

O GOOGLE COMO BARÓMETRO DA ATENÇÃO DO INVESTIDOR

PEDRO LATOEIFIRO E SOFIA B. RAMOS

1. INTRODUÇÃO

Na semana de Setembro de 2008 em que o banco Lehman Brothers' declarou falência o volume de pesquisas no Google sobre "Lehman" superou em 24 vezes a média histórica. Cerca de um ano depois, em Novembro de 2009, o Dubai World procurou um reescalonamento da dívida. Nesse mês, em que as suas acções acumularam perdas de 26%, a procura de informação no Google sobre "Dubai World" sextuplicou. No primeiro semestre do ano seguinte o vazamento de uma plataforma da British Petroleum desencadeou um desastre ambiental no Golfo do México. Em Londres, o volume de títulos da empresa transaccionado em Maio de 2010 duplicou a média dos 12 meses anteriores. No Google a frequência de pesquisas sobre "BP" quintuplicou nesse período.

Os motores de pesquisa tornaram-se numa popular e poderosa fonte de informação na Internet por funcionarem como atalhos que procuram satisfazer de forma rápida e eficiente as intenções do utilizador. Mas será que pesquisar implica decidir? A relação entre a procura de informação na Internet e a tomada de decisões dos investidores é o tema desta investigação.

Estudos recentes validam o motor de busca do Google como barómetro fiável da atenção do investidor. Da, Engelberg e Gao (2011) assumem que por detrás das pesquisas no Google estão sobretudo investidores particulares e

Bank, Larch e Peter (2011) associam um aumento da frequência de consultas sobre uma empresa a maior liquidez.

A presente investigação analisa a relação entre a procura de informação medida pela frequência de pesquisas no Google e a evolução em bolsa das cotadas do índice Euro Stoxx 50 entre 2004 e 2011. As principais novidades face à literatura são a focagem em relações não lineares – diferenciando-se o impacte de uma subida do volume de consultas do efeito de uma descida -, e o estudo das variações extremas na atenção do investidor – construindo-se indicadores para captar picos de concentração e momentos de hibernação.

Em primeiro lugar procede-se a uma avaliação da relação contemporânea entre a frequência de pesquisas no Google sobre as cotadas e a evolução do volume e da volatilidade. Constroem-se regressões *piecewise* para segmentar as empresas consoante a sua notoriedade no Google e logo aí se encontra os primeiros sinais de assimetria. O poder explicativo das pesquisas na Internet sobre o volume e a volatilidade manifesta menor robustez no grupo de cotadas de maior notoriedade. Esse mesmo poder revela contudo maior significância nos quartis que reúnem as empresas mais negligenciadas, onde uma subida da procura de informação na Internet está associada a um aumento do volume e da volatilidade.

O segundo tópico alvo de análise é a previsibilidade, que comporta especial importância teórica por se relacionar com a eficiência do mercado. Encontra-se evidência de que uma subida na frequência de consultas no Google antecipa menor rentabilidade e maior volatilidade e volume na semana seguinte. Mas também aqui se depara com sinais de não linearidade. Se um aumento da procura de informação no Google tem um efeito positivo sobre a volatilidade e um impacto negativo sobre o preço da acção, o inverso não se verifica. Já no caso do volume a resposta negativa a uma diminuição da procura de informação revela-se muito mais intensa que o efeito contrário.

Por último analisa-se que tipo de investidor estará por detrás do indicador de atenção gerado pela frequência de pesquisas no Google, aprofundando uma questão que tem suscitado o interesse da literatura. Encontra-se sinais de que as buscas na Internet captam sobretudo a atenção dos investidores particulares e, ao mesmo tempo, de que o fluxo de informação gerado pelo Google não está alheado da agenda dos institucionais.

As perguntas que motivam este estudo são relevantes em várias importantes frentes da investigação em Economia e Finanças. As respostas podem contribuir para perceber como a atenção do investidor é incorporada no preço e na liquidez de um activo e assim acrescentar valor ao estudo da eficiência do mercado. Podem relevar a importância do grau de atenção do investidor na formação do preço de um activo e assim legitimar estratégias de investimento construídas com base nesse critério. Podem ajudar a entender o comportamento dos investidores particulares e assim melhorar o entendimento sobre os efeitos da sua entrada no mercado. Podem, ao

testar a Internet como fonte privilegiada de informação dos investidores, promover o estudo das pesquisas do Google como uma variável avançada da evolução do mercado. E podem também relevar a importância de os reguladores com deveres de protecção sobre os investidores menos sofisticados dirigirem parte da sua acção para a Internet.

2. MEDINDO A ATENÇÃO

Em 1987 Robert C. Merton publicou um artigo dedicado à *investor recognition hypothesis* (IRH), argumentando que o custo de obter informação impede os investidores de conhecer todas as empresas disponíveis no mercado. O facto de cada carteira de investimento ser constituída apenas por uma fracção do universo de empresas existentes prova que cada investidor conhece e transacciona somente sobre um subconjunto de activos (Merton, 1987), o que conduz a uma diversificação imperfeita na alocação de activos¹. Para compensar esse desequilíbrio as empresas menos conhecidas têm de oferecer retornos mais elevados. Essa premissa, associada ao custo de produção e difusão de informação sobre uma empresa, sustenta o argumento de que o número de investidores conhecedores de um activo – a notoriedade de uma acção –, influencia o seu valor de mercado. Deste modo, e uma vez que os recursos dos investidores, nomeadamente o tempo, são limitados, as acções com maior notoriedade terão maior probabilidade de integrar o conjunto de oportunidades de investimento, da mesma forma que os títulos com menor visibilidade acarretarão maior risco de ser ignorados (Barber e Odean, 2008).

Mas como medir a atenção do investidor em cada momento? Nos últimos anos a literatura

1- Este resultado contrasta com o que um investidor racional deve obter na formação de uma carteira eficiente.

começou a procurar na Internet variáveis capazes de medir e acompanhar a notoriedade de uma acção no mercado. Nesse caminho, as pesquisas concretizadas no Google emergiram como variável de interesse devido a três propriedades fundamentais: a escala, a relevância² e o imediatismo. Quantitativamente, os serviços do Google são utilizados mensalmente por mais de mil milhões de indivíduos de múltiplas geografias³, uma escala sem paralelo nas variáveis testadas anteriormente⁴. Em adição, as pesquisas no Google só se materializam mediante uma impulsão do utilizador, pelo que “if you search for a stock in Google, you are undoubtedly paying attention to it” (Da et al., 2011: 2). As buscas são, portanto, dotadas de uma relevância que a publicação de artigos de imprensa, por exemplo, variável em que o investidor é um receptor passivo e relativamente impotente de informação, não dispõe. Além disso, em contraste com o desfasamento e a retrospectividade que dominam a informação financeira, o Google disponibiliza dados diários e semanais exportáveis com um clique para tratamento estatístico, uma vantagem temporal e metodológica relevante dado o imediatismo que caracteriza os mercados financeiros.

Estudos de outras áreas científicas mostraram a grande capacidade de previsão das pesquisas em motores de busca. Um trabalho publicado na *Nature* demonstrou ser possível prever surtos de gripe uma a duas semanas antes do Departamento de Saúde dos EUA através das variações no volume de pesquisas no Google sobre 45 termos coloquialmente associados ao vírus (Ginsberg, Mohebill, Patel, Brammer, Smolinskil e Brilliant, 2009). Também Choi e

Varian (2009) utilizaram o volume de buscas no Google, mas para prever as vendas da Chevrolet e da Toyota, as vendas de habitações nos EUA e ainda os fluxos de turismo em nove países. Na mesma linha, Goel, Hofman, Lahaie, Pennock e Watts (2010) usaram o Google para antecipar as receitas de bilheteira no fim-de-semana de estreia de 119 filmes nos EUA, as vendas do primeiro mês de vida de 106 videogames e ainda o *ranking* da revista *Billboard* sobre as 100 músicas mais populares no país. Já numa moldura macroeconómica, Askitas e Zimmermann (2009) detectaram uma correlação positiva entre o volume de pesquisas em alemão sobre os nomes dos sites de emprego mais populares na Alemanha e a evolução da taxa de desemprego germânica.

Da et al. (2011) terão sido os primeiros a testar a IRH de Merton no mercado accionista através do Google. Usando as acções do índice Russell 300 como painel, os autores concluem que uma subida no volume de buscas antecipa em duas semanas a pressão compradora atribuída a investidores particulares, que se revela mais intensa nas cotadas com menor valor de mercado. Os autores examinam depois todas as ofertas públicas iniciais, operações que habitualmente suscitam a atenção do retalho, concretizadas nos EUA entre 2004 e 2007 e deduzem que o Google comporta poder preditivo sobre a evolução da cotação na estreia em bolsa.

Olhando para 30 cotadas da *New York Stock Exchange*, Vlastakis e Markellos (2012) também encontraram relações positivas entre o volume de pesquisas na Internet e a liquidez e a volatilidade das acções. No mesmo sentido,

2- No sentido de que “os processos cognitivos humanos são dirigidos para a finalidade de conseguirem o maior efeito cognitivo possível com o menor esforço de processamento” (Sperber e Wilson, 2001: 23).

3- Dados da ComScore referentes a Maio de 2011 (<http://www.comscore.com>).

4- O grau de acompanhamento dos analistas profissionais (Arbel e Streble, 1983), o investimento publicitário (Frieder e Subrahmanyam, 2005), a cobertura mediática (Fang e Peress, 2009), a actividade nos fóruns de investimento na Internet (Antweiler e Frank, 2004) e a frequência de edição da Wikipedia (Rubin e Rubin, 2010) são algumas das variáveis propostas para captar a atenção do investidor.

Bank et al. (2011) utilizam o Google para estudar as cotadas do índice Xetra. Ao aumento do volume de pesquisas é associada uma apreciação temporária da cotação e uma subida da liquidez da acção, relação que os investigadores atribuem à redução de custos de assimetria de informação, deduzindo a partir daí que o Google serve sobretudo para medir a atenção dos investidores particulares.

Em estudos mais recentes, Dimpfl e Jank (2012) detectam comovimento e causalidade de Granger entre a frequência de pesquisas no Google sobre o *Dow Jones Industrial Average* e a volatilidade do índice. Smith (2012) conclui que um aumento das consultas sobre “crise económica” e “crise financeira” prediz em uma semana um incremento na volatilidade associada a sete divisas. Drake, Roulstone e Thornock (2012) encontram níveis anormais de pesquisas nas duas semanas que antecedem a publicação de contas de empresas do índice S&P 500 e argumentam que quanto maior é a procura antecipada por informação menor é a variação da cotação no dia do anúncio dos resultados.

3. MÉTODO E VARIÁVEIS

Utiliza-se como base do estudo os dados recolhidos e fornecidos pelo Google Trends⁵ para analisar a eventual relação entre o volume de pesquisas na Internet sobre uma empresa e a sua evolução no mercado. O Google Trends é um serviço gratuito disponibilizado ao público desde 2008 que regista a evolução do número de pesquisas sobre uma determinada palavra, ou conjunto de palavras, em termos relativos, apresentando um histórico. Usa-se o Google Trends por ser um serviço único prestado pelo

líder do mercado global de pesquisas com uma quota de 83%⁶. Procurar-se-á demonstrar que a variação do volume de pesquisas é um barómetro fiável da atenção do investidor, tal como definida na IRH de Merton (1987). Seguindo estudos anteriores, designa-se doravante a frequência de buscas por *Google Search Volume Indicator* (GSVI). A construção é a seguinte:

$$GSVI_{k,t} = \frac{PS_{k,t}}{PS_T} \quad (1)$$

com $\overline{PS_T}$ a designar a média das pesquisas sobre o termo particular k concretizadas durante todo o período analisado, T . A escala do GSVI varia entre zero (período em que o número de pesquisas é nulo ou insuficiente para gerar um coeficiente) e 100 (momento em que ocorreu o maior número de pesquisas), sendo que qualquer acréscimo ou decréscimo consiste sempre numa relação face à média histórica do período analisado.

A informação disponibilizada pelo Google Trends impõe três condicionantes à metodologia de análise a aplicar. Há desde logo uma limitação temporal, uma vez que só existem dados posteriores a Janeiro de 2004. Há também uma limitação de periodicidade porque o Google apenas permite a exportação de dados semanais. Por último, o facto de os GSVI serem apresentados como variações face à média impede de trabalhar econometricamente com valores absolutos individuais das empresas, limitando a exploração de eventuais relações entre as cotadas e entre os sectores.

A decisão sobre os termos de pesquisa a considerar é outra questão basilar na definição do método a aplicar. Uma pesquisa sobre

5- A 27 de Setembro de 2012 o Google anunciou sem aviso algumas mudanças no Google Trends (www.google.com/trends). Embora as características e a finalidade do serviço se mantenham, a normalização, disponibilização e exportação dos dados sofreram alterações que tornam única a base de dados em que esta investigação se sustenta. A descrição do método e das variáveis referem-se sempre às regras originais.

6- Dados da NetMarketShare para Julho de 2012 (<http://www.netmarketshare.com>).

“Santander” refere-se ao banco ou à cidade no norte de Espanha? Por detrás de uma pesquisa sobre “Carrefour” estará um potencial investidor ou apenas um utilizador que procura pelo supermercado mais próximo? Com a tecnologia actual disponível é impossível conhecer com exactidão a intenção final de todos os pesquisadores em todos os momentos. Não havendo forma de distinguir o “pesquisar para saber” do “pesquisar para decidir”, utiliza-se, tal como Vlastakis et al. (2012) e Chen (2011), o nome coloquial da empresa (apenas “Santander”).

Escolheu-se como objecto de análise o mercado accionista europeu através do Euro Stoxx 50, índice que reúne 50 cotadas de 12 países da zona euro⁷ e que serve de subjacente a um conjunto alargado de produtos de investimento. Este índice agrupa *blue chips* que partilham uma divisa, o euro, mas que têm sede em diferentes geografias e operam em sectores distintos, diluindo-se assim eventuais efeitos específicos derivados da evolução cambial, da nacionalidade, da dimensão e da natureza de actividade. O Euro Stoxx 50 nunca foi alvo de análise semelhante, até porque os estudos anteriores comparáveis adoptaram fronteiras mais restritas ao incidir sobre mercados accionistas domésticos.

Das 50 cotadas que compõem o índice actualmente só se conserva as empresas que integraram ininterruptamente o Euro Stoxx 50⁸ no período em análise, de acordo com dados gentilmente fornecidos pelo STOXX. Com este critério anula-se o efeito que a inclusão ou exclusão de um índice causa na notoriedade de um título. Em adição, para aumentar a precisão das conclusões segue-se o critério adoptado por

Chen (2011) e exclui-se as cotadas com GSVI igual a zero durante pelo menos oito semanas consecutivas, considerando que essas observações são infrutuosas para a investigação e podem mesmo distorcer os resultados. Da aplicação das duas regras resultou o objecto de análise, que consiste em 36 cotadas com origem em seis países e operando em nove sectores de actividade, apresentado no Quadro 3.1..

O período alvo de análise inicia-se, por imposição do Google Trends, na primeira semana de Janeiro de 2004 e finda na última semana de Junho de 2011.

3.1. Variáveis financeiras

Os dados sobre o volume e o preço das empresas foram importados da Bloomberg. Não se impõe restrições sobre os indicadores financeiros por se reputar que esse controlo é garantido pelos critérios definidores da revisão da composição do Euro Stoxx 50⁹.

As variáveis de volume são quantidades desprovidas do factor preço. As variações logarítmicas do volume são calculadas da seguinte forma:

$$(2) \quad \begin{aligned} V_{xt} &= \ln(v_{xt}) \\ \Delta V_{xt} &= v_{xt} - v_{xt-1} \end{aligned}$$

com ΔV_{xt} a ilustrar a variação do número de títulos da empresa x transaccionados na semana t face à semana anterior. Por se desconhecer com exactidão a natureza dos propulsores da atenção do investidor, nomeadamente se é conduzida por notícias positivas ou negativas, ou por ambas na mesma ponderação, também se utiliza o valor absoluto das variações do volume, obtido por:

7- Alemanha, Áustria, Bélgica, Finlândia, França, Grécia, Holanda, Irlanda, Itália, Luxemburgo, Portugal e Espanha.

8- A composição do índice é revista anualmente em Setembro.

9- Integram o índice as 50 cotadas da zona euro com maior valor de mercado, calculado a partir do *free float*, presentes no STOXX Europe 600.

$$absV_{x,t} = |\Delta V_{x,t}| \quad (3)$$

com $absV_{x,t}$ a significar a variação em módulo do número de acções da companhia x negociadas na semana t .

Tal como Chen (2011) e Dimpfl et al. (2012) usa-se a volatilidade realizada. Para a estimar determina-se em primeiro lugar o logaritmo natural do retorno, dado por:

$$r_{x,t} = \ln\left(\frac{P_{x,t}}{P_{x,t-1}}\right) \quad (4)$$

onde $r_{x,t}$ representa o retorno da acção x no dia t com base nas cotações P . Obtém-se posteriormente a volatilidade realizada a partir do somatório do quadrado dos retornos:

$$RV_{x,t} = \sum_{i=0}^N r_{x,t,i}^2 \quad (5)$$

com $r_{x,t,i}^2$ a corresponder ao retorno da cotada x no dia i da semana t . Implementa-se $N=4$ uma vez que os dados extraídos do Google são semanais. Ross (1989) sugeriu a variabilidade do retorno como indicador do fluxo de informação no mercado.

A partir da equação (4) determina-se as variações semanais do valor da acção para calcular o retorno absoluto de cada empresa, através de:

$$absr_{x,t} = |r_{x,t}| \quad (6)$$

com $absr_{x,t}$ a representar o módulo do retorno da acção x na semana t . Esta variável tem sido utilizada pela literatura para medir a atenção do investidor (Corwin e Coughenour, 2008).

No que diz respeito ao Stoxx 50, usa-se nas estimações, além do logaritmo natural do retorno

semanal, a evolução também em base semanal do VSTOXX, índice que mede a volatilidade implícita do índice a partir do preço das opções negociadas no mercado Eurex.

3.2. Variáveis Google

Através do GSVI estimam-se outras variáveis. A primeira é obtida pelo processo:

$$\lnGSVI_{x,t} = \ln(GSVI_{x,t}) \quad (7)$$

com $\lnGSVI_{x,t}$ a simbolizar o logaritmo natural do GSVI sobre a empresa x na semana t , tal como fizeram Bank et al. (2011), Da et al. (2011) e Chen (2011). Deste modo, GSVI inferiores a 1, um sinal de menor volume de pesquisas, aparecerão em logaritmo com sinal negativo. Essa é a principal motivação da variável.

Seguindo Da et al. (2011) e Smith (2012) constrói-se também o indicador *Jumps in Search Volume Indicator* (JSVI) para tentar capturar choques ou variações exógenas na atenção do investidor. Para tal, toma-se a diferença entre o GSVI e a média das quatro observações anteriores:

$$JSVI_{x,t} = GSVI_{x,t} - \sum_{i=1}^4 GSVI_{x,t-i} \quad (8)$$

São ainda criados indicadores que permitam identificar eventuais efeitos assimétricos da procura de informação no Google sobre as variáveis financeiras. Sabe-se que leituras de GSVI superiores a 1 e que valores positivos de \lnGSVI indicam um aumento da atenção do investidor. E que, contrariamente, uma descida no nível de atenção é descrita por registos de GSVI inferiores a 1 e valores negativos de \lnGSVI . Assim sendo, define-se:

$$\ln GSVI p_t = \max(0, \ln GSVI_t) \quad (9)$$

e

$$\ln GSVI n_t = \min(0, \ln GSVI_t) \quad (10)$$

Seguindo o mesmo raciocínio para variações acentuadas na atenção do investidor define-se:

$$\ln JSVI p_t = \max(0, \ln JSVI_t) \quad (11)$$

e

$$\ln JSVI n_t = \min(0, \ln JSVI_t) \quad (12)$$

4. VOLUME

Da caderneta de provérbios de Wall Street faz parte o adágio “it takes volume to make prices move”. Por ser uma medida indirecta de liquidez, por ser utilizado como previsor da evolução do preço de uma acção em estratégias de análise técnica, por reflectir o grau de notoriedade de um título, o volume é a primeira variável a ser testada em função do GSVI. A expectativa é a de que haverá uma relação entre as variáveis pois estudos anteriores estabeleceram uma ligação entre volume e atenção do investidor (Barber et al., 2008). Utiliza-se a análise de regressão com dados em painel para testar se as variáveis Google contêm poder explicativo sobre as variações no volume das cotadas do Stoxx 50 numa base contemporânea. O modelo estimado é o seguinte:

$$\Delta V_{x,t} = \alpha + \beta_0 \Delta V_{x,t-1} + \beta_1 R_{x,t} + \beta_2 RV_{x,t} + \beta_3 RS_t + \beta_4 RVS_t + \beta_5 GSVI_{x,t} + \epsilon_t$$

com $\Delta V_{x,t}$ a ilustrar a variação do número de acções da empresa x negociadas durante a semana t , α uma constante, $\Delta V_{x,t-1}$ a variável dependente desfasada em um período, $R_{x,t}$ o retorno da empresa x na mesma semana t , $RV_{x,t}$ a volatilidade realizada sobre a empresa x na semana t , RS_t o retorno do índice Stoxx 50 na semana t e RVS_t o valor do índice VSTOXX no mesmo período. O regressor $GSVI_{x,t}$ representa a frequência de pesquisas

no Google sobre a empresa x na semana t e constitui a variável independente de maior interesse. O último termo da regressão, ϵ_t é um processo de ruído branco. São estimadas quatro versões do modelo. Na primeira não se inclui qualquer variável Google. Já nas outras três usa-se isoladamente o GSVI, o lnGSVI e o JSVI, respectivamente. A partir do modelo (13) constrói-se uma regressão *piecewise*. Para tal divide-se as empresas em quartis em função do GSVI. O primeiro quartil reúne as cotadas de maior visibilidade e o último as de menor. Do modelo *piecewise* são calculados quatro conjuntos de resultados num total de 16 estimações. No primeiro conjunto não se utilizam variáveis Google. Nos restantes usa-se isoladamente o GSVI, o lnGSVI e o JSVI, respectivamente, para cada um dos quartis.

Conclui-se que a uma subida (descida) na frequência de pesquisas no Google está associada um acréscimo (decrécimo) do número de acções transaccionadas, dado o sinal dos coeficientes das variáveis Google apresentados no Quadro 4.1. ser positivo. O Google é portanto um barómetro da atenção do investidor porque pesquisar implica consciência sobre o termo pesquisado; o Google é também um reflector da notoriedade de uma acção, uma vez que toda a pesquisa é uma demonstração de interesse; e o Google é ainda um indicador indirecto de liquidez, pela frequência de pesquisas ter um efeito positivo na variação do volume de uma acção, com maior dimensão e robustez na variável JSVI. As regressões *piecewise* com o indicador JSVI expostas no Quadro 4.2. demonstram, contudo, que esta relação é não linear, na medida em que quanto maior (menor) é a notoriedade de uma empresa, menor (maior) o efeito da frequência de pesquisas no Google sobre as variações no volume da acção, um resultado novo para a literatura.

5. VOLATILIDADE

Depois da análise sobre o volume, estuda-se agora a relação entre a volatilidade e a frequência de pesquisas no Google. Por se referir à variabilidade do retorno, a volatilidade funciona como um indicador de risco e é por isso frequentemente lida como um índice reflector de incerteza. A incerteza pode gerar medo, que na Psicologia é uma emoção primária despertadora de atenção, por vezes exageradamente. Nos mercados financeiros não é diferente. E em ambos os casos a procura de informação é uma das possíveis manifestações dessa atenção. Espera-se por isso que o Google possua poder explicativo sobre a volatilidade, embora em menor grau do que o observado no volume, pois esta não é uma variável directamente observável e incorpora maior ruído e muitos outros factores além da atenção do investidor. Replicando a regressão com dados em painel utilizada para o volume, o modelo é estimado da seguinte forma:

$$RV_{x,t} = \alpha + \beta_0 RV_{x,t-1} + \beta_1 R_{x,t} + \beta_2 RS_t + \beta_3 RVS_t + \beta_4 GSVI_{x,t} + \epsilon_t \quad (14)$$

com $RV_{x,t}$ a representar a volatilidade realizada sobre a empresa x registada na semana t , α uma constante, $RV_{x,t-1}$ a variável dependente desfasada em um período, $R_{x,t}$ o retorno da empresa x na semana t , RS_t o retorno do índice Stoxx 50 na semana t , RVS_t o índice VSTOXX no mesmo período e, por último, $GSVI_{x,t}$ a designar a frequência de consultas no Google sobre a empresa x na semana t . A regressão inclui ainda um processo de ruído branco, ϵ_t . O modelo volta a ser usado em quatro versões: a primeira sem qualquer variável Google e as restantes com o GSVI, o lnGSVI e o JSVI, respectiva-

mente. Para aferir se o Google explica parte da volatilidade e se esse eventual poder explicativo varia consoante a visibilidade das empresas constrói-se, a partir da regressão (14), um modelo *piecewise* com parâmetros idênticos aos utilizados no estudo sobre o volume.

Deduz-se dos resultados da regressão (14) apresentados no Quadro 5.1. que o Google, enquanto fonte de informação, contém poder explicativo sobre a volatilidade das cotadas do Stoxx 50, ou seja, que parte da informação consultada é repercutida na variabilidade dos preços. Os dados confirmam que a relação é bastante mais fraca do que a verificada com o volume e estão em linha com o trabalho de Dimplf et al. (2012) e com o argumento de que “in an free arbitrage economy, the volatility of prices is directly related to the rate of flow information” (Ross, 1989).

Da análise gráfica patenteada na Figura 5.1. retiram-se indicações complementares que sustentam esta relação: o aumento da volatilidade parece estar sobretudo relacionado com leituras de GSVI superiores a 1, ou seja, com períodos em que a procura de informação subiu face à média histórica. Isso significa, aceitando que “the demand for information is positively related to risk aversion” (Vlastakis et al., 2010), que a frequência de pesquisas no Google pode ser lida como um reflexo invertido do apetite por risco dos investidores.

Adicionalmente, os resultados das regressões *piecewise* expostos no Quadro 5.2. voltam a denunciar, embora com um diferencial e uma robustez mais reduzidos dos que no volume, que é sobre os últimos quartis que as variáveis GSVI e lnGSVI manifestam maior poder explicativo em relação à volatilidade.

6. PREVISIBILIDADE

Ocasional ou intencional, extemporânea ou racional, a procura de informação no Google tem duas finalidades possíveis que podem coexistir: pesquisar para saber ou pesquisar para decidir. Ambas nascem de um impulso do utilizador, exigindo por isso consciência sobre o termo consultado, e em ambas há transferência de informação que, no último caso, motiva, justifica ou reforça uma decisão. Cronologicamente, tanto o conhecimento obtido como a decisão tomada em sua base são sempre posteriores ao clique da pesquisa. Aplicada a um contexto financeiro, esta sequência temporal significa que as necessidades de informação de um investidor, a serem satisfeitas na Internet, são-no num momento anterior à concretização de uma ordem de compra e/ou de venda. Por serem baseados em dados semanais, diluindo o histórico diário, os resultados obtidos nas relações entre os indicadores Google e o volume e a volatilidade podem até já reflectir este desfasamento entre pesquisar e decidir.

Impõe-se assim testar se a frequência passada das pesquisas no Google explica a evolução presente do mercado. Para concretizar essa análise sobre o volume e a volatilidade replica-se as regressões (13) e (14) com dados em painel, mas agora com todas as variáveis independentes desfasadas em um período. Por ser o retorno a ditar o resultado final de um investimento também se estima uma regressão com dados em painel para esta variável, na seguinte forma:

$$R_{x,t} = \alpha + \beta_0 R_{x,t-1} + \beta_1 RS_{t-1} + \beta_2 RV_{x,t-1} + \beta_3 RVS_{t-1} + \beta_4 \Delta V_{x,t-1} + \beta_5 GSVI_{x,t-1} + \epsilon_t \quad (15)$$

com $R_{x,t}$ a representar o retorno da empresa x na semana t , α uma constante, $R_{x,t-1}$ a variável

dependente desfasada em um período, RS_{t-1} o retorno do índice Stoxx 50 durante a semana $t-1$, $RV_{x,t-1}$ a volatilidade realizada sobre a empresa x na semana $t-1$, RVS_{t-1} a cotação do índice VSTOXX na semana $t-1$, $\Delta V_{x,t-1}$ a variação do volume da acção x na semana $t-1$ e $GSVI_{x,t-1}$ a frequência de consultas no Google sobre a empresa x na semana $t-1$. O termo ϵ_t designa um processo de ruído branco. São estimadas quatro versões para cada uma das regressões: a primeira sem variáveis Google e as restantes com GSVI, lnGSVI e JSVI, respectivamente.

Pela análise preditiva descobre-se que a procura de informação no Google sobre uma cotada influencia as variações do volume, da volatilidade e também do retorno registados na semana seguinte. Ou seja, que a frequência passada das pesquisas poderá servir de previsor sobre a evolução da empresa no mercado. No caso do volume, variável onde é mais nítida a associação com a notoriedade de uma cotada, observa-se nos resultados do Quadro 6.1. que o poder explicativo dos desfasamentos das variáveis Google mais do que duplica em relação aos modelos contemporâneos. Já os valores dos coeficientes das variáveis Google na análise preditiva sobre a volatilidade são praticamente idênticos aos verificados na estimação contemporânea (Quadro 6.2.).

Em relação à rendibilidade, verifica-se pelos resultados expostos no Quadro 6.3. que um aumento (diminuição) das consultas concretizadas no Google sobre uma empresa contribui para uma depreciação (apreciação) do valor da acção na semana seguinte, evidência que contraria a lógica da IRH e as conclusões do trabalho de Da et al. (2011). Todavia, os resultados obtidos por estes autores incidem sobre empresas negligenciadas que destoam desta amostra.

7. ASSIMETRIA

Comprovado o efeito que a atenção captada pelo Google tem sobre as variáveis financeiras, importa avaliar o grau de assimetria dessa relação. Intuitivamente, percebe-se que a atenção funciona como uma medida indirecta da actividade do investidor: hibernação significa inacção e atenção ilustra, no mínimo, maior disposição para agir. Ainda assim não se diferenciou até agora o efeito do Google numa e noutra situação.

As análises contemporânea e preditiva validaram a hipótese de que a frequência de pesquisas no Google, ao reflectir o grau de atenção do investidor, comporta poder explicativo sobre a evolução de uma empresa no mercado. No entanto, nenhum dos modelos permitiu avaliar o nível de assimetria dessa relação, isto é, se esse poder é mais robusto durante os picos ou durante os fossos de atenção ou se, pelo contrário, responde de forma relativamente análoga em ambas as situações. Por se desconhecer com exactidão os despertadores da atenção do investidor, nomeadamente o peso das boas e das más notícias, usa-se o volume absoluto para aferir se o poder explicativo do Google é assimétrico num modelo definido por:

$$\begin{aligned} \text{absV}_{x,t} = & \alpha + \beta_0 \text{absV}_{x,t-1} + \beta_1 \text{absr}_{x,t} \\ & + \beta_2 \text{RV}_{x,t} + \beta_3 \text{RS}_t + \beta_4 \text{RVS}_t + \beta_5 \text{GSVIP}_{x,t} \\ & + \beta_6 \text{GSVIN}_{x,t} + \epsilon_t \end{aligned} \quad (16)$$

com $\text{absV}_{x,t}$ a representar as variações em módulo do número de acções da empresa x negociadas na semana t e $\text{absr}_{x,t}$ o retorno absoluto da cotada x durante a semana t . Os regressores $\text{GSVIP}_{x,t}$ e $\text{GSVIN}_{x,t}$ são as variáveis independentes de maior interesse e designam, respectivamente, um aumento e uma diminuição na

atenção do investidor em relação à empresa x captados a partir da frequência de pesquisas na semana t . Toda a diferença entre $\text{GSVIP}_{x,t}$ e $\text{GSVIN}_{x,t}$ significa portanto assimetria.

Usando a volatilidade realizada como variável dependente, o modelo torna-se no seguinte:

$$\begin{aligned} \text{RV}_{x,t} = & \alpha + \beta_0 \text{RV}_{x,t-1} + \beta_1 \beta_1 \text{absr}_{x,t} + \beta_2 \text{RS}_t \\ & + \beta_3 \text{RVS}_t + \beta_4 \text{GSVIP}_{x,t} + \beta_5 \text{GSVIN}_{x,t} + \epsilon_t \end{aligned} \quad (17)$$

Para avaliar a existência de assimetria na previsibilidade recalcula-se as regressões (16) e (17) mas agora com todas as variáveis independentes desfasadas em um período. Adicionalmente, recupera-se o modelo (15) para a rendibilidade, substituindo apenas o último regressor pelas variáveis $\text{GSVIP}_{x,t-1}$ e $\text{GSVIN}_{x,t-1}$, ficando:

$$\begin{aligned} \text{R}_{x,t} = & \alpha + \beta_0 \text{R}_{x,t-1} + \beta_1 \text{RS}_{t-1} + \beta_2 \text{RV}_{x,t-1} \\ & + \beta_3 \text{RVS}_{t-1} + \beta_4 \Delta V_{x,t-1} + \beta_5 \text{GSVIP}_{x,t-1} \\ & + \beta_6 \text{GSVIN}_{x,t-1} + \epsilon_t \end{aligned} \quad (18)$$

Todos os modelos incluem um processo de ruído branco, ϵ_t . São estimadas quatro versões para cada um dos modelos: na primeira incluem-se os indicadores GSVIp e GSVIn, na segunda o logaritmo natural dessas duas variáveis, na terceira utiliza-se o JSVIP e o JSVIN, e na última versão os seus logaritmos naturais.

A análise de assimetria prova que o efeito dos indicadores Google sobre as variáveis financeiras não é linear. Observa-se nos resultados do Quadro 7.1. que a pressão negativa sobre o volume associada a um decréscimo na frequência de pesquisas é muito mais intensa do que no processo contrário. Em linha com resultados anteriores, o Quadro 7.2. reitera que o poder explicativo do Google sobre a volatilidade é

pouco significativo tanto nos picos como nos fossos de atenção.

Já o modelo estimado para o retorno, exposto no Quadro 7.3., sugere que um aumento da procura de informação no Google tem um efeito negativo sobre o preço da acção mas que o inverso não se verifica. Para tal poderá estar a contribuir o período alvo de análise - que inclui as depressões originadas pela crise do *subprime* e pelo início da crise europeia das dívidas soberanas -, e também a amostra escolhida, constituída em um terço por cotadas financeiras, um sector tipicamente cíclico.

8. PESQUISAR PARA DECIDIR

Uma questão que tem ocupado a literatura é a de saber quem está por detrás dos indicadores de atenção do Google, com Da et al. (2011) e Bank et al. (2011) a alegar que a frequência de pesquisas capta sobretudo a atenção dos investidores particulares, até porque: “Intuitively, people who search financial information related to a stock in Google are more likely to be individual or retail investors since institutional investors have access to more sophisticated information services such as Reuters or Bloomberg terminals” (Da et al., 2011). É por isso especialmente relevante se ter encontrado evidência de que as variáveis Google influenciam positivamente o volume. Isto porque num índice de referência como o Stoxx 50 não é crível que os investidores particulares possuam poder nem dimensão para impor variações no volume. Daí não se pode contudo validar o Google como descritor da atenção dos institucionais *a fortiori*; mas deve-se pelo menos inferir que o fluxo de informação formador do GSVI é relevante para o mercado e não é portanto alheio aos investidores mais sofisticados, nomeadamente em semanas de maior agitação.

Paralelamente, e embora seja impossível decompor a variável GSVI consoante o tipo de investidor que a origina, detecta-se vários sinais de que a frequência de pesquisas capta sobretudo a atenção dos investidores particulares. Desde logo na evidência de que o passado recente das pesquisas é melhor descritor que o seu presente no que ao volume diz respeito, pela força dos coeficientes das regressões do modelo preditivo e pela maior robustez dos R². É que se é duvidoso que os institucionais utilizem o Google como fonte privilegiada de informação, é menos crível ainda que, a fazê-lo, demorassem uma semana a repercutir a informação em ajustes de carteira, sobretudo num índice de referência como o Stoxx 50.

No desenho da microestrutura do mercado a literatura aponta três fontes primordiais de liquidez: diminuição dos custos de transacção (Stoll e Whaley, 1983), minoração do risco de inventário dos *market makers* (Garman, 1978) e redução dos custos de assimetria de informação (Glosten e Milgrom, 1985). Não se compreendendo como poderia o Google estar relacionado com os dois primeiros factores atribui-se a sua influência indirecta sobre a liquidez ao minguar dos custos de assimetria de informação, tal como fez Bank et al. (2011).

O domínio dos institucionais face aos investidores menos sofisticados também poderá justificar o fraco poder explicativo das variáveis Google sobre a volatilidade. De facto, olhando para as fontes clássicas de volatilidade que têm sido consensualizadas na literatura¹⁰ é difícil perceber como poderia o comportamento dos investidores particulares explicar de forma significativa a variabilidade dos preços num índice como o Stoxx 50. Ainda assim, a relação positiva entre a frequência de pesquisas e a volatilidade vai no sentido de que “anything that changes

10- Entre as fontes clássicas de volatilidade constam: risco político (Mei, 1999), liquidez (Domowitz, Glen e Madhavan, 2002) e vencimento de instrumentos derivados (Feinstein e Goetzmann, 1988).

the amount or character of noise trading will change the volatility of price” (Black, 1996). Os resultados sugerem portanto que a entrada do retalho no mercado acrescenta-lhe volatilidade, tal como concluíram Dimpfl et al. (2012).

Na base da relação negativa observada entre notoriedade e retorno, que contraria a lógica da IRH e as conclusões de Da et al. (2011), também poderá estar esse ofuscamento imposto pela influência dos institucionais sobre os particulares, que serão os principais condutores do GSVI. De qualquer forma a hipótese lançada por Barber et al. (2008) neste campo é insuficiente. Os autores argumentam que quanto maior (menor) a notoriedade de uma cotada, maior (menor) será a probabilidade de entrar no radar de atenção de cada investidor e assim integrar a sua carteira, o que, a verificar-se, gerará pressão compradora. Ignoram contudo que há notoriedade positiva e notoriedade negativa e que ambas são passíveis de influenciar a evolução de uma acção. Os casos da British Petroleum, Lehman e Dubai World expostos na Introdução são três exemplos disso mesmo.

9. CONCLUSÃO

Na última década a Internet tornou-se numa fonte privilegiada de procura de informação para consumidores e num meio preferencial de comunicação para empresas, Governos e reguladores. Nesse fluxo de informação os motores de busca como o Google cumprem uma função fundamental, orientando cada utilizador para o conteúdo pretendido com base no termo pesquisado. O volume dessas consultas pode por isso reflectir o grau de atenção de um investidor em relação a uma cotada e assim afectar o seu valor de mercado à luz da IRH (Merton, 1987). Sob essa premissa analisa-se o efeito da frequência de pesquisas no Google sobre 36 empresas do

Stoxx 50 no volume, volatilidade e retorno de cada cotada entre Janeiro de 2004 e Junho de

Além do carácter inovador da amostra – o Stoxx 50 nunca fora alvo de análise semelhante –, a investigação produziu resultados até agora desconhecidos na literatura, sobretudo ao nível da não linearidade na relação entre a procura de informação na Internet e o comportamento de uma cotada.

Nos modelos econométricos encontra-se evidência de que um aumento da procura de informação no Google sobre uma empresa tem um efeito contemporâneo positivo sobre o volume e a volatilidade. O efeito mantém-se nos modelos preditivos onde também se demonstra que um aumento da frequência de pesquisas no Google está associado a menor rendibilidade na semana seguinte. Infere-se contudo em todos os casos que se trata de uma relação não linear. Um fraquejar dos níveis de atenção não conduz a menor volatilidade nem a maior rendibilidade e tem um impacte negativo muito mais intenso sobre a liquidez que o efeito contrário. Além disso, no caso do volume a resposta na semana em que se realizam as consultas no Google é mais fraca do que o efeito registado na semana seguinte, um sinal da presença de investidores menos sofisticados.

Ao longo da investigação detectaram-se repetidas evidências de que por detrás das pesquisas concretizadas no Google estão sobretudo investidores particulares, tal como aclamaram Da et al. (2011) e Bank et al. (2011). Nesse quadro atribui-se o efeito positivo sobre o volume à redução de custos de assimetria de informação, o impacte positivo sobre a volatilidade ao *noise trading* e a resposta negativa da rendibilidade, que destoa dos trabalhos de Barber et al. (2008) e de Da et al. (2011), a um ofuscamento dos

dos institucionais num índice de referência como o Stoxx 50.

As conclusões supracitadas contribuem para várias frentes da investigação em Economia e Finanças: revelam que a atenção, enquanto capacidade cognitiva, é um recurso escasso e uma condição para decidir; demonstram que o grau de atenção do investidor sobre uma cotada influencia a sua liquidez, volatilidade e rendibilidade de forma assimétrica; sustentam que os investidores particulares actuam em desvantagem face aos institucionais; contrariam os modelos que reclamam a incorporação imediata no mercado de toda a informação disponível; e

sinalizam a potencialidade da Internet como reflexo do sentimento dos consumidores e da atenção dos investidores.

Adicionalmente, este trabalho providencia novos dados sobre o comportamento de pequenos investidores e de como a recolha de informação na Internet é relevante na sua tomada de decisão de investimento. As conclusões da investigação são por isso relevantes para os reguladores com deveres de supervisão da informação das cotadas e da protecção de pequenos investidores e sugerem que têm de estar também atentos à informação divulgada na Internet, pois esta é relevante para a tomada de decisão do investidor.



REFERÊNCIAS BIBLIOGRÁFICAS

- Antweiler, Werner e Murray Z. Frank (2004), “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards”, *The Journal of Finance*, 59 (3), pp.1259-1294.
- Arbel, Avner e Paul Strebler (1983), “Pay attention to neglected firms”, *The Journal of Portfolio Management*, 9 (2), pp.37-42.
- Askatas, Nikolaos e Klaus F. Zimmermann (2009), “Google Econometrics and Unemployment Forecasting”, *Applied Economics Quarterly*, 55 (2), pp.107-120.
- Bank, Matthias, Martin Larch e Georg Peter (2011), “Google search volume and its influence on liquidity and return of German stocks”, *Financial Markets and Portfolio Management*, 25 (3), pp.239-264.
- Barber, Brad e Terrace Odean (2008), “All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors”, *The Review of Financial Studies*, 21 (2), pp.785–818.
- Black, Fischer (1986), “Noise”, *The Journal of Finance*, 41(3), pp.529-543.
- Chen, Sissy (2011), *Google Search Volume: Influence and Indication for the Dutch Stock Market*, Dissertação de Licenciatura em Econometria e Investigação Operacional, Roterdão, Erasmus University Rotterdam.
- Choi, Hyunyoung e Hal Varian (2009), “Predicting the Present with Google Trends”, Working paper, Google Inc.
- Corwin, Shane A. e Jay F. Coughenour (2008), “Limited Attention and the Allocation of Effort in Securities Trading”, *The Journal of Finance*, 63 (6), pp.3031-3036.
- Da, Zhi, Joseph Engelberg e Pengjie Gao (2011), “In Search of Attention”, *The Journal of Finance*, 45(5), pp.1461-1499.
- Dimpfl, Thomas e Stephan Jank (2012), “Can internet search queries help to predict stock market volatility?”, Working paper, University of Tübingen.
- Domowitz, Ian, Jack Glen e Ananth Madhavan (2002), “Liquidity, Volatility and Equity Trading Costs Across Countries and Over Time”, *International Finance*, 4 (2), pp.221-255.
- Drake, Michael S., Darren T. Roulstone e Jacob R. Thornock (2012), “Investor Information Demand: Evidence from Google Searches around Earnings Announcements”, *Journal of Accounting Research*, 50 (4), pp.1001-1040.
- Fang, Lily e Joel Peress (2009), “Media coverage and the cross-section of stock returns”, *Journal of Financial and Quantitative Analysis*, 64 (5), pp. 2023–2052.
- Feinstein, Steven P. e William N. Goetzmann (1988), “The effect of the triple witching hour on stock market volatility”, *Federal Reserve Bank of Atlanta Economic Review*, 73 (5), pp.2-18.
- Frieder, Laura e Avanidhar Subrahmanyam (2005), “Brand perceptions and the market for common stock”, *Journal of Financial and Quantitative Analysis*, 40 (1), pp.57–85.
- Garman, Mark B. (1976), “Market microstructure”, *Journal of Financial Economics*, 3 (1), pp. 257-275.
- Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski e Larry Brilliant (2009), “Detecting influenza epidemics using search engine query data”, *Nature*, 457 (7232), pp.1012-1015.

REFERÊNCIAS BIBLIOGRÁFICAS

- Glosten, Lawrence R. e Paul R. Milgrom (1985), “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders”, *Journal of Financial Economics*, 14 (1), pp. 71-100.
- Goel, Sharad, Jake M. Hofman, Sébastien Lahaie, David M. Pennock e Duncan J. Watts (2010), “Predicting consumer behaviour with Web search”, Yahoo! Research., *Proceedings of the National Academy of Sciences*, 7 (41), pp.17486-17490.
- Mei, Jianping (1999), “Political Risk, Financial Crisis, and Market Volatility”, Working paper, New York University.
- Merton, Robert C. (1987), “A Simple Model of Capital Market Equilibrium with Incomplete Information”, *The Journal of Finance*, 42 (3), pp.483-510.
- Ross, Stephen A. (1989), “Information and Volatility: The No-Arbitrage Martingale Approach to Timing and Resolution Irrelevancy”, *The Journal of Finance*, 44 (1), pp.1-17.
- Rubin, Amir e Eran Rubin (2010), “Informed investors and the Internet”, *Journal of Business Finance and Accounting*, 37 (7/8), pp.841-865.
- Smith, Geoffrey Peter (2012), “Google Internet search activity and volatility prediction in the market of foreign currency”, *Finance Research Letters*, 9 (2), pp.103-110.
- Sperber, Dan e Deirdre Wilson (2001), *Revelância: comunicação e cognição*, Lisboa, Fundação Calouste Gulbenkian.
- Stoll, Hans R. e Robert E. Whaley (1983), “Transaction Costs and The Small Firm Effect”, *Journal of Financial Economics*, 12 (1), pp.57-59.
- Vlastakis, Nikolaos e Raphael N. Markellos (2012), “Information Demand and Stock Market Volatility”, *Journal of Banking & Finance*, 36 (6), pp.1808-1821.

ANEXOS

Quadro 3.1. Amostra

Descrição das empresas analisadas pelo nome, país sede, sector de actividade e termo de pesquisa considerado.

	Empresa	País sede	Sector	Termo pesquisado
1	AIR LIQUIDE SA	França	Materiais básicos	"Air Liquide"
2	ALLIANZ SE-REG	Alemanha	Financeiro	"Allianz"
3	AXA SA	França	Financeiro	"AXA"
4	BANCO SANTANDER SA	Espanha	Financeiro	"Santander"
5	BASF SE	Alemanha	Materiais básicos	"Bast"
6	BAYER AG-REG	Alemanha	Materiais básicos	"Bayer"
7	BANCO BILBAO VIZCAYA ARGENTARIA	Espanha	Financeiro	"BBVA"
8	BNP PARIBAS	França	Financeiro	"BNP Paribas"
9	CARREFOUR SA	França	Bens de consumo não cíclicos	"Carrefour"
10	DAIMLER AG-REGISTERED SHARES	Alemanha	Bens de consumo cíclicos	"Daimler"
11	DANONE	França	Bens de consumo não cíclicos	"Danone"
12	DEUTSCHE BANK AG-REGISTERED	Alemanha	Financeiro	"Deutsche Bank"
13	DEUTSCHE TELEKOM AG-REG	Alemanha	Telecomunicações	"Deutsche Telekom"
14	E.ON AG	Alemanha	Financeiro	"EON"
15	ENEL SPA	Itália	Utilities	"Enel"
16	ENI SPA	Itália	Energia	"ENI"
17	FRANCE TELECOM SA	França	Telecomunicações	"France Telecom"
18	ASSICURAZIONI GENERALI	Itália	Financeiro	"Generali"
19	IBERDROLA SA	Espanha	Utilities	"Iberdrola"
20	ING GROEP NV-CVA	Holanda	Financeiro	"ING"
21	KONINKLIJKE PHILIPS ELECTRON	Holanda	Industrial	"Philips"
22	L'OREAL	França	Bens de consumo não cíclicos	"L'Oreal"
23	LVMH MOET HENNESSY LOUIS VUITTON	França	Diversified	"LVMH"
24	NOKIA OYJ	Finlândia	Telecomunicações	"Nokia"
25	REPSOL YPF SA	Espanha	Energia	"Repsol"
26	RWE AG	Alemanha	Utilities	"RWE"
27	COMPAGNIE DE SAINT-GOBAIN	França	Industrial	"Saint-Gobain"
28	SANOFI	França	Bens de consumo não cíclicos	"Sanofi"
29	SIEMENS AG-REG	Alemanha	Industrial	"Siemens"
30	SOCIETE GENERALE	França	Financeiro	"Societe Generale"
31	TELECOM ITALIA SPA	Itália	Telecomunicações	"Telecom Italia"
32	TELEFONICA SA	Espanha	Telecomunicações	"Telefonica"
33	TOTAL SA	França	Energia	"Total"
34	UNICREDIT SPA	Itália	Financeiro	"Unicredit"
35	UNILEVER NV-CVA	Holanda	Bens de consumo não cíclicos	"Unilever"
36	VIVENDI	França	Telecomunicações	"Vivendi"

ANEXOS

Quadro 4.1. Modelo contemporâneo para o volume

*Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre as variações logarítmicas do volume, a variável dependente. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$\Delta V_{x,t-1}$	-0.24154*** 0.000	-0.24181*** 0.000	-0.24175*** 0.000	-0.24919*** 0.000
$R_{x,t}$	0.20568 0.301	0.21143 0.286	0.21081 0.287	0.19986 0.314
$RV_{x,t}$	5.96767** 0.032	5.80421** 0.035	5.77384** 0.036	5.79876** 0.034
RS_t	-0.60097 0.431	-0.61173 0.424	-0.61267 0.423	-0.62207 0.424
RVS_t	0.40167* 0.073	0.40127* 0.073	0.40048* 0.074	0.39666* 0.082
$GSVI_{x,t}$		0.03340*** 0.002		
$\ln GSVI_{x,t}$			0.03463*** 0.003	
$JSVI_{x,t}$				0.25060*** 0.000
Constante	-0.01126 0.200	-0.04442*** 0.002	-0.00927 0.285	-0.01131 0.180
Observações	14,004	14,004	14,004	13,930
Empresas	36	36	36	36
R2	0.088	0.089	0.089	0.098

ANEXOS

Quadro 4.2. Modelo *piecewise* para o volume

Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre as variações logarítmicas do volume, a variável dependente. Divide-se as cotadas em quartis consoante a sua notoriedade.

O primeiro quartil reúne as cotadas de maior visibilidade (colunas 1, 5, 9 e 13) e o último as de menor (colunas 4, 8, 12 e 16). Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.21394*** 0.000	-0.26640*** 0.000	-0.26012*** 0.000	-0.23801*** 0.000	-0.21435*** 0.000	-0.26833*** 0.000	-0.26054*** 0.000	-0.23785*** 0.000
$R_{x,t}$	1.71357*** 0.000	2.56690*** 0.000	2.85825*** 0.000	3.04232*** 0.000	1.71691*** 0.000	2.58624*** 0.000	2.79937*** 0.000	3.00220*** 0.000
$RV_{x,t}$	-0.31313 0.836	-4.10514** 0.042	-2.31549 0.592	-3.74721 0.113	-0.30695 0.839	-4.26131** 0.036	-2.81998 0.517	-4.05009* 0.086
RS_t	0.41112 0.638	0.06587 0.937	-0.18305 0.824	-0.26290 0.732	0.41003 0.640	0.05123 0.952	-0.20684 0.804	-0.25971 0.734
RVS_t	0.48096** 0.037	0.45909* 0.051	0.34814 0.139	0.33166 0.143	0.48234** 0.037	0.45877* 0.054	0.33997 0.152	0.33255 0.140
$GSVI_{x,t}$					0.03561 0.144	0.23186** 0.028	0.26316** 0.031	0.08081 0.167
$\ln GSVI_{x,t}$								
$JSVI_{x,t}$								
Constante	-0.05189*** 0.000	-0.06584*** 0.000	-0.07223*** 0.000	-0.08134*** 0.000	-0.10033*** 0.007	-0.31486*** 0.006	-0.30443*** 0.004	-0.13352*** 0.001
Observações	3,621	3,500	3,508	3,375	3,621	3,500	3,508	3,375
Empresas	36	36	36	36	36	36	36	36
R2	0.096	0.133	0.125	0.116	0.096	0.137	0.130	0.117

ANEXOS

Variáveis	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
$\Delta V_{x,t-1}$	-0.21429*** 0.000	-0.26841*** 0.000	-0.26061*** 0.000	-0.23779*** 0.000	-0.21972*** 0.000	-0.27283*** 0.000	-0.26937*** 0.000	-0.24578*** 0.000
$R_{x,t}$	1.71483*** 0.000	2.58670*** 0.000	2.78934*** 0.000	3.00795*** 0.000	1.72432*** 0.000	2.54536*** 0.000	2.81373*** 0.000	3.00784*** 0.000
$RV_{x,t}$	-0.31464 0.835	-4.30319** 0.034	-2.88271 0.509	-4.01017* 0.089	-0.37520 0.799	-4.14760** 0.032	-1.94486 0.648	-3.65901 0.119
RS_t	0.40854 0.641	0.03757 0.965	-0.23029 0.784	-0.26601 0.728	0.40463 0.647	0.04708 0.956	-0.27988 0.745	-0.27219 0.729
RVS_t	0.48187** 0.037	0.45453* 0.058	0.33120 0.166	0.33120 0.143	0.48095** 0.038	0.44954* 0.060	0.31795 0.195	0.34573 0.138
$GSVI_{x,t}$								
$\ln GSVI_{x,t}$	0.05304 0.170	0.27838** 0.018	0.26125** 0.015	0.03814 0.259				
$JSVI_{x,t}$					0.12867*** 0.010	0.26360*** 0.000	0.35493*** 0.001	0.56729*** 0.002
Constante	-0.06719*** 0.000	-0.08432*** 0.000	-0.03758* 0.058	-0.06308*** 0.004	-0.05798*** 0.000	-0.06256*** 0.000	-0.06724*** 0.000	-0.06886*** 0.000
Observações	3,621	3,500	3,508	3,375	3,603	3,482	3,488	3,357
Empresas	36	36	36	36	36	36	36	36
R2	0.096	0.138	0.131	0.117	0.100	0.142	0.139	0.138

ANEXOS

Quadro 5.1. Modelo contemporâneo para a volatilidade

*Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre a volatilidade realizada, a variável dependente. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$RV_{x,t-1}$	0.60130*** 0.000	0.59877*** 0.000	0.59827*** 0.000	0.60145*** 0.000
$R_{x,t}$	-0.01005 0.155	-0.00993 0.158	-0.00994 0.158	-0.01014 0.153
$RS_{x,t}$	0.00711 0.386	0.00689 0.399	0.00686 0.400	0.00717 0.380
RVS_t	0.00188 0.280	0.00186 0.282	0.00185 0.285	0.00184 0.285
$GSVI_{x,t}$		0.00057*** 0.001		
$\ln GSVI_{x,t}$			0.00059*** 0.001	
$JSVI_{x,t}$				0.00106** 0.011
Constante	0.00078*** 0.000	0.00021** 0.037	0.00081*** 0.000	0.00078*** 0.000
Observações	14,004	14,004	14,004	13,930
Empresas	36	36	36	36
R2	0.374	0.376	0.376	0.375

ANEXOS

Quadro 5.2. Modelo *piecewise* para a volatilidade

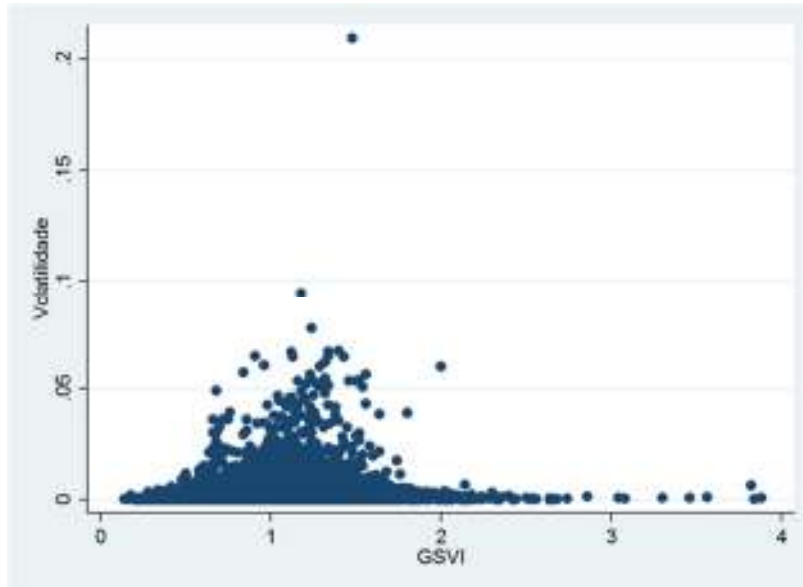
*Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre a volatilidade realizada, a variável dependente. Divide-se as cotadas em quartis consoante a sua notoriedade. O primeiro quartil reúne as cotadas de maior visibilidade (colunas 1, 5, 9 e 13) e o último as de menor (colunas 4, 8, 12 e 16). Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$RV_{x,t-1}$	0.68521*** 0.000	0.52508*** 0.000	0.49457*** 0.000	0.59224*** 0.000	0.68514*** 0.000	0.52505*** 0.000	0.49099*** 0.000	0.58694*** 0.000
$R_{x,t}$	-0.02349 0.111	-0.00840 0.211	-0.00068 0.919	0.00337** 0.046	-0.02351 0.111	-0.00841 0.211	-0.00066 0.921	0.00340** 0.042
$RS_{x,t}$	0.03277 0.222	0.00097 0.946	-0.00402 0.621	-0.00905 0.159	0.03279 0.221	0.00095 0.947	-0.00414 0.607	-0.00898 0.157
RVS_t	0.00422 0.254	0.00017 0.917	0.00180 0.158	0.00085 0.454	0.00422 0.254	0.00018 0.915	0.00175 0.162	0.00085 0.449
$GSVI_{x,t}$					-0.00009 0.647	0.00069 0.231	0.00144* 0.078	0.00104*** 0.006
$\ln GSVI_{x,t}$								
$JSVI_{x,t}$								
Constante	0.00086*** 0.000	0.00090*** 0.000	0.00077*** 0.000	0.00064*** 0.000	0.00098*** 0.004	0.00016 0.762	-0.00051 0.426	-0.00004 0.862
Observações	3,621	3,500	3,508	3,375	3,621	3,500	3,508	3,375
Empresas	36	36	36	36	36	36	36	36
R2	0.384	0.435	0.281	0.343	0.384	0.435	0.284	0.346

ANEXOS

Figura 5.1. Relação entre volatilidade e GSVI

Demonstração gráfica da relação entre a volatilidade realizada e a frequência normal de pesquisas no Google (GSVI) em toda a amostra.



ANEXOS

Quadro 6.1. Modelo preditivo para o volume

Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre as variações logarítmicas do volume, a variável dependente, na semana seguinte.

*Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$\Delta V_{x,t-1}$	-0.26683*** 0.000	-0.26870*** 0.000	-0.26882*** 0.000	-0.29294*** 0.000
$R_{x,t-1}$	-0.41768** 0.010	-0.40530** 0.012	-0.40465** 0.013	-0.42260*** 0.006
$RV_{x,t-1}$	-2.19704 0.112	-2.63912* 0.063	-2.75673* 0.053	-2.42596* 0.061
RS_{t-1}	2.12227*** 0.000	2.09924*** 0.000	2.09242*** 0.000	2.11408*** 0.000
RVS_{t-1}	0.90059*** 0.000	0.90138*** 0.000	0.89883*** 0.000	0.90110*** 0.000
$GSVI_{x,t-1}$		0.08405*** 0.000		
$\ln GSVI_{x,t-1}$			0.08737*** 0.000	
$JSVI_{x,t-1}$				0.62419*** 0.000
Constante	0.00315* 0.068	-0.08130*** 0.000	0.00487*** 0.008	0.00538*** 0.001
Observações	14,004	14,003	14,003	13,894
R2	0.094	0.098	0.099	0.142

ANEXOS

Quadro 6.2. Modelo preditivo para a volatilidade

*Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre a volatilidade realizada, a variável dependente, na semana seguinte. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$\Delta V_{x,t-1}$	-0.00037*** 0.000	-0.00038*** 0.000	-0.00038*** 0.000	-0.00040*** 0.000
$R_{x,t-1}$	-0.01078* 0.050	-0.01070* 0.052	-0.01069* 0.052	-0.01082** 0.050
$RV_{x,t-1}$	0.56709*** 0.000	0.56392*** 0.000	0.56275*** 0.000	0.56644*** 0.000
$RS_{x,t-1}$	-0.01990*** 0.000	-0.02006*** 0.000	-0.02013*** 0.000	-0.01994*** 0.000
RVS_{t-1}	-0.00118* 0.091	-0.00118* 0.090	-0.00120* 0.086	-0.00121* 0.087
$GSVI_{x,t-1}$		0.00060** 0.045		
$\ln GSVI_{x,t-1}$			0.00067** 0.017	
$JSVI_{x,t-1}$				0.00107*** 0.000
Constante	0.00052***	-0.00008	0.00053***	0.00053***
Observações	0.000	0.781	0.000	0.000
Observações	14,004	14,003	14,003	13,894
R2	0.406	0.408	0.408	0.407

ANEXOS

Quadro 6.3. Modelo preditivo para a rentabilidade

*Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre a rentabilidade, a variável dependente, na semana seguinte. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$\Delta V_{x,t-1}$	-0.00197*** 0.004	-0.00194*** 0.005	-0.00193*** 0.005	-0.00233*** 0.001
$R_{x,t-1}$	-0.07210*** 0.002	-0.07229*** 0.002	-0.07233*** 0.002	-0.07167*** 0.002
$RV_{x,t-1}$	-0.93721*** 0.002	-0.92997*** 0.002	-0.92667*** 0.002	-0.92965*** 0.002
$RS_{x,t-1}$	-0.11332*** 0.008	-0.11295*** 0.008	-0.11277*** 0.008	-0.10896** 0.010
RVS_{t-1}	-0.01904*** 0.002	-0.01905*** 0.002	-0.01900*** 0.002	-0.01703*** 0.007
$GSVI_{x,t-1}$		-0.00137 0.132		
$\ln GSVI_{x,t-1}$			-0.00164** 0.049	
$JSVI_{x,t-1}$				-0.00589** 0.013
Constante	0.00291***	0.00429***	0.00288***	0.00294***
Observações	0.000	0.000	0.000	0.000
Observações	14,004	14,003	14,003	13,894
R2	0.021	0.021	0.021	0.021

ANEXOS

Quadro 7.1. Modelo preditivo para o volume absoluto com assimetria

*Regressão com dados em painel para avaliar efeitos não lineares da frequência de pesquisas no Google sobre o volume absoluto, a variável dependente, na semana seguinte. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$absV_{x,t-1}$	0.21230*** 0.000	0.21233*** 0.000	0.20409*** 0.000	0.19985*** 0.000
$absr_{x,t-1}$	-0.37142*** 0.000	-0.37218*** 0.000	-0.36244*** 0.000	-0.34420*** 0.000
$RV_{x,t-1}$	2.69477*** 0.000	2.68234*** 0.001	2.44059*** 0.003	2.53343*** 0.002
RS_{t-1}	0.93339*** 0.000	0.93268*** 0.000	1.00770*** 0.000	1.05401*** 0.000
RVS_{t-1}	0.23676*** 0.000	0.23621*** 0.000	0.26464*** 0.000	0.27689*** 0.000
$GSVIP_{x,t-1}$	0.03514* 0.073			
$GSVIN_{x,t-1}$	-0.08363*** 0.009			
$lnGSVIP_{x,t-1}$		0.03894 0.114		
$lnGSVIN_{x,t-1}$		-0.05615** 0.018		
$JGSVIP_{x,t-1}$			0.22218*** 0.000	
$JGSVIN_{x,t-1}$			-0.48554*** 0.000	
$lnJGSVIP_{x,t-1}$				0.36544*** 0.000
$lnJGSVIN_{x,t-1}$				-0.66222*** 0.000
Constante	0.23666*** 0.000	0.18984*** 0.000	0.16470*** 0.000	0.14813*** 0.000
Observações				
Empresas	14,004	14,004	14,004	14,004
R2	0.070	0.070	0.091	0.102

ANEXOS

7.2. Modelo preditivo para a volatilidade com assimetria

*Regressão com dados em painel para avaliar efeitos não lineares da frequência de pesquisas no Google sobre a volatilidade realizada, a variável dependente, na semana seguinte. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$\Delta V_{x,t-1}$	-0.00039*** 0.000	-0.00039*** 0.000	-0.00040*** 0.000	-0.00041*** 0.000
$R_{x,t-1}$	-0.01068* 0.053	-0.01068* 0.052	-0.01081** 0.050	-0.01079** 0.050
$RV_{x,t-1}$	0.56222*** 0.000	0.56214*** 0.000	0.56635*** 0.000	0.56668*** 0.000
RS_{t-1}	-0.02019*** 0.000	-0.02019*** 0.000	-0.01983*** 0.000	-0.01991*** 0.000
RVS_{t-1}	-0.00124* 0.073	-0.00123* 0.074	-0.00116* 0.092	-0.00119* 0.088
$GSVIp_{x,t-1}$	-0.00003 0.926			
$GSVIn_{x,t-1}$	0.00151*** 0.000			
$\ln GSVIp_{x,t-1}$		0.00012 0.825		
$\ln GSVIn_{x,t-1}$		0.00103*** 0.000		
$JGSVIp_{x,t-1}$			0.00168*** 0.003	
$JGSVIn_{x,t-1}$			0.00040* 0.075	
$\ln JGSVIp_{x,t-1}$				0.00182*** 0.001
$\ln JGSVIn_{x,t-1}$				0.00057*** 0.006
Constante	-0.00083 0.142	0.00061*** 0.000	0.00046*** 0.000	0.00046*** 0.000
Observações	14,004	14,004	14,004	14,004
R2	0.408	0.408	0.407	0.407

ANEXOS

Quadro 7.3. Modelo preditivo para a rentabilidade com assimetria

*Regressão com dados em painel para avaliar efeitos não lineares da frequência de pesquisas no Google sobre a rentabilidade, a variável dependente, na semana seguinte. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)
$\Delta V_{x,t-1}$	-0.00192*** 0.005	-0.00192*** 0.005	-0.00180*** 0.009	-0.00165** 0.017
$R_{x,t-1}$	-0.07239*** 0.002	-0.07231*** 0.002	-0.07193*** 0.002	-0.07198*** 0.001
$RV_{x,t-1}$	-0.91924*** 0.002	-0.92161*** 0.002	-0.93301*** 0.002	-0.93405*** 0.002
RS_{t-1}	-0.11215*** 0.008	-0.11232*** 0.008	-0.11377*** 0.007	-0.11349*** 0.007
RVS_{t-1}	-0.01868*** 0.003	-0.01873*** 0.003	-0.01916*** 0.002	-0.01908*** 0.002
$GSVIp_{x,t-1}$	0.00258 0.184			
$GSVIn_{x,t-1}$	-0.00713*** 0.004			
$lnGSVIp_{x,t-1}$		0.00290 0.262		
$lnGSVIn_{x,t-1}$		-0.00455*** 0.007		
$JGSVIp_{x,t-1}$			-0.00957** 0.037	
$JGSVIn_{x,t-1}$			-0.00208 0.544	
$lnJGSVIp_{x,t-1}$				-0.01532** 0.010
$lnJGSVIn_{x,t-1}$				-0.00360 0.298
Constante	0.00662*** 0.000	0.00224*** 0.000	0.00324*** 0.000	0.00345*** 0.000
Observações	14,004	14,004	14,004	14,004
R2	0.021	0.021	0.021	0.022