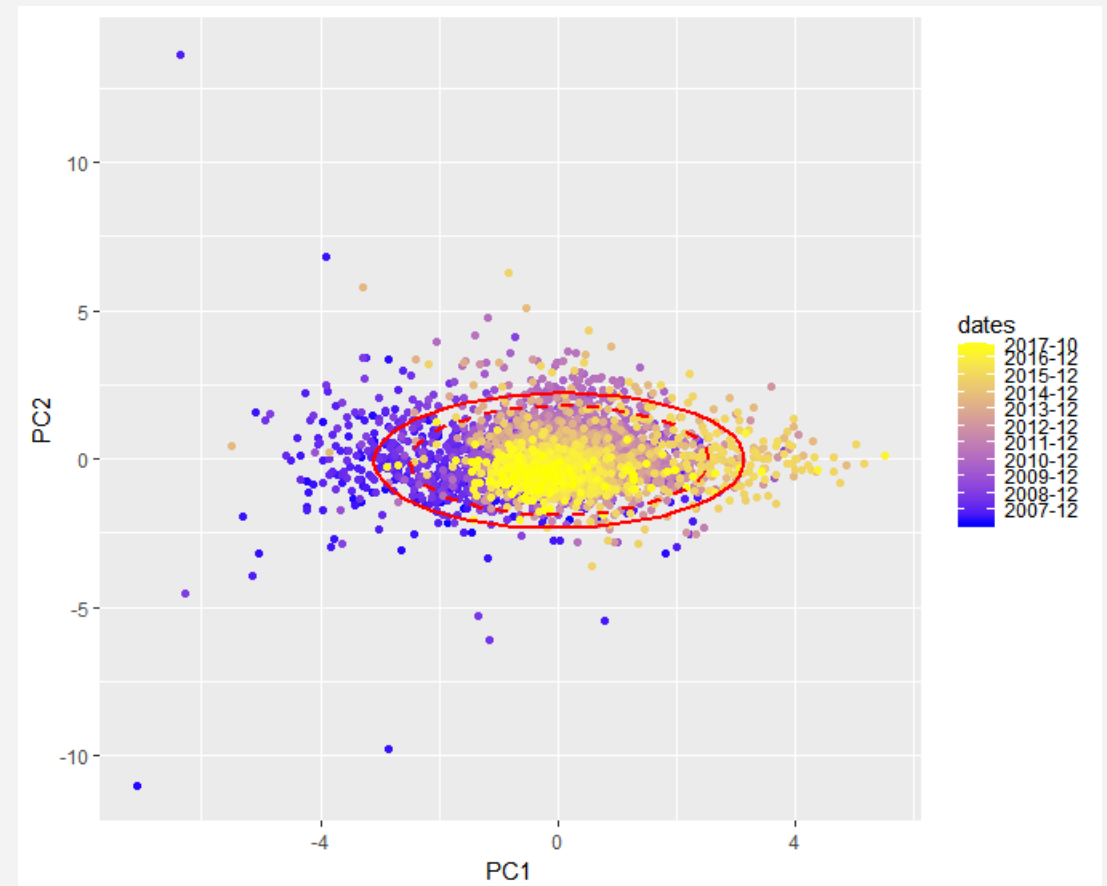
Trabalhos futuros

TRABALHOS FUTUROS

Clusters



Hotelling T²



AGENDA

Introdução / motivação

Revisão de literatura

Dados e metodologia

Resultados e conclusões

Trabalhos futuros

Perguntas e Dúvidas