



Escola de Ciências Sociais e Humanas

Departamento de Economia Política

Pesquisar para decidir: O Google como barómetro
da atenção do investidor

Pedro Jorge Carvalho Latoeiro

Dissertação submetida como requisito parcial para obtenção do grau de

Mestre em Economia Monetária e Financeira

Orientadora:

Doutora Sofia Ramos, Professora Auxiliar com Agregação
ISCTE-Instituto Universitário de Lisboa

Outubro, 2012

**Pesquisar para decidir: O Google como
barómetro da atenção do investidor**
Pedro Latoeiro

**Outubro
2012**

À Mariana, por ter existido.
À Quitéria e ao Eduardo, por existirem.

Agradecimentos

Agradeço à Prof.^a Sofia Ramos pela inspiração constante, o apoio incansável e a orientação ímpar desde a primeira hora.

À Sandra Almeida Simões pelo *never give up*.

Ao Pedro Sousa Carvalho por tornar menos impossível o estatuto jornalista-estudante.

Ao Bruno Gaminha pela companhia em tantas jornadas académicas.

Resumo

A procura de informação na Internet tem sido validada enquanto indicador fiável da atenção dos investidores particulares com reflexos significativos na liquidez do mercado. Este estudo analisa se há alguma relação entre a atenção do investidor medida pelo volume de pesquisas no Google e um conjunto de acções do índice Euro Stoxx 50. Na análise de uma amostra de acções europeias encontra-se evidência de que essa relação não é linear. Elevados níveis de atenção estão associados a maior liquidez, mas o efeito de uma diminuição da procura de informação revela-se muito mais intenso.

Infere-se ainda que a frequência de pesquisas no Google contém poder preditivo na medida em que um aumento da procura de informação antecipa maior volume e volatilidade e menor rendibilidade. Mas os resultados sugerem uma vez mais tratar-se de uma relação não linear, dado que uma diminuição na frequência de consultas na Internet não prediz uma queda na volatilidade nem uma subida da rendibilidade.

Embora se detecte algum impacte na semana em que são realizadas as pesquisas, o efeito sobre a liquidez é superior na semana seguinte, o que constitui um indício da presença de investidores particulares.

Palavras-chave: atenção, investidores particulares, procura de informação, pesquisas na Internet.

JEL Classification: G12, G14.

Abstract

Internet demand information has been considered a good indicator of small investors' attention with impact on liquidity. This study analyses if there is relation between investors' attention measured by Google search queries and a set of stocks from Euro Stoxx 50 index. Analyzing a sample of European stocks we find evidence that this relation is nonlinear. An increase in Internet information demand is associated with higher liquidity, but the opposite effect appears to be much more intense.

We find also evidence of predictability ability of Internet demand for information. A large volume of information demand on the Google engine is indicative of increasing volatility and volume and decreasing returns. But similarly the evidence suggests a nonlinear relation. A lower level of Google search does not predict a future fall in volatility and an increase in returns.

Although liquidity changes on the same week of Google variations, the impact is higher on the following week (week +1), suggesting less sophisticated investors.

Keywords: attention, individual investor, information demand, web search.

JEL Classification: G12, G14.

Índice Geral

Introdução	1
1. Revisão da Literatura	5
1.1. A atenção como condição	5
1.2. Medindo a atenção	7
1.3. Pesquisar para decidir	9
2. Método e Variáveis	13
2.1. O indicador Google	13
2.2. Euro Stoxx 50	16
2.3. Variáveis financeiras	17
2.4. Variáveis Google	19
2.5. Estacionariedade e defasamentos	20
3. Portfólios	21
3.1. Resultados	22
4. Volume	24
4.1. Regressões	24
4.2. Resultados	25
5. Volatilidade	29
5.1. Regressões	29
5.2. Resultados	30
6. Previsibilidade	32
6.1. Regressões	32
6.2. Resultados	33
7. Interdependência	37
7.1. O modelo VAR	38
7.2. Resultados	38
8. Assimetria	40
8.1. Regressões	41
8.2. Resultados	42
9. Robustez	44
Conclusão	45
Referências bibliográficas	47
Anexos	53

Índice de Quadros

2.1. Amostra	53
2.2. Estatísticas descritivas da variável volume	54
2.3. Estatísticas descritivas da variável volatilidade realizada	55
2.4. Estatísticas descritivas da frequência de pesquisas no Google	56
2.5. Resultados do teste de raiz unitária Zivot e Andrews	57
3.1. Correlação entre os indicadores Google e as variáveis financeiras	59
3.2. Portfólios por GSVI	60
3.3. Portfólios por JSVI	63
4.1. Correlação entre volume e GSVI	64
4.2. Modelo contemporâneo para o volume	65
4.3. Modelo <i>piecewise</i> para o volume	66
5.1. Correlação entre volatilidade e GSVI	68
5.2. Modelo contemporâneo para a volatilidade	70
5.3. Modelo <i>piecewise</i> para a volatilidade	72
6.1. Modelo preditivo para o volume	74
6.2. Modelo preditivo para a volatilidade	75
6.3. Modelo preditivo para a rendibilidade	76
7.1. Resultados do teste de causalidade de Granger	77
8.1. Modelo contemporâneo para o volume com assimetria	83
8.2. Modelo preditivo para o volume com assimetria	84
8.3. Modelo contemporâneo para o volume absoluto com assimetria	85
8.4. Modelo preditivo para o volume absoluto com assimetria	86
8.5. Modelo contemporâneo para a volatilidade com assimetria	87
8.6. Modelo preditivo para a volatilidade com assimetria	88
8.7. Modelo preditivo para a rendibilidade com assimetria	89
9.1. Modelo preditivo para o volume com winsorização	90
9.2. Modelo preditivo para a volatilidade com winsorização	91
9.3. Modelo preditivo para a rendibilidade com winsorização	92
9.4. Modelo preditivo para o volume até Novembro de 2007	93
9.5. Modelo preditivo para o volume após Novembro de 2007	94
9.6. Modelo preditivo para a volatilidade até Novembro de 2007	95
9.7. Modelo preditivo para a volatilidade após Novembro de 2007	96
9.8. Modelo preditivo para a rendibilidade até Novembro de 2007	97
9.9. Modelo preditivo para a rendibilidade após Novembro de 2007	98
9.10. Modelo preditivo para o volume com <i>dummies</i>	99
9.11. Modelo preditivo para a volatilidade com <i>dummies</i>	100
9.12. Modelo preditivo para a rendibilidade com <i>dummies</i>	101

Índice de Figuras

3.1. Evolução da frequência de pesquisas no Google por carteira	58
3.2. Rendibilidade acumulada por portfólio	61
3.3. Rendibilidade acumulada de estratégia baseada no GSVI	62
5.1. Volatilidade	69
5.2. Relação entre volatilidade e GSVI	71
7.1. Função impulso resposta – Air Liquide	79
7.2. Função impulso resposta – ING	80
7.3. Função impulso resposta – Sanofi	81
8.1. Relação entre volume e GSVI	82

Glossário de siglas e acrónimos

DOW – Dow Jones Industrial Average

FIR – Função impulso resposta

GSVI – Google Search Volume Indicator

IPO – Initial Public Offering

IRH – Investor Recognition Hypothesis

JSVI – Jumps in Search Volume Indicator

NYSE – New York Stock Exchange

WSJ – The Wall Street Journal

VAR – Vector Auto-Regressivo

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 as decisões dos investidores é o tema central desta investigação. Os exemplos supracitados constituem um sinal de que as pesquisas na Internet parecem oscilar em concordância com o fluxo de informação relevante para o mercado, não sendo portanto registos aleatórios e espúrios desprovidos de contexto.

Essa concordância foi comprovada em investigações recentes que validam o motor de busca do Google enquanto barómetro fiável da atenção¹ do investidor. Da, Engelberg e Gao (2011) assumem que por detrás das pesquisas no Google estão sobretudo investidores particulares². E Bank, Larch e Peter (2011) associam um aumento da frequência de consultas na Internet sobre uma cotada a maior liquidez. O quadro teórico de ambos os *papers* assenta em dois pilares estruturantes: antes de decidir sobre um activo o investidor³ tem de, no

¹ A atenção tem sido amplamente estudada pela Psicologia. A definição de William James coaduna-se com o sentido que se pretende vincar: "attention is the taking possession of the mind, in clear and vivid form, of one out of what may seem several simultaneously possible objects or trains of thoughts. It implies withdrawal from some things in order to deal effectively with others" (James, 1890: 403-404).

² "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: 11).

³ Referência ao investidor activo *per se*, em oposição ao investidor passivo, que entrega o seu capital à gestão de terceiros, demitindo-se da formação e gestão do portfolio, não tendo necessariamente consciência dos seus componentes.

mínimo, saber, ou estar consciente⁴, da existência do mesmo⁵; e o Google, ou outro motor de pesquisa, é uma fonte privilegiada de informação nessa tomada de consciência.

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. As principais novidades face à literatura são a focagem em interacções 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 entre o Google e outras variáveis que têm sido usadas para medir a atenção do investidor. Para tal constrói-se quatro carteiras a partir da frequência de pesquisas sobre as cotadas e analisa-se a sua evolução em termos de volume, volatilidade e rendibilidade. Logo aí encontra-se os primeiros sinais de assimetria com as empresas de maior notoriedade a revelarem menor rendibilidade e maior volatilidade. Essa não linearidade fica também patente nas regressões *piecewise* segmentadas com as mesmas quatro carteiras. O poder explicativo do Google 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 numa base contemporânea.

O segundo tópico alvo de análise é a previsibilidade, que comporta especial importância teórica por se relacionar com a eficiência do mercado no sentido de que os preços incorporam toda a informação disponível. Encontra-se evidência de que uma subida na frequência de consultas no Google antecipa menor rendibilidade e maior volatilidade e volume na semana seguinte, um indício de que as pesquisas são conduzidas sobretudo por investidores particulares. Mas também aqui se depara com sinais de não linearidade. Se um aumento da procura por informação no Google tem um efeito positivo sobre a volatilidade e um impacte negativo sobre o preço da acção, o inverso não se verifica. Já no caso do volume

⁴ Consciência como “um estado mental em que temos conhecimento da nossa própria existência e da existência daquilo que nos rodeia” (Damásio, 2010: 199) e no sentido de que “na ausência de consciência, a visão pessoal suspende-se, não temos conhecimento da nossa existência; e não sabemos que existe mais alguma coisa” (Damásio, 2010: 21).

⁵ Este é o argumento basilar da *investor recognition hypothesis* (IRH) formulada por Robert C. Merton: “an investor uses security k in construction his optimal portfolio only if the investor knows about security k” (Merton, 1987: 488).

a resposta negativa a uma diminuição da procura de informação revela-se muito mais intensa que o efeito contrário.

Por último analisa-se se pelo menos parte das pesquisas no Google não é gerada por informação já descontada no mercado, algo que a literatura não acautelou. Os testes de causalidade de Granger concretizados dão conta da existência de relações em ambos os sentidos – entre o indicador de pesquisas no Google e as variáveis financeiras mas também o contrário -, um sinal de que as consultas são sobretudo um impulso de investidores particulares.

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 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 também, 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.

Com cerca de 119 mil milhões de pesquisas concretizadas mensalmente em todo o mundo⁶, a singular, extensa e incomparável informação que vai engordando a base de dados de intenções⁷ do Google já começou a despertar o interesse de bancos centrais⁸, da banca de

⁶ Dados de Junho de 2012 da ComScore (<http://www.comscore.com>).

⁷ Para John Battelle “a busca está possivelmente a criar o mais duradouro, importante e significativo apetrecho cultural na história da humanidade, a base de dados de intenções, que é simplesmente isto: os resultados agregados de todas as listas de resultados até agora obtidas e todas as pistas seguidas em consequência disso”. É uma “base de dados de desejos, necessidades, vontades e preferências que podem ser descobertos, intimados, arquivados, seguidos e explorados para toda a espécie de fins (Battelle, 2005: 20).

⁸ Os departamentos de investigação do Banco de Inglaterra e do Banco de Itália rastream o volume de pesquisas sobre determinadas palavras para prever a evolução da taxa de desemprego. No Banco de Espanha seguiu-se método equivalente para antecipar em um mês fluxos de turistas britânicos. Também o Banco do Chile usou o Google para prever o comportamento das vendas automóveis. Nos EUA economistas da Reserva Federal (Fed) de Nova Iorque e de São Francisco utilizaram o Google para, respectivamente, prever subidas nos pedidos de refinanciamento de hipotecas e para avaliar a percepção do risco soberano de vários países.

investimento⁹ e também da comunidade científica. Contudo, na investigação em Finanças a utilização dessa informação é ainda muito escassa e no que a mercados financeiros diz respeito praticamente nula. A investigação que se segue é um contributo para inverter essa tendência. A estrutura é a seguinte: no primeiro capítulo procede-se à revisão da literatura; no segundo descreve-se o objecto de análise e a metodologia a aplicar; no terceiro analisa-se a evolução do volume, volatilidade e retorno em quatro carteiras criadas em função da frequência de pesquisas no Google; no quarto avalia-se a relação contemporânea entre a procura de informação na Internet e o volume de uma cotada; no quinto repete-se essa análise mas para a volatilidade; no sexto capítulo avalia-se o poder preditivo do Google para o comportamento de uma cotada; no sétimo estuda-se as interdependências lineares entre o Google e as variáveis financeiras; no oitavo estuda-se o grau de assimetria entre a procura de informação na Internet e a evolução do mercado; no nono verifica-se a robustez dos resultados obtidos; e no décimo conclui-se.

⁹ Brian Jacobsen, *chief portfolio strategist* no Wells Fargo, utiliza as pesquisas no Google para tentar antecipar a evolução do consumo privado.

1. Revisão da Literatura

Neste capítulo procede-se à revisão da literatura para definir o quadro teórico da análise e também para a contextualizar no estado actual da investigação neste tópico. Começa-se por visitar a IRH de Merton e o fenómeno *home bias* que lhe está associado. Enumera-se de seguida as variáveis (indirectas) que têm sido propostas para testar e comprovar a IRH. Para terminar percorre-se os trabalhos de investigação já publicados que utilizam a base de dados de intenções do Google como variável (directa) para estudar temas económicos e financeiros.

1.1. A atenção como condição

Publicada em 1987, a IRH de Merton distancia-se do *Capital Asset Pricing Model* (CAPM) por defender que o custo de obter informação impede os investidores de conhecer todas as empresas disponíveis no mercado. Para Merton o facto de cada carteira de investimento ser constituída apenas por uma fracção do universo de empresas existentes prova que cada investidor conhece e opera somente sobre um subconjunto de activos (Merton, 1987: 488), o que conduz a uma diversificação imperfeita na alocação de activos¹⁰. Para compensar esse desequilíbrio as empresas menos conhecidas têm de oferecer retornos mais elevados.

Essa premissa, associada ao custo inerente à 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: “*ceteris paribus*, an increase in the relative size of the firm’s investor base will reduce firm’s cost of capital and increase the market value of the firm” (Merton, 1987: 500).

Além de chamar a atenção para a importância da informação na definição do preço de um activo e do custo de financiamento de uma empresa, o modelo de Merton sugere que a notoriedade terá maior influência explicativa na pressão compradora. Isto porque, na aquisição, o investidor pode escolher entre um conjunto imenso de activos, enquanto que na venda, contando com restrições ao *short-selling*, estará limitado aos activos que detém em carteira e de que, portanto, em algum momento foi conhecedor. E como o tempo dos investidores é limitado, as acções com maior notoriedade terão maior probabilidade de

¹⁰ A teoria da alocação de activos estuda a distribuição da atenção do investidor, considerando-a um recurso limitado. A escolha de um activo implica uma alocação de atenção. Sobre este tópico ver e Corwin e Coughenour (2008) e Van Nieuwerburgh e Veldkamp (2009).

integrar os subconjuntos de Merton, da mesma forma que os títulos com menor visibilidade acarretarão maior risco de ser ignorados (Barber e Odean, 2008: 785-786). É também relevante notar que, para se concretizar, a ordem de compra exige uma contraparte vendedora, pelo que haverá investidores menos sensíveis à notoriedade dos títulos. Barber et al. (2008) sugerem que os investidores institucionais actuam frequentemente como reverso da medalha dos investidores particulares. Por venderem a descoberto com regularidade, pelo maior poder de escolha na hora de alienar activos, por dedicarem mais tempo a analisar o mercado, vigiando e actuando frequentemente em sectores específicos, e por habitualmente subordinarem informaticamente as suas decisões a complexas métricas financeiras, os institucionais estarão mais disponíveis que os particulares para vender nas marés cheias de notoriedade e comprar durante as marés vazias. Nessa linha Yuan (2012), analisando a cobertura mediática do Dow Jones Industrial Average (Dow), conclui que os eventos mediáticos têm maior influência sobre o comportamento dos investidores particulares em comparação com os institucionais. Nos picos de atenção os investidores vendem de forma maciça, “obrigando” os institucionais a agir, ou compram de forma modesta consoante o mercado transacionar a um nível elevado ou baixo, respectivamente.

Outra questão deixada em aberto pela IRH é a forma como os investidores seleccionam subconjuntos de activos e os critérios seguidos nessa construção de portfólios. Pelo menos parte desse processo pode ser explicado pela tendência dos investidores para apostar em activos domésticos e em acções de empresas que lhes estão próximas e lhes são familiares. Tal como os futebolistas que revelam maior confiança nos jogos disputados em casa¹¹, ou os espectadores que preferem filmes dobrados, ou os consumidores mais sensíveis à nacionalidade dos produtos, os investidores tendem a preferir empresas que lhes são familiares, estando disponíveis para assumir maior risco nas cotadas que conhecem melhor, até porque o custo de obter informação credível sobre essas companhias é mais baixo do que nas restantes.

Este fenómeno, rotulado na literatura por *home bias*, é um contrasenso financeiro à luz dos modelos que encontram na diversificação a consequência natural da aversão ao risco. Inicialmente a propensão doméstica do investimento foi justificada por barreiras governamentais, elevados custos de transacção e efeitos fiscais que penalizavam ou

¹¹ Num estudo apresentado na conferência anual da British Psychological Society, Neave e Wolfson (2002) provaram que os níveis de testosterona dos futebolistas aumentam “significativamente” nos jogos em casa e que esse incremento varia consoante a reputação do rival.

desmoralizavam o investimento no exterior. No entanto, apesar de nas últimas décadas algumas dessas fricções terem diminuído, e outras desaparecido, continuam a proliferar estudos que provam a persistência do fenómeno *home bias*. Os trabalhos de French e Poterba (1991), que deram conta de que os *traders* norte-americanos alocavam quase 94% dos seus fundos em activos domésticos, e de Kilka e Weber (2000), que mostraram o maior optimismo dos estudantes alemães de economia e finanças sobre títulos germânicos em detrimento de acções norte-americanas, e que a conclusão contrária era válida para os estudantes dos EUA, são alguns exemplos.

Ao estudar o fenómeno *home bias* através de dados da Morningstar, Falkenstein (1996) propõe cinco critérios que influenciam positivamente a notoriedade da acção e o investimento dos gestores de fundos: valor de mercado, senioridade num índice de referência, liquidez, volatilidade e acompanhamento mediático. Coval e Moskowitz (1999) acrescentam que a proximidade geográfica das empresas influencia a formação do portfólio e que a um nível local os investidores tendem a preferir produtoras de bens não transaccionáveis com baixo valor de mercado e com elevada alavancagem. Num estudo para o mercado finlandês, Grinblatt e Keloharju (2001) adicionam factores culturais aos critérios de selecção do investidor, argumentando que a língua em que as empresas comunicam e o perfil cultural dos seus executivos também afectam as opções de investimento.

1.2. Medindo a atenção

Revisitada a IRH de Merton e explorada a importância da familiaridade nas escolhas do investidor, importa agora olhar para a literatura que procurou variáveis capazes de medir e acompanhar a notoriedade de uma acção no mercado. A questão implícita fundamental nessa busca é saber como, onde e em que momento o investidor procura e obtém informação e de que forma isso influencia o valor de uma empresa. Sem o intuito de testar a IRH, Arbel e Strebel (1983), ao distinguirem títulos “genéricos” de títulos “negligenciados” consoante o acompanhamento regular dos analistas profissionais, acabaram por deixar uma sugestão. Já num quadro assumidamente mertoniano, Grullon, Kanatas e Weston (2004) argumentam que a publicidade influencia a visibilidade da marca, seduzindo consumidores, mas também a visibilidade financeira do título, atraindo investidores ao reduzir a assimetria de informação. Apoiando-se no fenómeno *home bias* e relevando que tanto investidores profissionais como amadores tomam decisões com base em informação não financeira, os autores concluem para parte significativa da amostra que um aumento de 10% no investimento publicitário

incrementa em 2,7% o número de accionistas institucionais. No mesmo espírito, Chemmanur e Yan (2009) e Lou (2011) reclamam que a publicidade, ao promover a firma no mercado, pode conduzir, temporariamente, a que as ofertas de compra sobre o título superem as de venda, apreciando assim a cotação. Um dos trabalhos mais citados na literatura que procura estudar a construção do portfólio via publicidade é o *paper* de Frieder e Subrahmanyam (2005) que dá conta, em sintonia com a IRH, da propensão dos investidores particulares para adquirirem empresas com marcas fortes e produtos que lhes são muito familiares – a Apple é um exemplo.

Sem aprofundar o tópico, Merton sugere especificamente que a cobertura mediática de uma empresa poderá servir para testar a sua IRH, acabando assim por relevar a presença das acções em índices de referência. Esse trilha foi percorrido por muitos investigadores e as conclusões estão longe da unanimidade e até incorrem em contradição. Roll (1984), na tentativa de explicar a variação de preço dos futuros sobre sumo de laranja, estudou o impacte das notícias na imprensa financeira norte-americana sobre meteorologia, o factor mais determinante na evolução do valor do contrato, e deduziu a sua quase irrelevância. Ainda na década de 80, Cutler, Poterba e Summers (1989) inferem que as notícias com informação quantitativa macroeconómica explicam apenas um quinto da variação do preço de uma acção. Em sentido contrário, Ederington e Lee (1993) responsabilizam a divulgação nos media de indicadores macroeconómicos calendarizados, como relatórios sobre emprego, índices de preços no consumidor e no produtor, pelo aumento da volatilidade no mercado de futuros de dívida pública norte-americana, de eurodólares e do marco alemão, entretanto extinto.

Mas é sobretudo na literatura mais recente que emerge o elo econométrico entre mediatismo de uma acção e a sua evolução no mercado. Tetlock (2007) conclui que elevado pessimismo latente na coluna *Abreast of the Market* do *Wall Street Journal* (WSJ) prenuncia pressão vendedora sobre as acções do Dow, de maior intensidade nas cotadas com menor valor de mercado. E Fang e Peress (2009), analisando as cotadas do New York Stock Exchange (NYSE) e 500 outras firmas com assento no índice Nasdaq, atentam que uma carteira de acções sem cobertura mediática rende ao ano 3% mais do que uma carteira de títulos com elevada presença nos media¹². No mesmo sentido, Meschke e Kim (2004) provam

¹² O resultado está em linha com o *small firm effect*, teoria segundo a qual as cotadas com menor valor de mercado tendem a gerar retornos superiores às grandes capitalizações (Lustig e Leinbach, 1983); e com a explicação proposta por Elfakhani e Zaher (1998): para investir e conservar títulos de cotadas com baixo valor de mercado, e portanto com fraco ou nulo acompanhamento da parte dos analistas, os

que as entrevistas a presidentes executivos de empresas à cadeia de televisão CNBC geram pressão compradora sobre os títulos, que se revela mais persistente nas cotadas com menor acompanhamento da parte dos analistas.

1.3. Pesquisar para decidir

Para encerrar esta revisão da literatura propõe-se uma leitura da investigação que nos últimos anos começou a procurar repostas na Internet, e na mudança por ela imposta nos fluxos de informação e nas decisões de investimento, para alguns puzzles financeiros irresolutos, como o dos determinantes da variação de uma acção no mercado de capitais.

Antweiler e Frank (2004) debruçaram-se sobre as mensagens partilhadas nos fóruns do Yahoo! Finance e da Raging Bull, catalogando os conteúdos dos utilizadores como recomendações de comprar, vender e manter, e detectando uma correlação positiva entre o volume de *posts* e a liquidez e volatilidade das acções transaccionadas no Dow. Rubin e Rubin (2010) propuseram por seu turno a frequência de edição da Wikipedia como critério de avaliação do grau de conhecimento de um investidor sobre uma acção. E Mondria, Wu e Zhang (2010) utilizaram o volume de pesquisas no site AOL (America Online) como barómetros de atenção para estudar o fenómeno *home bias*.

Parece ser consensual que, ao reduzir custos de transacção, ao facilitar o acesso a informação, ao exponenciar a velocidade de execução e, no fundo, ao democratizar a negociação, a Internet contribuiu para a eficiência do mercado¹³. Mas terá tido um efeito positivo na rendibilidade? Barber et al. (2002) acompanharam 1.607 investidores particulares que trocaram o telefone pela Internet durante a década de 90, concluindo que na era telefónica registavam desempenhos mais lucrativos. Os autores argumentam que o imediatismo e a abundância de informação geram uma sobrestimação de conhecimento, uma autoconfiança despropositada e uma ilusão de controlo que levam os investidores *online* a assumir maior risco e a negociar mais activamente, mas, na amostra estudada, com menor competência.

Outro dos caminhos seguidos pela literatura mais recente tem sido o de usar o Google como fonte de informação, sublinhando algumas vantagens empíricas face aos barómetros

investidores exigem um retorno adicional para compensar a escassez de informação pública sobre a empresa.

¹³ Eficiência no sentido de que "the ultimate goal of a well-functioning stock market is to bring together all possible buyers and sellers, so that the market price reflects the combined preferences of all participants" (Barber et al., 2001: 43).

testados anteriormente. Quantitativamente, os serviços do Google são utilizados mensalmente por mais de mil milhões¹⁴ de indivíduos de múltiplas geografias, enquanto que a versão impressa do WSJ, um dos títulos mais citados na literatura, chega em língua inglesa a cerca de dois milhões de leitores¹⁵. Este ponto foi em parte relevado por Brewster Kahle, fundador do Internet Archive, quando alegou que “a Biblioteca de Alexandria assinala a primeira vez que a Humanidade tentou reunir a totalidade do saber humano num só lugar e num só momento. A última tentativa? O Google” (Kahle citado por Battelle, 2005: 13). O próprio Google define a sua missão no seu site como sendo “to organize the world’s information and make it universally accessible and useful”.

Comparando o uso dos media para medir a atenção do investidor, e consequente decisão de investimento, Da et al. (2011) questionam se a publicação de uma notícia num jornal significa que o investidor efectivamente a leu, crítica que pode ser extrapolada para um anúncio publicitário numa era em que “a wealth of information creates a poverty of attention” (Simon, 1971: 40). Esta questão de prova não se coloca no Google, dado que a busca tem origem num impulso pessoal de procura de sentido e “a soma de todas essas buscas constitui a base de dados das nossas intenções” (Battelle, 2005: 17). Na imprensa o investidor é receptor passivo e relativamente impotente de informação. Já no Google a informação só se concretiza mediante uma impulsão, uma escolha, do investidor, pelo que “if you search for a stock in Google, you are undoubtedly paying attention to it” (Da et al., 2011: 2).

Há, além disso, um benefício temporal e metodológico relevante dado o imediatismo que caracteriza os mercados financeiros e os problemas do desfasamento e retrospectividade de alguma informação económica. O Google disponibiliza dados diários exportáveis com um clique para tratamento estatístico, possibilitando que a informação produzida e absorvida por investigadores e outros interessados se concretize em tempo mais útil. Foi nesse sentido que um estudo publicado na *Nature* em 2009 propôs o Google como detector e preditor de surtos de gripe¹⁶, considerando que “harnessing the collective intelligence of millions of users, Google web search logs can provide one of the most timely, broad-reaching influenza monitoring systems available today” (Ginsberg, Mohebbi, Patel, Brammer, Smolinski e Brilliant, 2009: 1014). Com as variações no volume de pesquisas sobre 45 termos associados

¹⁴ Dados da ComScore referentes a Maio de 2011 (<http://www.comscore.com>).

¹⁵ Dados para o mercado norte-americano da Audit Bureau of Circulations relativos aos seis meses terminados em Setembro de 2011.

¹⁶ O artigo gerou o Google Flu Trends (<http://www.google.org/flutrends/>), site que monitoriza o risco de propagação de surtos de gripe em algumas regiões do globo.

ao vírus da gripe, os autores conseguiram prever surtos uma a duas semanas antes do Departamento de Saúde da Administração norte-americana.

Choi e Varian (2009) lançam proposta idêntica quando argumentaram que os dados recolhidos pelo Google podem melhorar a qualidade de previsão sobre algumas variáveis económicas. Dado o número crescente de consumidores que procura e compara informação na Internet antes de decidir sobre a aquisição de um automóvel, um imóvel ou uma viagem, os dois investigadores utilizam as pesquisas no Google para prever as vendas da Chevrolet e da Toyota, as vendas de casas nos EUA e também os fluxos de turismo em nove países. Choi et al. (2009) destacam nomeadamente a utilidade eventual do Google em antecipar pontos de inflexão nas séries estatísticas - uma subida no volume de pesquisas sobre imóveis numa determinada cidade, por exemplo, poderá indiciar um incremento nas vendas de casas nessa região -, ponto também avultado no trabalho de Schmidt e Vosen (2009) sobre o consumo privado norte-americano. Em relação ao mercado imobiliário dos EUA destaca-se ainda dois outros estudos: Wu e Brynjolfsson (2009) conseguem antecipar as variações nos preços e no volume de casas vendidas através do Google; e Kulkarni, Haynes, Stough e Paelinck (2009) detectam causalidade de Granger entre a frequência de pesquisas sobre vários conjuntos de termos – como “preços de casas”, “vendas de casas” e “mercado imobiliário” -, e a evolução do índice Case-Schiller para 20 cidades.

Na mesma linha, Goel, Hofman, Lahaie, Pennock e Watts (2010) utilizam as variações das pesquisas no Google para prever as receitas de bilheteira no fim-de-semana de estreia de 119 filmes nos EUA, as vendas no primeiro mês de vida de 106 videojogos também nos EUA e ainda o *ranking* da revista *Billboard* com as 100 músicas mais populares no país.

Com premissas semelhantes, mas numa moldura macroeconómica, Askitas e Zimmermann (2009) detectam uma correlação positiva entre o volume de pesquisas em alemão sobre “centros de emprego”, “taxa de desemprego”, “consultor pessoal” e os nomes dos sites de emprego mais populares na Alemanha, com a evolução da taxa de desemprego germânica. No mesmo campo D’Amuri e Marcucci (2009) concluem que o Google acrescenta precisão, nomeadamente em termos de erro quadrático médio, às previsões sobre o desemprego nos EUA, e Anvik e Gjelstad (2010) validam o Google como previsor das variações de curto prazo na taxa de desemprego norueguesa.

Já Da et al. (2011) terão sido os primeiros a testar a IRH de Merton no mercado accionista através do Google. Usando o índice norte-americano Russell 3000 como painel, os autores concluem que uma subida no volume de buscas antecipa em duas semanas pressão

compradora atribuída a investidores particulares, que se revela mais intensa nas cotadas com menor valor de mercado. Para explorar essa conclusão, Da et al. (2011) examinam depois todas as ofertas públicas iniciais (IPO), operações que habitualmente suscitam a atenção do retalho, concretizadas nos EUA entre Janeiro de 2004 e Dezembro de 2007 e daí deduzem que o Google tem poder preditivo sobre a evolução da cotação na estreia em bolsa. Num outro estudo, os mesmos autores partem da premissa de que “when aggregate search volume for a particular product is high, demand for that product is likely to be high” (Da et al., 2010: 1) para inferir que o Google é também um previsor de receitas no mercado de produto, podendo responder, por exemplo, sobre se as vendas acumuladas num feriado vão superar as do anterior.

Olhando para 30 cotadas da (NYSE), Vlastakis e Markellos (2012) também encontraram relações positivas entre o volume de pesquisas na Internet e a liquidez e volatilidade das acções. A principal conclusão é contudo a de que a procura de informação aumenta com a aversão ao risco.

Na mesma linha, Bank et al. (2011) utilizam o Google para estudar as cotadas do índice alemão 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. Chen (2011) replica a metodologia para o mercado holandês e apresenta resultados sintónicos.

Em artigos publicados já no decorrer deste ano, Dimpfl e Jank (2012) detectam comovimento e causalidade de Granger entre a frequência de pesquisas no Google sobre o Dow e a volatilidade do índice norte-americano; 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, à luz do modelo GARCH (1,1); e Drake, Roulstone e Thornock (2012) encontram níveis anormais de pesquisas nas duas semanas que antecedem a publicação de contas de empresas do S&P 500 e argumentam que quanto maior é a procura antecipada por informação menor é a variação da cotação no dia do anúncio.

2. Método e Variáveis

Neste capítulo procede-se à caracterização dos dados importados do Google e à apresentação teórica do índice Euro Stoxx 50 para expor e justificar as opções tomadas em cada momento na definição do objecto de análise e da metodologia a aplicar. Segue-se a exposição dos indicadores financeiros seleccionados e das variáveis construídas a partir do Google para levar a cabo a investigação.

2.1. O indicador Google

Utiliza-se os dados recolhidos e fornecidos pelo Google Trends¹⁷ para estudar a eventual relação entre o volume de pesquisas na Internet sobre uma empresa e a sua evolução no mercado de capitais. O Google Trends é um serviço gratuito disponibilizado ao público pelo Google 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 83,06%¹⁸. 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).

O *modus operandi* de um motor de busca como o Google consiste em ligar as palavras consultadas a uma base de dados de páginas de Internet para produzir sumários de conteúdo de endereços que sejam relevantes para satisfazer as intenções do utilizador. Existem três peças essenciais neste processo: o *crawl*, *software* que folheia e captura páginas de Internet com base no conteúdo pesquisado; o índice, que organiza e rotula essa informação; e o processador de consulta, que liga o *crawl* ao índice para fornecer os resultados de cada pesquisa. O Google Trends opera como um subsistema neste universo, registando em cada momento o interesse dos internautas sobre todos os tópicos com base no volume de pesquisas efectuadas no Google.

¹⁷ 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.

¹⁸ Dados da NetMarketShare para Julho de 2012 (<http://www.netmarketshare.com>).

A construção deste indicador tem algumas particularidades. Para evitar um empolamento artificial dos resultados, o Google exclui a repetição de pesquisas produzidas por um único utilizador durante uma pequena janela temporal. Além disso, para acautelar o crescimento do número de pesquisas ao longo do tempo, os dados do Google Trends são normalizados da seguinte forma:

$$GC_{k,t} = \frac{PS_{k,t}}{TS_{k,t}} \quad (1)$$

em que $GC_{k,t}$ designa o coeficiente Google sobre o termo particular k registado no período t , $PS_{k,t}$ representa o volume de pesquisas sobre o termo k em t e $TS_{k,t}$ o total de consultas concretizadas no Google no mesmo período t . Cada valor de $GC_{k,t}$ reflecte portanto a probabilidade de em t um qualquer utilizador ter concretizado uma consulta sobre o termo particular k . Os coeficientes variam 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).

O histórico do indicador só pode ser extraído para tratamento estatístico sob duas formas. No modo “fixo” os dados são escalonados como variações face ao volume de pesquisas registado num ponto específico, predefinido como Janeiro de 2004 e representado pelo número 1. Uma leitura de 3, por exemplo, significa que o volume de consultas nesse momento foi superior em três vezes às pesquisas concretizadas em Janeiro de 2004. Já no modo “relativo” os dados são tratados como variações em relação à média, representada pelo número 1, do volume de pesquisas sobre o termo em análise no período selecionado. Nesta abordagem, a mesma leitura de 3 já ilustra por isso o triplo das consultas face à média de todo o período considerado. Tal como fizeram Da et al. (2011), Bank et al. (2011) e Chen (2011), adopta-se o modo “relativo” por uma questão de equidade temporal. Considera-se que o modo “fixo”, ao relevar fracções temporais face às demais, poderia nublar as estimações e conclusões. Seguindo estudos anteriores, designa-se doravante o coeficiente produzido pelo modo “relativo” por *Google Search Volume Indicator* (GSVI). A construção é a seguinte:

$$GSVI_{k,t} = \frac{PS_{k,t}}{\overline{PS_T}} \quad (2)$$

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 também varia entre zero e 100, mas aqui qualquer acréscimo ou decréscimo consiste sempre numa relação face à média histórica.

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, apesar de o serviço ser actualizado diariamente. 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.

Outra das particularidades do Google Trends é permitir filtrar as pesquisas pela sua origem geográfica. Mas contrariando Bank et al. (2011), que só contabilizaram pesquisas concretizadas na Alemanha para examinar o mercado accionista germânico, importa-se todos os dados disponíveis por o objecto de análise ser um índice multinacional. Ao reduzir a informação às pesquisas originadas no país sede de cada uma das empresas poderia-se excluir a participação estrangeira na formação e evolução da notoriedade da acção e, portanto, estar a estudar um conjunto diferenciado de realidades nacionais em vez de o mercado como um todo.

A decisão sobre os termos de pesquisa a considerar é outra questão basilar na definição do método a aplicar. Uma pesquisa sobre “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 momentos¹⁹. Não havendo forma de distinguir o “pesquisar para saber” do “pesquisar para decidir”, emergem três opções: empregar a firma (“Banco Santander SA”, por exemplo); utilizar, como Da et al. (2011), o *ticker* da acção (“SAN”); ou considerar - como Vlastakis et al. (2012) e Chen (2011) -, o nome coloquial da empresa (apenas “Santander”).

Exclui-se o primeiro filtro por nunca ter sido usado na literatura e por o número de observações obtido da sua aplicação parecer insuficiente para viabilizar uma investigação. No segundo critério reconhece-se a vantagem de eliminar algum ruído por os GSVI se referirem a códigos eruditos que identificam as empresas num contexto de investimento. Escolhe-se contudo o terceiro critério com base em cinco argumentos. Primeiro: o facto de ter sido utilizado em estudos anteriores cujos resultados foram recebidos como sendo válidos e

¹⁹ Ao identificar esta limitação, John Battelle escreve que “o Santo Graal de todos os motores de busca é conseguir decifrar a verdadeira intenção – o que se procura e em que contexto” (Battelle, 2005: 20).

consistentes pela comunidade científica. Segundo: com a produção de GSVI através dos *tickers* arrisca-se excluir os investidores particulares, menos sofisticados, que sustentam parte das conclusões de Barber et al. (2008) e de Bank et al. (2011). Esse risco não existe ao adoptar o nome coloquial das empresas, o mais universal dos três critérios. Terceiro: a normalização dos dados pelo Google reduz algum do ruído gerado pela abrangência de intenções permitidas pelo grau de liberdade dos resultados. O ruído remanescente não deve influenciar de forma determinante nem sistemática a evolução dos GSVI. Quarto: a pesquisa na Internet é um acto informal dominado por consultas com uma e duas palavras²⁰, características que condizem com a decisão de pautar os resultados pelo nome coloquial das cotadas. Quinto: o critério adoptado respeita os princípios fundamentais da Pragmática de que “as elocuições criam expectativas de relevância” (Sperber e Wilson, 2001: 11) e 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 et al., 2001: 23).

2.2. Euro Stoxx 50

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 euro²¹ e que serve de subjacente a um conjunto alargado de produtos de investimento, como contratos de futuros, opções, produtos estruturados e *Exchange Traded Funds*. 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 em mercados accionistas domésticos, Bank et al. (2011) sobre o alemão e Chen (2011) sobre o holandês.

Das 50 cotadas que compõem o índice actualmente só se conserva as empresas que integraram ininterruptamente o Euro Stoxx 50²² no período em análise, de acordo com dados gentilmente fornecidos pelo STOXX. Com este critério anula-se o comprovado efeito que a

²⁰ Em Maio de 2012 as pesquisas de uma e duas palavras representavam 53,23% do total de consultas concretizadas nos EUA, segundo um estudo da Experian Hitwise (<http://www.experian.com/hitwise>). Pesquisas mais complexas, com cinco ou seis palavras, por exemplo, representavam apenas 6,93% e 3,79% do total, respectivamente.

²¹ Alemanha, Áustria, Bélgica, Finlândia, França, Grécia, Holanda, Irlanda, Itália, Luxemburgo, Portugal e Espanha.

²² A composição do índice é revista anualmente em Setembro.

inclusão ou exclusão de um índice causa na notoriedade de um título²³. Em adição, para aumentar a precisão das conclusões segue-se o critério adoptado por Chen (2001) 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. A GFD Suez e a Munich Re são excluídas da amostra após a aplicação deste filtro. Deixa também de ser válido o histórico de pesquisas sobre o índice Stoxx 50.

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 2.1.. A maioria das cotadas, 16, negocia em plataformas da NYSE Euronext. As restantes transacionam nos sistemas da Deutsche Börse, Nasdaq, London Stock Exchange Group e Bolsas y Mercados Españoles.

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, data em que a presente investigação deu os primeiros passos. Abarca acontecimentos marcantes como o boom dos IPO²⁴ a crise do *subprime* e a tensão europeia das dívidas soberanas, e compreende períodos de *bull market* e de *bear market*²⁵, diversidade que se julga poder enriquecer a solidez das estimações, além de respeitar a duração de um ciclo económico tal como definida nos trabalhos de Kitchin (1923) e de Burns e Mitchell (1946).

2.3. Variáveis financeiras

Os dados sobre o volume e o preço das empresas foram retirados 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 50²⁶.

As variáveis de volume são quantidades desprovidas do factor preço. O volume é dado por:

²³ Ler por exemplo Shleifer (1986), Kaul, Mehrotra e Morck (200), Baker, Nofsinger e Weaver (2002) ou Hedge e McDermott (2003).

²⁴ De acordo com dados gentilmente fornecidos pela Dealogic, entre 2005 e 2007 concretizaram-se 1529 IPO na Europa no valor de 206,73 mil milhões de euros, igualando-se tanto em valor como em número o registo entre 1998 e 2000, período marcado pela bolha tecnológica.

²⁵ Expressões comuns no mercado de acções para designar períodos em que as cotações apresentam tendências de subida (*bull*) ou descida (*bear*) prolongadas face à média histórica.

²⁶ 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.

$$V_{x,t} = \ln(v_{x,t}) \quad (3)$$

com $V_{x,t}$ a representar o logaritmo natural do número de acções da empresa x transaccionadas na semana t . As variações logarítmicas do volume são calculadas da seguinte forma:

$$\Delta V_{x,t} = v_{x,t} - v_{x,t-1} \quad (4)$$

com $\Delta V_{x,t}$ 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 dimensão, também se utiliza o valor absoluto das variações do volume, obtido por:

$$absV_{x,t} = |\Delta V_{x,t}| \quad (5)$$

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:

$$r_{x,t} = \ln\left(\frac{P_{x,t}}{P_{x,t-1}}\right) \quad (6)$$

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:

$$RV_{x,t} = \sum_{i=0}^N r_{x,t,i}^2 \quad (7)$$

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 reflexo do fluxo de informação no mercado.

A partir da equação (6) determinam-se as variações semanais do valor da acção para calcular o retorno absoluto de cada empresa, através de:

$$absr_{x,t} = |r_{x,t}| \quad (8)$$

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 et al., 2008).

No Quadros 2.2. e 2.3. apresenta-se as estatísticas descritivas média, desvio padrão, máximo, mínimo e soma do volume ($V_{x,t}$) e da volatilidade realizada ($RV_{x,t}$) de cada uma das cotadas.

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.

2.4. Variáveis Google

Através do GSVI estimam-se outras 10 variáveis. A primeira é obtida pelo processo:

$$\ln GSVI_{x,t} = \ln(GSVI_{x,t}) \quad (9)$$

com $\ln GSVI_{x,t}$ 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:

$$JSVI_{x,t} = GSVI_{x,t} - \sum_{i=1}^4 GSVI_{x,t-i} \quad (10)$$

São ainda criados indicadores que permitam identificar eventuais efeitos assimétricos da procura por informação no Google sobre as variáveis financeiras. Sabe-se que leituras de GSVI superiores a 1 e que valores positivos de $\ln GSVI$ 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 $\ln GSVI$. Assim sendo, define-se:

$$GSVIp_t = \max(1, GSVI_t) \quad (11)$$

e

$$GSVIN_t = \min(1, GSVI_t) \quad (12)$$

Em logaritmo fica:

$$\ln GSVIp_t = \max(0, \ln GSVI_t) \quad (13)$$

e

$$\ln GSVI n_t = \min(0, \ln GSVI_t) \quad (14)$$

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

$$JGSVI p_t = \max(1, JSVI_t) \quad (15)$$

e

$$JGSVI n_t = \min(1, JSVI_t) \quad (16)$$

que em logaritmo ficam:

$$\ln JSVI p_t = \max(0, \ln JSVI_t) \quad (17)$$

e

$$\ln JSVI n_t = \min(0, \ln JSVI_t) \quad (18)$$

O Quadro 2.4. contém as estatísticas descritivas sobre o GSVI de todas as empresas, destacando-se em favor da consistência da dispersão dos dados que o desvio padrão é sempre inferior a 1 e que apenas quatro cotadas – Iberdrola, Sanofi, Unicredit e Vivendi -, registaram leituras acima de 3 no período analisado.

2.5. Estacionariedade e defasamentos

Para avaliar se as variáveis podem ser descritas como um processo estacionário realiza-se o teste de raiz unitária com quebras estruturais de Zivot e Andrews (Zivot e Andrews, 1992), subordinando o número de defasamentos a considerar à indicação do critério *Schwarz*. Dos resultados descritos no Quadro 2.5. rejeita-se em quase todos os casos a hipótese nula de existência de raiz unitária com um nível de significância de 99%, concluindo-se que as variáveis são estacionárias.

A construção dos modelos de análise e o estudo do valor preditivo do GSVI exige o cálculo de defasamentos para todas as variáveis, o que implica uma redução progressiva no número de observações.

3. Portfólios

Inicia-se a investigação com uma análise genérica sobre a relação entre a procura de informação na Internet e o comportamento de uma cotada em bolsa. Para tal agrupa-se as 36 empresas em quatro carteiras, consoante apresentem maior ou menor GSVI, e examina-se a sua relação com as variáveis volume, volatilidade e retorno. O estudo também permitirá assim aferir a utilidade da frequência de pesquisas no Google como critério definidor de uma estratégia de investimento.

Uma das limitações desta análise prende-se com a impossibilidade de conhecer com precisão que tipo de informação flui em cada consulta no Google em termos do seu efeito sobre o mercado. Presume-se que as notícias sejam uma das fontes de informação mas, uma vez mais, desconhece-se se a atenção do investidor é conduzida por boas ou más notícias, ou se por ambas. Intuitivamente, percebe-se que a reacção esperada à divulgação de uma má notícia é tentar vender um activo, pelo que nesse contexto é expectável uma diminuição da procura, um aumento da pressão vendedora e, conseqüentemente, uma descida do retorno. Da mesma forma, a difusão de uma boa notícia funcionará como um convite à compra ou ao reforço num activo, estando por isso associada a um aumento da procura com provável reflexo positivo na rendibilidade.

Definido o comportamento esperado em ambas as situações, avança-se para uma primeira avaliação sobre a forma como a atenção, medida pela frequência de pesquisas no Google, se relaciona com a evolução de uma cotada no mercado. Para fazê-lo divide-se as empresas em quartis consoante o registo semanal de GSVI e constrói-se a partir daí quatro carteiras de investimento. A primeira reúne as cotadas de maior visibilidade e a última agrega as de menor. Calcula-se depois a média de cada uma das variáveis de interesse: retorno, retorno absoluto, volume, volume absoluto e volatilidade. Da mesma maneira serão construídas outras quatro carteiras em função do indicador JSVI, que capta variações mais extremas da procura por informação na Internet. Em nenhum dos casos os dados são winsorizados – essa análise será desenvolvida no capítulo dedicado à robustez.

A Figura 3.1. prenuncia que as cotadas assumirão comportamentos diferenciados consoante o grau de atenção do investidor. Isto porque o GSVI do primeiro quartil apresenta uma tendência inversa à do último quartil ao longo de quase todo o período. O mesmo acontece entre o segundo e terceiro quartis, embora com um diferencial bastante mais estreito. A principal excepção é o pico de Novembro de 2007, um ponto comum de inflexão que terá

sido despoletado pela crise do *subprime*. O primeiro quartil, que até aí revelara uma tendência decrescente, inicia a partir do final de 2007 uma trajetória ascendente. No GSVI do último quartil constata-se justamente o contrário.

Na análise de correlação apresentada no Quadro 3.1. recolhe-se indícios adicionais que associam um aumento da atenção do investidor sobre uma empresa a uma subida no volume transaccionado e na volatilidade, no caso das carteiras definidas pelo indicador GSVI. O efeito sobre a rendibilidade deverá contudo ser negativo.

3.1. Resultados

Os resultados expostos no Quadro 3.2. demonstram que a carteira onde se reúne as cotadas com GSVI mais elevados regista rendibilidade negativa em todas as semanas, tal como indicara a análise de correlação. Esta relação é dissonante dos resultados de Barber et al. (2008) e de Da et al. (2011), embora este último se foque nos IPO. De forma inversa, tal como denunciara a Figura 3.1., a carteira que agrupa as empresas com menor notoriedade no Google apresenta rendibilidade positiva, facto que está em sintonia com o *neglected firm effect*²⁷ presente no modelo de Merton (1987), que é evidente na Figura 3.2.. Verifica-se também que as variações negativas do retorno no primeiro quartil são mais intensas que as variações positivas observadas no último quartil e regista-se que em nenhum dos casos há pontos de inversão. Observa-se ainda que o retorno absoluto é sempre mais elevado no primeiro quartil, tanto numa base contemporânea como desfasada. E perante essa evidência recebe-se sem surpresa que o primeiro quartil revele índices de volatilidade superiores aos demais, o que vai aliás ao encontro do trabalho de Falkenstein (1996). Na Figura 3.3. apresenta-se a rendibilidade acumulada de uma estratégia de investimento baseada no primeiro e último quartis.

Tal como no retorno, a evolução do volume denuncia uma relação inversa entre o primeiro e o último quartis definidos a partir do indicador GSVI. Até à terceira semana as variações positivas no volume de um coincidem com as variações negativas do outro, sendo que a segunda semana é um momento de inversão para ambos. No volume continua a ser válido que as cotadas de maior notoriedade apresentam leituras mais intensas face aos restantes quartis, facto confirmado nas variações absolutas.

²⁷ Tendência para as cotadas de menor visibilidade registarem retornos mais elevados, um prémio exigido pelos investidores para compensar o risco acrescido associado a menor monitorização institucional e menor acompanhamento da parte dos analistas. Ver Arbel et al. (1983).

Um outro importante foco de análise reside no diferencial entre o primeiro e o último quartis ao longo do tempo em relação a todas as variáveis, que parece ser mais forte no ponto zero e sobretudo na primeira semana. Este “atraso” pode indiciar a presença de investidores menos sofisticados, tal como defendido no estudo de Da et al. (2011).

A maioria destes comportamentos mantém-se nas carteiras definidas a partir do indicador JSVI (Quadro 3.3.). Picos na atenção captados pela frequência de consultas no Google voltam a ser associados a pressão vendedora. No entanto, aqui essa pressão está também patente no último quartil de cotadas até à terceira semana, altura em que se assiste a uma inversão positiva na rendibilidade. A partir desse momento, uma descida significativa na frequência de pesquisas prenuncia maior retorno na semana seguinte.

Em suma, a conclusão fundamental da análise das carteiras é a de que a evolução de uma cotada no mercado não é neutra à frequência de pesquisas realizadas no Google, o que é um bom ponto de partida para o estudo dos efeitos específicos sobre cada uma das variáveis financeiras que se desenvolverá nos próximos capítulos. Nesta primeira análise verificou-se que o GSVI está associado a outros indicadores mais clássicos de atenção dos investidores, tais como o volume transacionado (Barber et al., 2008), as rendibilidades absolutas (Li e Yu, 2011) e os fluxos de informação (Ross, 1989). Os resultados demonstram que uma frequência elevada de pesquisas no Google está relacionada com maior rendibilidade absoluta, volume transacionado e volatilidade.

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 técnica²⁸, por reflectir o grau de notoriedade de um título, o volume é a primeira variável a ser testada em função do GSVI com dados em painel.

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). Esta expectativa é sustentada pelos resultados do capítulo anterior e também pela análise de correlação exposta no Quadro 4.1., reveladora de uma intensidade muito significativa na relação entre GSVI e volume. Dentro dos tradicionais limites de aceitação são apenas duas as variáveis a registar coeficiente de correlação negativo. Além disso um terço da amostra, 12 cotadas, apresenta coeficientes superiores a 0,2. De relevar que as cotadas com maior coeficiente de correlação - Unicredit (0,80), ING (0,61) e Soci t  G n rale (0,53) -, actuam todas no sector financeiro, de onde se interpreta que tamb m ao n vel do GSVI a natureza da actividade ter  valor estat stico, o que n o   evidente nas conclus es de Bank et al. (2011) nem de Chen (2011) e constitui um sinal em prol da credibilidade do indicador do Google.

4.1. Regress es

Utiliza-se a an lise de regress o com dados em painel para testar se as vari veis Google cont m poder explicativo sobre as varia es no volume das cotadas do Stoxx 50 numa base contempor nea. Inclui-se o primeiro desfasamento das varia es no volume como regressor para controlar eventuais rela es entre os valores passados da vari vel dependente e os registos contempor neos das vari veis independentes. A volatilidade e o retorno semanais, das cotadas e do Stoxx 50, s o usados como vari veis de controlo do mercado. O modelo estimado   o seguinte:

$$\Delta V_{x,t} = \alpha + \beta_0 \Delta V_{x,t-1} + \beta_1 R_{x,t} + \beta_2 RV_{x,t} + \beta_3 RS_t + \beta_4 RVS_t + \beta_5 GSVI_{x,t} + \epsilon_t \quad (19)$$

²⁸ On Balance Volume (OBV), de Joseph Granville (Granville, 1963), foi uma das primeiras estrat gias de investimento baseadas na premissa de que a varia o do volume precede o movimento do pre o e pode por isso antecipar pontos de inflex o na evolu o das cota es.

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 da regressão (19) constrói-se um segundo modelo de análise, conhecido na literatura econométrica por *piecewise regression*. Para tal divide-se as empresas em quartis em função do GSVI, tal como se fez para os portfólios. Quanto maior a frequência de pesquisas no Google sobre uma empresa, maior a sua visibilidade. O primeiro quartil volta assim a reunir as cotadas de maior visibilidade e o último as de menor. O primeiro objectivo continua a ser o de descobrir se as variáveis Google explicam variações no volume. Este modelo permite contudo acrescentar um segundo objectivo, que é o de saber se o eventual poder explicativo do Google varia consoante a notoriedade da empresa, indício levantado pela análise de portfólios que pode ser relevante para, por exemplo, aferir que tipo de investidor conduz o GSVI. Do modelo *piecewise* são calculados quatro conjuntos de resultados num total 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.

4.2. Resultados

A uma subida (descida) na frequência de pesquisas no Google está associada um acréscimo (decrécimo) do número de acções transaccionadas na mesma semana. Esta é a primeira conclusão fundamental, extraída do sinal dos coeficientes das variáveis Google apresentados no Quadro 4.2.. Os resultados confirmam, com um grau de confiança de 99%, a expectativa gerada pela análise de correlação e estão em linha com as ilacções de Bank et al. (2011) e de Chen (2011). O primeiro desfasamento do volume é negativo por a variável derivar das suas próprias variações. As relações para com a volatilidade realizada das cotadas e a volatilidade implícita do índice revelam-se positivas, o que está amplamente comprovado na literatura, como por exemplo em Kamara e Koski (2001), Chordia, Huh e Subrahmanyam (2007) e em Bjursell, Frino, Tse e Wang (2010).

Do sinal positivo dos coeficientes das variáveis GSVI, lnGSVI e JSVI infere-se ainda que o Google é um barómetro da atenção do investidor dado que pesquisar implica consciência sobre o termo pesquisado; que o Google é reflector da notoriedade de uma acção uma vez que toda a pesquisa é uma demonstração de interesse; e que o Google é também um indicador indirecto de liquidez pela frequência de pesquisas ter um efeito positivo na variação do volume de uma acção, ou seja, por maior (menor) notoriedade estar associada a maior (menor) liquidez.

Embora todos os coeficientes das variáveis Google sejam positivos, a sua intensidade não é uniforme: o valor explicativo do JSVI sobre o volume é muito superior ao do GSVI e o R² é também mais robusto. A diferença não é, *per se*, uma surpresa por se ter já deduzido que quanto maior (menor) for a frequência de pesquisas, maior (menor) será a variação no volume de uma acção na mesma semana. Será portanto natural que uma variável construída para captar variações anormais de pesquisas contenha maior poder explicativo que o GSVI. A dimensão dessa diferença – de 3 para 25% -, estava todavia fora das expectativas e sinaliza a utilidade acrescida do Google como reflector de agitações na atenção do investidor.

Nos resultados das regressões *piecewise*, expostos no Quadro 4.3., verifica-se que o coeficiente não é constante, na medida em que quanto maior (menor) a notoriedade de uma empresa, menor (maior) o efeito da frequência de pesquisas no Google sobre as variações no volume da acção. Esta segunda conclusão fundamental deriva de nas regressões para o JSVI (colunas 13-16) existir uma relação inversa entre notoriedade e poder explicativo sobre o volume, com o coeficiente a aumentar sequencialmente de 13% até 57% para a fracção que agrupa as empresas de menor visibilidade, sempre com um grau de confiança de 99%.

Uma das questões fundamentais é saber de que forma a evolução da notoriedade captada pelo Google influencia o valor da acção. Isto é, se essa notoriedade é sobretudo geradora de pressão compradora, como sugere a IRH, ou se, pelo contrário, induz pressão vendedora, como indicaram os portfólios. Pelas variações contemporâneas do volume é impossível encontrar uma resposta. A literatura concernente à relação entre o preço e o volume está longe da unanimidade e não é por isso esclarecedora. Granger e Morgenstern (1963) são exemplos de autores com conclusões no sentido de que “no connection could be found between the prices series and the corresponding volume of transactions series” (Granger et al., 1963: 25), enquanto Crouch (1970), Copeland (1976) e Rogalski (1978) sugerem por seu turno existir interdependência entre as duas variáveis.

A hipótese lançada por Barber et al. (2008) também é insuficiente. Os autores argumentam que a atenção enquanto capacidade cognitiva com limitações impõe que cada

investidor actue e seja conhecedor de apenas um subconjunto do universo de empresas no mercado. Deste modo, 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. O raciocínio pode até explicar parte dos resultados dos Quadros 4.2. e 4.3. mas não de forma normativa, pois ignora que há notoriedade positiva e notoriedade negativa e que ambas são passíveis de aumentar o volume de uma acção. Os casos da British Petroleum, Lehman Brohters e Dubai World expostos na Introdução são três exemplos disso mesmo.

Outra motivação que tem ocupado a literatura é saber quem está por detrás do GSVI, com Da et al. (2011) e Bank et al. (2011) a alegar que o Google capta sobretudo a atenção dos investidores particulares. A primeira conclusão fundamental – as variáveis Google influenciam positivamente o volume -, é por isso especialmente relevante, dado referir-se ao Stoxx 50, índice accionista de referência que, como tal, conta com forte presença e acompanhamento da parte de investidores institucionais. Num índice com as características do Stoxx – em média 906 milhões de acções trocaram de mãos diariamente no período analisado -, não é crível que os particulares possuam poder nem dimensão significativa em termos de variações no volume. Daí não se pode contudo validar o Google como descritor da atenção dos institucionais *a fortiori*: primeiro porque se o fosse os coeficientes das variáveis GSVI e lnGSVI, e respectivos R², teriam de ter outra força na regressão (19); e depois porque se considera que esses investidores consomem informação com maior frequência e profundidade e por canais inacessíveis aos particulares, além de automatizarem habitualmente decisões a sofisticadas métricas financeiras. Mas desta primeira conclusão fundamental pode-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.

Já a segunda conclusão fundamental – quanto maior a notoriedade de uma empresa menor o poder explicativo do Google sobre o volume -, sugere que a frequência de pesquisas capta sobretudo a atenção dos investidores particulares. Ofuscadas pela presença e ditames dos institucionais nas cotadas de maior visibilidade, o efeito das variáveis Google sobre o primeiro quartil torna-se quase residual. Esta relação parece assim ir ao encontro das conclusões de Da et al. (2011) e de Bank et al. (2011) de que o poder explicativo do Google, no primeiro caso sobre o retorno e no segundo sobre a liquidez, é mais robusto nas pequenas capitalizações.

Em suma, concluiu-se que um aumento na frequência de pesquisas no Google sobre uma empresa tem um efeito positivo no número de acções transaccionadas na mesma semana, contribuindo assim para a sua liquidez. Por se referir ao índice Stoxx 50 este resultado ilustra que o fluxo de informação por detrás das consultas do Google não é irrelevante para os investidores institucionais. Conclui-se também que o poder explicativo do Google sobre as variações no volume é tanto maior quanto menor for a notoriedade da cotada, um novo indício de que a frequência de pesquisas na Internet capta sobretudo a atenção dos investidores particulares. Constata-se ainda que a variável JSVI revela-se mais útil que as demais em todas as interacções.

5. Volatilidade

Dedica-se este capítulo ao estudo da relação entre a volatilidade e o GSVI. 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 medo²⁹, 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 por informação é uma das possíveis manifestações dessa atenção.

Nessa perspectiva espera-se que o Google, enquanto fonte de informação, possua poder explicativo sobre a volatilidade. Esse poder deverá contudo ser menor do que o identificado no caso do volume, dado que a volatilidade não é uma variável directamente observável e incorpora maior ruído e muitos outros factores além da atenção do investidor. A análise de correlação partilhada no Quadro 5.1. alimenta essa expectativa, com a volatilidade sobre 11 cotadas a mostrar coeficientes negativos com o GSVI. Verifica-se ao mesmo tempo que das remanescentes que revelam coeficientes positivos, 11 têm leituras superiores a 0,2 dentro dos parâmetros de aceitação. Quando existe relação, ela é intensa.

De sublinhar que os cinco maiores coeficientes pertencem todos, uma vez mais, a firmas do sector financeiro: Deutsche Bank (0,39) Soci t  G n rale (0,38), UniCredit (0,36) Generali (0,35) e Axa (0,35). A an lise da Figura 5.1. ajuda a perceber porqu  ao deixar evidente que todas as empresas registam um aumento significativo da volatilidade na mesma data, em 2008, o que ser  reflexo da crise do *subprime*. Esse agravamento   exponencial no sector financeiro, sobretudo nos casos do franc s Soci t  G n rale (30), do alem o Deutsche Bank (12) e do holand s ING (20), e est  em sintonia com a evolu o volatilidade impl cita sobre o  ndice Stoxx 50, que atingiu o m ximo hist rico de 87,51 pontos na sess o de 16 de Outubro de 2008.

5.1. Regress es

Para estudar a rela o contempor nea entre a frequ ncia de consultas no Google sobre uma empresa e a evolu o da volatilidade volta-se a recorrer a uma regress o com dados em painel. Replicando o m todo da an lise sobre as varia es no volume, o modelo   estimado da seguinte forma:

$$RV_{x,t} = \alpha + \beta_0 RV_{x,t-1} + \beta_1 R_{x,t} + \beta_2 RS_t + \beta_3 RVS_t + \beta_4 GSVI_{x,t} + \epsilon_t \quad (20)$$

²⁹ Ver Armony e Dolan (2002).

com $RV_{x,t}$ a representar a volatilidade realizada sobre a empresa x registada na semana t (variável dependente), 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, respectivamente.

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 (20), um modelo *piecewise* com parâmetros idênticos aos utilizados no estudo sobre o volume. Ou seja, as empresas são segmentadas em quartis consoante a sua visibilidade medida pelo GSVI. Deste procedimento derivam 16 estimações agrupadas em quatro conjuntos de resultados: o primeiro sem variáveis Google e os restantes com GSVI, lnGSVI e JSVI, respectivamente, a responder perante cada um dos quartis, ordenados de forma decrescente consoante a visibilidade das empresas.

5.2. Resultados

Deduz-se dos resultados da regressão (20) apresentados no Quadro 5.2. que a frequência de pesquisas no Google explica parte da volatilidade das cotadas do Stoxx 50 com um grau de confiança de 95%. Ou seja, que parte da informação consultada no Google é repercutida na variabilidade dos preços na mesma semana. Os dados estão em linha com os trabalhos de Chen (2011) e de Dimplf et al. (2012) e confirmam a supracitada expectativa de que a relação é bastante mais fraca do que a verificada com as variações no volume devido aos determinantes da volatilidade. O primeiro desfasamento da variável dependente é naturalmente positivo e muito significativo, elevando o R².

Um pedaço importante destes resultados mostra que as agitações capturadas pelo JSVI voltam a evidenciar maior poder explicativo face às restantes variáveis Google, o que está em sintonia com o teorema financeiro segundo o qual “in an free arbitrage economy, the volatility of prices is directly related to the rate of flow of information” (Ross, 1989: 16). Da análise gráfica patenteada na Figura 5.2. retiram-se indicações complementares que sustentam esta relação: o aumento da volatilidade parece estar sobretudo relacionado com leituras de GSVI acima de 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: 34), 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 5.3. voltam a denunciar, embora com um diferencial e uma robustez mais reduzidos dos que no volume, que é sobre os últimos quartis (colunas 7, 8, 11 e 12) que as variáveis GSVI e lnGSVI manifestam maior poder explicativo em relação à volatilidade. Esse facto é uma vez mais indiciador de que por detrás dos GSVI estarão sobretudo investidores particulares que utilizam o Google como fonte de informação privilegiada. Se fossem institucionais, e dado que o objecto de análise é o índice Stoxx 50, os coeficientes das variáveis Google certamente não seriam tão fracos, mesmo considerando todo o ruído que a volatilidade consubstancia. A elevada visibilidade do Stoxx 50 enquanto índice de referência explica por isso porque Chen (2011) alcançou resultados mais robustos no estudo da volatilidade sobre as cotadas do holandês AEX.

Face ao exposto, e sendo tecnicamente impossível decompor a variável GSVI consoante o tipo de investidor que a origina, os resultados dos Quadros 5.2. e 5.3. vão no sentido de que “anything that changes the amount or character of noise trading will change the volatility of price” (Black, 1996: 533). Assim, aceitando que o GSVI é mormente um barómetro da atenção do investidor particular, cujas decisões não dependem somente dos fundamentais, o sinal positivo dos coeficientes das variáveis Google evidenciam que a entrada do retalho no mercado acrescenta-lhe volatilidade, tal como concluíram Lux e Marchesi (1999), Foucault, Sraer e Thesmar (2011) e Dimpfl et al. (2012)

Concluindo, o Google, enquanto fonte de informação, contém poder explicativo sobre a volatilidade das cotadas do Stoxx 50. Esta propriedade valida a frequência de pesquisas na Internet como um indicador do apetite por risco no mercado, dado que a literatura provou existir uma relação entre a procura de informação e a aversão ao risco. Tal como aconteceu no volume, o poder explicativo do Google é menor perante as empresas de maior visibilidade, o que se interpreta como sendo um indício adicional de que o GSVI capta sobretudo a atenção dos investidores particulares. Prova-se também assim que a entrada desses investidores no mercado acrescenta-lhe volatilidade, mesmo num índice de referência como o Stoxx 50.

6. Previsibilidade

A descoberta de que as consultas no Google sobre uma empresa influenciam positivamente tanto o volume como a volatilidade numa base contemporânea criou a expectativa de que essa relação possa também existir de forma desfasada. Ou seja, que a frequência passada de pesquisas sobre uma empresa possa explicar, e assim antecipar, os registos presentes do volume e da volatilidade. Neste capítulo efectua-se uma análise preditiva para perseguir essa expectativa. Inclui-se o retorno semanal das empresas no leque de variáveis a testar em função do Google.

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 nos capítulos 4 e 5 podem até já reflectir este desfasamento entre pesquisar e decidir. Ainda assim, e dado os indícios encontrados de que por detrás do GSVI estão sobretudo investidores particulares, é importante avaliar se também os desfasamentos das variáveis Google contêm poder explicativo sobre o comportamento de uma cotada em bolsa.

6.1. Regressões

Para testar se a frequência passada das pesquisas no Google explica as variações presentes do volume de uma acção replica-se a regressão (19) com dados em painel mas agora com todas as variáveis independentes desfasadas em um período. O modelo torna-se no seguinte:

$$\Delta V_{x,t} = \alpha + \beta_0 \Delta V_{x,t-1} + \beta_1 R_{x,t-1} + \beta_2 RV_{x,t-1} + \beta_3 RS_{t-1} + \beta_4 RVS_{t-1} + \beta_5 GSVI_{x,t-1} + \epsilon_t \quad (21)$$

com $GSVI_{x,t-1}$ a ilustrar a frequência de pesquisas sobre a empresa x na semana anterior, $t-1$,

à variação do número de acções transaccionadas da empresa x representada na variável dependente $\Delta V_{x,t}$.

Seguindo o mesmo procedimento para a volatilidade realizada estima-se, a partir da regressão (20), o seguinte modelo:

$$RV_{x,t} = \alpha + \beta_0 RV_{x,t-1} + \beta_1 R_{x,t-1} + \beta_2 RS_{t-1} + \beta_3 RVS_{t-1} + \beta_4 \Delta V_{x,t-1} + \beta_5 GSVI_{x,t-1} + \epsilon_t \quad (22)$$

com $GSVI_{x,t-1}$ a ser uma vez mais a variável independente de maior interesse. Inclui-se no modelo o regressor $\Delta V_{x,t-1}$ por a análise contemporânea sobre o volume ter mostrado uma relação positiva muito forte para com a volatilidade realizada.

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:

$$R_{x,t} = \alpha + \beta_0 R_{x,t-1} + \beta_1 RS_{t-1} + \beta_2 RV_{x,t-1} + \beta_3 RVS_{t-1} + \beta_4 \Delta V_{x,t-1} + \beta_5 GSVI_{x,t-1} + \epsilon_t \quad (23)$$

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.

6.2. Resultados

Dos resultados expostos no Quadro 6.1. deduz-se que as pesquisas no Google sobre uma empresa contêm poder explicativo sobre a variação do número de títulos transaccionados na semana seguinte com um nível de confiança de 99%. Em adição, e tal como verificado nos modelos contemporâneos, os coeficientes de todas as variáveis Google apresentam sinais positivos (colunas 2, 3 e 4). Deste modo, a maior (menor) frequência de pesquisas é associado um aumento (diminuição) do volume de acções negociadas na semana seguinte. Estas inferências são válidas para todas as variáveis construídas a partir do Google, embora o JSVI volte a evidenciar um poder explicativo muito superior ao do GSVI e ao do lnGSVI.

A conclusão fundamental extraída do Quadro 6.1. é contudo a de que o efeito do primeiro desfasamento das variáveis Google sobre as variações do volume é muito mais forte do que o verificado na análise contemporânea. Ou seja, que o passado recente das pesquisas é melhor descritor que o seu presente no que ao volume diz respeito. E é-o tanto ao nível dos R^2 , que são ligeiramente mais robustos do que no modelo do Quadro 4.2., como, e principalmente, ao nível da força dos coeficientes, que mais do que duplica. No caso das agitações captadas pelo JSVI o coeficiente aumenta de 25% na análise contemporânea para 62% na análise preditiva.

Se é duvidoso que os investidores institucionais utilizem o Google como fonte privilegiada de informação, é menos crível ainda que, a fazê-lo, demorassem uma semana a repercutir a informação obtida em ajustes de carteira, sobretudo num índice de referência como o Stoxx 50. A análise preditiva oferece por isso uma pista adicional de que o GSVI reflecte sobretudo a atenção dos investidores particulares, que não actuam ao microssegundo nem decidem unicamente com base nos fundamentais.

Se o Google capta mormente a atenção dos particulares e se a frequência de pesquisas antecipa maior volume de negociação, importa perceber como podem esses investidores contribuir positivamente para a liquidez de um título. No desenho da microestrutura do mercado a literatura aponta três fontes primordiais de liquidez: diminuição dos custos de transacção (Stoll et. al, 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).

Em relação à volatilidade também se conclui dos resultados apresentados no Quadro 6.2. que a frequência passada das consultas no Google é portadora de poder explicativo sobre a magnitude da variabilidade dos preços na semana seguinte com um grau de confiança de 95%. Os coeficientes das variáveis Google voltam a ser todos positivos, o que já acontecera na análise contemporânea. Deduz-se assim que maior (menor) notoriedade medida pelo GSVI antecipa maior (menor) volatilidade na semana seguinte, em linha com as conclusões de Dimpfl et al. (2012) e de Smith (2012).

Todavia, e ao contrário do que aconteceu no estudo sobre o volume, os valores dos coeficientes das variáveis Google na análise preditiva sobre a volatilidade são praticamente idênticos aos verificados no modelo contemporâneo. O efeito das consultas do Google continua a ser limitado devido por um lado à diversidade de factores e de ruído que a

volatilidade incorpora e, por outro, à forte presença de investidores institucionais no Stoxx 50. De facto, olhando para as fontes mais clássicas de volatilidade que têm sido consensualizadas na literatura – risco político (Mei, 1999), alterações regulatórias (Hsieh e Miller, 2012), evolução e perspectivas sobre indicadores macroeconómicos (Diebold e Yilmaz, 2009), gestão da empresa (Kothari, Li e Short, 2009), liquidez (Domowitz, Glen e Madhavan, 2002), efeito da taxa de câmbio (Bodart e Reding, 1999) e vencimento de instrumentos derivados (Feinstein e Goetzmann, 1988) - é difícil perceber como poderia o comportamento dos investidores particulares explicar de forma significativa a variabilidade dos preços num índice como o Stoxx 50.

Uma limitação dos estudos sobre o volume e a volatilidade reside em não permitirem aferir se a notoriedade medida pelo Google influencia de forma positiva o valor de mercado de uma cotada, tal como sugere a IRH de Merton (1987). Essa constitui pois a principal motivação da análise sobre o retorno semanal das empresas do Stoxx 50 em função das pesquisas realizadas no Google na semana anterior.

Dessa análise, cujos resultados estão patenteados no Quadro 6.3., conclui-se que as variáveis $\ln\text{GSVI}$ e JSVI explicam parte da evolução do retorno na semana seguinte com um grau de confiança de 95%. Esse poder é em ambos os casos mais robusto do que aquele verificado para a volatilidade, embora os modelos para a rendibilidade apresentem R^2 muito baixos.

A inferência fundamental resulta contudo de os sinais dos coeficientes das variáveis Google serem todos negativos (colunas 3 e 4). Ou seja, de pelo modelo estimado um aumento (diminuição) das consultas concretizadas no Google sobre uma empresa contribuir para uma depreciação (apreciação) do valor da acção na semana seguinte. Na base desta relação, que se encontrara no capítulo 3 e que contraria a lógica da IRH e as conclusões do trabalho de Da et al. (2011), poderá estar um ofuscamento imposto pelo domínio dos investidores institucionais num índice de referência como o Stoxx 50 sobre os investidores particulares, que serão os principais condutores do GSVI.

Em suma, pela análise preditiva descobre-se que a procura por 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, o poder explicativo dos desfasamentos das variáveis Google mais do que duplica em relação aos modelos contemporâneos. Ao contrário do que acontece no volume e na volatilidade, a frequência de

pesquisas no Google contribui de forma negativa para o retorno de uma acção na semana seguinte, relação que se entende reflectir um ofuscamento dos investidores particulares imposto pelo domínio dos institucionais no índice Stoxx 50.

7. Interdependência

Neste capítulo estima-se um Vector Auto-Regressivo (VAR) para tentar captar interdependências lineares entre o GSVI, as variações no volume, a volatilidade realizada e o retorno de cada acção, mas também a autocorrelação que se verificou nas variáveis. A partir dos resultados do VAR realiza-se testes de causalidade de Granger para avaliar o poder preditivo incremental do GSVI sobre os indicadores financeiros. Para terminar analisa-se, através de uma função impulso resposta (FIR), o comportamento dinâmico das variáveis aquando de um choque exógeno sobre cada uma das variáveis endógenas do sistema, tendo o GSVI como principal foco de interesse.

As análises contemporânea e preditiva demonstraram que a atenção do investidor, medida pela frequência de pesquisas no Google, contém poder explicativo sobre a evolução do volume, da volatilidade e também do retorno de uma cotada. Todos os modelos estimados até agora foram todavia unidireccionais ao servirem para estudar o efeito do regressor GSVI sobre uma variável dependente financeira. Validou-se uma relação e não uma interdependência.

Desconhecendo-se com precisão os despertadores da atenção do investidor, seguiu-se nos capítulos anteriores uma linha de raciocínio que assume a pesquisa no Google sobre uma cotada como um acto temporalmente coincidente ou desfasado (*a priori*) do seu comportamento no mercado. Contudo, um dos despertadores da atenção poderá ser justamente a evolução do mercado, ainda para mais sendo o GSVI um indicador que parece captar a atenção dos investidores particulares, descritos na literatura como sendo menos informados, menos atentos, menos sofisticados e menos competentes que os institucionais.

A condição “investors trade only after paying attention to a stock” (Da et al., 2011: 11) mantém-se. Mas acrescenta-se agora, no estudo das interdependências entre o GSVI e as variáveis financeiras, a hipótese de que a consulta no Google também possa ser despoletada por eventos já descontados em bolsa - uma queda abrupta no valor de uma acção, uma fraude num banco ou o reporte de lucros históricos, por exemplo -, sendo portanto a procura por informação na Internet um acto *a posteriori* neste cenário. A confirmação desta hipótese constituirá um forte indício de que por detrás do GSVI estão investidores particulares e não beliscará os resultados obtidos anteriormente.

7.1. O modelo VAR

Dada a robustez dos termos autorregressivos nos modelos contemporâneos, utiliza-se os modelos VAR para avaliar o grau de interdependência entre a procura por informação na Internet sobre uma cotada e a sua evolução em bolsa. Garantida a estacionariedade pelo teste de Zivot e Andrews (Quadro 2.5.), sobra encontrar o número óptimo de defasamentos do modelo para cada uma das cotadas, o que será feito com base no critério *Schwartz*. O modelo VAR (1) é definido pela forma matricial:

$$\begin{bmatrix} \ln GSVI_{x,t} \\ \Delta V_{x,t} \\ RV_{x,t} \\ R_{x,t} \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} \ln GSVI_{x,t-1} \\ \Delta V_{x,t-1} \\ RV_{x,t-1} \\ R_{x,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \\ \epsilon_{4,t} \end{bmatrix} \quad (24)$$

em que o valor de cada uma das variáveis endógenas na semana t depende dos seus valores passados mas também dos defasamentos das outras variáveis. Assume-se que os erros (ϵ_t) têm média igual a zero e variância constante.

Seguidamente realiza-se o teste de causalidade de Granger com dois objectivos: avaliar se as pesquisas no Google são também acontecimentos *a posteriori*; e aferir se os valores passados do $\ln GSVI$ comportam poder preditivo incremental sobre as variáveis financeiras. Toma-se como hipótese nula a não existência de causalidade entre o $\ln GSVI$ e a variável financeira e como hipótese alternativa a não existência de causalidade entre essa variável e o $\ln GSVI$.

Por fim, e mantendo o indicador Google como principal foco de interesse, examina-se através de uma FIR se, como e durante quanto tempo as variáveis endógenas do sistema VAR respondem a choques exógenos.

7.2. Resultados

Os resultados dos testes de causalidade de Granger expostos no Quadro 7.1. comprovam o que os modelos preditivos já sugeriam: a frequência de pesquisas no Google sobre uma empresa comporta poder preditivo incremental sobre o seu comportamento em bolsa. A principal novidade reside todavia na evidência de que o contrário é igualmente válido: o passado das variáveis financeiras influencia o volume de pesquisas de uma acção no motor de busca do Google.

A interdependência detectada a partir do modelo VAR sinaliza que parte do volume de pesquisas pode ser induzido por informação já descontada no mercado, o que destoa das conclusões de Dimpfl et. al (2012) para a volatilidade. Essa reacção *a posteriori* é o indício

mais robusto dos até agora encontrados de que por detrás do GSVI estarão sobretudo investidores particulares. Isto porque ilustra uma fracção da procura de informação na Internet originada pelo mercado que reconhece, no fundo, o “atraso” dos particulares face aos institucionais encontrado na análise dos portfólios (Quadros 3.2. e 3.3.).

Do Quadro 7.1. regista-se ainda que em apenas cinco cotadas – L’Oreal (22), Nokia (24), Telefonica (32), Total (33) e Vivendi (36) - se verificou a não existência de causalidade de Granger em todas as interações, o que poderá estar relacionado com o grau de liberdade das pesquisas consideradas.

O comportamento dinâmico das variáveis examinado através das FIR também confirma as conclusões da análise preditiva. A transmissão ao sistema de um choque exógeno na procura por informação na Internet tem um impacte significativo sobre o volume, que se desvanece a partir da primeira semana, e um efeito quase residual sobre a volatilidade e o retorno. É esse o efeito padrão. Nas Figuras 7.1., 7.2. e 7.3. apresenta-se os casos da Air Liquide, ING e Sanofi, respectivamente. No primeiro *lag* a resposta padrão do volume é positiva - um aumento da frequência de pesquisas prediz maior liquidez (Bank et al., 2011) -, o efeito sobre a volatilidade é positivo – o *noise trading* acrescenta incerteza ao mercado (Foucault et al., 2011) -, enquanto o impacte sobre a rendibilidade é insignificante – a presença dos institucionais ofusca o comportamento do retalho.

Já o efeito contrário, dos choques exógenos nas variáveis financeiras sobre o indicador Google, é praticamente nulo e no primeiro *lag* a resposta dominante sobre a frequência de pesquisas revela-se maioritariamente negativa. A periodicidade semanal dos dados poderá estar a distorcer esta relação que intuitivamente se esperava ser positiva, com o comportamento do mercado a despertar a atenção dos investidores e a gerar um aumento da frequência de pesquisas no Google.

Em suma, nos testes de causalidade de Granger encontrou-se evidência de que a procura de informação na Internet contém poder preditivo incremental sobre a evolução de uma cotada em bolsa. Constatou-se ainda que o Google é também uma variável *a posteriori*. Isto é, que a frequência de pesquisas pode ser induzida por informação já incorporada no mercado. Esta novidade é lida como um forte indício de que o GSVI é sobretudo alimentado por investidores particulares. Adicionalmente, as interdependências dinâmicas dos modelos VAR permitiram concluir, através das FIR, que um choque exógeno sobre o Google tem um efeito positivo sobre o volume e a volatilidade na semana seguinte, mas afecta a rendibilidade de forma negativa, corroborando assim os resultados da análise preditiva.

8. Assimetria

Comprovado o efeito que a atenção captada pelo Google tem sobre as variáveis financeiras, impõe-se avaliar o grau de assimetria dessa relação. Intuitivamente, percebe-se que a atenção funciona como uma medida indirecta de 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. É esse o estudo que se desenvolverá neste nono capítulo da investigação, olhando uma vez mais para o volume, a volatilidade e o retorno.

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. Já foi destacado que na maioria das análises concretizadas o indicador JSVI revelou-se o melhor descritor e também o melhor predictor das variáveis financeiras. Mas por se desconhecer a natureza dos choques e o sentido das variações exógenas de atenção captadas pelo JSVI, daí não foi possível extrair qualquer conhecimento sobre assimetria. Na dispersão de observações das Figuras 5.2. e 8.1. também não se encontraram indícios sólidos em relação a esta matéria.

A literatura sugere que nos picos de atenção os investidores vendem de forma maciça ou compram de forma modesta consoante o mercado transaccionar a um nível elevado ou baixo, respectivamente (Yuan, 2008); que a atenção tem maior influência sobre o comportamento dos investidores particulares em comparação com os institucionais (Yuan, 2008); que fracos níveis de atenção conduzem a uma sub-reacção do preço porque a informação disponível não é imediatamente incorporada (Hou, Peng e Xiong, 2006); e que os investidores mais entorpecidos tendem a privilegiar informação sobre a generalidade do mercado, ou sobre um sector, a especificidades de empresas (Peng e Xiong, 2006). Estes elementos vão enriquecer a interpretação dos resultados mas são insuficientes para sustentar uma expectativa sobre se o poder explicativo do Google revelará maior robustez na agitação ou na hibernação do investidor.

8.1. Regressões

Para aferir se o poder explicativo do Google sobre as variáveis financeiras é assimétrico utiliza-se uma vez mais a análise de regressão com dados em painel. O modelo estimado para as variações no volume é o seguinte:

$$\begin{aligned} \Delta V_{x,t} = & \alpha + \beta_0 \Delta V_{x,t-1} + \beta_1 R_{x,t} + \beta_2 RV_{x,t} + \beta_3 RS_t + \beta_4 RVS_t \\ & + \beta_5 GSVIp_{x,t} + \beta_6 GSVIn_{x,t} + \epsilon_t \end{aligned} \quad (25)$$

onde os regressores $GSVIp_{x,t}$ e $GSVIn_{x,t}$ 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 $GSVIp_{x,t}$ e $GSVIn_{x,t}$ significa assimetria.

Por se desconhecer com exactidão os despertadores da atenção do investidor, nomeadamente o peso das boas e das más notícias, usa-se também o volume absoluto como variável dependente num modelo definido por:

$$\begin{aligned} absV_{x,t} = & \alpha + \beta_0 absV_{x,t-1} + \beta_1 absr_{x,t} + \beta_2 RV_{x,t} + \beta_3 RS_t \\ & + \beta_4 RVS_t + \beta_5 GSVIp_{x,t} + \beta_6 GSVIn_{x,t} + \epsilon_t \end{aligned} \quad (26)$$

com $absV_{x,t}$ a representar as variações em módulo do número de acções da empresa x negociadas na semana t e $absr_{x,t}$ o retorno absoluto da cotada x durante a semana t .

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

$$\begin{aligned} RV_{x,t} = & \alpha + \beta_0 RV_{x,t-1} + \beta_1 \beta_1 absr_{x,t} + \beta_2 RS_t + \beta_3 RVS_t \\ & + \beta_4 GSVIp_{x,t} + \beta_5 GSVIn_{x,t} + \epsilon_t \end{aligned} \quad (27)$$

Para avaliar a existência de assimetria na previsibilidade recalcula-se as regressões (25), (26) e (27) mas agora com todas as variáveis independentes desfasadas em um período. Adicionalmente, recupera-se o modelo (23) para a rendibilidade, substituindo apenas o último regressor pelas variáveis $GSVIp_{x,t-1}$ e $GSVIn_{x,t-1}$, ficando:

$$\begin{aligned} R_{x,t} = & \alpha + \beta_0 R_{x,t-1} + \beta_1 RS_{t-1} + \beta_2 RV_{x,t-1} + \beta_3 RVS_{t-1} \\ & + \beta_4 \Delta V_{x,t-1} + \beta_5 GSVIp_{x,t-1} + \beta_6 GSVIn_{x,t-1} + \epsilon_t \end{aligned} \quad (28)$$

Todos os modelos estimados incluem um processo de ruído branco, ϵ_t . São estimadas quatro versões para cada um dos modelos: na primeira inclui-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.

8.2. Resultados

Se dos modelos contemporâneos e preditivos se inferiu que um aumento na procura por informação no Google sobre uma empresa tem um efeito positivo sobre o volume e que, portanto, a um alheamento está associado uma diminuição no número de acções transaccionadas, na análise sobre assimetria só se encontra evidência do primeiro comportamento (Quadros 8.1. e 8.2.). As duas relações são contudo validadas na análise contemporânea relativa às variações absolutas do volume (Quadro 8.3.), que oferece uma novidade: a descida no volume associada a um decréscimo na atenção do investidor na mesma semana é mais forte do que o efeito contrário, com um grau de confiança de 95% (colunas 3 e 4). Isto é, a inactividade do investidor, descrita por uma diminuição acentuada na frequência de pesquisas no Google, tem um efeito contemporâneo negativo sobre o volume absoluto que se revela mais intenso do que o efeito positivo observado nos picos de atenção.

Esta conclusão é reforçada pelos resultados da análise preditiva apresentados no Quadro 8.4., onde é ainda maior a diferença entre a influência negativa sobre o volume absoluto associada a uma ausência do investidor e o efeito positivo relacionado com um crescimento anormal na procura por informação no Google na semana anterior (colunas 3 e 4). Além disso, nas estimações expostas nas colunas 1 e 2, os coeficientes das variáveis que descrevem uma diminuição na atenção do investidor (GSVIn e lnGSVIn) são mais robustos do que os dos indicadores que captam pulos na frequência de pesquisas (GSVIp e lnGSVIp).

Estas conclusões constituem mais um sinal de que por detrás das pesquisas realizadas no Google estarão sobretudo investidores particulares. No sentido de que a variável JSVIp descreve picos na atenção do investidor, ou seja, momentos com fluxo de informação mais intenso e, portanto, com maior ruído. Num índice com as características do Stoxx 50 isso implica a acção de investidores institucionais, cuja presença condiciona o poder explicativo do Google, tanto ao nível do valor do coeficiente como ao nível da sua robustez.

No que diz respeito à volatilidade, a análise de assimetria confirma que o poder explicativo do Google é reduzido nos picos e também nos fossos de atenção (Quadros 8.5. e 8.6.). Dos resultados dos modelos contemporâneos e de previsibilidade infere-se que tanto uma subida como uma descida na frequência de pesquisas no Google estão relacionadas com um aumento dos efeitos não lineares da volatilidade.

Relativamente ao retorno, na análise de previsibilidade verificou-se que a uma subida do GSVI está associada uma queda da rendibilidade da acção na semana seguinte e que o inverso – uma descida no volume de pesquisas no Google prenuncia um maior retorno –, será portanto igualmente válido. Todavia, na análise de assimetria, cujos resultados são expressos

no Quadro 8.7., só se encontra evidência estatística do primeiro processo com um grau de confiança de 95%.

Esta relação negativa entre visibilidade e retorno ficou patente logo na construção das carteiras, sendo que uma das fontes da pressão vendedora poderá ser a propensão dos investidores para tomar mais-valias de forma precipitada e conservar acções com menos-valias potenciais durante demasiado tempo. Esse comportamento, baptizado na literatura por *disposition effect*³⁰, tem sido associado aos investidores particulares (Yuan, 2008) que se julga serem os principais condutores do indicador de atenção do Google.

Outra explicação poderá estar relacionada com as características históricas do período alvo de análise - durante o qual o índice Stoxx 50 registou a pior prestação anual de sempre, em 2008, com perdas de 44,37% -, e também com as características do objecto de investigação, formado em um terço por cotadas do sector financeiro. A intensidade e persistência da notoriedade negativa originada pela crise do *subprime* e pela crise das dívidas soberanas europeias, cujos maiores reflexos se sentiram no sector financeiro, podem dominar de forma significativa o período examinado. Aceitando esta hipótese, a maioria dos picos de frequência de pesquisas no Google terá por base informação com impacte negativo na cotação da acção, o que explica os resultados do Quadro 8.7..

Concluindo, a análise de assimetria prova que o efeito dos indicadores Google sobre as variáveis financeiras não é linear, o que explica a menor robustez de alguns dos resultados obtidos nos capítulos anteriores. Observa-se que a pressão negativa sobre o volume associada a um decréscimo na frequência de pesquisas é muito mais intensa do que no processo contrário e que o poder explicativo do Google sobre a volatilidade é pouco significativo tanto nos picos como nos fossos de atenção. O modelo sugere que um aumento da procura por informação no Google tem um efeito negativo sobre o preço da acção mas que o inverso não se verifica.

³⁰ Ver Frazzini (2006) e Shefrin e Statman (2012).

9. Robustez

Dedica-se o último capítulo da investigação à verificação da robustez dos resultados obtidos em termos de previsibilidade. Para tal recorre-se à winsorização dos dados para controlar a presença de *outliers*, à divisão do período sob análise em dois subgrupos para verificar o impacte da crise do *subprime* e ainda à determinação de variáveis *dummy* para acautelar efeitos específicos do país sede e do sector de actividade das cotadas.

Uma das ameaças ao tratamento consistente de dados estatísticos deriva da existência de *outliers*. Os estimadores utilizados ao longo da investigação não garantem resistência à influência de estes valores extremos ou mesmo aberrantes que podem ter afectado os resultados ao induzirem efeitos espúrios capazes de envesgar a interpretação. Para testar a robustez das conclusões face a *outliers* procede-se à winsorização dos dados, uma transformação estatística que limita os valores extremos das series. Os resultados são apresentados nos Quadros 9.1., 9.2. e 9.3. e não beliscam as conclusões da investigação.

Admite-se para além disso que a crise financeira do *subprime* tenha motivado alterações significativas nos padrões “normais” de comportamento do mercado. Essa suspeita é sustentada pelos picos anormais de pesquisas no Google registados nas semanas terminadas a 16, 23 e 30 de Novembro de 2007 (Figura 3.1.), altura em que os bancos europeus começaram a reportar prejuízos derivados da crise de crédito. Impõe-se por isso também uma avaliação sobre se o impacte da crise não enviesou as conclusões da investigação. Para tal divide-se a amostra em dois subperíodos: o primeiro entre 2004 e 16 de Novembro de 2007 e o segundo entre 30 de Novembro de 2007 e o final do período sob análise. Os resultados (Quadros 9.4.-9.9.) mantêm-se, sendo o único facto relevante a registar o de que a previsibilidade das rendibilidades parece apenas existir no segundo subperíodo.

Por último há ainda que verificar se efeitos específicos sobre as cotadas, motivados por notícias ou por outras informações transversais a indústrias e a países, minaram a robustez dos resultados obtidos nos capítulos anteriores. Para lidar com essa influência adiciona-se *dummies* que identificam países e sectores de actividade como variáveis de controlo nas regressões. Os resultados (Quadros 9.10., 9.11. e 9.12.) também não afectam a consistência das conclusões da investigação.

Conclusão

A presente investigação contribui para a literatura que analisa o impacto da atenção do investidor sobre a liquidez e a rendibilidade de activos financeiros. Mais concretamente segue trabalhos recentes que propõem a frequência de pesquisas no Google como barómetro da atenção do investidor. Na última década a Internet tornou-se numa fonte privilegiada 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 2011 numa base semanal.

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, e da endogeneidade da frequência de pesquisas no Google.

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 impacto 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 detectou-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). O mais sólido dos indícios foi a inferência, pelos testes de causalidade de Granger, que parte das pesquisas no Google é concretizada *a posteriori* do mercado, o que é um dado novo na literatura. Nesse quadro atribui-se o efeito

positivo sobre o volume à redução de custos de assimetria de informação, o impacto 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 investidores particulares derivado do domínio dos institucionais num índice de referência como o Stoxx 50.

As conclusões supracitadas contribuem para várias frentes da investigação em Economia e Finanças: relevam 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; dão conta da tendência para as cotadas de menor visibilidade registarem retornos mais elevados; 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.

As principais limitações da investigação residiram em duas impossibilidades incontornáveis: a de aceder a valores absolutos do volume de pesquisas sobre cada cotada, que o Google não disponibiliza; e a de distinguir o “pesquisar para saber” do “pesquisar para decidir”, ou seja, conhecer em todos os momentos qual a motivação que conduz cada pesquisa, algo que nenhuma tecnologia conseguiu até hoje descortinar.

Partilha-se para terminar a expectativa de que a presente investigação estimule a prossecução de estudos que recorram à Internet para captar a atenção dos investidores e dos consumidores e assim elucidar o seu efeito nos mercados financeiros e de produto. Como direcções futuras de investigação sugere-se uma análise sectorial sobre o impacto das pesquisas no Google sobre o mercado accionista, presumindo-se que os diversos ramos de actividade responderão de forma díspar à atenção do investidor. Embora de complexa execução, seria igualmente pertinente testar a IRH a partir dos conteúdos do Facebook e do Twitter, que se afirmam como um meio de comunicação cada vez mais popular na Internet.

Referências bibliográficas

- Ahearne, Alan G., William L. Grier e Francis E. Warnock (2004), “Information Costs and Home Bias: An Analysis of U.S. Holdings of Foreign Equities”, *Journal of International Economics*, 62 (2), pp.313-336.
- 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.
- Anvik, Christian e Kristoffer Gjelstad (2010), “Just Google it. Forecasting Norwegian unemployment figures with web queries”, Dissertação de Mestrado em Economia e Gestão, Oslo, Norwegian School of Management.
- Arbel, Avner e Paul Strebel (1983), “Pay attention to neglected firms”, *The Journal of Portfolio Management*, 9 (2), pp.37-42.
- Armony, Jorge L. e Raymond J. Dolan (2002), “Modulation of spatial attention by fear-conditioned stimuli: an event related fMRI study”, *Neuropsychologia*, 40 (7), pp.817-826.
- Askitas, Nikolaos e Klaus F. Zimmermann (2009), “Google Econometrics and Unemployment Forecasting”, *Applied Economics Quarterly*, 55 (2), pp.107-120.
- Baker, H. Kent, John R. Nofsinger e Daniel G. Weaver (2002), “International Cross-Listing and Visibility”, *Journal of Financial and Quantitative Analysis*, 37 (3), pp.495-521.
- 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 (2001), “The Internet and the Investor”, *Journal of Economic Perspectives*, 15 (1), pp.41-54.
- Barber, Brad e Terrace Odean (2002), “Online Investors: Do the Slow Die First”, *Review of Financial Studies*, 15 (2), pp.455-488.
- 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.
- Battelle, John (2005), *The Search – Como o Google mudou as regras do negócio e revolucionou a cultura*, Cruz Quebrada, Casa das Letras.
- Bjursell, Johan, Alex Frino, Yiuman Tse e George H. K. Wang (2010), “Volatility and trading activity following changes in the size of futures contracts”, *Journal of Empirical Finance*, 17 (5), pp.967-980.
- Black, Fischer (1986), “Noise”, *The Journal of Finance*, 41(3), pp.529-543.
- Bodart, Vincent e Paul Reding (1999), “Exchange rate regime, volatility and international correlations on bond and stock markets”, *Journal of International Money and Finance*, 18 (1), pp.133-151.
- Burns, Arthur F. e Wesley C. Mitchell (1946), *Measuring Business Cycles*, Nova Iorque, NBER.
- Chemmanur, Thomas J. e An Yan (2009), “Advertising, Attention and Stock Returns”, Working paper, Boston College and Fordham University.
- 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.
- Chordia, Tarun, Sahn-Wook Huh e Avanidhar Subrahmanyam (2007), “The Cross-Section of Expected Trading Activity”, *Review of Financial Studies*, 20 (3), pp.709-740.
- Copeland, Thomas E. (1976), “A Model of Asset Trading Under the Assumption of Sequential Information Arrival”, *The Journal of Finance*, 31 (4), pp. 1149-1168.
- 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.
- Coval, Joshua D. e Tobias J. Moskowitz (1999), “Home Bias at Home: Local Equity Preference in Domestic Portfolios”, *The Journal of Finance*, 54 (6), pp.2045-2073.
- Crouch, Robert L. (1970), “The Volume of Transactions And Price Changes on The New York Stock Exchange”, *Financial Analysts Journal*, 26 (4), pp.104-109.
- Cutler, David M., James M. Poterba e Lawrence H. Summers (1989), “What moves stock prices”, *Journal of Portfolio Management*, 15 (3), pp.4-12.
- D’Amuri, Francesco e Juri Marcucci (2009), “Google it! Forecasting the US unemployment rate with Google job search index”, Working paper, University of Essex.
- Da, Zhi, Joseph Engelberg e Pengjie Gao (2011), “In Search of Attention”, *The Journal of Finance*, 45(5), pp.1461-1499.
- Da, Zhi, Joseph Engelberg e Pengjie Gao (2010), “In Search of Fundamentals”, Working paper, University of Notre Dame e University of North Carolina.
- Damásio, António (2010), *O Livro da Consciência – A Construção do Cérebro Consciente*, Lisboa, Círculo de Leitores.
- Diebold, Francis X. e Kamil Yilmaz (2009), “Macroeconomic Volatility and Stock Market Volatility, World-Wide”, em T. Bollerslev, J. Russell and M. Watson (org.), *Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle*, Oxford, Oxford University Press.
- 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.
- Driscoll, John C. e Aart C. Kraay (1998), “Consistent covariance matrix estimation with spatially dependent panel data”, *The Review of Economics and Statistics*, 80 (4), pp.549-560.
- Eco, Umberto (1991), *Apocalípticos e Integrados*, Lisboa, Difel.
- Ederington, Louis H. e Jae Ha Lee (1993) ”How Markets Process Information - News Releases and Volatility”, *The Journal of Finance*, 48 (4), pp.1161-1191.
- Elfakhani, Said e Tarek Zaher (1998), “Differential Information Hypothesis, Firm Neglect and the Small Firm Size Effect”, *Journal of Financial and Strategic Decisions*, 11 (2), pp.29-40.

- Falkenstein, Eric G. (1996), “Preferences for Stock Characteristics As Revealed by Mutual Fund Portfolio Holdings”, *The Journal of Finance*, 51 (1), pp.111-135.
- 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.
- Foucault, Thierry, David Sraer e David J. Thesmar (2011), “Individual Investors and Volatility”, *The Journal of Finance*, 64 (4), pp.1369-1406.
- Frazzini, Andrea (2006), “The Disposition Effect and Underreaction to News”, *The Journal of Finance*, 61 (4), pp.2017-2046.
- French, Kenneth R. e James M. Poterba (1991), “Investor diversification and international equity markets”, *American Economic Review*, 81 (2), pp.222-226.
- 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.
- 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.
- Granger, Clive W. J. e Oskar Morgenstern (1964), “Spectral Analysis of New York Stock Market Prices”, *Kyklos*, 16 (1), pp.1-27.
- Granville, Joseph E. (1963), *Granville’s New Key to Stock Market Profits*, Englewood Cliffs, Prentice-Hall.
- Grinblatt, Mark e Matti Keloharju (2001), “How Distance, Language, and Culture Influence Stockholdings and Trades”, *The Journal of Finance*, 54 (3), pp.1053-1073.
- Grullon, Gustavo, George Kanatas e James P. Weston (2004), “Advertising, breadth of ownership, and liquidity”, *Review of Financial Studies*, 17 (2), pp.439–461.
- Hedge, Shantaram P. e John B. McDermott (2003), “The liquidity effects of revisions to the S&P 500 index: an empirical analysis”, *Journal of Financial Markets*, 6 (3), pp.413-439.
- Hoechle, Daniel (2007), “Robust standard errors for panel regressions with cross-sectional dependence”, *The Stata Journal*, 7 (3), pp.281-312.
- Hou, Kewei, Lin Peng e Wei Xiong (2006), “A Tale of Two Anomalies: The Implications of Investor Attention for Prices and Earnings Momentum”, Working paper, Ohio State University.
- Hsieh, David A. e Merton H. Miller (2012), “Margin Regulation and Stock Market Volatility”, *The Journal of Finance*, 45 (1), pp.3-29.

- Huberman, Gur (2001), “Familiarity Breeds Investment”, *The Review of Financial Studies*, 14 (3), pp.659-680.
- James, William (1890), *Principles of Psychology*, Nova Iorque, Henry Holt.
- Kamara, Avraham e Jennifer Lynch Koski (2001), “Volatility, autocorrelations, and trading activity after stock splits”, *Journal of Financial Markets*, 4 (2), pp.163-184.
- Katz, James E. e Ronald E. Rice (2002), *Social Consequences of Internet Use – Access, Involvement, and Interaction*, Londres, The MIT Press.
- Kaul, Aditya e Vikas Mehrotra e Randall Morck (2000), “Demand curves for stocks slope down: New evidence from an index weight adjustment”, *The Journal of Finance*, 55 (2), pp.893-912.
- Kilka, Michael e Martin Weber (2000), “Home Bias in International Stock Return Expectations”, *Journal of Psychology and Financial Markets*, 1(3/4), pp.176-193.
- Kitchin, Joseph (1923), “Cycles and Trends in Economic Factors”, *The Review of Economics and Statistics*, 5 (1), pp.10-16.
- Kothari, S. P., Xu Li e James E. Short (2009), “The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecast: A Study Using Content Analysis”, *The Accounting Review*, 84 (5), pp.1639-1670.
- Kulkarni, Rajendra, Kingsley E. Haynes, R.R. Stough e Jean H.P. Paelinck (2009), “Forecasting Housing Prices with Google Econometrics”, Working paper, George Mason University.
- Levy, Steven (2011), *In the Plex: how Google thinks, works, and shapes our lives*, Nova Iorque, Simon & Schuster.
- Lewis, Karen K. (1999), “Trying to Explain Home Bias in Equities and Consumption”, *Journal of Economic Literature*, 37 (2), pp.571-608.
- Li, Jun e Jianfeng Yu (2011), “Investor Attention, Psychological Anchors, and Stock Return Predictability”, *Journal of Financial Economics*, Forthcoming paper.
- Lou, Dong (2011), “Attracting Investor Attention through Advertising”, Working paper, London School of Economics.
- Lustig, Ivan L. e Philip A. Leinbach (1983), “The Small Firm Effect”, *Financial Analysts Journal*, 39 (3), pp.46-49.
- Lux, Thomas e Michele Marchesi (1999), “Scaling and criticality in a stochastic multi-agent model of a financial market”, *Nature*, 397 (498), pp.498-500.
- 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.
- Meschke, Felix e Young Han Kim (2004), “CEO Interviews on CNBC”, Working paper, Arizona State University.
- Mondria, Jordi, Thomas Wu (2011), “Asymmetric Attention and Stock Returns”, Working paper, University of Toronto e University of California.
- Mondria, Jordi, Thomas Wu e Yi Zhang (2010), “The Determinants of International Investment and Attention Allocation: Using Internet Search Query Data”, *Journal of International Economics*, 82 (1), pp.85-95.

- Nieuwerburgh, Stijn Van e Laura Veldkamp (2009), “Information Immobility and the Home Bias Puzzle”, *The Journal of Finance*, 64 (3), pp.1187-1215.
- Peng, Lin e Wei Xiong (2006), “Investor Attention, Overconfidence and Category Learning”, Working Paper, *Journal of Financial Economics*, 80 (3), pp.563-602.
- Posner, Michael I. e Steven E. Petersen (1990), “The Attention System of the Human Brain”, *Annual Review of Neurosciences*, 13, pp.25-42.
- Rogalski, Richard J. (1978), “The Dependence of Prices and Volume”, *The Review of Economics and Statistics*, 18 (1), pp.268-274.
- Roll, Richard (1984), “Orange Juice and Weather”, *American Economic Review*, 74 (5), pp.861-880.
- 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.
- Schmidt, Torsten e Simeon Vosen (2009), “Forecasting Private Consumption: Survey-Based Indicators vs. Google Trends”, *Journal of Forecasting*, 30 (6), pp.565-578.
- Shefrin, Hersh e Meir Statman (2012), “The Disposition to Sell Winners Too Early and Ride Losers to Long: Theory and Evidence”, *The Journal of Finance*, 40 (3), pp.777-790.
- Shleifer, Andrei (1986), “Do demand curves for stocks slope down?”, *The Journal of Finance*, 41 (3), pp.579-590.
- Simon, Herbert A. (1971), “Designing Organizations for an Information-Rich World” em Martin Greenberger (org.), *Computers, Communication, and the Public Interest*, Baltimore, The Johns Hopkins University Press.
- 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.
- Tesar, Linda L. e Ingrid M. Werner (1995), “Home bias and high turnover”, *Journal of International Money and Finance*, 14 (4), pp.467-492.
- Tetlock, Paul C. (2007), “Giving Content to Investor Sentiment: The Role of Media in the Stock Market”, *The Journal of Finance*, 62 (3), pp.1139-1168.
- Vlastakis, Nikolaos e Raphael N. Markellos (2012), “Information Demand and Stock Market Volatility”, *Journal of Banking & Finance*, 36 (6), pp.1808-1821.
- Wooldridge, Jeffrey M. (2002), *Econometric Analyses of Cross Section and Panel Data*, Cambridge, The MIT Press.
- Wu, Lynn e Erik Brynjolfsson (2009), “The Future of Prediction: How Google Searches Foreshadow Housing Prices and Quantities”, Investigaç o apresentada na International Conference on Information Systems, 15 a 18 de Dezembro de 2009, Phoenix.
- Yuan, Yu (2012), “Attention and Trading”, Working Paper, University of Iowa.

Zivot, Eric e Donald K. Andrews (1992), "Further Evidence On The Great Crash, The Oil Price Shock, and The Unit Root Hypothesis", *Journal of Business and Economic Statistics*, 10 (10), pp.251-270.

Anexos

Quadro 2.1. Amostra

Descrição das cotadas 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	"BASF"
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"

Quadro 2.2. Estatísticas descritivas da variável volume*Média, desvio padrão, máximo e mínimo das variações logarítmicas do volume por cotada.*

	Empresa	Média	Desvio padrão	Máximo	Mínimo
1	AIR LIQUIDE SA	15.439	0.359	16.604	14.145
2	ALLIANZ SE-REG	16.495	0.433	18.087	14.781
3	AXA SA	17.619	0.400	19.078	16.205
4	BANCO SANTANDER SA	19.764	0.543	21.181	17.968
5	BASF SE	17.048	0.405	18.393	15.396
6	BAYER AG-REG	16.839	0.428	18.061	14.957
7	BANCO BILBAO VIZCAYA ARGENTARIA	19.178	0.545	20.613	17.406
8	BNP PARIBAS	16.891	0.410	18.258	15.304
9	CARREFOUR SA	16.520	0.392	17.872	14.971
10	DAIMLER AG-REGISTERED SHARES	17.271	0.457	18.578	15.441
11	DANONE	16.313	0.403	18.140	14.897
12	DEUTSCHE BANK AG-REGISTERED	17.167	0.496	18.718	15.416
13	DEUTSCHE TELEKOM AG-REG	18.515	0.418	19.828	16.773
14	E.ON AG	17.691	0.395	19.004	15.972
15	ENEL SPA	19.186	0.481	20.768	17.310
16	ENI SPA	18.497	0.531	19.968	16.789
17	FRANCE TELECOM SA	17.722	0.398	18.880	16.171
18	ASSICURAZIONI GENERALI	17.354	0.391	18.664	15.687
19	IBERDROLA SA	18.874	0.515	20.530	17.524
20	ING GROEP NV-CVA	18.209	0.515	19.946	16.426
21	KONINKLIJKE PHILIPS ELECTRON	17.260	0.386	18.299	15.761
22	L'OREAL	15.580	0.385	16.736	14.075
23	LVMH MOET HENNESSY LOUIS VUITTON	15.671	0.381	16.951	14.161
24	NOKIA OYJ	18.720	0.462	19.993	16.918
25	REPSOL YPF SA	17.797	0.485	19.705	16.441
26	RWE AG	16.464	0.399	17.798	14.848
27	COMPAGNIE DE SAINT-GOBAIN	16.193	0.456	17.721	14.734
28	SANOFI	16.793	0.397	18.241	14.871
29	SIEMENS AG-REG	16.974	0.435	18.299	15.252
30	SOCIETE GENERALE	16.549	0.579	18.250	14.768
31	TELECOM ITALIA SPA	20.002	0.508	21.693	18.290
32	TELEFONICA SA	19.214	0.490	20.861	17.573
33	TOTAL SA	17.591	0.373	18.628	16.101
34	UNICREDIT SPA	18.485	0.798	20.232	15.940
35	UNILEVER NV-CVA	17.470	0.383	18.598	15.898
36	VIVENDI	17.186	0.355	18.329	15.679

Quadro 2.3. Estatísticas descritivas da variável volatilidade realizada*Média, desvio padrão, máximo e mínimo da volatilidade realizada por cotada.*

	Empresa	Média	Desvio padrão	Máximo	Mínimo
1	AIR LIQUIDE SA	0.001	0.002	0.015	0.000
2	ALLIANZ SE-REG	0.002	0.006	0.067	0.000
3	AXA SA	0.004	0.008	0.053	0.000
4	BANCO SANTANDER SA	0.002	0.005	0.054	0.000
5	BASF SE	0.002	0.004	0.042	0.000
6	BAYER AG-REG	0.002	0.002	0.023	0.000
7	BANCO BILBAO VIZCAYA ARGENTARIA	0.002	0.004	0.047	0.000
8	BNP PARIBAS	0.003	0.007	0.061	0.000
9	CARREFOUR SA	0.001	0.002	0.028	0.000
10	DAIMLER AG-REGISTERED SHARES	0.003	0.005	0.051	0.000
11	DANONE	0.001	0.002	0.015	0.000
12	DEUTSCHE BANK AG-REGISTERED	0.003	0.008	0.067	0.000
13	DEUTSCHE TELEKOM AG-REG	0.001	0.003	0.039	0.000
14	E.ON AG	0.002	0.004	0.054	0.000
15	ENEL SPA	0.001	0.003	0.053	0.000
16	ENI SPA	0.001	0.003	0.030	0.000
17	FRANCE TELECOM SA	0.001	0.002	0.017	0.000
18	ASSICURAZIONI GENERALI	0.001	0.002	0.019	0.000
19	IBERDROLA SA	0.002	0.005	0.049	0.000
20	ING GROEP NV-CVA	0.005	0.016	0.210	0.000
21	KONINKLIJKE PHILIPS ELECTRON	0.002	0.003	0.023	0.000
22	L'OREAL	0.001	0.002	0.025	0.000
23	LVMH MOET HENNESSY LOUIS VUITTON	0.002	0.003	0.025	0.000
24	NOKIA OYJ	0.001	0.003	0.038	0.000
25	REPSOL YPF SA	0.002	0.003	0.036	0.000
26	RWE AG	0.001	0.003	0.042	0.000
27	COMPAGNIE DE SAINT-GOBAIN	0.003	0.005	0.065	0.000
28	SANOFI	0.001	0.002	0.022	0.000
29	SIEMENS AG-REG	0.002	0.005	0.049	0.000
30	SOCIETE GENERALE	0.004	0.007	0.065	0.000
31	TELECOM ITALIA SPA	0.002	0.003	0.033	0.000
32	TELEFONICA SA	0.001	0.002	0.022	0.000
33	TOTAL SA	0.001	0.003	0.038	0.000
34	UNICREDIT SPA	0.003	0.007	0.060	0.000
35	UNILEVER NV-CVA	0.001	0.002	0.018	0.000
36	VIVENDI	0.001	0.002	0.026	0.000

Quadro 2.4. Estatísticas descritivas da frequência de pesquisas no Google

Média, desvio padrão, máximo, mínimo e soma da frequência de pesquisas no Google por cotada.

	Empresa	Média	Desvio padrão	Máximo	Mínimo	Soma
1	AIR LIQUIDE SA	1.010	0.216	2.040	0.460	395.020
2	ALLIANZ SE-REG	0.998	0.177	2.180	0.640	390.040
3	AXA SA	0.999	0.158	1.820	0.670	390.770
4	BANCO SANTANDER SA	0.949	0.451	2.300	0.320	371.200
5	BASF SE	1.008	0.174	1.520	0.470	394.280
6	BAYER AG-REG	1.009	0.176	1.590	0.530	394.520
7	BANCO BILBAO VIZCAYA ARGENTARIA	0.981	0.271	1.780	0.420	383.420
8	BNP PARIBAS	0.979	0.398	2.500	0.380	382.920
9	CARREFOUR SA	0.996	0.230	2.200	0.500	389.420
10	DAIMLER AG-REGISTERED SHARES	1.016	0.339	1.870	0.400	397.320
11	DANONE	1.016	0.279	2.240	0.420	397.140
12	DEUTSCHE BANK AG-REGISTERED	1.000	0.130	1.670	0.690	390.830
13	DEUTSCHE TELEKOM AG-REG	1.022	0.439	2.640	0.380	399.460
14	E.ON AG	1.012	0.262	2.540	0.560	393.780
15	ENEL SPA	1.000	0.259	2.060	0.380	391.000
16	ENI SPA	0.988	0.266	2.000	0.420	386.180
17	FRANCE TELECOM SA	1.030	0.526	2.280	0.260	402.840
18	ASSICURAZIONI GENERALI	0.995	0.174	1.760	0.560	389.060
19	IBERDROLA SA	1.006	0.265	3.080	0.480	393.340
20	ING GROEP NV-CVA	0.983	0.296	1.560	0.450	384.360
21	KONINKLIJKE PHILIPS ELECTRON	1.014	0.259	1.980	0.590	396.550
22	L'OREAL	0.997	0.098	1.530	0.640	389.960
23	LVMH MOET HENNESSY LOUIS VUITTON	1.009	0.279	2.740	0.500	394.340
24	NOKIA OYJ	0.997	0.127	1.790	0.760	389.930
25	REPSOL YPF SA	1.002	0.192	1.630	0.520	391.650
26	RWE AG	1.004	0.156	1.890	0.610	392.730
27	COMPAGNIE DE SAINT-GOBAIN	1.020	0.356	1.810	0.280	398.870
28	SANOFI	1.013	0.273	3.820	0.460	396.160
29	SIEMENS AG-REG	1.025	0.551	2.340	0.340	400.700
30	SOCIETE GENERALE	0.993	0.356	1.690	0.340	388.280
31	TELECOM ITALIA SPA	1.017	0.415	2.680	0.360	397.460
32	TELEFONICA SA	1.020	0.251	1.540	0.440	398.750
33	TOTAL SA	1.002	0.088	1.470	0.880	391.720
34	UNICREDIT SPA	0.964	0.557	2.140	0.140	377.060
35	UNILEVER NV-CVA	1.008	0.312	3.880	0.500	394.060
36	VIVENDI	1.022	0.508	3.840	0.380	399.660

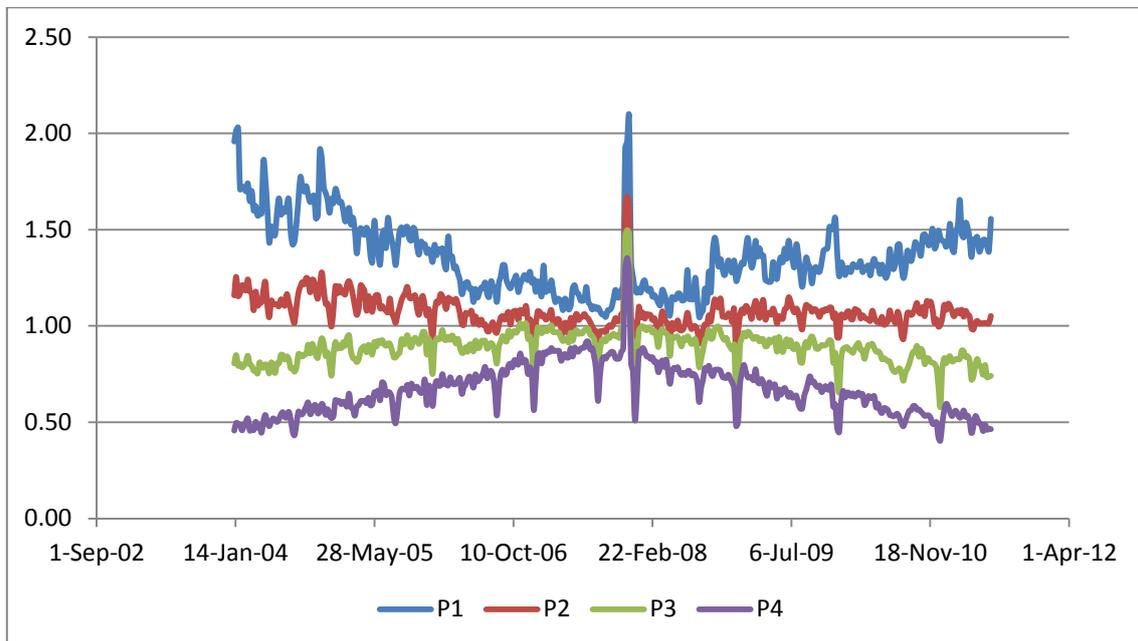
Quadro 2.5. Resultados do teste de raiz unitária Zivot e Andrews

Estatística do teste de raiz unitária com quebras estruturais de Zivot e Andrews. Os símbolos ** e *** denotam níveis de significância de 5% e 1%, respectivamente.

	Empresa	Volume	Retorno	GSVI	Volatilidade
1	AIR LIQUIDE SA	-14.447***	-22.553***	-12.825***	-9.133***
2	ALLIANZ SE-REG	-14.599***	-21.389***	-10.638***	-6.225***
3	AXA SA	-15.753***	-19.938***	-10.968***	-6.445***
4	BANCO SANTANDER SA	-14.75***	-23.278***	-4.439	-8.223***
5	BASF SE	-14.943***	-21.196***	-11.992***	-7.251***
6	BAYER AG-REG	-14.828***	-20.136***	-12.214***	-7.936***
7	BANCO BILBAO VIZCAYA ARGENTARIA	-18.878***	-22.662***	-5.304**	-8.005***
8	BNP PARIBAS	-13.598***	-8.364***	-7.756***	-6.581***
9	CARREFOUR SA	-13.719***	-21.584***	-7.049***	-13.407***
10	DAIMLER AG-REGISTERED SHARES	-21.987***	-20.761***	-10.558***	-6.658***
11	DANONE	-16.019***	-24.453***	-9.299***	-8.593***
12	DEUTSCHE BANK AG-REGISTERED	-14.195***	-22.432***	-6.917***	-6.201***
13	DEUTSCHE TELEKOM AG-REG	-13.778***	-26.349***	-9.938***	-6.885***
14	E.ON AG	-15.841***	-24.219***	-7.331***	-6.532***
15	ENEL SPA	-13.182***	-20.833***	-8.911***	-8.895***
16	ENI SPA	-16.363***	-13.619***	-8.511***	-6.931***
17	FRANCE TELECOM SA	-15.135***	-25.311***	-7.493***	-10.281***
18	ASSICURAZIONI GENERALI	-15.701***	-20.027***	-10.012***	-8.584***
19	IBERDROLA SA	-19.44***	-24.682***	-8.364***	-13.461***
20	ING GROEP NV-CVA	-15.855***	-19.691***	-10.987***	-9.613***
21	KONINKLIJKE PHILIPS ELECTRON	-16.447***	-19.589***	-5.268**	-8.921***
22	L'OREAL	-15.019***	-21.024***	-8.405***	-10.063***
23	LVMH MOET HENNESSY LOUIS VUITTON	-13.822***	-21.645***	-10.698***	-6.781***
24	NOKIA OYJ	-14.333***	-21.965***	-7.101***	-7.156***
25	REPSOL YPF SA	-14.884***	-22.538***	-6.004***	-7.814***
26	RWE AG	-15.932***	-23.084***	-11.733***	-8.139***
27	COMPAGNIE DE SAINT-GOBAIN	-15.478***	-20.319***	-12.944***	-9.718***
28	SANOFI	-14.108***	-25.692***	-14.309***	-8.362***
29	SIEMENS AG-REG	-14.792***	-20.357***	-5.172**	-7.769***
30	SOCIETE GENERALE	-14.334***	-22.469***	-6.009***	-7.812***
31	TELECOM ITALIA SPA	-13.172***	-23.672***	-10.727***	-8.259***
32	TELEFONICA SA	-16.123***	-24.876***	-4.442	-8.716***
33	TOTAL SA	-15.284***	-14.311***	-7.256***	-6.633***
34	UNICREDIT SPA	-14.149***	-13.984***	-9.456***	-7.782***
35	UNILEVER NV-CVA	-15.376***	-22.599***	-7.743***	-9.721***
36	VIVENDI	-14.811***	-22.912***	-9.463***	-7.813***
	Retorno do Stoxx 50	-23.144***			
	VSTOXX	-22.452***			

Figura 3.1. Evolução da frequência de pesquisas no Google por carteira

A primeira carteira (P1) reúne as cotadas com maior frequência de consultas e a última carteira (P4) agrupa as empresas com menor notoriedade no Google.



Quadro 3.1. Correlação entre os indicadores Google e as variáveis financeiras

*Análise de correlação entre as frequências normal (GSVI) e anormal (JSVI) de pesquisas no Google, e as variáveis financeiras. Os símbolos ** e *** denotam níveis de significância de 5% e 1%, respectivamente.*

	GSVI	<i>p value</i>	JSVI	<i>p value</i>
Valor de mercado	-0.0195**	0.02	0.0109	0.20
Rendibilidade	-0.0133	0.11	-0.0071	0.40
Retorno do Stoxx 50	0.0022	0.79	-0.012	0.16
VSTOXX	-0.0229***	0.01	0.011	0.19
Volume	0.0311***	0.00	0.0884***	0.00
Volume absoluto	-0.0035	0.68	-0.0021	0.81
Retorno absoluto	0.0719***	0.00	0.0243***	0.00
Volatilidade	0.0807***	0.00	-0.0248***	0.00

Quadro 3.2. Portfólios por GSVI

Evolução de quatro carteiras construídas a partir da frequência normal de pesquisas no Google (GSVI). A primeira carteira (P1) agrupa as cotadas de maior visibilidade e a última carteira (P4) reúne as de menor notoriedade.

	P1	P2	P3	P4	
GSVI médio	1.36	1.07	0.89	0.67	

Panel A: Rendibilidade					
semana	P1	P2	P3	P4	P4-P1
-1	-0.03%	-0.06%	0.07%	0.06%	0.09%
0	-0.10%	-0.05%	0.09%	0.06%	0.16%
1	-0.15%	0.03%	0.07%	0.05%	0.20%
2	-0.19%	0.07%	0.06%	0.03%	0.21%
3	-0.20%	0.04%	0.06%	0.07%	0.26%
4	-0.14%	0.04%	0.08%	0.01%	0.15%

Panel A: Rendibilidade Absoluta					
semana	P1	P2	P3	P4	P4-P1
-1	3.27%	2.95%	2.71%	2.70%	-0.58%
0	3.30%	2.95%	2.70%	2.68%	-0.62%
1	3.37%	2.91%	2.69%	2.67%	-0.70%
2	3.27%	2.97%	2.68%	2.73%	-0.54%
3	3.30%	2.96%	2.69%	2.71%	-0.59%
4	3.31%	2.93%	2.69%	2.72%	-0.59%

Panel C: Volume					
semana	P1	P2	P3	P4	P4-P1
-1	0.78%	0.40%	0.23%	-0.04%	-0.82%
0	0.58%	0.49%	0.29%	-0.13%	-0.71%
1	1.43%	-0.09%	-0.27%	-1.20%	-2.63%
2	-1.06%	0.28%	-0.04%	0.77%	1.82%
3	-0.63%	0.10%	0.18%	0.39%	1.02%
4	-0.26%	0.21%	0.16%	-0.15%	0.11%

Panel D: Volume Absoluto					
semana	P1	P2	P3	P4	P4-P1
-1	29.88%	29.72%	29.33%	29.21%	-0.67%
0	30.03%	29.73%	29.03%	29.38%	-0.65%
1	30.28%	28.85%	29.39%	28.79%	-1.49%
2	29.95%	28.97%	29.34%	29.10%	-0.86%
3	29.57%	29.37%	29.06%	29.44%	-0.14%
4	29.56%	29.40%	29.13%	29.46%	-0.10%

Panel E: Volatilidade Realizada					
semana	P1	P2	P3	P4	P4-P1
-1	0.26%	0.21%	0.15%	0.16%	-0.10%
0	0.26%	0.20%	0.15%	0.16%	-0.10%
1	0.27%	0.20%	0.15%	0.16%	-0.11%
2	0.26%	0.20%	0.15%	0.17%	-0.10%
3	0.27%	0.19%	0.15%	0.17%	-0.10%
4	0.27%	0.19%	0.15%	0.17%	-0.10%

Figura 3.2. Rendibilidade acumulada por portfólio

Evolução da rendibilidade acumulada de quatro carteiras construídas a partir da frequência normal de pesquisas no Google (GSVI). A primeira carteira (P1) agrupa as cotadas de maior visibilidade e a última carteira (P4) reúne as de menor notoriedade.

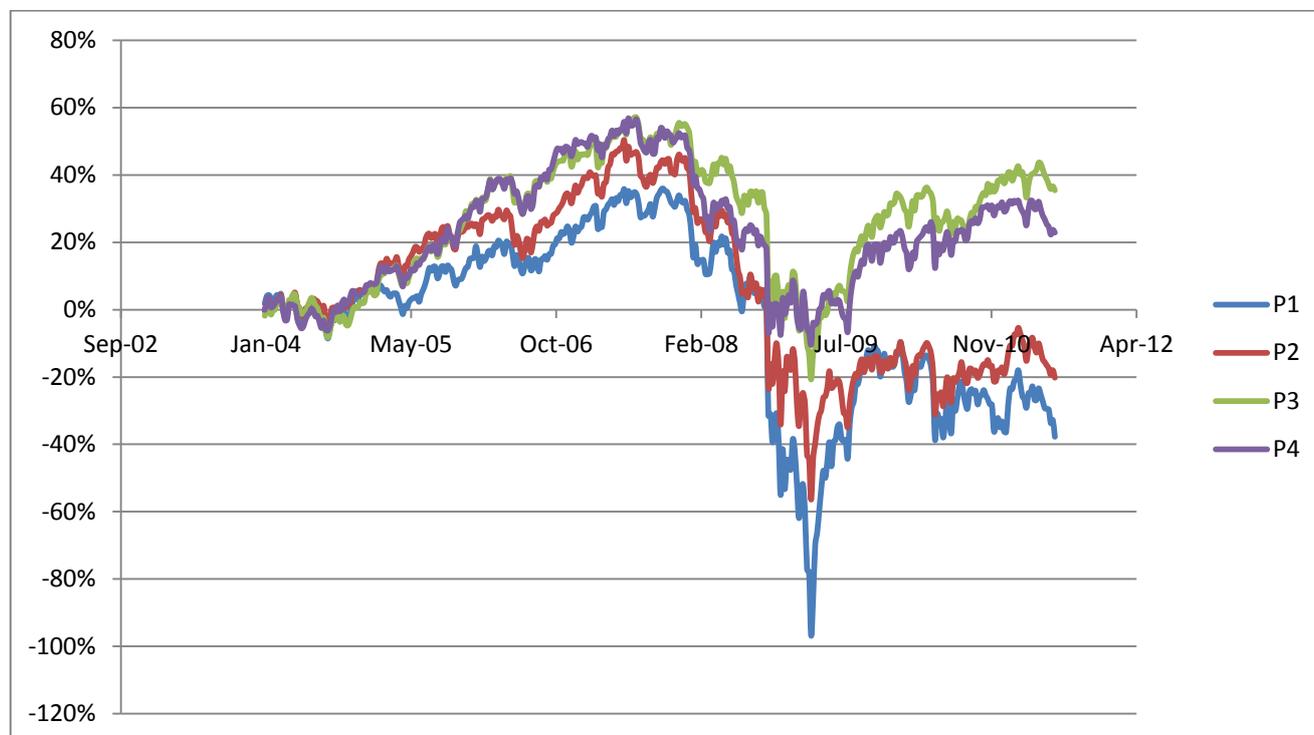
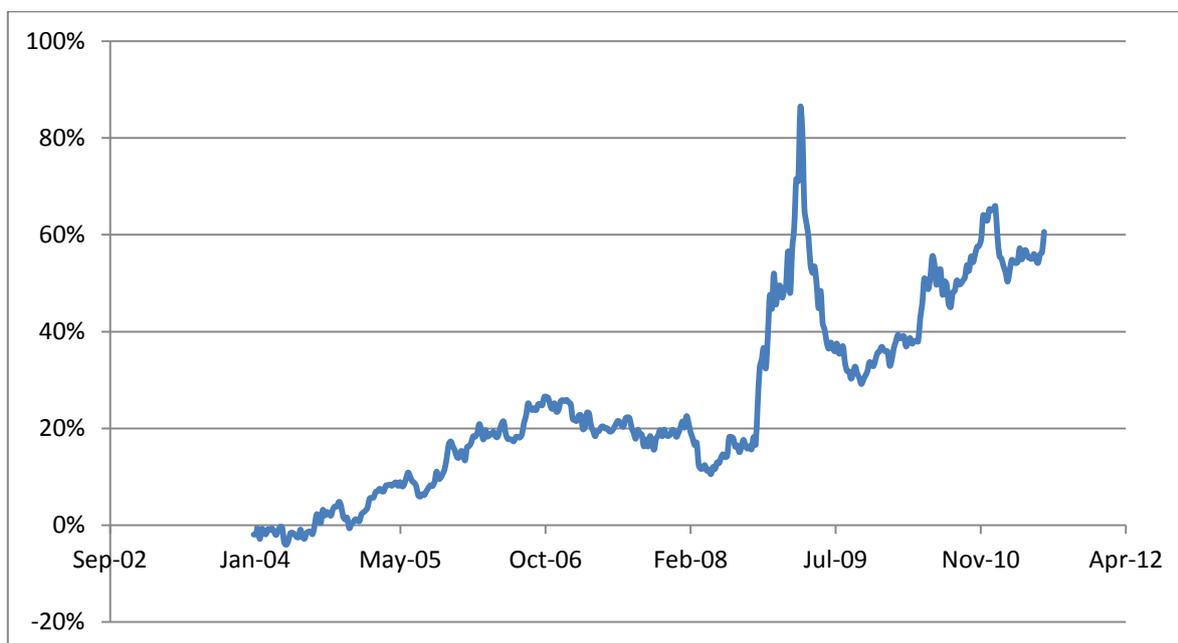


Figura 3.3. Rendibilidade acumulada de estratégia baseada no GSVI

Evolução da rentabilidade acumulada das estratégias de investimento em cotadas de menor notoriedade (P4) e de maior notoriedade (P1) medidas pela frequência normal de pesquisas no Google (GSVI).



Quadro 3.3. Portfólios por JSVI

Evolução de quatro carteiras construídas a partir da frequência anormal de pesquisas no Google (JSVI). A primeira carteira (P1) agrupa as cotadas de maior visibilidade e a última carteira (P4) reúne as de menor notoriedade.

	P1	P2	P3	P4	
JSVI médio	0.12	0.02	-0.03	-0.12	

Panel A: Rendibilidade					
semana	P1	P2	P3	P4	P4-P1
-1	-0.03%	0.05%	-0.03%	0.01%	0.03%
0	-0.03%	-0.04%	0.05%	-0.02%	0.01%
1	-0.01%	0.03%	-0.01%	-0.02%	-0.01%
2	-0.03%	0.03%	-0.02%	0.01%	0.05%
3	-0.06%	0.04%	-0.01%	0.01%	0.07%
4	-0.08%	0.01%	0.01%	0.03%	0.11%

Panel A: Rendibilidade Absoluta					
semana	P1	P2	P3	P4	P4-P1
-1	2.88%	2.92%	2.87%	2.98%	0.10%
0	2.95%	2.93%	2.90%	2.88%	-0.07%
1	3.13%	2.89%	2.85%	2.80%	-0.33%
2	2.92%	2.89%	2.93%	2.94%	0.02%
3	2.88%	2.88%	2.95%	2.98%	0.10%
4	2.86%	2.98%	2.92%	2.94%	0.07%

Panel C: Volume					
semana	P1	P2	P3	P4	P4-P1
-1	0.20%	0.33%	-0.56%	0.18%	-0.02%
0	1.12%	0.51%	0.50%	-2.11%	-3.23%
1	4.34%	0.72%	-1.61%	-3.61%	-7.95%
2	-1.53%	-0.73%	0.72%	1.75%	3.28%
3	-1.47%	-0.45%	1.04%	1.00%	2.47%
4	-0.41%	0.98%	-0.37%	-0.10%	0.32%

Panel D: Volume Absoluto					
semana	P1	P2	P3	P4	P4-P1
-1	29.74%	28.74%	28.57%	30.28%	0.54%
0	29.60%	28.47%	29.28%	30.07%	0.47%
1	30.97%	29.01%	28.75%	28.80%	-2.17%
2	30.65%	28.68%	28.82%	29.41%	-1.24%
3	29.89%	28.40%	29.22%	30.06%	0.18%
4	29.57%	28.63%	29.32%	30.10%	0.53%

Panel E: Volatilidade Realizada					
semana	P1	P2	P3	P4	P4-P1
-1	0.21%	0.19%	0.19%	0.20%	-0.01%
0	0.21%	0.20%	0.19%	0.18%	-0.03%
1	0.21%	0.20%	0.19%	0.19%	-0.03%
2	0.20%	0.19%	0.19%	0.20%	0.00%
3	0.20%	0.19%	0.20%	0.19%	-0.01%
4	0.20%	0.19%	0.20%	0.20%	0.00%

Quadro 4.1. Correlação entre volume e GSVI

Análise de correlação entre as variações logarítmicas do volume e a frequência normal (GSVI) de pesquisas no Google por cotada. Os símbolos ** e *** denotam níveis de significância de 5% e 1%, respectivamente.

	Empresa	Correlação	p value
1	AIR LIQUIDE SA	0.1073**	0.0339
2	ALLIANZ SE-REG	0.0735	0.1468
3	AXA SA	0.3541***	0
4	BANCO SANTANDER SA	0.3610***	0
5	BASF SE	0.3790***	0
6	BAYER AG-REG	0.2424***	0
7	BANCO BILBAO VIZCAYA ARGENTARIA	0.3776***	0
8	BNP PARIBAS	0.2472***	0
9	CARREFOUR SA	0.0092	0.8564
10	DAIMLER AG-REGISTERED SHARES	-0.0396	0.4353
11	DANONE	0.1002**	0.0477
12	DEUTSCHE BANK AG-REGISTERED	0.3847***	0
13	DEUTSCHE TELEKOM AG-REG	0.1176**	0.02
14	E.ON AG	0.1552***	0.0021
15	ENEL SPA	0.2900***	0
16	ENI SPA	0.0765	0.1312
17	FRANCE TELECOM SA	-0.0026	0.9593
18	ASSICURAZIONI GENERALI	0.1381***	0.0063
19	IBERDROLA SA	0.2425***	0
20	ING GROEP NV-CVA	0.6172***	0
21	KONINKLIJKE PHILIPS ELECTRON	0.1908***	0.0001
22	L'OREAL	-0.0128	0.8014
23	LVMH MOET HENNESSY LOUIS VUITTON	0.1436***	0.0044
24	NOKIA OYJ	-0.3143***	0
25	REPSOL YPF SA	-0.0707	0.1627
26	RWE AG	0.0438	0.388
27	COMPAGNIE DE SAINT-GOBAIN	-0.3841***	0
28	SANOFI	0.1183**	0.0193
29	SIEMENS AG-REG	0.0659	0.1935
30	SOCIETE GENERALE	0.5369***	0
31	TELECOM ITALIA SPA	0.0026	0.9594
32	TELEFONICA SA	0.0192	0.7054
33	TOTAL SA	0.0548	0.2795
34	UNICREDIT SPA	0.8015***	0
35	UNILEVER NV-CVA	0.0343	0.4983
36	VIVENDI	0.0994**	0.0495

Quadro 4.2. 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***	-0.24181***	-0.24175***	-0.24919***
	0.000	0.000	0.000	0.000
$R_{x,t}$	0.20568	0.21143	0.21081	0.19986
	0.301	0.286	0.287	0.314
$RV_{x,t}$	5.96767**	5.80421**	5.77384**	5.79876**
	0.032	0.035	0.036	0.034
RS_t	-0.60097	-0.61173	-0.61267	-0.62207
	0.431	0.424	0.423	0.424
RVS_t	0.40167*	0.40127*	0.40048*	0.39666*
	0.073	0.073	0.074	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	-0.04442***	-0.00927	-0.01131
	0.200	0.002	0.285	0.180
Observações	14,004	14,004	14,004	13,930
Cotadas	36	36	36	36
R2	0.088	0.089	0.089	0.098

Quadro 4.3. 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
Cotadas	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

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
Cotadas	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

Quadro 5.1. Correlação entre volatilidade e GSVI

Análise de correlação entre a volatilidade realizada e a frequência normal (GSVI) de pesquisas no Google por cotada. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

	Empresa	Correlação	p value
1	AIR LIQUIDE SA	-0.1165**	0.0213
2	ALLIANZ SE-REG	0.1450***	0.0041
3	AXA SA	0.3482***	0.0000
4	BANCO SANTANDER SA	0.2130***	0.0000
5	BASF SE	0.0268	0.5976
6	BAYER AG-REG	-0.0778	0.1252
7	BANCO BILBAO VIZCAYA ARGENTARIA	0.2388***	0.0000
8	BNP PARIBAS	0.1059**	0.0365
9	CARREFOUR SA	0.1854***	0.0002
10	DAIMLER AG-REGISTERED SHARES	-0.0713	0.1602
11	DANONE	-0.1044**	0.0394
12	DEUTSCHE BANK AG-REGISTERED	0.3988***	0.0000
13	DEUTSCHE TELEKOM AG-REG	-0.1314***	0.0094
14	E.ON AG	0.2585***	0.0000
15	ENEL SPA	0.2492***	0.0000
16	ENI SPA	0.2622***	0.0000
17	FRANCE TELECOM SA	-0.0767	0.1304
18	ASSICURAZIONI GENERALI	0.3505***	0.0000
19	IBERDROLA SA	0.0936*	0.0647
20	ING GROEP NV-CVA	0.2937***	0.0000
21	KONINKLIJKE PHILIPS ELECTRON	-0.1887***	0.0002
22	L'OREAL	0.1365* **	0.0069
23	LVMH MOET HENNESSY LOUIS VUITTON	-0.1461***	0.0038
24	NOKIA OYJ	-0.0276	0.5865
25	REPSOL YPF SA	-0.2504***	0.0000
26	RWE AG	0.0713	0.1598
27	COMPAGNIE DE SAINT-GOBAIN	-0.1952***	0.0001
28	SANOFI	0.0496	0.3284
29	SIEMENS AG-REG	-0.2009***	0.0001
30	SOCIETE GENERALE	0.3816***	0.0000
31	TELECOM ITALIA SPA	-0.2422***	0.0000
32	TELEFONICA SA	-0.1323***	0.0089
33	TOTAL SA	0.0993**	0.0500
34	UNICREDIT SPA	0.3633***	0.0000
35	UNILEVER NV-CVA	-0.0601	0.2365
36	VIVENDI	-0.1462***	0.0038

Figura 5.1. Volatilidade

Evolução da volatilidade realizada por cotada em todo o período sob análise.



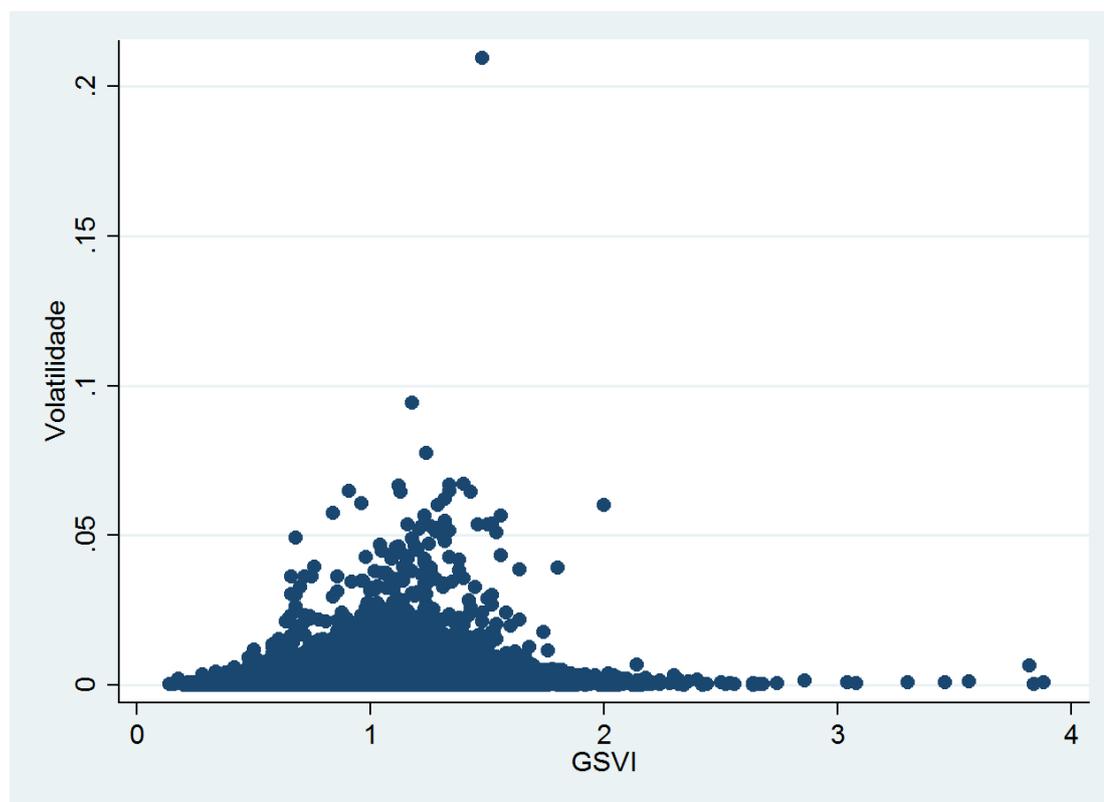
Quadro 5.2. Modelo contemporâneo para a volatilidade

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

Figura 5.2. 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.



Quadro 5.3. 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
$lnGSVI_{x,t}$								
$JSVI_{x,t}$								
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
Cotadas	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

Variáveis	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
$RV_{x,t-1}$	0.68521***	0.52497***	0.49077***	0.58586***	0.68527***	0.52615***	0.49480***	0.59348***
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$R_{x,t}$	-0.02349	-0.00841	-0.00063	0.00337**	-0.02363	-0.00846	-0.00074	0.00329*
	0.111	0.211	0.924	0.043	0.111	0.209	0.912	0.052
$RS_{x,t}$	0.03277	0.00090	-0.00427	-0.00906	0.03303	0.00110	-0.00399	-0.00889
	0.222	0.950	0.596	0.153	0.220	0.938	0.625	0.164
RVS_t	0.00422	0.00016	0.00171	0.00083	0.00422	0.00011	0.00179	0.00088
	0.254	0.921	0.170	0.460	0.254	0.946	0.167	0.435
$GSVI_{x,t}$								
$lnGSVI_{x,t}$	0.00001	0.00095	0.00129*	0.00066***				
	0.971	0.184	0.066	0.001				
$JSVI_{x,t}$					0.00074	0.00160**	0.00017	0.00075*
					0.123	0.014	0.597	0.068
Constante	0.00086***	0.00084***	0.00094***	0.00094***	0.00083***	0.00092***	0.00078***	0.00066***
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observações	3,621	3,500	3,508	3,375	3,603	3,482	3,488	3,357
Cotadas	36	36	36	36	36	36	36	36
R2	0.384	0.435	0.284	0.346	0.385	0.436	0.281	0.344

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

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		
$lnGSVI_{x,t-1}$			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

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

Quadro 7.1. Resultados do teste de causalidade de Granger

*Estatística do teste de causalidade de Granger sob a hipótese nula da não existência de causalidade de Granger entre a variável A e a variável B em cada cotada. O símbolo * denota a rejeição da hipótese nula.*

A	Rendibilidade	Volume	Volatilidade	GSVI	GSVI	GSVI
B	GSVI	GSVI	GSVI	Rendibilidade	Volume	Volatilidade
1	5.6521 0.059	9.2683* 0.01	11.375* 0.003	5.2254 0.073	84.52* 0	4.9034 0.086
2	5.2321 0.156	6.0389 0.11	0.74972 0.861	4.2743 0.233	54.571* 0	0.6368 0.888
3	3.1744 0.366	10.665 0.014	2.0328 0.566	1.6916 0.639	74.747* 0	8.5702* 0.036
4	0.17251 0.678	0.5344 0.465	0.85736 0.354	0.74232 0.389	0.28367* 0	12.171 0.594
5	6.2452 0.1	23.613* 0	4.4236 0.219	4.2121 0.239	175.7* 0	1.8082 0.613
6	2.7843 0.613	10.115* 0.006	2.6206 0.27	0.07525 0.963	47.017* 0	3.7926 0.15
7	5.7064 0.222	19.184* 0.001	1.7118 0.789	5.1219 0.275	20.742* 0	6.9672 0.138
8	8.2059* 0.004	0.2123 0.645	0.06173 0.804	0.02103 0.885	1.4045 0.236	4.8654* 0.027
9	0.08115 0.776	0.2725 0.602	0.00622 0.937	0.51084 0.475	0.01367 0.907	4.1011* 0.043
10	3.6719 0.159	23.722* 0	3.2648 0.195	0.33939 0.844	156.91* 0	2.9354 0.23
11	0.79195 0.673	6.544* 0.038	5.2813 0.071	0.93724 0.626	64.288* 0	14.23* 0.001
12	5.5375 0.236	7.281 0.122	2.1233 0.713	2.1576 0.707	42.755* 0	12.24* 0.016
13	3.772 0.152	7.5537* 0.023	0.6974 0.706	2.3541 0.308	102.14* 0	2.79 0.248
14	3.7059 0.295	12.567* 0.006	2.1678 0.538	8.5899* 0.035	117.57* 0	1.8522 0.604
15	0.54543 0.46	1.0091 0.315	3.314 0.069	0.31103 0.577	1.4756 0.224	9.0965* 0.003
16	0.70494 0.703	7.5559* 0.023	3.2506 0.197	0.78215 0.676	38.007* 0	1.5923 0.451
17	0.02529 0.874	9.1539* 0.002	0.73745 0.39	0.6452 0.422	0.62844 0.428	0.00107 0.974
18	2.0556 0.152	8.8132* 0.003	1.8786 0.17	0.09657 0.756	1.3543 0.245	14.012* 0

A	Rendibilidade	Volume	Volatilidade	GSVI	GSVI	GSVI
B	GSVI	GSVI	GSVI	Rendibilidade	Volume	Volatilidade
19	41243	1.5316	0.01743	5.4206*	9.857*	3.0205
	0.521	0.216	0.895	0.02	0.002	0.082
20	13.657*	8.1967*	3.9372	1.6081	12.05*	12.113*
	0.001	0.017	0.14	0.448	0.002	0.002
21	1.0842	3.8683	10.108*	1.2209	31.279*	4.8398
	0.582	0.145	0.006	0.543	0	0.089
22	0.81396	0.00044	0.14144	0.06044	0.23468	2.3071
	0.367	0.983	0.707	0.806	0.628	0.129
23	6.3677	10.23*	8.0564*	5.7693	52.509*	6.5496
	0.095	0.017	0.045	0.123	0	0.088
24	1.3295	2.5934	0.14192	0.21051	1.307	0.00227
	0.249	0.107	0.706	0.646	0.253	0.962
25	1.9162	1.6431	3.3286	3.0978	8.2855*	11.344*
	0.384	0.44	0.189	0.212	0.016	0.003
26	0.62903	6.0196*	0.02605	1.3926	79.289*	2.5698
	0.73	0.049	0.987	0.498	0	0.277
27	4.1956	23.072*	4.2099	1.1577	56.133*	1.9658
	0.123	0	0.122	0.561	0	0.374
28	0.36812	12.602*	11.214*	0.11007	177.11*	17.584*
	0.947	0.006	0.011	0.991	0	0.001
29	0.57248	0.17571	0.02001	1.0088	0.38381	3.5905
	0.449	0.675	0.887	0.315	0.536	0.058
30	2.6908	11.796*	2.7566	3.2911	6.1788	8.9864*
	0.442	0.008	0.431	0.349	0.103	0.029
31	1.1113	3.2462	2.9129	2.1563	8.9408*	7.8619*
	0.292	0.072	0.088	0.142	0.003	0.005
32	0.15103	0.80778	1.684	1.4305	5.5072	0.37993
	0.927	0.668	0.431	0.489	0.064	0.827
33	2.0769	0.02418	0.16448	0.11622	0.22456	1.1637
	0.15	0.876	0.685	0.733	0.636	0.281
34	0.31107	5.1421	2.5656	0.78655	25.136*	4.8623
	0.856	0.076	0.277	0.675	0	0.088
35	2.626	11.293*	5.3411	1.4498	141.54*	5.6792
	0.269	0.004	0.069	0.484	0	0.058
36	0.08072	0.03507	2.2985	1.8294	1.2608	1.5949
	0.776	0.851	0.129	0.176	0.262	0.207

Figura 7.1. Função impulso resposta – Air Liquide

Comportamento dinâmico de cada uma das variáveis relativas à Air Liquide quando de um choque exógeno sobre cada uma das variáveis endógenas do sistema. “InGSVI-Rendibilidade” indica a resposta da rendibilidade a um impulso na frequência de pesquisas no Google em logaritmo.

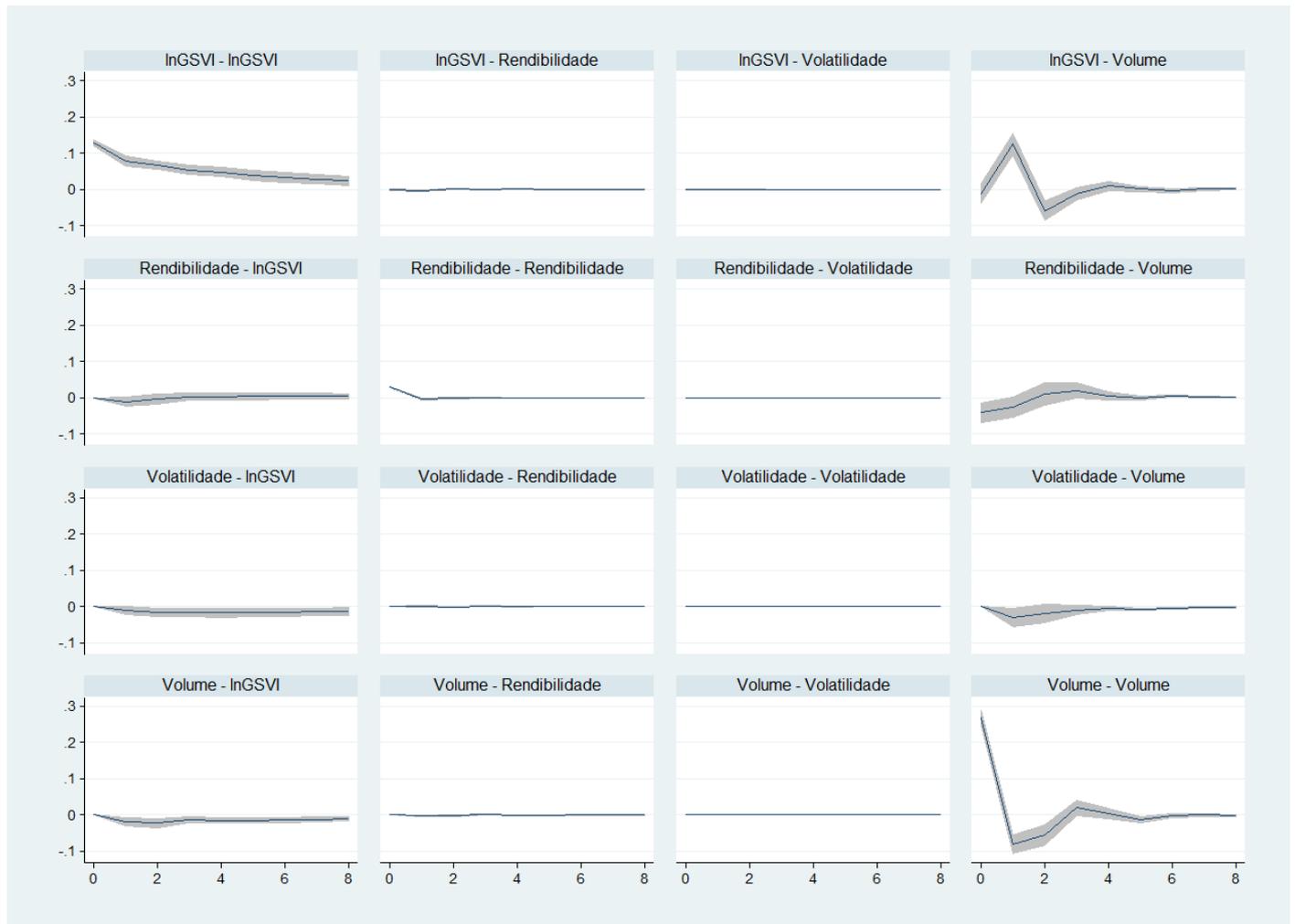


Figura 7.2. Função impulso resposta – ING

Comportamento dinâmico de cada uma das variáveis relativas ao ING quando de um choque exógeno sobre cada uma das variáveis endógenas do sistema. “InGSVI-Rendibilidade” indica a resposta da rendibilidade a um impulso na frequência de pesquisas no Google em logaritmo.

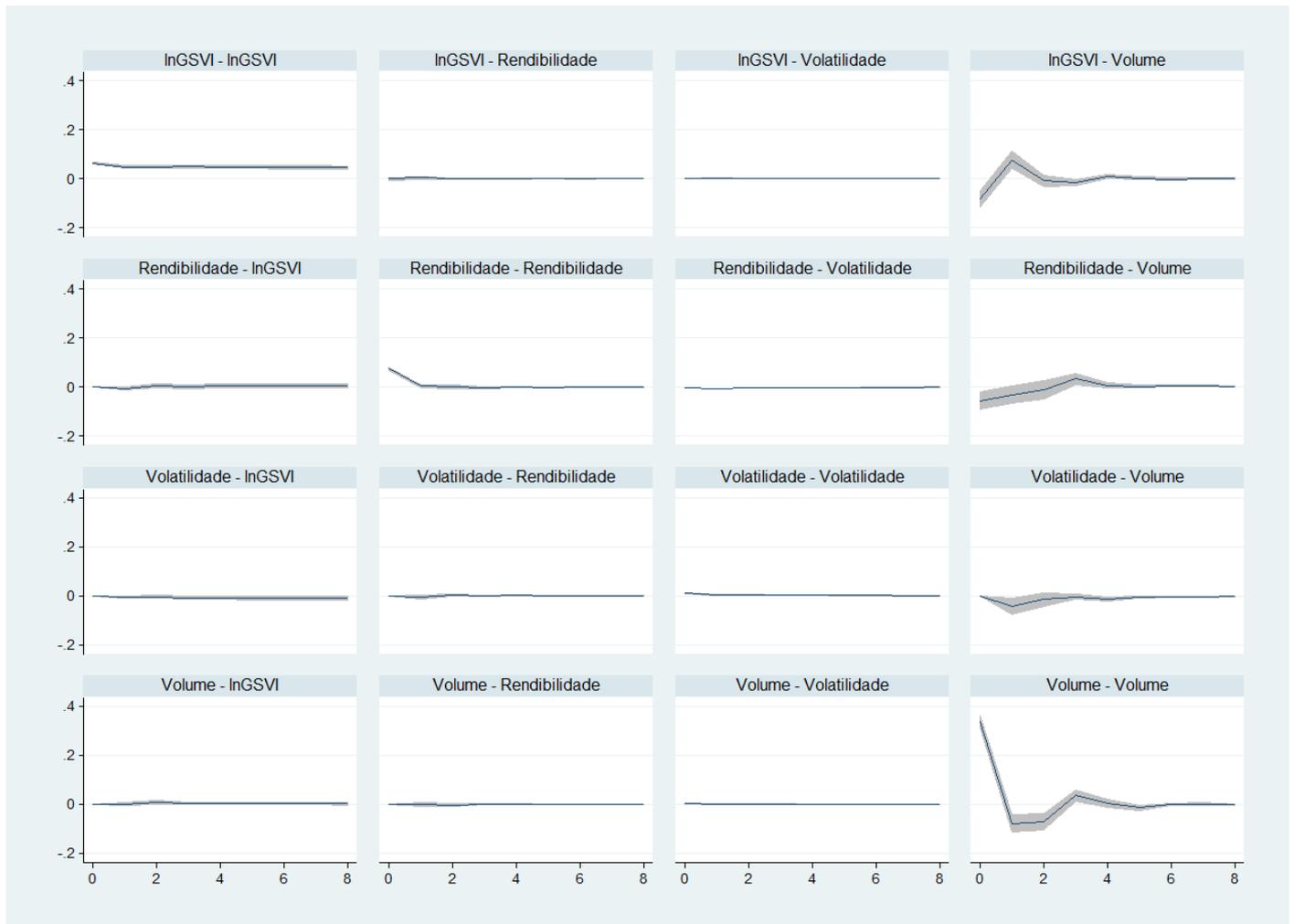


Figura 7.3. Função impulso resposta – Sanofi

Comportamento dinâmico de cada uma das variáveis relativas à Sanofi quando de um choque exógeno sobre cada uma das variáveis endógenas do sistema. “InGSVI-Rendibilidade” indica a resposta da rentabilidade a um impulso na frequência de pesquisas no Google em logaritmo.

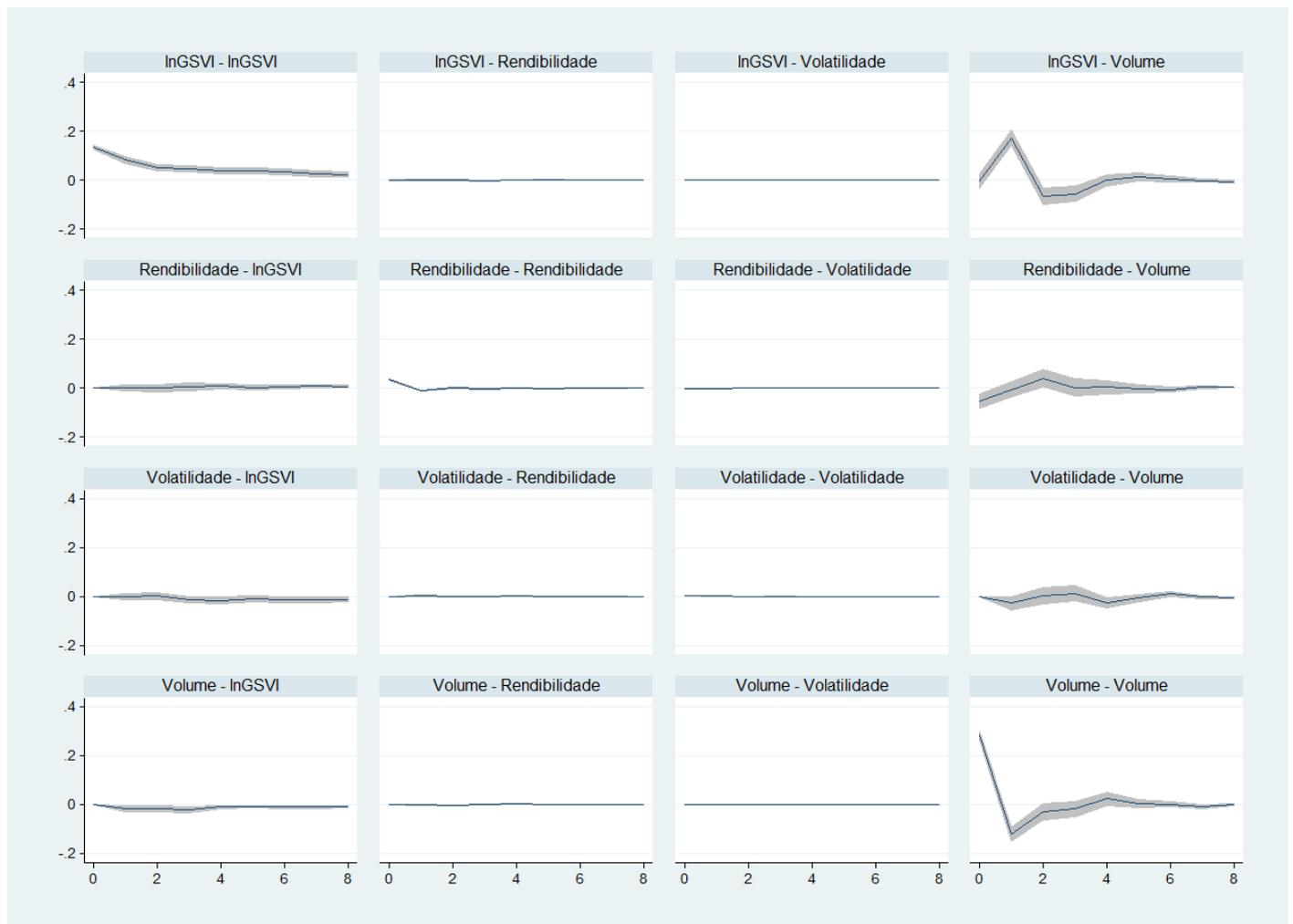
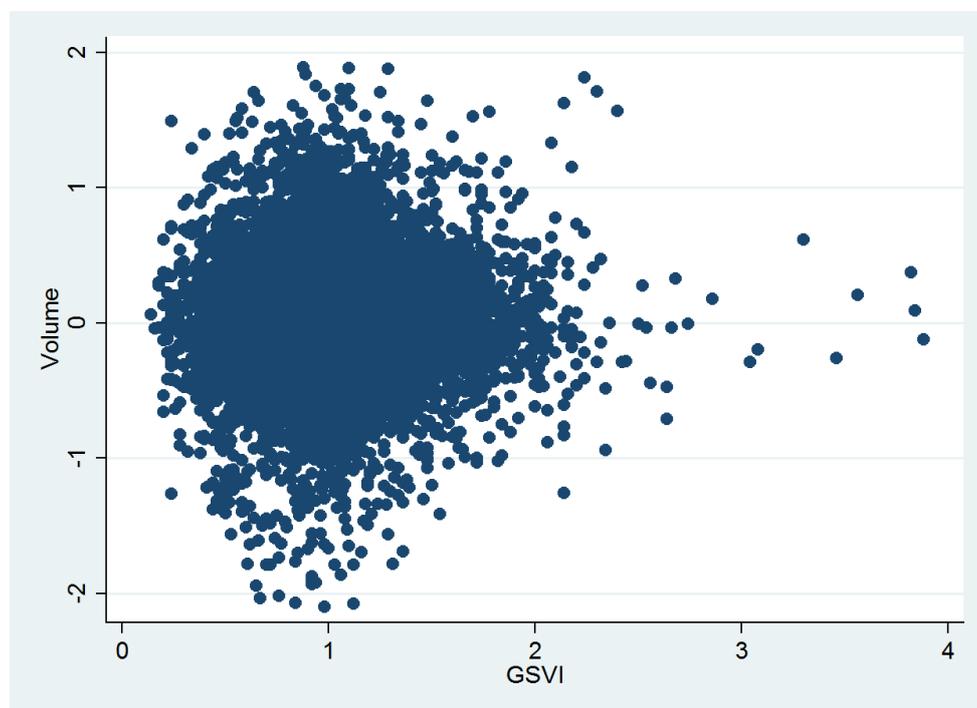


Figura 8.1. Relação entre volume e GSVI

Demonstração gráfica da relação entre as variações logarítmicas do volume e a frequência normal de pesquisas no Google (GSVI) em toda a amostra.



Quadro 8.1. Modelo contemporâneo para o volume com assimetria

Regressão com dados em painel para avaliar efeitos não lineares 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)
$\Delta V_{x,t-1}$	-0.24440*** 0.000	-0.24437*** 0.000	-0.24529*** 0.000	-0.24514*** 0.000
$absr_{x,t}$	2.35753*** 0.000	2.35848*** 0.000	2.36285*** 0.000	2.37068*** 0.000
$RV_{x,t}$	-2.32351 0.163	-2.31374 0.166	-2.36683 0.136	-2.39131 0.129
RS_t	0.00834 0.992	0.00899 0.991	-0.00537 0.995	-0.00294 0.997
RVS_t	0.40759* 0.073	0.40788* 0.072	0.39897* 0.081	0.39771* 0.083
$GSVIp_{x,t}$	0.00493 0.807			
$GSVIn_{x,t}$	0.05410 0.109			
$lnGSVIp_{x,t}$		0.00832 0.758		
$lnGSVIn_{x,t}$		0.03427 0.146		
$JGSVIp_{x,t}$			0.25766*** 0.007	
$JGSVIn_{x,t}$			0.23301*** 0.001	
$lnJGSVIp_{x,t}$				0.45919*** 0.001
$lnJGSVIn_{x,t}$				0.22506*** 0.005
Constante	-0.11725*** 0.000	-0.05971*** 0.000	-0.06449*** 0.000	-0.07324*** 0.000
Observações	14,004	14,004	14,004	14,004
Cotadas	36	36	36	36
R2	0.115	0.115	0.122	0.126

Quadro 8.2. Modelo preditivo para o volume com assimetria

Regressão com dados em painel para avaliar efeitos não lineares 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)
$\Delta V_{x,t-1}$	-0.26905*** 0.000	-0.26883*** 0.000	-0.28506*** 0.000	-0.29320*** 0.000
$R_{x,t-1}$	-0.40392** 0.013	-0.40565** 0.013	-0.42539*** 0.005	-0.40774*** 0.006
$RV_{x,t-1}$	-2.85332** 0.046	-2.83757** 0.046	-2.45995* 0.058	-2.53385** 0.047
RS_{t-1}	2.08356*** 0.000	2.08565*** 0.000	2.04356*** 0.000	1.96519*** 0.000
RVS_{t-1}	0.89369*** 0.000	0.89422*** 0.000	0.87060*** 0.000	0.85322*** 0.000
$GSVIp_{x,t-1}$	0.00557 0.771			
$GSVIn_{x,t-1}$	0.19824*** 0.000			
$lnGSVIp_{x,t-1}$		0.01582 0.523		
$lnGSVIn_{x,t-1}$		0.13316*** 0.000		
$JGSVIp_{x,t-1}$			0.45128*** 0.000	
$JGSVIn_{x,t-1}$			0.81043*** 0.000	
$lnJGSVIp_{x,t-1}$				0.61171*** 0.000
$lnJGSVIn_{x,t-1}$				1.05597*** 0.000
Constante	-0.18444*** 0.000	0.01492*** 0.002	0.02465*** 0.000	0.03013*** 0.000
Observações	14,004	14,004	14,004	14,004
R2	0.100	0.100	0.141	0.163

Quadro 8.3. Modelo contemporâneo 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. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)
$absV_{x,t-1}$	0.22294***	0.22292***	0.22123***	0.21895***
	0.000	0.000	0.000	0.000
$absr_{x,t}$	0.29457	0.29512	0.30349	0.31021
	0.148	0.147	0.132	0.126
$RV_{x,t}$	-0.31425	-0.31192	-0.37947	-0.33858
	0.702	0.705	0.644	0.680
RS_t	0.21557	0.21580	0.24999	0.26786
	0.708	0.708	0.664	0.639
RVS_t	0.14763	0.14759	0.16025	0.16454
	0.317	0.317	0.280	0.264
$GSVIp_{x,t}$	-0.00564			
	0.765			
$GSVIn_{x,t}$	-0.02437			
	0.328			
$lnGSVIp_{x,t}$		-0.00945		
		0.711		
$lnGSVIn_{x,t}$		-0.01715		
		0.342		
$JGSVIp_{x,t}$			0.09467	
			0.126	
$JGSVIn_{x,t}$			-0.18892***	
			0.003	
$lnJGSVIp_{x,t}$				0.17775**
				0.031
$lnJGSVIn_{x,t}$				-0.24130***
				0.001
Constante	0.24726***	0.21779***	0.20791***	0.20291***
	0.000	0.000	0.000	0.000
Observações	14,004	14,004	14,004	14,004
Cotadas	36	36	36	36
R2	0.056	0.056	0.059	0.061

Quadro 8.4. 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				
Cotadas	14,004	14,004	14,004	14,004
R2	0.070	0.070	0.091	0.102

Quadro 8.5. Modelo contemporâneo 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. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)
$RV_{x,t-1}$	0.46413*** 0.000	0.46398*** 0.000	0.46551*** 0.000	0.46565*** 0.000
$absr_{x,t}$	0.05016*** 0.000	0.05015*** 0.000	0.05026*** 0.000	0.05024*** 0.000
$RS_{x,t}$	0.00491 0.437	0.00491 0.438	0.00498 0.431	0.00491 0.435
RV_t	0.00151 0.364	0.00151 0.363	0.00150 0.365	0.00147 0.371
$GSVIp_{x,t}$	0.00038** 0.037			
$GSVIn_{x,t}$	0.00033** 0.037			
$lnGSVIp_{x,t}$		0.00065** 0.021		
$lnGSVIn_{x,t}$		0.00017* 0.077		
$JGSVIp_{x,t}$			0.00096 0.114	
$JGSVIn_{x,t}$			0.00055* 0.056	
$lnJGSVIp_{x,t}$				0.00097 0.110
$lnJGSVIn_{x,t}$				0.00078* 0.052
Constante	-0.00114*** 0.000	-0.00045*** 0.000	-0.00044*** 0.000	-0.00043*** 0.000
Observações	14,004	14,004	14,004	14,004
Cotadas	36	36	36	36
R2	0.459	0.459	0.459	0.459

Quadro 8.6. 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			
$lnGSVIp_{x,t-1}$		0.00012 0.825		
$lnGSVIn_{x,t-1}$		0.00103*** 0.000		
$JGSVIp_{x,t-1}$			0.00168*** 0.003	
$JGSVIn_{x,t-1}$			0.00040* 0.075	
$lnJGSVIp_{x,t-1}$				0.00182*** 0.001
$lnJGSVIn_{x,t-1}$				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

Quadro 8.7. 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			
$\ln GSVIp_{x,t-1}$		0.00290 0.262		
$\ln GSVIn_{x,t-1}$		-0.00455*** 0.007		
$JGSVIp_{x,t-1}$			-0.00957** 0.037	
$JGSVIn_{x,t-1}$			-0.00208 0.544	
$\ln JGSVIp_{x,t-1}$				-0.01532** 0.010
$\ln JGSVIn_{x,t-1}$				-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

Quadro 9.1. Modelo preditivo para o volume com winsorização

*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. Todas as variáveis são winsorizadas. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.26640*** 0.000	-0.26817*** 0.000	-0.26831*** 0.000	-0.29331*** 0.000	-0.26852*** 0.000	-0.26833*** 0.000	-0.28551*** 0.000	-0.29282*** 0.000
$R_{x,t-1}$	-0.41898** 0.012	-0.40415** 0.015	-0.40299** 0.015	-0.42131*** 0.007	-0.40174** 0.015	-0.40350** 0.015	-0.42395*** 0.006	-0.40652*** 0.008
$RV_{x,t-1}$	-2.41873 0.135	-2.91024* 0.081	-3.05365* 0.068	-2.71673* 0.071	-3.18249* 0.058	-3.16522* 0.059	-2.75615* 0.068	-2.82351* 0.058
RS_{t-1}	2.09550*** 0.000	2.06891*** 0.000	2.06110*** 0.000	2.07450*** 0.000	2.05036*** 0.000	2.05266*** 0.000	2.00247*** 0.000	1.94136*** 0.000
RVS_{t-1}	0.89510*** 0.000	0.89558*** 0.000	0.89308*** 0.000	0.89175*** 0.000	0.88737*** 0.000	0.88806*** 0.000	0.86143*** 0.000	0.85012*** 0.000
$GSVIp_{x,t-1}$		0.08285*** 0.000						
$lnGSVIp_{x,t-1}$			0.08679*** 0.000					
$JSVIp_{x,t-1}$				0.67654*** 0.000				
$GSVIp_{x,t-1}$					-0.00385 0.820			
$GSVIn_{x,t-1}$					0.20416*** 0.000			
$lnGSVIp_{x,t-1}$						0.00888 0.707		
$lnGSVIn_{x,t-1}$						0.13644*** 0.000		
$JGSVIp_{x,t-1}$							0.48124*** 0.000	
$JGSVIn_{x,t-1}$							0.87632*** 0.000	
$lnJGSVIp_{x,t-1}$								0.62230*** 0.000
$lnJGSVIn_{x,t-1}$								1.05776*** 0.000
Constante	0.00341* 0.087	-0.07978*** 0.000	0.00520** 0.013	0.00601*** 0.001	-0.17923*** 0.000	0.01615*** 0.001	0.02703*** 0.000	0.02942*** 0.000
Observações	14,004	14,003	14,003	13,894	14,004	14,004	14,004	14,004
R2	0.094	0.098	0.099	0.144	0.100	0.099	0.143	0.161

Quadro 9.2. Modelo preditivo para a volatilidade com winsorização

*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. Todas as variáveis são winsorizadas. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{-x,t-1}$	-0.00036*** 0.000	-0.00037*** 0.000	-0.00037*** 0.000	-0.00037*** 0.000	-0.00037*** 0.000	-0.00037*** 0.000	-0.00039*** 0.000	-0.00040*** 0.000
$R_{x,t-1}$	-0.00808** 0.030	-0.00798** 0.031	-0.00797** 0.031	-0.00797** 0.031	-0.00795** 0.032	-0.00796** 0.032	-0.00810** 0.029	-0.00809** 0.029
$RV_{x,t-1}$	0.57288*** 0.000	0.56943*** 0.000	0.56813*** 0.000	0.56813*** 0.000	0.56740*** 0.000	0.56731*** 0.000	0.57192*** 0.000	0.57242*** 0.000
RS_{t-1}	-0.01925*** 0.000	-0.01944*** 0.000	-0.01951*** 0.000	-0.01951*** 0.000	-0.01958*** 0.000	-0.01957*** 0.000	-0.01918*** 0.000	-0.01925*** 0.000
RVS_{t-1}	-0.00110* 0.097	-0.00110* 0.094	-0.00112* 0.090	-0.00112* 0.090	-0.00116* 0.076	-0.00115* 0.077	-0.00108* 0.097	-0.00110* 0.093
$GSVIp_{x,t-1}$		0.00058** 0.045						
$lnGSVIP_{x,t-1}$			0.00065** 0.017					
$JSVIP_{x,t-1}$				0.00065** 0.017				
$GSVIP_{x,t-1}$					-0.00007 0.822			
$GSVIn_{x,t-1}$					0.00149*** 0.000			
$lnGSVIP_{x,t-1}$						0.00006 0.908		
$lnGSVIn_{x,t-1}$						0.00102*** 0.000		
$JGSVIP_{x,t-1}$							0.00193*** 0.003	
$JGSVIn_{x,t-1}$							0.00035 0.200	
$lnJGSVIP_{x,t-1}$								0.00191*** 0.002
$lnJGSVIn_{x,t-1}$								0.00052** 0.031
Constante	0.00051*** 0.000	-0.00007 0.801	0.00052*** 0.000	0.00052*** 0.000	-0.00078 0.158	0.00060*** 0.000	0.00044*** 0.000	0.00044*** 0.000
Observações	14,004	14,003	14,003	14,003	14,004	14,004	14,004	14,004
R2	0.409	0.410	0.411	0.411	0.411	0.411	0.410	0.410

Quadro 9.3. Modelo preditivo para a rentabilidade com winsorização

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

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00194*** 0.005	-0.00192*** 0.006	-0.00191*** 0.006	-0.00230*** 0.001	-0.00189*** 0.007	-0.00190*** 0.006	-0.00177** 0.012	-0.00162** 0.021
$R_{x,t-1}$	-0.07356*** 0.002	-0.07379*** 0.002	-0.07384*** 0.001	-0.07317*** 0.002	-0.07394*** 0.001	-0.07386*** 0.001	-0.07343*** 0.002	-0.07349*** 0.001
$RV_{x,t-1}$	-1.03458*** 0.000	-1.02665*** 0.001	-1.02284*** 0.001	-1.02581*** 0.000	-1.01331*** 0.001	-1.01603*** 0.001	-1.02958*** 0.000	-1.03096*** 0.000
RS_{t-1}	-0.10822** 0.011	-0.10780** 0.012	-0.10759** 0.012	-0.10373** 0.015	-0.10687** 0.012	-0.10706** 0.012	-0.10851** 0.011	-0.10822** 0.011
RVS_{t-1}	-0.01813*** 0.002	-0.01813*** 0.002	-0.01809*** 0.002	-0.01609*** 0.007	-0.01774*** 0.003	-0.01779*** 0.003	-0.01819*** 0.002	-0.01812*** 0.002
$GSVIp_{x,t-1}$		-0.00133 0.134						
$lnGSVIp_{x,t-1}$			-0.00160** 0.047					
$JSVIp_{x,t-1}$				-0.00614** 0.016				
$GSVIp_{x,t-1}$					0.00294 0.146			
$GSVIn_{x,t-1}$					-0.00731*** 0.003			
$lnGSVIp_{x,t-1}$						0.00322 0.216		
$lnGSVIn_{x,t-1}$						-0.00467*** 0.005		
$JGSVIp_{x,t-1}$							-0.00981** 0.044	
$JGSVIn_{x,t-1}$							-0.00259 0.497	
$lnJGSVIp_{x,t-1}$								-0.01484** 0.012
$lnJGSVIn_{x,t-1}$								-0.00419 0.231
Constante	0.00302*** 0.000	0.00436*** 0.000	0.00299*** 0.000	0.00305*** 0.000	0.00651*** 0.000	0.00231*** 0.000	0.00334*** 0.000	0.00351*** 0.000
Observações	14,004	14,003	14,003	13,894	14,004	14,004	14,004	14,004
R2	0.022	0.022	0.022	0.022	0.022	0.022	0.022	0.022

Quadro 9.4. Modelo preditivo para o volume até Novembro de 2007

*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, entre o início de 2004 e 16 de Novembro de 2007. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.26054*** 0.000	-0.26617*** 0.000	-0.26526*** 0.000	-0.29446*** 0.000	-0.26563*** 0.000	-0.26499*** 0.000	-0.27786*** 0.000	-0.28098*** 0.000
$R_{x,t-1}$	-0.56066** 0.013	-0.51814** 0.019	-0.50628** 0.023	-0.56966*** 0.009	-0.50251** 0.025	-0.50907** 0.023	-0.59530*** 0.007	-0.56123*** 0.010
$RV_{x,t-1}$	-30.13661** 0.021	-29.91716** 0.025	-30.70503** 0.020	-26.59326** 0.037	-30.87831** 0.019	-30.72184** 0.019	-27.78366** 0.025	-27.79341** 0.025
RS_{t-1}	3.02883*** 0.000	2.93158*** 0.000	2.84486*** 0.000	3.07434*** 0.000	2.82227*** 0.000	2.84354*** 0.000	2.86400*** 0.000	2.85065*** 0.000
RVS_{t-1}	0.77570*** 0.000	0.77744*** 0.000	0.76549*** 0.000	0.83702*** 0.000	0.76337*** 0.000	0.76460*** 0.000	0.78753*** 0.000	0.80994*** 0.000
$GSVIp_{x,t-1}$		0.17895*** 0.000						
$lnGSVIp_{x,t-1}$			0.18341*** 0.000					
$JSVIp_{x,t-1}$				0.66548*** 0.000				
$GSVIp_{x,t-1}$					0.12241*** 0.006			
$GSVIn_{x,t-1}$					0.31346*** 0.000			
$lnGSVIp_{x,t-1}$						0.18132*** 0.001		
$lnGSVIn_{x,t-1}$						0.18492*** 0.001		
$JGSVIp_{x,t-1}$							0.51147*** 0.000	
$JGSVIn_{x,t-1}$							0.84743*** 0.000	
$lnJGSVIp_{x,t-1}$								0.62995*** 0.000
$lnJGSVIn_{x,t-1}$								0.99453*** 0.000
Constante	0.01409* 0.093	-0.19092*** 0.000	-0.00796 0.387	0.01378* 0.098	-0.43565*** 0.000	-0.00761 0.452	0.03646*** 0.000	0.03561*** 0.001
Observações	7,200	7,199	7,199	7,090	7,200	7,200	7,200	7,200
R2	0.093	0.103	0.103	0.142	0.104	0.103	0.138	0.148

Quadro 9.5. Modelo preditivo para o volume após Novembro de 2007

*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, entre 30 de Novembro de 2007 e o final de Junho de 2011. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.25960*** 0.000	-0.26305*** 0.000	-0.26591*** 0.000	-0.28179*** 0.000	-0.26897*** 0.000	-0.26845*** 0.000	-0.28144*** 0.000	-0.29599*** 0.000
$R_{x,t-1}$	-0.33727** 0.047	-0.33167** 0.049	-0.33545** 0.046	-0.33789** 0.036	-0.34213** 0.043	-0.34241** 0.043	-0.33826** 0.036	-0.33251** 0.036
$RV_{x,t-1}$	-2.12305 0.138	-2.54625* 0.056	-2.83213** 0.032	-2.68928** 0.040	-2.95028** 0.030	-2.94566** 0.030	-2.67105** 0.043	-2.78891** 0.030
RS_{t-1}	2.43272*** 0.000	2.42376*** 0.000	2.42613*** 0.000	2.23075*** 0.000	2.43627*** 0.000	2.43523*** 0.000	2.21756*** 0.000	2.06801*** 0.000
RVS_{t-1}	1.18061*** 0.000	1.17738*** 0.000	1.17173*** 0.000	1.08246*** 0.000	1.16158*** 0.000	1.16427*** 0.000	1.07593*** 0.000	1.01176*** 0.000
$GSVIp_{x,t-1}$		0.21101*** 0.000						
$lnGSVIp_{x,t-1}$			0.27602*** 0.000					
$JSVIp_{x,t-1}$				0.68052*** 0.000				
$GSVIp_{x,t-1}$					0.02847 0.379			
$GSVIn_{x,t-1}$					0.59739*** 0.000			
$lnGSVIp_{x,t-1}$						0.07870 0.136		
$lnGSVIn_{x,t-1}$						0.40413*** 0.000		
$JGSVIp_{x,t-1}$							0.59695*** 0.000	
$JGSVIn_{x,t-1}$							0.74413*** 0.000	
$lnJGSVIp_{x,t-1}$								0.82610*** 0.000
$lnJGSVIn_{x,t-1}$								1.03086*** 0.000
Constante	0.00661** 0.015	-0.17418*** 0.000	0.05296*** 0.000	0.01290*** 0.000	-0.52796*** 0.000	0.07738*** 0.000	0.01887*** 0.004	0.02489*** 0.000
Observações	6,696	6,696	6,696	6,696	6,696	6,696	6,696	6,696
R2	0.107	0.117	0.123	0.159	0.127	0.127	0.159	0.191

Quadro 9.6. Modelo preditivo para a volatilidade até Novembro de 2007

*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, entre o início de 2004 e 16 de Novembro de 2007. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00004	-0.00004*	-0.00004*	-0.00006**	-0.00004*	-0.00004	-0.00005*	-0.00005*
	0.119	0.088	0.086	0.039	0.086	0.103	0.050	0.053
$R_{x,t-1}$	-0.00062	-0.00059	-0.00058	-0.00058	-0.00055	-0.00055	-0.00070	-0.00066
	0.282	0.305	0.317	0.333	0.335	0.334	0.223	0.245
$RV_{x,t-1}$	0.20489***	0.20514***	0.20431***	0.20617***	0.20340***	0.20212***	0.20884***	0.20809***
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RS_{t-1}	-0.00285***	-0.00294***	-0.00305***	-0.00306***	-0.00314***	-0.00321***	-0.00287***	-0.00290***
	0.007	0.006	0.005	0.004	0.004	0.004	0.007	0.006
RVS_{t-1}	0.00027*	0.00027*	0.00026	0.00030*	0.00025	0.00024	0.00027*	0.00028*
	0.094	0.099	0.119	0.071	0.132	0.153	0.093	0.087
$GSVIP_{x,t-1}$		0.00017						
		0.228						
$lnGSVIP_{x,t-1}$			0.00020*					
			0.096					
$JSVIP_{x,t-1}$				0.00056**				
				0.020				
$GSVIP_{x,t-1}$					0.00007			
					0.707			
$GSVIN_{x,t-1}$					0.00041**			
					0.016			
$lnGSVIP_{x,t-1}$						0.00002		
						0.946		
$lnGSVIN_{x,t-1}$						0.00033***		
						0.002		
$JGSVIP_{x,t-1}$							0.00106**	
							0.033	
$JGSVIN_{x,t-1}$							-0.00006	
							0.714	
$lnJGSVIP_{x,t-1}$								0.00096**
								0.047
$lnJGSVIN_{x,t-1}$								0.00000
								0.998
Constante	0.00051***	0.00031*	0.00049***	0.00051***	0.00002	0.00052***	0.00045***	0.00046***
	0.000	0.067	0.000	0.000	0.927	0.000	0.000	0.000
Observações								
Observações	7,200	7,199	7,199	7,090	7,200	7,200	7,200	7,200
R2	0.086	0.088	0.088	0.092	0.089	0.089	0.096	0.092

Quadro 9.7. Modelo preditivo para a volatilidade após Novembro de 2007

*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, entre 30 de Novembro de 2007 e o final de Junho de 2011. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00055*** 0.002	-0.00057*** 0.002	-0.00058*** 0.001	-0.00060*** 0.001	-0.00059*** 0.001	-0.00061*** 0.001	-0.00005* 0.050	-0.00064*** 0.001
$R_{x,t-1}$	-0.01435** 0.047	-0.01432** 0.048	-0.01434** 0.048	-0.01435** 0.047	-0.01437** 0.047	-0.01435** 0.047	-0.00070 0.223	-0.01435** 0.047
$RV_{x,t-1}$	0.52447*** 0.000	0.52241*** 0.000	0.52083*** 0.000	0.52306*** 0.000	0.52053*** 0.000	0.52279*** 0.000	0.20884*** 0.000	0.52312*** 0.000
RS_{t-1}	-0.01310** 0.020	-0.01314** 0.020	-0.01313** 0.019	-0.01360** 0.016	-0.01309** 0.020	-0.01341** 0.018	-0.00287*** 0.007	-0.01351** 0.017
RVS_{t-1}	0.00186** 0.024	0.00184** 0.024	0.00181** 0.026	0.00162** 0.049	0.00177** 0.029	0.00171** 0.037	0.00027* 0.093	0.00166** 0.045
$GSVIp_{x,t-1}$		0.00103** 0.043		0.00169*** 0.000				
$\ln GSVIp_{x,t-1}$			0.00142*** 0.002					
$JSVIp_{x,t-1}$				0.00112*** 0.000				
$GSVIp_{x,t-1}$					0.00018 0.782			
$GSVIn_{x,t-1}$					0.00282*** 0.000			
$\ln GSVIp_{x,t-1}$						0.00294*** 0.007		
$\ln GSVIn_{x,t-1}$						0.00075** 0.043		
$JGSVIp_{x,t-1}$							0.00106** 0.033	
$JGSVIn_{x,t-1}$							-0.00006 0.714	
$\ln JGSVIp_{x,t-1}$								0.00350*** 0.005
$\ln JGSVIn_{x,t-1}$								0.00072* 0.057
Constante	0.00079*** 0.000	-0.00009 0.833	0.00103*** 0.000	0.00081*** 0.000	-0.00178** 0.012	0.00112*** 0.000	0.00072*** 0.000	0.00067*** 0.000
Observações	6,696	6,696	6,696	6,696	6,696	6,696	6,696	6,696
R2	0.399	0.400	0.401	0.400	0.401	0.401	0.401	0.401

Quadro 9.8. Modelo preditivo para a rentabilidade até Novembro de 2007

Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre a rentabilidade, a variável dependente, na semana seguinte, entre o início de 2004 e 16 de Novembro de 2007. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00049	-0.00040	-0.00043	-0.00142*	-0.00041	-0.00041	-0.00052	-0.00055
	0.484	0.567	0.540	0.052	0.561	0.562	0.452	0.427
$R_{x,t-1}$	-0.01801	-0.01868	-0.01870	-0.01420	-0.01845	-0.01835	-0.01829	-0.01854
	0.185	0.168	0.168	0.318	0.171	0.174	0.181	0.173
$RV_{x,t-1}$	-1.03625*	-1.03935*	-1.02947*	-0.98197*	-1.04669*	-1.06203*	-1.02250*	-1.00680*
	0.079	0.074	0.077	0.096	0.070	0.067	0.081	0.087
RS_{t-1}	-0.33834***	-0.33690***	-0.33618***	-0.32340***	-0.33772***	-0.33855***	-0.33817***	-0.33808***
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RVS_{t-1}	-0.05637***	-0.05641***	-0.05627***	-0.05181***	-0.05649***	-0.05659***	-0.05639***	-0.05647***
	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$GSVIp_{x,t-1}$		-0.00263*						
		0.066						
$lnGSVIp_{x,t-1}$			-0.00214					
			0.103					
$JSVIp_{x,t-1}$				0.00165				
				0.638				
$GSVIp_{x,t-1}$					-0.00305*			
					0.073			
$GSVIn_{x,t-1}$					-0.00165			
					0.535			
$lnGSVIp_{x,t-1}$						-0.00487**		
						0.045		
$lnGSVIn_{x,t-1}$						-0.00025		
						0.863		
$JGSVIp_{x,t-1}$							0.00386	
							0.603	
$JGSVIn_{x,t-1}$							-0.00155	
							0.721	
$lnJGSVIp_{x,t-1}$								0.00920
								0.242
$lnJGSVIn_{x,t-1}$								-0.00481
								0.278
Constante	0.00392***	0.00694***	0.00418***	0.00399***	0.00909***	0.00462***	0.00362***	0.00322***
	0.000	0.000	0.000	0.000	0.008	0.000	0.000	0.000
Observações								
Observações	7,200	7,199	7,199	7,090	7,200	7,200	7,200	7,200
R2	0.023	0.023	0.023	0.021	0.023	0.023	0.023	0.023

Quadro 9.9. Modelo preditivo para a rentabilidade após Novembro de 2007

Regressão com dados em painel para avaliar o efeito da frequência de pesquisas no Google sobre a rentabilidade, a variável dependente., na semana seguinte, entre 30 de Novembro de 2007 e o final de Junho de 2011. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00055*** 0.002	-0.00057*** 0.002	-0.00058*** 0.001	-0.00060*** 0.001	-0.00059*** 0.001	-0.00061*** 0.001	-0.00005* 0.050	-0.00064*** 0.001
$R_{x,t-1}$	-0.01435** 0.047	-0.01432** 0.048	-0.01434** 0.048	-0.01435** 0.047	-0.01437** 0.047	-0.01435** 0.047	-0.00070 0.223	-0.01435** 0.047
$RV_{x,t-1}$	0.52447*** 0.000	0.52241*** 0.000	0.52083*** 0.000	0.52306*** 0.000	0.52053*** 0.000	0.52279*** 0.000	0.20884*** 0.000	0.52312*** 0.000
RS_{t-1}	-0.01310** 0.020	-0.01314** 0.020	-0.01313** 0.019	-0.01360** 0.016	-0.01309** 0.020	-0.01341** 0.018	-0.00287*** 0.007	-0.01351** 0.017
RVS_{t-1}	0.00186** 0.024	0.00184** 0.024	0.00181** 0.026	0.00162** 0.049	0.00177** 0.029	0.00171** 0.037	0.00027* 0.093	0.00166** 0.045
$GSVIp_{x,t-1}$		0.00103** 0.043		0.00169*** 0.000				
$\ln GSVIp_{x,t-1}$			0.00142*** 0.002					
$JSVIp_{x,t-1}$				0.00112*** 0.000				
$GSVIp_{x,t-1}$					0.00018 0.782			
$GSVIn_{x,t-1}$					0.00282*** 0.000			
$\ln GSVIp_{x,t-1}$						0.00294*** 0.007		
$\ln GSVIn_{x,t-1}$						0.00075** 0.043		
$JGSVIp_{x,t-1}$							0.00106** 0.033	
$JGSVIn_{x,t-1}$							-0.00006 0.714	
$\ln JGSVIp_{x,t-1}$								0.00350*** 0.005
$\ln JGSVIn_{x,t-1}$								0.00072* 0.057
Constante	0.00079*** 0.000	-0.00009 0.833	0.00103*** 0.000	0.00081*** 0.000	-0.00178** 0.012	0.00112*** 0.000	0.00072*** 0.000	0.00067*** 0.000
Observações								
Observações	6,696	6,696	6,696	6,696	6,696	6,696	6,696	6,696
R2	0.399	0.400	0.401	0.400	0.401	0.401	0.401	0.401

Quadro 9.10. Modelo preditivo para o volume com dummies

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. Incluem-se dummies que identificam países e sectores de actividade. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.26683*** 0.000	-0.26870*** 0.000	-0.26882*** 0.000	-0.29294*** 0.000	-0.26905*** 0.000	-0.26883*** 0.000	-0.28506*** 0.000	-0.29320*** 0.000
$R_{x,t-1}$	-0.41768** 0.010	-0.40530** 0.012	-0.40465** 0.013	-0.42260*** 0.006	-0.40392** 0.013	-0.40565** 0.013	-0.42539*** 0.005	-0.40774*** 0.006
$RV_{x,t-1}$	-2.19704 0.112	-2.63912* 0.063	-2.75673* 0.053	-2.42596* 0.061	-2.85332** 0.046	-2.83757** 0.046	-2.45995* 0.058	-2.53385** 0.047
RS_{t-1}	2.12227*** 0.000	2.09924*** 0.000	2.09242*** 0.000	2.11408*** 0.000	2.08356*** 0.000	2.08565*** 0.000	2.04356*** 0.000	1.96519*** 0.000
RVS_{t-1}	0.90059*** 0.000	0.90138*** 0.000	0.89883*** 0.000	0.90110*** 0.000	0.89369*** 0.000	0.89422*** 0.000	0.87060*** 0.000	0.85322*** 0.000
$GSVIp_{x,t-1}$		0.08405*** 0.000						
$lnGSVIp_{x,t-1}$			0.08737*** 0.000					
$JSVIp_{x,t-1}$				0.62419*** 0.000				
$GSVIp_{x,t-1}$					0.00557 0.771			
$GSVIn_{x,t-1}$					0.19824*** 0.000			
$lnGSVIp_{x,t-1}$						0.01582 0.523		
$lnGSVIn_{x,t-1}$						0.13316*** 0.000		
$JGSVIp_{x,t-1}$							0.45128*** 0.000	
$JGSVIn_{x,t-1}$							0.81043*** 0.000	
$lnJGSVIp_{x,t-1}$								0.61171*** 0.000
$lnJGSVIn_{x,t-1}$								1.05597*** 0.000
Constante	0.00390 0.105	-0.08016*** 0.000	0.01572*** 0.000	0.00664*** 0.003	-0.17900*** 0.000	0.01492*** 0.002	0.03041*** 0.000	0.03511*** 0.000
Observações	14,004	14,003	14,003	13,894	14,004	14,004	14,004	14,004
R2	0.094	0.098	0.099	0.142	0.100	0.100	0.141	0.163

Quadro 9.11. Modelo preditivo para a volatilidade com dummies

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. Incluem-se dummies que identificam países e sectores de actividade. Todas as variáveis são winsorizadas. Os símbolos *, ** e *** denotam níveis de significância de 10%, 5% e 1%, respectivamente.

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00037*** 0.000	-0.00038*** 0.000	-0.00038*** 0.000	-0.00040*** 0.000	-0.00039*** 0.000	-0.00039*** 0.000	-0.00040*** 0.000	-0.00041*** 0.000
$R_{x,t-1}$	-0.01078* 0.050	-0.01070* 0.052	-0.01069* 0.052	-0.01082** 0.050	-0.01068* 0.053	-0.01068* 0.052	-0.01081** 0.050	-0.01079** 0.050
$RV_{x,t-1}$	0.56709*** 0.000	0.56392*** 0.000	0.56275*** 0.000	0.56644*** 0.000	0.56222*** 0.000	0.56214*** 0.000	0.56635*** 0.000	0.56668*** 0.000
RS_{t-1}	-0.01990*** 0.000	-0.02006*** 0.000	-0.02013*** 0.000	-0.01994*** 0.000	-0.02019*** 0.000	-0.02019*** 0.000	-0.01983*** 0.000	-0.01991*** 0.000
RVS_{t-1}	-0.00118* 0.091	-0.00118* 0.090	-0.00120* 0.086	-0.00121* 0.087	-0.00124* 0.073	-0.00123* 0.074	-0.00116* 0.092	-0.00119* 0.088
$GSVIp_{x,t-1}$		0.00060** 0.045						
$lnGSVIp_{x,t-1}$			0.00067** 0.017					
$JSVIp_{x,t-1}$				0.00107*** 0.000				
$GSVIp_{x,t-1}$					-0.00003 0.926			
$GSVIn_{x,t-1}$					0.00151*** 0.000			
$lnGSVIp_{x,t-1}$						0.00012 0.825		
$lnGSVIn_{x,t-1}$						0.00103*** 0.000		
$JGSVIp_{x,t-1}$							0.00168*** 0.003	
$JGSVIn_{x,t-1}$							0.00040* 0.075	
$lnJGSVIp_{x,t-1}$								0.00182*** 0.001
$lnJGSVIn_{x,t-1}$								0.00057*** 0.006
Constante	0.00074*** 0.000	0.00014 0.649	0.00105*** 0.000	0.00074*** 0.000	-0.00058 0.304	0.00061*** 0.000	0.00066*** 0.000	0.00067*** 0.000
Observações	14,004	14,003	14,003	13,894	14,004	14,004	14,004	14,004
R2	0.406	0.408	0.408	0.407	0.408	0.408	0.407	0.407

Quadro 9.12. Modelo preditivo para a rentabilidade com dummies

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

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta V_{x,t-1}$	-0.00197*** 0.004	-0.00194*** 0.005	-0.00193*** 0.005	-0.00233*** 0.001	-0.00192*** 0.005	-0.00192*** 0.005	-0.00180*** 0.009	-0.00165** 0.017
$R_{x,t-1}$	-0.07210*** 0.002	-0.07229*** 0.002	-0.07233*** 0.002	-0.07167*** 0.002	-0.07239*** 0.002	-0.07231*** 0.002	-0.07193*** 0.002	-0.07198*** 0.001
$RV_{x,t-1}$	-0.93721*** 0.002	-0.92997*** 0.002	-0.92667*** 0.002	-0.92965*** 0.002	-0.91924*** 0.002	-0.92161*** 0.002	-0.93301*** 0.002	-0.93405*** 0.002
RS_{t-1}	-0.11332*** 0.008	-0.11295*** 0.008	-0.11277*** 0.008	-0.10896** 0.010	-0.11215*** 0.008	-0.11232*** 0.008	-0.11377*** 0.007	-0.11349*** 0.007
RVS_{t-1}	-0.01904*** 0.002	-0.01905*** 0.002	-0.01900*** 0.002	-0.01703*** 0.007	-0.01868*** 0.003	-0.01873*** 0.003	-0.01916*** 0.002	-0.01908*** 0.002
$GSVIp_{x,t-1}$		-0.00137 0.132						
$lnGSVIp_{x,t-1}$			-0.00164** 0.049					
$JSVIp_{x,t-1}$				-0.00589** 0.013				
$GSVIp_{x,t-1}$					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	0.00341*** 0.000	0.00478*** 0.000	0.00271*** 0.000	0.00341*** 0.000	0.00690*** 0.000	0.00224*** 0.000	0.00383*** 0.000	0.00404*** 0.000
Observações	14,004	14,003	14,003	13,894	14,004	14,004	14,004	14,004
R2	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.022