



**LISBOA  
SCHOOL OF  
ECONOMICS &  
MANAGEMENT**

**MESTRADO**

**ECONOMETRIA APLICADA E PREVISÃO**

**EFEITOS DO CONSUMO DIGITAL NA IMPRENSA  
ESCRITA: EVIDÊNCIA EMPÍRICA PARA OS  
JORNALIS DIÁRIOS PORTUGUESES**

**NUNO DUARTE FIALHO SANCHES BORGES DOS  
SANTOS**

**OUTUBRO – 2016**



**LISBOA  
SCHOOL OF  
ECONOMICS &  
MANAGEMENT**

**MESTRADO**

**ECONOMETRIA APLICADA E PREVISÃO**

**EFEITOS DO CONSUMO DIGITAL NA IMPRENSA  
ESCRITA: EVIDÊNCIA EMPÍRICA PARA OS  
JORNALIS DIÁRIOS PORTUGUESES**

NUNO DUARTE FIALHO SANCHES BORGES DOS  
SANTOS

**ORIENTAÇÃO:**

PROFESSOR DOUTOR ANÍBAL JORGE DA COSTA  
CRISTÓVÃO CAIADO

OUTUBRO – 2016

## Resumo

O mercado dos media em geral é um dos sectores mais sensível ao choque tecnológico da digitalização.

No caso particular da imprensa escrita, as investigações existentes, não sendo concordantes, apontam pistas para a existência de efeitos positivos (ou de complementaridade), negativos (ou de substituição) ou mesmo neutrais entre os consumos digital e impressos. Dada a escassez de dados e investigações no mercado Português, o presente trabalho constitui uma investigação de natureza assumidamente empírica na pesquisa de efeitos de causalidade entre o consumo digital e o consumo em papel no mercado da imprensa escrita diária em Portugal.

Na perspectiva de cariz aplicado da investigação, a investigação adopta a metodologia VAR como veículo para o estudo da (não-) causalidade *à Granger*. No caso, tanto numa perspectiva individual através de modelos VAR independentes para cada um dos títulos, bem como numa perspectiva global e conjunta com modelos VAR para painéis de diferentes segmentos de mercado (mercado total, desporto, economia e negócios e informação geral).

Um dos pontos centrais da investigação prende-se com a utilização de variáveis exógenas. A incorporação destes factores revela-se particularmente importante na eliminação de efeitos e causalidades espúrias.

Ao não encontrar fortes evidências estatísticas de causalidade e de dinâmicas de longo-prazo, a investigação acaba por não partilhar o pessimismo dos que defendem a canibalização das edições impressas pelas edições digitais (ou o optimismo dos que defendem a sua complementaridade), orientando antes para a importância e relevância dos conteúdos e dos seus diferentes efeitos nos dois tipos de consumo.

**Palavras-chave:** VAR; Raízes Unitárias; Cointegração, Media; Imprensa; Jornais; Digital; Causalidade.

### **Abstract**

The media market is one of the most sensitive industries to the technological shock of digitalization.

In the case of print media, the existing research, not being consistent, point clues to the existence of positive effects (or complementary), negative (or substitution) or even neutral between digital and print consumption. Given the lack of data and research in the Portuguese market, this work is an assumedly empirical research of causal effects between digital and print consumption in the market of daily newspapers in Portugal.

In the applied-oriented research perspective, the research adopts the VAR methodology as a vehicle for the study of Granger (non-)causality. Both from an individual perspective through independent VAR models for each of the titles, as well as a general and joint approach with VAR models for panels of different market segments (total market, sports, economics and business and general information).

One of the central points of the research is related to the use of exogenous variables. The incorporation of these factors is particularly important in eliminating spurious effects and causalities.

Not finding strong statistical evidence of causality and long-term dynamics, the research does not share the pessimism of those who defend the cannibalization of printed editions by digital editions (or the optimism of those who defend their complementarity), guiding instead to the importance and relevance of content and their different effects on both types of consumption.

**Key-words: VAR; Unit-roots; Cointegration, Media; Print; Newspapers; Digital; Causality.**

### **Agradecimentos:**

A todo o corpo docente do Mestrado em Econometria Aplicada e Previsão. São fantásticos e conquistaram um amigo e acérrimo adepto da equipa e do programa.

Muito em particular à Professora Isabel Proença, e ao Professor orientador Jorge Caiado. As suas ajuda, apoio, disponibilidade, compreensão e ensinamentos foram centrais para a execução do trabalho.

Ao Professor Artur Silva Lopes. Pela sua personalidade e oportunidade de ter uma experiência única no processo de (como ele diz) nos “arturarmos” mutuamente.

A todos os parceiros e amigos. Particularmente , à Cristina Soares, ao Diogo Homem, ao Edgar Madeira, ao Jorge Graça, ao Luís Amorim, ao Marco Gonçalves, ao Paulo Cardoso, ao Pedro Araújo e Sá, ao Raul Carvalho das Neves, ao Vasco Faustino, e a tantos outros e todos aqueles que ao longo das nossas relações foram e são uma fonte de aprendizagem e ensinamentos constantes.

À Maria do Carmo Menano e à APCT, à Maria João Valente Rosa e à FMMS pela cordial e pronta disponibilidade de informação.

A todos os colegas de profissão e à PSE sem os quais esta oportunidade não existia.

Ao Zé Borges e à Maria Luisa, ao Zé Carvalho e à Maria Fernanda, por serem como são, e nos fazerem como somos.

Às Madalenas, à Rita, à Maria e à Inês.

À Rita e ao João. São a força, a orientação e a motivação das aventuras em que me meto. Eu sou eles.

## Índice

1 – Introdução .....	1
1.1 – Enquadramento e Objectivo .....	1
1.2 – Visão global da imprensa escrita .....	3
1.3 – Jornais diários em Portugal .....	5
2 – Resumo da Literatura .....	7
2.1 - Efeitos de substituição .....	7
2.2 – Independência ou ausência/neutralidade de efeitos .....	10
2.3 – Efeitos de complementaridade .....	11
2.4 – Principais limitações .....	13
3 – Aplicação Empírica .....	14
3.1 – Os Dados .....	14
3.2 – Metodologia .....	15
3.3 – Resultados .....	17
4 – Conclusões .....	31
Referências .....	34
Apêndice I – Gráficos e Figuras .....	39
Apêndice II – Testes de Raízes Unitárias e Cointegração .....	41
Apêndice III – Sumário dos VAR Independentes .....	47
Apêndice IV – Glossário de Variáveis .....	48

# 1 – Introdução

## 1.1 – Enquadramento e Objectivo

Um dos sectores de actividade onde a digitalização está mais na agenda é o sector de media, e em particular, a imprensa escrita. A análise dos impactos da digitalização da economia neste sector tem suscitado profundo debate e discordância. Perante o decréscimo de vendas em banca e circulação, é comum o sentimento da existência de um efeito negativo (e de canibalização) do consumo digital no consumo tradicional. Porém, as diferentes evidências empíricas não são concordantes, tendo sido identificados em diferentes mercados efeitos positivos, negativos e neutrais, quer do consumo digital no consumo tradicional, como vice-versa. Neste contexto, a presente pesquisa, sendo assumidamente uma pesquisa de cariz aplicado, pretende analisar as relações de causalidade entre os consumos digital e tradicional no mercado nacional das publicações diárias.

A pesquisa surge como uma sequência natural e tem como motivação trabalho prévio desenvolvido ao longo da disciplina de Macroeconometria II para o segmento dos diários desportivos nacionais. Neste trabalho, não foi identificada nenhuma relação de causalidade entre o consumo digital e o consumo em papel. Se esse trabalho prévio explora as relações de causalidade individualmente para cada um dos títulos no segmento desportivo, o presente trabalho alarga o âmbito ao universo dos diários nacionais e procura identificar os efeitos globais neste mercado, bem como efeitos em cada um dos segmentos de mercado respectivos, nomeadamente: desporto, economia e negócios, e informação geral.

Mais do que a mera correlação, o presente trabalho foca-se na identificação e na evidência empírica de relações de causalidade estatística entre o consumo *online* e *offline* (e vice-versa). Assumindo que tanto o consumo em papel pode influenciar o consumo digital, como o consumo digital pode influenciar o consumo tradicional, existem diferentes efeitos que podem ser identificados:

- Efeitos positivos (onde o consumo *online/offline* promove impactos positivos no consumo *offline/online*, isto é, existem efeitos de complementaridade);

- Efeitos negativos (os que popularmente são aceites como um sentimento unânime, isto é, a existência de efeitos de substituição);
- Efeitos neutros (ou a não identificação de quaisquer efeitos, o que conduz à não identificação de evidência estatística de dependência entre o mercado digital e tradicional).

O sector de media tradicional (e da imprensa escrita em particular) é tipicamente um *two-sided market*, uma vez que actua em dois mercados: o mercado dos conteúdos e da publicidade. Sendo que naturalmente ambos os lados do mercado estão relacionados, a inclusão da dimensão digital levanta inúmeras questões quanto ao modelo de negócio. Serão os canais digitais apenas uma nova forma de alcançar a mesma audiência? Ou antes uma oportunidade de chegar a novos mercados e criar novas ofertas e modelos de negócio? A compreensão destas questões e da relação entre o consumo digital e tradicional mostram a relevância das hipóteses em teste no presente trabalho para a tomada de decisão, nomeadamente nas opções de produção de conteúdos e de modelos de distribuição.

É neste contexto que o recurso à evidência empírica pode contribuir para a clarificação do fenómeno, tanto a nível nacional, como numa óptica de complementaridade de trabalhos anteriores (como Kaiser e Kongsted, 2005, Filistrucchi, 2005, ou Gentzkow, 2003) permitindo uma visão mais integrada do fenómeno da imprensa escrita a nível global.

De modo a controlar possíveis enviesamentos na análise de impactos e de causalidade, vão-se procurar os factores que podem condicionar quaisquer dos consumos. Factores como o contexto macroeconómico, a existência de novas demografias e estruturas sociais e económicas, os choques tecnológicos e os novos canais de comunicação, os novos modelos de negócio (como jornais gratuitos, *syndication*, agregadores de notícias e conteúdos, etc), a própria natureza dos conteúdos, são alguns dos diferentes factores que podem influenciar a dinâmica dos consumos. Num quadro de crescente penetração de canais digitais, da internet das coisas (IOT) e da internet de tudo (IOE), esta autêntica revolução digital tem acompanhado contemporaneamente uma tendência de decréscimo no negócio tradicional, fazendo com que o “negócio digital” assuma um

cada vez maior papel de destaque no sector. Independentemente das técnicas utilizadas (preditivas ou não), os resultados devem ser vistos numa perspectiva decisional e constituir conhecimento complementar e um elemento de debate no sector. Efectivamente, dada a natureza do fenómeno, o objectivo do presente documento não é fornecer previsões para o sector, mas sim analisar a causalidade entre as duas formas de consumo.

Numa perspectiva de análise empírica dos efeitos existentes entre o consumo digital e o consumo tradicional, é importante compreender o negócio e o fenómeno, quer a nível local, quer a nível global.

## 1.2 – Visão global da imprensa escrita

A ideia de uma quebra geral nas vendas e na circulação não pode ser generalizada se olharmos para o fenómeno em diferentes territórios. Para proporcionarmos uma visão panorâmica do sector, importa caracterizar a sua dimensão, e as suas especificidades locais.

Relativamente à sua dimensão, falamos de uma indústria que tem uma circulação de cerca de 534 milhões de cópias que alcançam aproximadamente 2.5 mil milhões de adultos no planeta<sup>1</sup>. E (para eventual surpresa de alguns) uma circulação que se mantém estável quando comparada com 2010 (+2%)<sup>2</sup>. Ao mesmo tempo, em termos de receitas é uma indústria que gerou 187 mil milhões de dólares (USD) em 2014, com receitas estáveis da circulação de cerca de 92 mil milhões de USD, mas com receitas publicitárias em queda para aproximadamente 87 mil milhões de USD<sup>3</sup>. Em oposição, as receitas digitais têm revelado crescimento, com a circulação digital a aumentar 1400% desde 2010 (para 2.5 mil milhões de USD), e as receitas de publicidade digital a aumentarem 60% também desde 2010 (para 9.5 mil milhões de USD)<sup>4</sup>. Apesar da existência de alguns estudos que sugerem a migração do *offline* para o *online*<sup>5</sup>, e que suportam assim o

---

<sup>1</sup> Fonte: Wan-Ifra – World Press Trends 2014

<sup>2</sup> Fontes: Wan-Ifra – World Press Trends 2014, Wan-Ifra – World Press Trends 2015

<sup>3</sup> Fonte: Zenith Optimedia and PWC Global Entertainment & Media Outlook 2013

<sup>4</sup> Fonte: Wan-Ifra – World Press Trends 2015

<sup>5</sup> Como é o caso do *Pew Research Center's Journalism Project – The State of News Media 2012*

sentimento global da existência de efeitos de substituição, mostram também que apesar do aumento da importância do digital, este assume ainda uma importância reduzida não sendo ainda capaz de compensar a perda de receita do negócio tradicional.

Todavia, esta dimensão e estabilidade da circulação não são transversais em diferentes territórios. Efectivamente, existem diferenças culturais, sociais e económicas que proporcionam diferentes modelos de negócio e estádios de maturidade face ao digital. Analisando os diferentes modelos de negócio, diga-se que circulação e publicidade representam diferentes importâncias na receita em diferentes territórios. Por exemplo, nos EUA o peso da circulação é de 27% (e a publicidade 73%), na China 44%, na Alemanha e no RU cerca de 60%, e no caso extremo do Japão, 70% (onde a publicidade apenas representa 30% da receita total)<sup>6</sup>. De facto, recorde-se que *“os jornais diários Japoneses Yomiuri Shimbun, Asahi Shimbun e Mainichi Shimbun com 9.69, 7.45 e 3.32 milhões de exemplares são os maiores diários do mundo. No top 20 de diários, encontramos títulos europeus e americanos famosos (como o BILD-Zeitung/BZ, o Wall Street Journal, The Sun e o New York Times, com 2.66, 2.38, 2.17 e 1.87 milhões de exemplares diários respectivamente. Neste conjunto de títulos, vemos claramente que os mercados Japonês, Chinês e Indiano são claramente dominantes<sup>7</sup>”*.

Se existem diferenças de dimensão entre territórios, as próprias dinâmicas são distintas e o sentimento geral de que a circulação está em queda não é transversal. De 2008 a 2013, apenas os mercados europeu, norte-americano e australiano apresentaram quedas de circulação de 23, 10 e 20% respectivamente. Territórios como o africano, latino-americano ou o asiático apresentam crescimentos de 7.5, 6.3 e 6.7% no mesmo período<sup>8</sup>.

Paralelamente às diferentes dimensões e dinâmicas existentes, os diferentes territórios também mostram diferentes estádios de envolvimento e adopção de conteúdos digitais. Comparando a adopção do digital (através do peso relativo do digital face ao papel e do nível de audiência de sites de notícias), se a nível global existem claramente diferenças

---

<sup>6</sup> Fonte: Wan-Ifra – World Press Trends 2014

<sup>7</sup> Fonte: Wan-Ifra – World Press Trends 2014

<sup>8</sup> Fonte: Wan-Ifra – World Press Trends 2014

de estágio de adoção, mesmo a nível do mercado europeu podemos identificar diferenças entre países escandinavos, anglófonos, germânicos e latinos (sendo os primeiros os que se encontram num estágio mais avançado de adoção do consumo digital)<sup>9</sup>.

No que diz respeito à circulação no espaço europeu, e se por um lado o sentimento de queda generalizada da circulação não pode ser constatada a nível global, já a nível europeu (mesmo com diferentes estágios de adoção do digital), os principais mercados europeus<sup>10</sup> apresentam fortes sinais de quedas de circulação que se reflectem tanto nas receitas de publicidade, como no emprego gerado pelo sector (ambos em queda)<sup>11</sup>.

### 1.3 – Jornais diários em Portugal

No sector da imprensa diária Portuguesa, três segmentos de mercado representam praticamente 100% das vendas em banca: o segmento da **informação geral** (que representou em média 66% das vendas de jornais diários em 2016), o segmento **desportivo** (que representou 32% no mesmo período), e o segmento de **economia e negócios** (2% das vendas de diários em 2016)<sup>12</sup>.

Já em termos de consumo digital, estes segmentos de mercado têm importâncias distintas. Não sendo o primeiro em termos de exemplares vendidos, o segmento desportivo (constituído pelos títulos *A Bola*, *O Jogo*, *Record*), é de longe o segmento com maior audiência *online*: *A Bola*, *Record* e *O Jogo*, são os 2º, 3º, e 6º *websites* em termos de visitas em Portugal<sup>13</sup>. De entre os títulos analisados no presente trabalho (*A Bola*, *O Jogo* e o *Record* no segmento desportivo, *Diário Económico* e *Jornal de Negócios* no segmento de economia e negócios, e o *Correio da Manhã*, *Diário de Notícias*, *I*, *Jornal de Notícias* e *Público* no segmento de informação geral), o segmento desportivo representa mais de 74 milhões de visitas mensais em Junho de 2016. Em termos gerais,

---

<sup>9</sup> Fonte: Wan-Ifra – World Press Trends 2014

<sup>10</sup> Áustria, França, Dinamarca, Holanda, República Checa, Portugal, Suécia, Finlândia, Bélgica, Suíça, Noruega, Hungria, Alemanha, Irlanda, Polónia, Espanha, Itália, Grécia, Reino Unido, são os países que por ordem crescente apresentam maiores decréscimos de circulação, de -2% a -21% (OCDE, 2010)

<sup>11</sup> Fonte: Eurostat, Comissão Europeia - The Newspaper Publishing Industry (2012)

<sup>12</sup> Fonte: APCT, 1º Bimester 2016, Sociedade Vicra de Desporto

<sup>13</sup> Fonte: Marktest - Ranking netScope de tráfego web, Junho 2016

no mês de Junho de 2016, os títulos analisados contabilizaram 145 milhões de visitas (51% destas no segmento de desporto, 42% em títulos de informação geral, e finalmente, 7% no segmento de economia e negócios<sup>14</sup>).

O presente estudo foca-se na existência de hipotéticas relações de causalidade entre os consumos *online* e *offline* dos diários destes três segmentos de mercado: economia e negócios (Diário Económico, Jornal de Negócios), desporto (A Bola, O Jogo, Record), e informação geral (Correio da Manhã, Diário de Notícias, I, Jornal de Notícias, Público).

Embora o negócio tradicional esteja claramente num estágio de considerável maturidade (excluindo o “I” que entrou no mercado em 2009, os títulos mais recentes são o Jornal de Negócios e o Público que foram lançados em 1998 e 1990 respectivamente), e destes diferentes segmentos terem diferentes importâncias nos consumos *online/offline*, qualquer destes segmentos apresenta uma tendência de decréscimo das suas vendas em banca. Analisando as vendas médias diárias nos últimos 12 anos podemos visualmente constatar a tendência decrescente do negócio tradicional<sup>15</sup>. Nesta perspectiva macro, o mercado dos diários nacionais (incluindo apenas os títulos em análise) caiu de mais de meio-milhão de exemplares vendidos em 2004 para cerca de 250 mil em 2016. No mesmo período, o segmento desportivo perdeu mais de metade das suas vendas diárias (cerca de 63% de 2004 a 2016). Em síntese, qualquer dos segmentos está confrontado com esta queda (drástica) das vendas em banca. No mesmo período, o comportamento do consumo *online* tem apresentado uma tendência sistematicamente crescente<sup>16</sup>. No segmento de informação geral, o número de visitas aumentou de uma média de 5 milhões para mais de 57 milhões de visitas por mês. Seguindo a mesma tendência, o segmento desportivo aumentou de menos de 20 milhões para cerca de 70 milhões de visitas.

Tal como noutros mercados, o sector de media enfrenta diferentes estruturas de mercado, com novos e não-convencionais concorrentes. Hoje, os principais *stakeholders* do sector em Portugal acreditam que a “*Era digital não é o principal problema do sector,*

---

<sup>14</sup> Fonte: Marktest - Ranking netScope de tráfego web, Junho 2016

<sup>15</sup> Apêndice 1, Figura 3: Vendas médias diárias (2004- Fev/2016)

<sup>16</sup> Apêndice 1, Figura 4: Visitas médias mensais (2007- Jun/2016)

*mas sim a fatia das receitas publicitárias digitais captada pelas plataformas que distribuem conteúdos (como o Google e o Facebook). Os publishers e agregadores globais têm 81% das receitas da publicidade digital”<sup>17</sup>.*

## 2 – Resumo da Literatura

O estudo do modelo de negócio tradicional da imprensa vem já de longa data, por exemplo com os trabalhos de Corden (1953) ou Reddaway (1963). O progresso tecnológico dos últimos anos veio levantar novas questões e estimulou nova pesquisa do sector.

Na investigação dos efeitos entre o negócio digital e o negócio tradicional, os resultados, apesar de muitas vezes discordantes, revelam utilidade na detecção de alguns efeitos, de metodologias de análise, e na identificação de factores que influenciam o consumo. Todavia, a dificuldade de obtenção de dados, tanto para *online* como para *offline* tornam ainda escassa a investigação existente.

Independentemente da natureza qualitativa ou quantitativa e das metodologias envolvidas, existem diversos resultados quer no sentido da existência de efeitos de substituição, quer de complementaridade e mesmo de independência. Para sublinhar a incoerência de resultados, agrupamos os principais resultados em função dos efeitos que identificaram entre os consumos *online* e *offline*.

### 2.1 - Efeitos de substituição

Apesar do sentimento geral (quase unânime e que em Portugal se traduz no facto de 73% dos gestores de media terem perspectivas e expectativas negativas sobre a evolução da receita no negócio tradicional<sup>18</sup>) de queda do negócio não significar a existência de efeitos de substituição, alguns estudos apontam efectivamente esse caminho de uma “evidente” canibalização. Por exemplo, Barsh et al. (1999) suportam que aproximadamente 16% dos utilizadores de internet gastam menos tempo de leitura

---

<sup>17</sup> Fonte: Jornal de Negócios, 14 de Julho de 2015; url: [http://www.jornaldenegocios.pt/empresas/detalhe/publishers e agregadores globais controlam 81 das receitas da publicidade digital em portugal.html](http://www.jornaldenegocios.pt/empresas/detalhe/publishers_e_agregadores_globais_controlam_81_das_receitas_da_publicidade_digital_em_portugal.html)

<sup>18</sup> Fonte: OberCom – Barómetro da Comunicação (2014)

de jornais e revistas e substituem este tempo por consumo na internet. Mitchell (2001), recorrendo no caso a entrevistas a mais de 250 editores dos EUA, conclui no sentido das tecnologias digitais terem um impacto negativo no longo-prazo no negócio tradicional. Numa perspectiva mais quantitativa, derivando um modelo estrutural de escolha discreta e comparando estimadores SML (*simulated maximum likelihood*) e MSM (*method of simulated moments*) para o mercado dos EUA, Gentzkow (2003) conclui que as edições *online* e em papel são substitutos claros, porém, são substitutos fracos. Apesar do *crowding out* no papel não ser negligenciável, é reduzido, e como tal, as edições *online* não ameaçariam as edições impressas. Já numa análise dos diários italianos de 1976 a 2001, Filistrucchi (2005) conclui que existe uma significativa substituição, e que mais do que a significância estatística, a sua pesquisa foca-se no elevado impacto negativo promovido pela existência de *websites* nestas publicações. Para Filistrucchi (2005), este elevado impacto negativo pode ser atribuído a três factores: ao segmento de mercado analisado, ao período alargado, e ao modelo de disponibilidade de *websites* onde a maioria se baseava na disponibilização dos conteúdos integrais gratuitamente. Por seu lado, Simon (2004) volta-se para um conjunto de segmentos de mercado mais abrangente com mais de 500 publicações de cerca de 40 segmentos de mercado, tendo estimado um decréscimo de 3% em média quando uma publicação também disponibiliza um website. A pesquisa mais radical talvez possa ser atribuída a Patel (2010), que conclui mesmo que a única estratégia de sobrevivência possível no sector passa em exclusivo pelo digital e pelo abandono da impressão e do modelo de negócio tradicional. Mais do que a prescrição deste modelo de negócio radical, os principais méritos de Patel (2010) prendem-se na identificação de cinco *drivers* para sustentar esta estratégia, nomeadamente a crise económica (2007-2009), a difusão e evolução da internet, o comportamento dos consumidores, as novas tendências na publicidade e o progresso tecnológico. Outra forma de medir as interações entre o digital e o consumo tradicional foi o contributo de Pino (2010) que, com dados de 1999 a 2007 quantifica o efeito da internet na circulação através da penetração de banda-larga e, levando em linha de conta factores exógenos como o rendimento e o desemprego nos EUA, conclui que a internet é responsável por 2/3 do

declínio de 24% da circulação dos jornais diários. Pino (2010) alerta ainda para a diferença entre jornais e revistas, dado que as últimas não foram afectadas pela internet, e acima de tudo, que quanto mais avançada tecnologicamente é a audiência de uma revista, maior a ameaça de substituição do papel em favor do digital. Também é relevante o trabalho de Seamans e Zhu (2011) que estudam um segmento muito particular através da análise do impacto do *Craigslist*, um website de classificados presente em mais de 300 mercados<sup>19</sup>. Também neste trabalho é verificada a substituição. O estudo encontra evidência estatística significativa que os jornais com maior receita publicitária no segmento dos classificados verificam um decréscimo significativo após a entrada do *Craigslist* no mercado. Embora seja um *side* muito específico do *multi-sided market*, o principal mérito do estudo prende-se com a análise do impacto da internet na circulação com a entrada de novos *players* no mercado digital. Em Portugal, Alturas e Fernandes (2013) questionaram aproximadamente 380 respondentes (numa população maioritariamente universitária), concluindo a favor de uma relação negativa entre a idade e o consumo digital, sendo a audiência mais jovem a mais propensa à mudança de hábitos de consumo para os conteúdos digitais. Aliás, o estudo de Alturas e Fernandes (2013) vai de encontro à análise de uma das principais fragilidades do sector: a captação de novos leitores e a penetração em públicos jovens para contrariar a diminuição dos hábitos de leitura (Comissão Europeia, 2012)<sup>20</sup>. Conclusões semelhantes às de Liebowitz e Zentner (2012), que identificam uma relação similar entre o consumo *online* e o consumo de televisão e a importância das gerações na natureza do consumo. Finalmente, também o trabalho de Li (2013) constitui um importante contributo. Apesar de não estar relacionado directamente com a imprensa escrita mas sim com o mercado dos livros e no impacto de vendas de *eBooks* nas de edições impressas da Amazon, Li estima o nível de canibalização e o efeito de expansão de mercado. No seu modelo conclui que 2/3 dos *eBooks* canibalizam as edições em papel, e que apenas 1/3 das vendas se devem à expansão do mercado e a novos

---

<sup>19</sup> Fonte: <https://www.craigslist.org/about/expansion>

<sup>20</sup> Diversos estudos apontam para uma maior adesão ao consumo de notícias através da internet principalmente por parte de um públicos jovens, tipicamente abaixo dos 25, 30 anos (Slot e Munniks de Jongh Luchsinger, 2011; Huysmans e de Haan, 2010; Eimeren e Vrees, 2011; Schrøder e Kobbernagel, 2010; Alturas, B., Fernandes, F. (2013); Schrøder e Larsen, 2010)

leitores/clientes. Apesar do trabalho de Li não medir directamente o impacto no consumo *offline* (uma vez que mesmo as edições em papel são vendidas *online* da Amazon), e de estar focado numa população muito particular já com hábitos digitais, o estudo revela a importância dos *devices* e da tecnologia no consumo e nos hábitos de leitura.

## 2.2 – Independência ou ausência/neutralidade de efeitos

Antes de mais, é importante alertar que a não observação de efeitos positivos ou negativos, pode não ser necessariamente um sinal de independência. É relevante ter consciência que pode dar-se a situação dos efeitos positivos e negativos se anularem. Mas também pode ser um sinal de independência, naturalmente, e de que não existe efectivamente relação causal entre o consumo *online* e *offline*, e que na verdade estamos a falar de audiências/consumos independentes, com as suas especificidades e características próprias, sendo o digital mais um potencial *side* do mercado.

Deleersnyders et al. (2002), testa a eventual alteração de estrutura com o lançamento de *websites* nos mercados britânico e holandês. Concluem que poucos jornais vêem a sua circulação sofrer alterações de estrutura com o lançamento de *websites*, rejeitando assim a hipótese de alteração de estrutura na circulação com a disponibilização de canais digitais. Rathmann (2002), no segmento dos diários alemães e em particular junto da população de assinantes *online*, identificou efeitos de complementaridade. Efectivamente, segundo Rathmann (2002), para além de, à data, as edições *online* não serem ainda um substituto das edições em papel, são os leitores frequentes de papel que apresentam também uma maior intenção de procura *online*, essencialmente com o objectivo de encontrarem conteúdos em maior detalhe e profundidade do que as edições impressas permitem adquirir. Os (ligeiros) efeitos de substituição só são verificados nos leitores menos frequentes, sendo de realçar, todavia, que todos os leitores envolvidos no estudo consideraram as edições em papel mais importantes que as edições *online*. Já num segmento de mercado diferente, Kaiser (2002) estuda as revistas femininas na Alemanha, concluindo que o lançamento ou a existência de *websites* não afecta significativamente a procura *offline* neste segmento de mercado. Lauf, Schönbach e Waal (2005), para o mercado holandês, apesar de verificarem uma

ligeira correlação negativa entre as edições *online* e *offline*, concluem que os jornais *online* não são substitutos das edições em papel. Aliás, o grande mérito da pesquisa acaba por ser a relevância da segmentação da audiência na identificação de diferentes propensões para o consumo *online*, nomeadamente para a importância da idade, sexo e educação. De entre este conjunto de trabalhos que apontam para a independência entre os consumos digitais e tradicionais, Ahlers (2006) segmenta as audiências para concluir que a migração do papel para o digital não ocorreu ou que pelo menos, não ocorreu numa magnitude que permita antecipar o fim do negócio tradicional. A maior contribuição de Ahlers é a importância atribuída à segmentação da audiência para analisar o consumo de notícias. Segundo Ahlers, existem audiências para diferentes meios (digital ou papel), audiências onde os efeitos de substituição serão mais claros, e outras onde a complementaridade pode identificar segmentos de genuínos leitores multi-canal. E nesse campo, ao analisar uma população específica dos “nativos-digitais”, Chyi (2012) reforça também a ideia que não são os meios (digitais ou papel) que são relevantes, mas sim outros factores como a perda de interesse em notícias e alguns conteúdos, e que o decréscimo de consumo está longe de ser uma questão a ser resolvida pela tecnologia. De igual forma, num estudo do mercado húngaro, Fodor et al. (2013) concluem igualmente que existem audiências fiéis ao papel, e que o interesse nos conteúdos digitais decresce com o preço, e que mesmo as elasticidades face ao preço são diferentes por tipo de conteúdo e segmento de mercado. Adicionalmente, a adopção do digital depende dos perfís socio-económicos da audiência (sendo as audiências urbanas e com escolaridades mais elevadas as mais propensas ao consumo digital).

### 2.3 – Efeitos de complementaridade

Finalmente, podem também existir efeitos positivos (de complementaridade) entre as edições *online* e *offline* na imprensa escrita. Tradicionalmente, estes efeitos podem ser atribuídos a múltiplos factores como a própria notoriedade da marca, a quota e penetração no mercado, assinaturas *online/offline*, ofertas e promoções, ou mesmo a própria política editorial e de conteúdos.

De entre as evidências empíricas, podemos destacar o contributo de Pauwels e Dans (2001) sobre a circulação de 12 jornais diários em Espanha, onde se conclui a existência de uma relação causal (à Granger) onde a circulação tem um impacto positivo no consumo digital (nas visitas aos *websites*). É importante referir que este caso se centrou na utilização de dados diários (e portanto com uma frequência elevada), o que pode ajudar a perceber a natureza deste efeito positivo identificado. Também Kaiser e Kongsted (2005), no mercado das revistas alemãs, identificam uma relação de causalidade (à Granger) com um efeito positivo das visitas a *websites* na circulação. Por seu turno, Lucena (2010) alarga o estudo a mais territórios (Canada, EUA, Holanda, Japão, Brasil e Índia) para analisar a penetração de internet e o impacto na circulação. Lucena encontra diferenças entre países: se nos países desenvolvidos não encontra evidências estatísticas, nos países em desenvolvimento já é verificável uma relação causal positiva entre a penetração de internet e o aumento da circulação (que diminui, todavia, com o aumento da penetração da internet). Já Knuff (2011) foca-se na análise de assinaturas em papel, incorporando diferentes modelos de negócio (de conteúdos gratuitos *online*, a conteúdos parciais, ou conteúdos integralmente pagos), e formas de pagamento. A sua principal conclusão é dos jornais *online* serem mais complementares que substitutos, e que para preservar as assinaturas em papel, a melhor estratégia é mesmo a disponibilização de conteúdos gratuitos *online*. E finalmente, Newman et al. (2012), olham para o fenómeno das redes e dos media sociais e do seu impacto sobre o modelo tradicional de media na Grã-Bretanha. Mais do que olhar para um modelo que encara os canais como potenciais concorrentes, os autores destacam essencialmente uma nova visão do mercado. Efectivamente, o mercado, a audiência ou os indivíduos de uma forma geral procuram actualmente formas de aceder à informação de forma independente e de se relacionarem com outros indivíduos expandindo a sua capacidade de comunicação. Esta visão comportamental dos indivíduos sugere que as oportunidades de complementaridade entre canais no sector de media é efectivamente o novo paradigma uma vez que *“they are not simply competing or substituting for one another, but involved in an ecology of media that is also enabling the two estates to be mutually complementary and reinforcing”* (Newman et al. , 2012).

## 2.4 – Principais limitações

Como referido anteriormente, a evidência empírica existente é muitas vezes contraditória. Apesar de contraditórias, as pesquisas anteriores têm o mérito de sugerir estratégias de análise e metodologias para identificação dos efeitos entre os consumos digital e tradicional e, em algumas situações, apontar para uma melhor caracterização do fenómeno e de eventuais factores que possam condicionar as procuras (tanto *online* como *offline*).

Por outro lado, uma das maiores dificuldades dos estudos quantitativos prende-se efectivamente com a dificuldade de recolha de dados. E potencialmente, esta limitação pode condicionar a incorporação de factores exógenos que influenciam a procura. De facto, a maior crítica às pesquisas existentes (principalmente as quantitativas) prende-se com a não incorporação de factores exógenos. Por exemplo, a maioria dos estudos não leva em linha de conta os progressos tecnológicos mais recentes. Focam-se num consumo tradicional de internet (via desktop), quando hoje o consumo é cada vez mais multi (ou omni) canal. Se nos lembrarmos que o iPhone foi lançado em 2007, o sistema Android em 2008, e que apenas em 2010/2011 estas tecnologias se tornaram *mainstream*, ou que a primeira geração de iPads foi lançada em 2010, que a penetração de *Tablets* e *phablets* está consideravelmente a aumentar, e que a emergência de agregadores e outros distribuidores alterou a configuração do mercado (em 2006 a Google lançou o *Google News*, ou que *Facebook* foi lançado apenas em 2004, ou o *Twitter* em 2006), então uma análise de impactos deve ter estes aspectos em linha de conta. Também é fundamental incorporar outros factores exógenos para além da tecnologia (factores económicos, conteúdos, etc) no sentido de evitar a identificação de relações de causalidade espúrias. Esta é uma das preocupações claras da presente pesquisa que, para além da análise da dinâmica dos consumos *online* e *offline* e da sua causalidade, procura eliminar eventuais efeitos promovidos por outros factores exógenos.

## 3 – Aplicação Empírica

### 3.1 – Os Dados

Dado que o principal objectivo do presente trabalho é analisar as relações de causalidade entre o consumo *online* e *offline* nos jornais diários portugueses, a principal informação de base são evoluções de vendas em banca e de visitas aos respectivos *websites* para as diferentes publicações em estudo, nomeadamente: A Bola, Record, e O Jogo no segmento desportivo, Correio da Manhã, Diário de Notícias, I, Jornal de Notícias e Público no segmento de informação geral, e Diário Económico e Jornal de Negócios no segmento de economia e negócios. Esta informação tem origem em três fontes: APCT (Associação Portuguesa para o Controlo de Tiragem e Circulação), Sociedade Vicra Desportiva SA (dona do título A Bola, uma vez que esta não é associada da APCT), e a informação dos Rankings Netscope da Marktest.

Foram recolhidas as vendas médias diárias mensais para o período de Janeiro de 2000 a Abril de 2016, bem como o número de visitas mensais de Janeiro de 2007 a Junho de 2016.

Para além destas fontes e dados centrais, para identificar potenciais factores que possam influenciar os consumos *online* e *offline*, foi recolhida informação de variáveis exógenas:

- Resultados desportivos da primeira liga de futebol profissional portuguesa desde a época de 2006/07 até à data actual (disponíveis em [www.lfpf.pt](http://www.lfpf.pt));
- Calendário com os principais eventos desportivos (1ª liga, taça, competições internacionais como campeonatos europeus, mundiais, principais jogos, e outros);
- Principais temas, conteúdos e eventos (por exemplo, eleições, principais acontecimentos e conteúdos políticos, económicos ou sociais, etc), extraídos das manchetes diárias de todas as publicações em estudo desde 2008;
- O número de clientes com internet fixa e/ou móvel (disponível em [www.anacom.pt](http://www.anacom.pt));
- Pesquisas Google dos *websites* em estudo (obtidas através do *Google Trends*);

- Séries mensais do Índice de Produção Industrial Total (IPI), IPI de bens de consumo, Indicador do sentimento económico, Índice de Confiança do Consumidor, Índice de Preços do Consumidor, Taxa de Desemprego, todas obtidas e disponíveis no sítio de internet do *Banco de Portugal*;
- Séries mensais da despesa e receita públicas (disponíveis na Direcção Geral do Orçamento);
- Outras variáveis que representam marcos de inovação tecnológica ou de alteração de modelo de negócio (introdução de smartphones, tablets, facebook, etc, ou alteração do modelo de preço *online*, por exemplo).

Na prática, estas (candidatas a) variáveis exógenas funcionarão como variáveis para capturar efeitos adicionais (de conteúdos e eventos, da economia, da tecnologia, de novos concorrentes e modelos de negócio) e mitigar eventuais efeitos indesejados de especificação.

### 3.2 – Metodologia

O presente trabalho é assumidamente uma pesquisa aplicada na área econométrica e que procura evidência empírica de relações de causalidade entre os consumos tradicionais e digitais na imprensa escrita. Nesse sentido, de um ponto de vista metodológico, para além de seguirmos a aproximação de Kaiser e Kongsted (2005) na identificação de relações de causalidade individualmente para cada um dos títulos, segmentos de mercado e mercado, alargamos a aproximação à identificação de efeitos globais de mercado e de segmentos de mercado.

De uma forma geral, as principais hipóteses que serão objecto de avaliação empírica serão as seguintes:

- *Hipótese 1*: As visitas aos *websites* têm uma relação causal com as vendas em banca;
- *Hipótese 2*: As vendas em banca têm uma relação causal com as visitas aos *websites*.

Dado este quadro onde qualquer dos consumos pode influenciar ou ser influenciado, onde os mercados (digital e tradicional) podem ser substitutos, complementares ou

independentes, o presente trabalho seguirá uma **metodologia VAR** como veículo para a identificação de **causalidade à Granger**. Trata-se de uma metodologia onde para além da simetria no tratamento das variáveis de interesse (no caso as vendas em banca e as visitas aos *websites*), permite que os mercados sejam explicados quer pelo seu comportamento passado, quer por efeitos cruzados dos mercados.

Com a utilização de uma aproximação VAR para testar a causalidade à Granger, em função do objectivo de estudo, seguimos diferentes estratégias de teste:

- a) Estimação de (14) modelos VAR independentes (um modelo para cada título, segmento e mercado total, para testar individualmente as hipóteses 1 e 2 a nível individual);
- b) Estimação de (4) modelos VAR num painel mensal para os 10 títulos (como veículo de teste global do mercado e de cada segmento de mercado para as hipóteses 1 e 2 a nível global para cada segmento de mercado e mercado total).

A estratégia de assumir independência entre as diferentes publicações (com um modelo VAR independente para cada título), embora tenha o custo implícito de ser uma hipótese altamente restritiva (por exemplo, pode o aumento de visitas de um dado título aumentar/diminuir as visitas aos *websites* de outros concorrentes?), tem o benefício de reduzir o número de parâmetros a estimar. A manutenção desta aproximação deve-se fundamentalmente a questões de parcimónia. Tal como Kaiser e Kongsted (2005), a utilização da metodologia VAR tem como principal finalidade ser o veículo base para a análise de (não-) causalidade no sentido da definição de Granger(1969). No caso, a utilização de uma aproximação VAR não-estrutural, para além de ser uma das mais convencionais, é também apropriada para os nossos propósitos dado que permite, por um lado, que os choques num dado mercado estejam contemporaneamente correlacionados, e por outro, que os mercados sejam explicados tanto pelo seu comportamento passado como pelos efeitos cruzados entre mercados (no caso, tradicional e digital).

Lembramos que, no caso, estaremos a lidar com 10 títulos, 3 segmentos de mercado, e o mercado total de diários (no nosso caso, consideramos que os 10 títulos correspondem

ao total de mercado). Se a análise independente de cada título (a) permite identificar efeitos isolados para cada uma das publicações, segmentos de mercado e mercado total, a análise em painel (b) permitirá captar efeitos globais

Em termos concretos, tratando-se de informação (de vendas e visitas) mensal (e não sendo o objecto de estudo a sazonalidade), as variáveis foram previamente ajustadas para remoção da sazonalidade (evitando o mero aumento de parâmetros determinísticos a estimar). Para este ajustamento sazonal recorreremos ao método multiplicativo de rácio de médias móveis.

Neste contexto, consideramos as variáveis endógenas :  $y_t = \begin{bmatrix} v_t \\ w_t \end{bmatrix}$ , onde, (1)

- $v_t$ , é o logaritmo natural das vendas médias diárias
- $w_t$ , é o logaritmo natural das visitas aos *websites*

### 3.3 – Resultados

As séries de vendas logaritmizadas e já corrigidas da sazonalidade <sup>21</sup> apresentam na globalidade a tendência decrescente que o sector da imprensa escrita diária enfrenta também no mercado nacional. Se a evolução das vendas em banca apresenta este comportamento generalizado, a evolução das visitas aos respectivos *websites* apresenta um comportamento contrário<sup>22</sup>. Visualmente (exceptuando principalmente publicações como o I ou o Diário Económico), podemos ver a diferença face às evoluções das vendas em banca. Na maioria das publicações vemos um crescimento significativo ao longo dos últimos anos, sendo de salientar todavia o (aparente) amortecimento deste crescimento no passado mais recente. Constatamos também que as séries (ajustadas da sazonalidade e logaritmizadas) apresentam-se como séries candidatas a séries não-estacionárias. Efectivamente, como podemos ver no Apêndice II – Testes de Raízes Unitárias e Cointegração, as séries são I(1) e não-cointegradas.

Neste quadro de não-estacionaridade e ausência de cointegração, foca-mo-nos em primeiro lugar na estimação para as taxas de crescimento mensais ( $\Delta y_t \sim I(0)$ ) como

---

<sup>21</sup> Apêndice 1, Figura 3: Evolução do logaritmo das vendas em banca

<sup>22</sup> Apêndice 1, Figura 4: Evolução do logaritmo das visitas aos websites

veículo para estudo da causalidade o que se traduzirá num modelo VAR não-estrutural baseado no seguinte sistema:

$$\Delta y_t = A_0 + \sum A_p \Delta y_{t-p} + \sum B_j x_{t-j} + \varepsilon_t \quad , p=1, 2, \dots, T, j=0, 1, \dots, T \quad (2)$$

No caso,  $x_t$  representa o vector de factores exógenos (regressores exógenos ou outras componentes determinísticas),  $j$  é o desfasamento das variáveis exógenas, e  $p$  a ordem do modelo VAR.

Como critério de selecção da ordem VAR ( $p$ ), apesar da orientação pelo Critério de Informação de Schwarz (SC) e de outros critérios de selecção<sup>23</sup>, a principal preocupação será garantir a estabilidade dos VAR e que os resíduos são aproximadamente ruídos brancos.

A metodologia VAR revela-se de particular interesse para o objectivo de análise. A título de exemplo, um VAR(2) seria um sistema:

$$\begin{bmatrix} \Delta v_t \\ \Delta w_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11,1} & a_{12,1} \\ a_{21,1} & a_{22,1} \end{bmatrix} \begin{bmatrix} \Delta v_{t-1} \\ \Delta w_{t-1} \end{bmatrix} + \begin{bmatrix} a_{11,2} & a_{12,2} \\ a_{21,2} & a_{22,2} \end{bmatrix} \begin{bmatrix} v_{t-2} \\ w_{t-2} \end{bmatrix} + \begin{bmatrix} b_{11} & \dots & b_{1k} \\ b_{21} & \dots & b_{2k} \end{bmatrix} \begin{bmatrix} x_{1t} \\ \vdots \\ x_{kt} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (3)$$

Com  $k$  variáveis exógenas.

Esta formulação permite testar as duas hipóteses. No exemplo, as matrizes  $A_1$  e  $A_2$  assumem maior relevância uma vez que contêm os parâmetros para testar a causalidade à Granger. No caso,  $a_{12,1}$  e  $a_{12,2}$  podem ser identificados como os coeficientes dos efeitos cruzados das visitas nas vendas, e  $a_{21,1}$  e  $a_{21,2}$  os efeitos cruzados das vendas nas visitas. Adicionalmente,  $a_{12,1}$  pode ser considerado o efeito de curto-prazo das visitas nas vendas (e  $a_{21,1}$  das vendas nas visitas), enquanto os efeitos de longo-prazo podem ser dados por  $\frac{a_{12,1} + a_{12,2}}{1 - a_{11,1} - a_{11,2}}$  e  $\frac{a_{21,1} + a_{21,2}}{1 - a_{22,1} - a_{22,2}}$ , respectivamente para as vendas e visitas.<sup>24</sup>

<sup>23</sup> LR (sequential modified LR test statistic), FPE (Final prediction error), AIC (Akaike information criterion), SC (Schwarz information criterion) e HQ (Hannan-Quinn information criterion)

<sup>24</sup> Uma vez mais alertamos que para a informação disponível (dados mensais de 2007 a 2016), e para a incerteza associada a novos choques tecnológicos e inovações, quaisquer relações de longo-prazo ou conclusões devem ser tomadas com especial prudência.

Lembrando a equação (2), para além da definição da ordem dos VAR, a selecção das variáveis exógenas seguiu uma aproximação *GTS t-sig*, o que resultou na inclusão de diferentes variáveis exógenas para cada título, tendo presente o objectivo de alcançar especificações parcimoniosas garantam estabilidade, ausência de autocorrelação e/ou heterocedasticidade das componentes residuais. Assumindo esta aproximação com prudência, como sugere Lütkepohl (2006), com variáveis endógenas estacionárias, estabilidade dos VAR, e resíduos a seguirem ruídos brancos, a consistência tanto dos estimadores como da matriz de covariância dos erros implicam que os rácios-t têm uma distribuição assintoticamente normal, permitindo concluir que podemos usar os rácios-t no teste individual dos coeficientes.

Esta inclusão de factores exógenos é particularmente relevante quando é uma das principais críticas e omissões das pesquisas existentes no campo quantitativo que podem conduzir a relações de causalidade espúrias entre o consumo *online* e *offline*.

As diferentes especificações dos VAR independentes<sup>25</sup> são, como referimos atrás, um veículo para o nosso principal propósito: testar as relações de causalidade.

**Tabela I: VAR Granger Causality/Block Exogeneity Wald Tests**

	VAR Granger Causality/Block Exogeneity Wald Tests	
	$\Delta w_t \rightarrow \Delta v_t$	$\Delta v_t \rightarrow \Delta w_t$
A Bola	Chi-sq(2) = 2.492281 Prob.= 0.2876	Chi-sq(2) = 5.394586 Prob.= 0.0674
Record	Chi-sq(3) = 2.486855 Prob.= 0.4777	Chi-sq(3) = 4.804606 Prob.= 0.1867
O Jogo	Chi-sq(2) = 0.328344 Prob.= 0.8486	Chi-sq(2) = 2.151102 Prob.= 0.3411
Segmento Desportivo	Chi-sq(2) = 0.133472 Prob.= 0.9354	Chi-sq(2) = 3.553970 Prob.= 0.1691
Diário Económico	Chi-sq(2) = 3.142857 Prob.= 0.2077	Chi-sq(2) = 4.782179 Prob.= 0.0915
Jornal de Negócios	Chi-sq(5) = 7.340311 Prob.= 0.1965	Chi-sq(5) = 5.898880 Prob.= 0.3162
Economy and Business Segment	Chi-sq(2) = 0.783054 Prob.= 0.6760	Chi-sq(2) = 1.224720 Prob.= 0.5421
Correio da Manhã	Chi-sq(2) = 0.369900 Prob.= 0.8311	Chi-sq(2) = 4.404743 Prob.= 0.1105
Diário de Notícias	Chi-sq(3) = 5.655416 Prob.= 0.1296	Chi-sq(3) = 0.759281 Prob.= 0.8592
I	Chi-sq(2) = 2.427629 Prob.= 0.2971	Chi-sq(2) = 0.304792 Prob.= 0.8586
Jornal de Notícias	Chi-sq(2) = 1.964915 Prob.= 0.3744	Chi-sq(2) = 2.956962 Prob.= 0.2280
Público	Chi-sq(2) = 2.095034 Prob.= 0.3508	Chi-sq(2) = 1.834106 Prob.= 0.3997

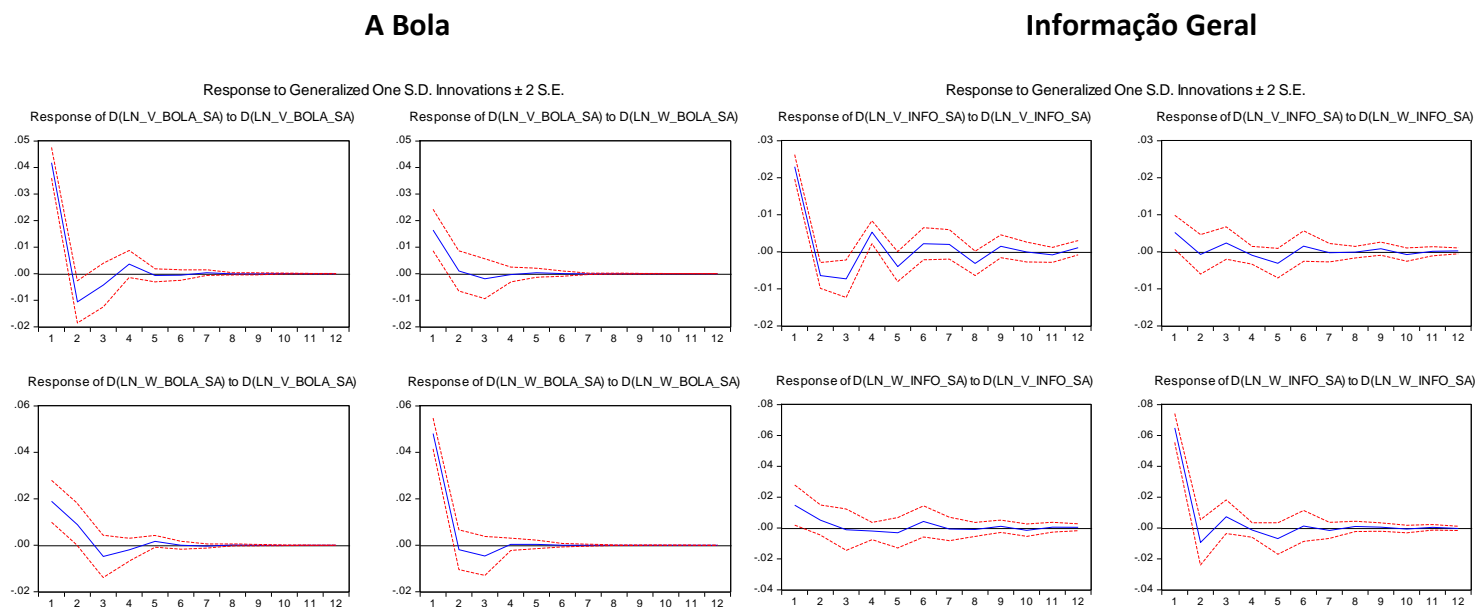
<sup>25</sup> Que podem ser consultadas no Apêndice III – Sumário dos VAR Independentes

	VAR Granger Causality/Block Exogeneity Wald Tests	
	$\Delta w_t \rightarrow \Delta v_t$	$\Delta v_t \rightarrow \Delta w_t$
Segmento Informação Geral	Chi-sq(3) = 10.60118 Prob.= 0.0599	Chi-sq(3) = 5.398477 Prob.= 0.3692
Total Market	Chi-sq(3) = 0.888202 Prob.= 0.8283	Chi-sq(3) = 5.514789 Prob.= 0.1378

Como podemos constatar na tabela anterior, não encontramos evidência significativa (aos níveis de significância habituais de 0.05) para rejeitar a hipótese nula de ausência de relação causal entre as taxas de crescimento de visitas a *websites* e de vendas, ou vice versa. Já a um nível de 0.1, rejeitamos a hipótese de não causalidade em dois casos:  $\Delta v_t \rightarrow \Delta w_t$  para a publicação “A Bola”, e  $\Delta w_t \rightarrow \Delta v_t$  para o segmento de “Informação Geral” no seu todo. Devemos referir como nota complementar a estes resultados que sem a inclusão de factores exógenos foram encontradas diversas relações causais potenciais. Porém, com a inclusão destes factores externos (com particular importância dos conteúdos editoriais), as evidências estatísticas de causalidade desaparecem. Uma vez mais, sublinhamos que a ausência de evidências estatísticas a favor da causalidade (positiva ou negativa) pode não significar que estas são independentes. O que podemos assumir é que para cada publicação, individualmente, não foi identificada relação causal entre o consumo tradicional e digital de um ponto de vista estritamente estatístico.

A análise das funções de resposta a impulsos dos dois casos mais significativos (a 0.1), temos uma melhor percepção do fenómeno.

**Figura 1 – Respostas a Impulsos para “A Bola” e Segmento de Informação Geral - Impulsos Generalizados (Pesaran e Shin, 1998)**



Note-se que as funções de resposta a impulsos anteriores são relativas às séries das publicações que mostraram maiores indícios de presença de causalidade (a 0.1). Naturalmente, para as restantes situações, as respectivas funções de resposta a impulsos revelaram-se praticamente nulas. As funções de resposta a impulsos são ilustrativas do comportamento nas restantes publicações cujas funções de resposta cruzadas têm intervalos de confiança que contêm a resposta “0” para quaisquer desfasamentos, isto é, as variáveis endógenas estatisticamente não respondem a impulsos cruzados. Em oposição (e igualmente comum às demais publicações), as variáveis endógenas respondem significativamente a impulsos “próprios”.

Apesar da presente pesquisa ter como alvo a análise das dinâmicas, de um ponto de vista macro, e da causalidade no sentido estrito de identificar se os consumos online/offline são indicadores avançados um do outro, é relevante analisar igualmente a causalidade instantânea. Como veículo para testar a causalidade instantânea (à Granger), utilizamos um sistema para cada publicação, segmento de mercado e mercado total, que genericamente pode ser representado por:

$$\Delta y_t = A_0 + C\Delta y_t + \sum A_p \Delta y_{t-p} + \sum B_j x_{t-j} + \varepsilon_t \quad , p=1, 2, \dots, T, j=0, 1, \dots, T \quad (4)$$

Onde,  $C = \begin{bmatrix} 0 & c_{12} \\ c_{21} & 0 \end{bmatrix}$

Por exemplo, no caso de um VAR (2), a causalidade instantânea (à Granger) poderá ser testada recorrendo ao habitual teste de Wald onde para :

- Ho: A variável excluída não causa à Granger a variável dependente

Será equivalente ao teste de:

- Ho:  $c_{12}=a_{12,1}=a_{12,2}= 0$ , para  $\Delta w_t \rightarrow \Delta v_t$
- Ho:  $c_{21}=a_{21,1}=a_{21,2}= 0$ , para  $\Delta v_t \rightarrow \Delta w_t$

Se em termos gerais da análise anterior concluímos que  $\Delta w \rightarrow \Delta v$  e  $\Delta v \rightarrow \Delta w$ , apresentamos os resultados dos parâmetros estimados da matriz C e os respectivos testes de causalidade instantânea:

**Tabela II: Causalidade Instantânea e parâmetros estimados**

		$\Delta v_t \rightarrow \Delta w_t$	$\Delta w_t \rightarrow \Delta v_t$
<b>A Bola</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.452956 0.000	0.342007 0.0002
<b>Record</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.349725 0.0130	0.216501 0.0372
<b>O Jogo</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.486881 0.0125	0.187263 0.0309
<b>Segmento Desportivo</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.530553 0.0016	0.210342 0.0095
<b>Diário Económico</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.085162 0.1804	0.037987 0.3517
<b>Jornal de Negócios</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.249330 0.0386	0.314222 0.0215
<b>Segmento Economia e Negócios</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.643953 0.0001	0.261144 0.0002
<b>Correio da Manhã</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.186408 0.1663	0.040176 0.7879
<b>Diário de Notícias</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	-0.022625 0.9428	-0.011652 0.2314
<b>I</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.219002 0.7600	0.066101 0.3492
<b>Jornal de Notícias</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	0.000633 0.4039	0.000311 0.5847
<b>Público</b>	Coef. <i>p-value (Granger causality / Wald test)</i>	-0.087108 0.4151	-0.132574 0.3747
	Coef.	0.646267	0.080941

		$\Delta v_t \Rightarrow \Delta w_t$	$\Delta w_t \Rightarrow \Delta v_t$
<b>Segmento Informação Geral</b>	<i>p-value (Granger causality / Wald test)</i>	0.1483	0.0211
<b>Total Mercado</b>	Coef.	0.467306	0.100834
	<i>p-value (Granger causality / Wald test)</i>	0.0493	0.3164

Na tabela anterior chegamos a conclusões deveras interessantes. Apesar de não ter sido verificada evidência de causalidade entre os consumos online e offline para a maioria dos títulos e mercados, verificamos causalidade instantânea em todo o segmento desportivo e no segmento de economia e negócios (em particular no Jornal de Negócios). Mais do que causalidade instantânea, se observarmos o sentido dos parâmetros estimados, podemos concluir que existe uma forma de feedback instantâneo, uma vez que tanto  $\Delta w_t \rightarrow \Delta v_t$  como  $\Delta v_t \rightarrow \Delta w_t$  e de forma positiva! Este resultado vem reforçar a eventual importância dos conteúdos, uma vez que tal significa que no mesmo momento de tempo, contemporaneamente, as procuras de conteúdos online e offline variam tendencialmente no mesmo sentido, ou seja, potencialmente conteúdos contemporâneos promovem aumentos/diminuições de procura em ambos os consumos. A existência desta forma de feedback instantâneo vem reforçar que estes segmentos de mercado serão os que mais podem beneficiar de alguma complementaridade na produção de conteúdos. Em oposição, o segmento da informação geral mostra também sinais de independência instantânea, isto é, não foram verificadas evidências estatísticas para rejeitar a não-causalidade instantânea.

Se a metodologia VAR apresenta boas propriedades na modelização de séries estacionárias, já a sua utilização em situações de não-estacionaridade é motivo de controvérsia desde que Sims (1980) introduziu a metodologia. Se por um lado existem correntes que seguindo ainda o trabalho de Granger e Newbold (1974) sobre regressões espúrias e de Granger e Engle (1987) sobre cointegração que sugerem a estimação de modelos VEC na presença de cointegração (dada a perda das relações de longo-prazo num VAR estacionário às primeiras diferenças), por outro lado, outros autores seguem o já clássico resultado de Sims, Stock e Watson (1990) que sugere que “*se o coeficiente de um regressor (mesmo que este seja  $I(1)$ ) puder ser escrito como coeficiente de uma variável estacionária com média nula, então o estimador OLS desse coeficiente converge*

à taxa ( $\sqrt{T}$ ) para uma distribuição normal<sup>26</sup>. Ou, como alerta Martin et al (2013), mesmo no caso da verdadeira natureza da não estacionaridade ser estocástica (dada por um passeio aleatório com deriva, mas os dados são modelizados com uma tendência determinística, os estimadores OLS são consistentes à taxa ( $\sqrt{T}$ ). No caso, sendo o modelo verdadeiro,  $y_t = \delta + y_{t-1} + v_t$ , ou reescrevendo,  $y_t = \delta t + u_t$ , e o modelo especificado:  $y_t = \beta_0 + \beta_1 t + e_t$ , com  $v_t$  e  $e_t \text{ iid}(0, \sigma^2)$  então,  $\sqrt{T}(\hat{\beta}_1 - \delta) \xrightarrow{d} N\left(0, \frac{6\sigma^2}{5}\right)$ .<sup>27</sup> Um resultado que requer que  $\delta \neq 0$ , ou seja, que efectivamente o processo tenha deriva.

Nesse sentido, apresentamos também os resultados dos modelos em níveis (tendo naturalmente as devidas reservas na sua utilização como veículo para teste da não-causalidade) através do sistema:

$$y_t = A_0 + \sum A_p y_{t-p} + \sum B_j x_{t-j} + \gamma t + \varepsilon_t, \quad p=1, 2, \dots, T, j=0, 1, \dots, T \quad (5)$$

**Tabela III: Testes de causalidade para VAR em Níveis**

	VAR Granger Causality/Block Exogeneity Wald Tests	
	$\Delta w_t \rightarrow \Delta v_t$	$\Delta v_t \rightarrow \Delta w_t$
A Bola	Chi-sq(2) = 1.707087 Prob.= 0.4259	Chi-sq(2) = 4.917502 Prob.= 0.0855
Record	Chi-sq(3) = 2.588017 Prob.= 0.4596	Chi-sq(3) = 5.174055 Prob.= 0.1595
O Jogo	Chi-sq(2) = 1.099755 Prob.= 0.5770	Chi-sq(2) = 10.50326 Prob.= 0.0052
Segmento Desportivo	Chi-sq(2) = 0.323618 Prob.= 0.8506	Chi-sq(2) = 4.604178 Prob.= 0.1000
Diário Económico	Chi-sq(2) = 3.514897 Prob.= 0.1725	Chi-sq(2) = 4.194279 Prob.= 0.1228
Jornal de Negócios	Chi-sq(5) = 8.299124 Prob.= 0.1405	Chi-sq(5) = 2.763765 Prob.= 0.7363
Economy and Business Segment	Chi-sq(2) = 7.833907 Prob.= 0.0199	Chi-sq(2) = 1.490584 Prob.= 0.4746
Correio da Manhã	Chi-sq(2) = 3.542357 Prob.= 0.1701	Chi-sq(2) = 4.758049 Prob.= 0.0926
Diário de Notícias	Chi-sq(3) = 3.899773 Prob.= 0.2725	Chi-sq(3) = 2.275498 Prob.= 0.5172
I	Chi-sq(2) = 1.484799 Prob.= 0.4760	Chi-sq(2) = 0.133505 Prob.= 0.9354
Jornal de Notícias	Chi-sq(2) = 4.270212 Prob.= 0.1182	Chi-sq(2) = 1.118709 Prob.= 0.5716
Público	Chi-sq(2) = 2.144375 Prob.= 0.3423	Chi-sq(2) = 0.645864 Prob.= 0.7240
Segmento Informação Geral	Chi-sq(3) = 0.763432 Prob.= 0.9793	Chi-sq(3) = 4.572450 Prob.= 0.4702
Total Market	Chi-sq(3) = 4.878092 Prob.= 0.1809	Chi-sq(3) = 4.663613 Prob.= 0.1982

<sup>26</sup> Lopes, A (2015)

Como podemos constatar (e com as devidas reservas), apenas verificamos evidências estatísticas de rejeição da ausência de causalidade à Granger a um nível de 0.05 em duas situações: no caso das vendas em banca no jornal “O Jogo” terem um impacto (negativo) nas visitas, e no caso das visitas no segmento de economia e negócios terem um impacto (positivo) nas vendas em banca.

Porém, a análise individual não permite (apesar da análise anterior para cada segmento de mercado e para o total de mercado) estimar efeitos e testar a causalidade a nível global. Se a estrita análise independente pode constituir uma limitação, uma perspectiva importante é a estimação e teste de efeitos e causalidades a nível global e para cada segmento de mercado. É exactamente neste quadro que a metodologia em painel desempenha um papel essencial. Após a introdução por Sims (1980) da metodologia VAR como alternativa a equações simultâneas no campo da macroeconometria, Holtz-Eakin, et al. (1988) expandiram esta aproximação a estruturas em painel, conduzindo a um vasto número de aplicações.

No nosso caso, consideramos um VAR num painel  $k$ -variado de ordem  $p$  representado pelo seguinte sistema de equações lineares:

$$\Delta Y_{it} = \sum_1^p \Delta Y_{it-j} A_j + X_{it} B + u_i + e_{it} \quad (6)$$

$$i \in \{1, 2, \dots, N\}, t \in \{1, 2, \dots, T\}$$

onde  $\Delta Y_{it}$  representa o vector das variáveis endógenas (taxa de crescimento das vendas em banca e visitas aos *websites*),  $X_{it}$  representa o vector de variáveis exógenas,  $u_i$  e  $e_{it}$  são termos de efeitos fixos e erros idiossincráticos respectivamente.  $A_j$  e  $B$  são as matrizes de parâmetros a estimar.

A aplicação desta metodologia deve ter em atenção alguns aspectos para estimação. De facto, dada a correlação entre  $\Delta Y_{it-j}$  (as variáveis endógenas desfasadas) e  $u_i$  (os efeitos individuais), mesmo respeitando as condições de estacionaridade, os coeficientes *OLS* serão enviesados (sobre-estimados) e inconsistentes, enquanto estimações *within* (*least squares dummy variables* – LSDV) resultarão numa sub-

estimação dos parâmetros<sup>28</sup>. Embora a maioria das pesquisas nesta área econométrica seja no campo uni-equacional (e para valores reduzidos de T), alguns avanços e estudos de simulação<sup>29</sup> fornecem um vasto conjunto de correcções do enviesamento e estratégias de estimação. De entre os mais usados, o estimador de Anderson e Hsiao (1982) – o estimador AH – onde a transformação em primeiras diferenças pode ser consistentemente estimada equação-a-equação instrumentando os desfasamentos das primeiras diferenças com diferenças e níveis de  $Y_{it}$ . Todavia, o estimador de Anderson e Hsiao tem a propriedade de amplificar diferenças em painéis não equilibrados<sup>30</sup>, o que é o nosso caso. Como alternativa, Arellano e Bond (1991) propõem o estimador standard GMM. Enquanto o estimador de Arellano e Bond é apenas aplicado às equações em diferenças, mais recentemente, Arellano e Bover (1995) e Blundell e Bond (1998) apresentaram estimadores GMM alternativos que usam condições de momentos adicionais onde as diferenças desfasadas são usadas como instrumentos nas equações em níveis. Blundell e Bond (1998) e Blundell, Bond, e Windmeijer (2000) mostram que estes estimadores produzem resultados com enviesamentos significativamente inferiores aos de Arellano e Bond (1991). No nosso caso seguiremos a metodologia sugerida por Arellano e Bover (1995). Essencialmente, o estimador minimiza a perda de observações e as fraquezas do estimador das primeiras diferenças utilizando a transformação “*forward orthogonal deviation*”. Em vez de calcular desvios face a observações passadas, esta transformação subtrai a média de todas as observações futuras. Neste caso, como não estamos a incluir observações passadas na transformação, estes valores podem ser utilizados como instrumentos válidos<sup>31</sup>.

Para estimação recorreu-se ao procedimento PVAR para o *software* Stata cuja última versão e actualização é apresentada por Abrigo e Love (2015).

---

<sup>28</sup> Bond, Nauges e Windmeijer (2002), “Unit Roots and Identification in Autoregressive Panel Data Models”

<sup>29</sup> Como é o exemplo do trabalho de Mutl(2009) – “Panel VAR Models with Spatial Dependence” ou Juessen and Linnemann (2010) – “Estimating panel VARs from macroeconomic data: Some Monte Carlo evidence and an application to OECD public spending shocks”, que fornece *insights* valiosos sobre a consistência dos estimadores de VAR em painel.

<sup>30</sup> Abrigo, M. e Love, I. (2015)

<sup>31</sup> Abrigo, M. e Love, I. (2015)

Desta forma, passando para uma estrutura com um painel mensal com as dez publicações em análise, e utilizando o estimador de Arellano e Bover (1995), os resultados da estimação podem ser encontrados na tabela seguinte:

**Tabela IV: Estimadores GMM para o VAR em painel para taxas de crescimento mensais**

	Dependent = $\Delta$ Sales							Dependent = $\Delta$ Website					
	Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]		Coef.	Std. Err.	z	P>z	[95% Conf.	Interval]
$\Delta$ Sales							$\Delta$ Sales						
L1.	-0.41531	0.057	-7.29	0.000	-0.527	-0.304	L1.	-0.04075	0.041	-0.98	0.326	-0.122	0.041
L2.	-0.34538	0.068	-5.07	0.000	-0.479	-0.212	L2.	-0.00966	0.043	-0.22	0.822	-0.094	0.075
L3.	-0.12181	0.055	-2.21	0.027	-0.230	-0.014	L3.	-0.07400	0.040	-1.86	0.064	-0.152	0.004
L4.	-0.17174	0.049	-3.52	0.000	-0.267	-0.076	L4.	-0.03167	0.037	-0.85	0.394	-0.105	0.041
$\Delta$ Website							$\Delta$ Website						
L1.	0.07026	0.037	1.88	0.060	-0.003	0.144	L1.	-0.13049	0.042	-3.08	0.002	-0.213	-0.048
L2.	0.02324	0.033	0.71	0.477	-0.041	0.087	L2.	-0.20769	0.039	-5.29	0.000	-0.285	-0.131
L3.	-0.05474	0.036	-1.53	0.125	-0.125	0.015	L3.	-0.03285	0.040	-0.83	0.408	-0.111	0.045
L4.	0.00581	0.037	0.16	0.874	-0.066	0.077	L4.	-0.07871	0.041	-1.92	0.055	-0.159	0.002
X_AJUDAEXTERNA	0.07434	0.042	1.79	0.074	-0.007	0.156	X_AJUDAEXTERNA	-0.02231	0.016	-1.37	0.169	-0.054	0.010
X_BANIF	0.06637	0.028	2.34	0.019	0.011	0.122	X_BANIF	0.01979	0.036	0.55	0.584	-0.051	0.091
X_BES	0.09876	0.025	3.94	0.000	0.050	0.148	X_BES	0.01867	0.027	0.68	0.495	-0.035	0.072
X_SECRETSERVICES	-0.04598	0.023	-1.96	0.050	-0.092	0.000	X_SECRETSERVICES	-0.00450	0.018	-0.25	0.800	-0.039	0.030
X_ELECTIONS	0.05715	0.016	3.49	0.000	0.025	0.089	X_ELECTIONS	0.04325	0.016	2.77	0.006	0.013	0.074
X_EUSEBIO	-0.01782	0.012	-1.48	0.139	-0.041	0.006	X_EUSEBIO	0.04167	0.017	2.48	0.013	0.009	0.075
X_MINISTERGRADUATION	-0.03876	0.014	-2.8	0.005	-0.066	-0.012	X_MINISTERGRADUATION	0.04322	0.016	2.73	0.006	0.012	0.074
X_BUDGET	-0.04972	0.025	-1.99	0.047	-0.099	-0.001	X_BUDGET	-0.05352	0.019	-2.77	0.006	-0.091	-0.016

X_CPCOURTSENT	-0.04312	0.013	-3.35	0.001	-0.068	-0.018	X_CPCOURTSENT	-0.02517	0.018	-1.38	0.168	-0.061	0.011
X_TERROR	0.03045	0.019	1.62	0.106	-0.006	0.067	X_TERROR	0.02525	0.025	1.02	0.309	-0.023	0.074
X_CONSTCOURTDECISION	0.04495	0.027	1.66	0.097	-0.008	0.098	X_CONSTCOURTDECISION	-0.06251	0.018	-3.54	0.000	-0.097	-0.028
X_TROIACRISIS	-0.10914	0.050	-2.17	0.030	-0.208	-0.010	X_TROIACRISIS	0.00575	0.035	0.16	0.871	-0.064	0.075
X_UNEMPLOYMENT	-0.00154	0.001	-2.27	0.023	-0.003	0.000	X_UNEMPLOYMENT	0.00169	0.001	1.95	0.052	0.000	0.003
X_FIXEDINTERNET	-0.00003	0.000	-2.52	0.012	0.000	0.000	X_FIXEDINTERNET	0.00000	0.000	0.1	0.919	0.000	0.000
X_ECON_SENT_IND	-0.00172	0.001	-2.01	0.044	-0.003	0.000	X_ECON_SENT_IND	0.00189	0.001	1.72	0.086	0.000	0.004

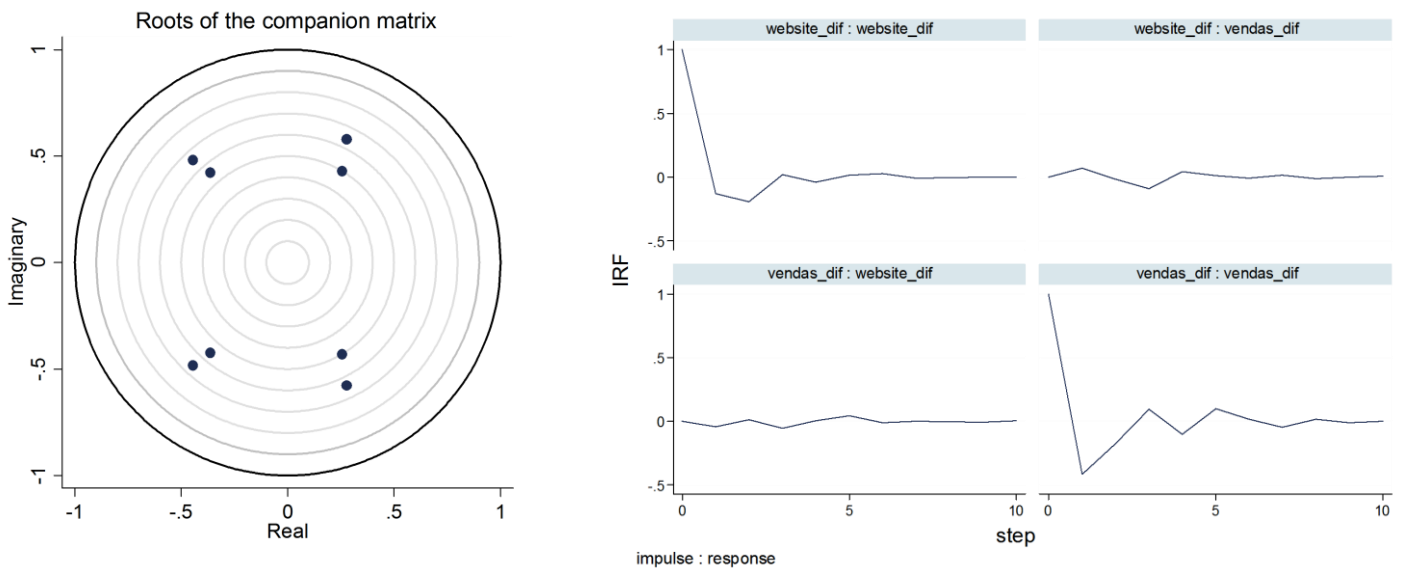
Na tabela anterior podemos ver os estimadores GMM. No nosso caso, é importante referir que os instrumentos são constituídos pelas variáveis que contêm o prefixo “X” (as variáveis exógenas), e desfasamentos (1 a 5) tanto de vendas como visitas.

Note-se como nenhum dos efeitos cruzados são significativos ao nível de 0.05.

O teste de Hansen de sobre-identificação ( Hansen's J  $\chi^2(4) = 1.2920987$  ( $p = 0.863$ )), não rejeita a hipótese nula que os instrumentos são válidos e não correlacionados com os erros.

Com esta especificação mais complexa, as condições de estabilidade são um aspecto relevante. Como podemos ver na figura abaixo, todos os valores próprios estão dentro do círculo unitário, satisfazendo assim o nosso VAR as condições de estabilidade.

**Figura 2 – Condições de Estabilidade e Funções de Resposta a Impulsos**



Com esta estrutura, analisemos o teste de causalidade à Granger em painel associado.

**Tabela V: Teste de Causalidade de Granger em Painel**

Equation \ Excluded	chi2	df	Prob
$\Delta$ Sales			
$\Delta$ Website	7.376	4	0.117
$\Delta$ Website			
$\Delta$ Sales	4.514	4	0.341

Uma vez que:

- Ho: A variável excluída não causa à Granger a variável dependente,
- Ha: A variável excluída causa à Granger a variável dependente,

Uma vez mais não rejeitamos a hipótese nula, isto é, nem as vendas nem as visitas causam à Granger as taxas de crescimento das visitas ou das vendas respectivamente.

Na análise dos factores exógenos que influenciam os consumos, podemos comparar os efeitos estimados. A título de exemplo, note-se como a variável X\_BES<sup>32</sup> que representa conteúdos editoriais, é uma das variáveis que representa um impacto com maior magnitude nas vendas em banca. No caso, estes conteúdos editoriais representaram em média um acréscimo na taxa de crescimento de quase 10 pontos percentuais. Em contraste, a mesma variável não revelou um efeito significativo no consumo *online*.

<sup>32</sup> Ver Apêndice IV – Glossário de Variáveis

Também a título de exemplo, veja-se o efeito contrário que os conteúdos editoriais associados à variável X\_MINISTERGRADUATION promoveram nas vendas e visitas: em média, motivaram o desinteresse dos leitores tradicionais (uma taxa de crescimento das vendas inferior em 3.9 pontos percentuais), mas promoveram mais actividade digital (com um aumento da taxa de crescimento das visitas em 4.3 pontos percentuais). Naturalmente, a fundamentação destes fenómenos e dos restantes cabe fundamentalmente ao conhecimento do negócio e aos *stakeholders* envolvidos. A leitura de alguns efeitos aparentemente contrários ao que seria esperado devem ter em linha de conta quer o tempo, quer o facto de estarmos a tratar informação mensal, quer o volume (número) de conteúdos e edições associados a cada tema.

Como resultado complementar aos VAR individuais de cada segmento de mercado, recorreremos a painéis independentes por segmento de mercado para identificar potenciais relações de causalidade em cada segmento. Mais do que a estimação dos parâmetros, apresentamos os respectivos testes de causalidade de Granger para cada um dos segmentos.

**Tabela VI–Testes de Causalidade de Granger para VAR em Painel para cada segmento de mercado**

	Desporto	Economia e Negócios	Informação Geral
Equation \ Excluded	chi2 df Prob.	df	Prob
$\Delta$ Website -> $\Delta$ Sales	4.442 4 0.349	9.676 4 0.046	5.967 4 0.202
$\Delta$ Sales -> $\Delta$ Website	13.997 4 0.007	9.412 4 0.052	3.187 4 0.527

Para um nível de 0.05 podemos rejeitar a hipótese nula de ausência de causalidade para três relações: no segmento desportivo as vendas em banca causam à Granger as visitas aos Websites, e no segmento de Economia e Negócios verificamos um caso de potencial Feedback. Neste último caso, os únicos parâmetros significativos são mesmo  $a_{12,1}$  e  $a_{12,2}$  (o efeito cruzado das visitas nas vendas e o efeito cruzado das vendas nas visitas respectivamente), que assumem os valores 0.24 e -0.134, indicando que as visitas têm um impacto positivo nas vendas em papel, enquanto as vendas em papel têm um

impacto negativo nas visitas online. De igual forma no segmento desportivo, todos os efeitos cruzados das vendas nas visitas são negativos.

## 4 – Conclusões

Em termos gerais, as implicações mais relevantes da pesquisa são tanto a ausência de evidências estatísticas de dinâmicas de longo-prazo entre as vendas e visitas no mercado dos diários em Portugal, como as (escassas) evidências de dinâmicas de curto-prazo serem diferentes em diferentes segmentos de mercado.

Efectivamente, encontramos evidências estatísticas de feedback instantâneo em todas as publicações do segmento desportivo e para o Jornal de Negócios em particular no segmento de economia e negócios, e por outro lado ausência de causalidade no segmento de informação geral. Numa análise global, são também identificadas evidências de feedback no segmento de economia e negócios. No caso, o painel para este segmento de mercado revela ainda uma realidade interessante: se por um lado o impacto das vendas em banca nas visitas é negativo, por outro, as visitas têm um impacto positivo nas vendas em banca (o que pode revelar que a venda em banca canibaliza o consumo online, mas que temas e conteúdos de interesse online promovem o interesse pelas edições em papel neste segmento em particular).

Os resultados, por um lado são convergentes com Deleersnyders et al. (2002), Kaiser (2002), e Fodor et al. (2013), orientando-nos para referir que tanto o consumo digital como o consumo em papel dependem essencialmente da sua **dinâmica própria**, bem como de outros **factores exógenos** (conteúdos, economia, tecnologia essencialmente), podendo tratar-se de mercados distintos com audiências distintas. Ao mesmo tempo, e no segmento de mercado dos jornais diários, os resultados divergem da substituição encontrada por Filistrucchi (2005) no mercado italiano ou de Pino (2010) no mercado americano, ou da complementaridade encontrada por Pauwels e Dans (2001) no mercado espanhol. É aliás a inclusão de factores exógenos que se revela determinante no nosso caso para remover hipotéticas relações de causalidade de natureza espúria. Efectivamente, os conteúdos assumem um papel determinante na procura de qualquer dos consumos. Ao identificarmos conteúdos editoriais que tanto promovem

simultaneamente os mesmos efeitos (positivos ou negativos) nos dois tipos de consumo, como provocam efeitos contrários, vemos o resultado ir de alguma forma ao encontro de Chyi (2012) que identifica que a natureza do acréscimo ou decréscimo das intenções de procura no sector estão muito mais ligadas ao interesse dos conteúdos e hábitos de leitura do que necessariamente à tecnologia. A identificação de um conjunto de conteúdos que provocam efeitos contrários no consumo em papel ou digital, vem dar força à necessidade de análise mais aprofundada de conteúdos editoriais e aos diferentes perfís de audiência existentes como sugere Ahlers (2006) ou Fodor et al. (2013). O próprio caso de feedback instantâneo no segmento desportivo e de economia e gestão, onde os efeitos instantâneos são positivos, orientam-se para as conclusões de Pauwels e Dans (2001) no mercado espanhol (cuja diferente frequência se aproxima de efeitos instantâneos). Estes efeitos voltam a colocar o tema dos conteúdos como o centro e o potencial principal *driver* de qualquer dos consumos.

Todavia, a ausência de mais fortes evidências estatísticas remete-nos para novas pistas e linhas de investigação. Por exemplo, se no caso os dados são mensais, futuras investigações devem orientar-se para a análise de duas frequências extremas onde os efeitos poderão ser diferentes: a alta frequência (quase instantaneidade) com, por exemplo dados diários ou com maior frequência, e a baixa frequência, por exemplo com períodos mais longos de tempo (20, 30 anos de dados), para que possa ser plausível estimar efeitos quer de muito curto-prazo, quer de longo-prazo. A frequência dos dados é de extrema importância uma vez que é razoável assumir que existirão, com certeza, diferentes efeitos em diferentes frequências. Ao mesmo tempo, é relevante sublinhar que a realidade digital é mais complexa do que a análise do eventual trade-off entre papel e *website* de uma publicação. Efectivamente, o cenário concorrencial é mais vasto, e será necessário investigar efeitos que outros *websites* e concorrentes digitais têm nas publicações (por exemplo, todos os *websites* de canais de TV dispõem de serviços informativos *online*, ou distinguir o consumo *via-desktop* ou *via-mobile*). Ao mesmo tempo, apesar de não partilharmos o entusiasmo de algumas pesquisas que identificam efeitos de complementaridade entre os consumos, nem de identificarmos a clara canibalização encontrada por outros, para fornecer recomendações sobre potenciais

modelos de negócio (que em alguns casos recomendam a estratégia radical de modelos exclusivamente digitais como sobrevivência para a imprensa escrita), as pesquisas devem ser completas e incluir aspectos do negócio adicionais como publicidade e assinaturas.

De qualquer forma, as evidências empíricas da presente pesquisa orientam no sentido de considerar o negócio digital como mais um *side* do *multi-sided market* deste sector.

Assumir que os consumos *online* e *offline* são independentes sugere que os canais digitais alcançam segmentos de mercado independentes, de novos leitores e não-convencionais, com diferentes características demográficas, sociais e económicas, sugerindo uma nova questão: Será esta “independência” verdadeira para diferentes segmentos demográficos, económicos ou sociais da população? Claramente, a segmentação da análise para grupos sociais e económicos distintos é outra perspectiva de investigação futura. Por outro lado, a ausência (hoje) de evidência estatística a favor da causalidade pode ser um reflexo do (ainda inicial) estágio de maturidade do mercado português.

Numa era de rápida inovação tecnológica e incerteza, tanto os novos equipamentos e canais digitais, como os instrumentos científicos para observar, medir e adequadamente identificar impactos e efeitos, constituem desafios e oportunidades para uma melhor compreensão das relações entre um mundo físico e digital em constante evolução...

*“There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.” Charles Darwin, On the Origin of Species, (1859)*

## Referências

- Abrigo, M. R. M., Love, I., 2015. "Estimation of Panel Vector Autoregression in Stata: a Package of Programs." manuscrito , Febr 2015 disponível em <http://paneldataconference2015.ceu.hu/Program/Michael-Abrigo.pdf>
- Ahlers, 2006, "News Consumption and the New Electronic Media", *Press/Politics* 11(1):29-52, DOI: 10.1177/1081180X05284317.
- Anderson, T. W. e Hsiao, Cheng, 1982. "Formulation and estimation of dynamic models using panel data," *Journal of Econometrics*, Elsevier, vol. 18(1), pages 47-82.
- Arellano e Bond, 1991, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", *Review of Economic Studies*, 1991, vol. 58, issue 2, pages 277-297.
- Arellano e Bover, 1995, "Another look at the instrumental variable estimation of error-components models", *Journal of econometrics* 68 (1), 29-51.
- Argentesi, E. e Filistrucchi, L., 2005, "Estimating Market Power In A Two-Sided Market: The Case Of Newspapers", *EUI Working Paper Eco N° 2005/7*. European University Institute.
- Barsh, J., Lee, G. e Miles, A., 1999, "Beyond Print: A Future For Magazines", *The Mckinsey Quarterly* 1999 (3), 122-130.
- Binder, M., Hsiao, C. and Pesaran, M., 2004, "Estimation and inference in short panel vector autoregressions with unit roots and cointegration", *Banco de España*.
- Blundell e Bond, 1998, "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics* 87, 115—143.
- Blundell, R., Bond, S., Windmeijer, F., 2000, "Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator", *IFS Working Papers W00/12*, Institute for Fiscal Studies: London, UK.
- Bond, Stephen, Nauges, Celine and Windmeijer, Frank, (2002), "Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests",

No C5-4, 10th International Conference on Panel Data, Berlin, July 5-6, 2002, International Conferences on Panel Data

Canova, F. e Ciccarelli, M., 2013, "Panel Vector Autoregressive Models A Survey", Working Paper Series Nº 1507. European Central Bank.

Chamberlain, G., 1987, "Asymptotic efficiency in estimation with conditional moment restrictions", *Journal of Econometrics*, 1987, vol. 34, issue 3, pages 305-334.

Chyi, H., 2012, "Are 'Digital Natives' Dropping Print Newspapers? A National Survey Of College Newspaper Advisers", The University Of Texas.

Corden, 1953, "The Maximisation of Profit by a Newspaper", *Review of Economic Studies* 20(3), 181.

Deleersnyder, B., Geyskens, I., Gielens, K., Dekimpe, M.G., 2002, "How Cannibalistic is the Internet Channel? A Study of the Newspaper Industry in the United Kingdom and The Netherlands", *International Journal of Research in Marketing*, Vol. 19, No. 4, pp. 337-348.

Eimeren, B. v., e Vrees, B., 2011, "Drei von vier Deutschen im Netz – ein Ende des digitalen Grabens in Sicht? Ergebnisse der ARD/ZDF Online Studie 2011", *Media Perspektiven*, pp. 334-349.

Leurdijk et al, 2012, "The Newspaper Publishing Industry", Joint Research Report, European Commission.

Alturas, B. e Fernandes, F., 2013, "Press Consumption In The Digital Age: Habits And Needs Regarding Online Press", University Of Glamorgan.

Filistrucchi, L., 2005, "The Impact Of Internet On The Market For Daily Newspapers In Italy", *EUI Working Paper Eco Nº 2005/12*. European University Institute.

Fodor et al., 2013, "Print Media Versus Digital Media and the Expected Results Based on the Results of a Comprehensive Trend Research", DOI: 10.7763/IPEDR. 2013. V59. 42.

Gentzkow, M., 2003, "Valuing New Goods In A Model With Complementarity: Online Newspapers", *The American Economic Review* 97 (3), 713-744.

Granger e Engle, 1987, “Co-Integration and Error Correction: Representation, Estimation, and Testing”, *Econometrica*, Vol. 55, No. 2, pp. 251-276.

Granger e Newbold, 1974, “Spurious regressions in econometrics”, *Journal of Econometrics*, vol. 2, issue 2, pages 111-120.

Hansen, 1982, “Large Sample Properties of Generalized Method of Moments Estimators”, *Econometrica*, vol. 50, issue 4, pages 1029-54.

Holtz-Eakin, D., Newey, W. e Rosen, H., 1988, “Estimating Vector Autoregressions with Panel Data”, *Econometrica*, Vol. 56, No. 6., pp. 1371-1395.

Huysmans, F. e J. de Haan, 2010, “Alle kanalen staan open; digitalisering van het mediagebruik”, Haia, Sociaal en Cultureel Planbureau.

Juessen, F. e Linnemann, L., 2010, “Estimating panel VARs from macroeconomic data: Some Monte Carlo evidence and an application to OECD public spending shocks”, SFB 823 Discussion Paper, TU Dortmund University.

Kaiser, U., 2002, “The Effects Of Website Provision On The Demand For German Women’s Magazines”, Working Paper 8806. National Bureau Of Economic Research.

Kaiser, U. e Kongsted, H., 2005, “Do Magazines’ *Companion Websites* Cannibalize The Demand For The Print Version?”, Centre For Economic And Business Research, Copenhagen & Centre For Applied Microeconometrics (CAM), Copenhagen.

Knuff, C., 2011, “Has The Development Of Online Newspaper Components Positively Or Negatively Impacted Newspaper Print Sales?”, College Of The Holy Cross.

Lauf, E., Schönbach, K., e Waal, E., 2005, “Communications”, *The European Journal of Communication Research*, 30 (2005), 55-72.

Li, H., 2013, “The Impact Of Ebooks On Print Book Sales: Cannibalization And Market Expansion”, University Of Pennsylvania.

Liebowitz, Stan J., e Alejandro Zentner, 2012, “Clash of the Titans: Does Internet Use Reduce Television Viewing?”, *Review of Economics and Statistics* 94 (1): 234–45.

Lopes, A, 2015, “Raízes Unitárias-Uma Introdução”, Coleção Económicas, Série II, nº 26, Almedina.

Lucena, A., 2010, “The Print Newspaper In The Information Age: An Analysis Of Trends And Perspectives”, University Of Alberta.

Lutkepohl, H., 2005, “New Introduction To Multiple Time Series Analysis”, Springer.

Martin et al, 2013, “Econometric Modelling with Time Series”, Cambridge University Press.

Mitchell, 2001, “The year of living dangerously”, Editor & Publisher 134(4), 20-26.

Mutl, J., 2009, “Panel VAR models with spatial dependence”.

Newman et al., 2012, “Social Media in the Changing Ecology of News: The Fourth and Fifth Estates in Britain”, International Journal of Internet Science, 2012, 7 (1), 6–22, ISSN 1662-5544.

Obercom, 2014, Barómetro Da Comunicação, 10ª Ed.

OECD, 2010, “The evolution of news and the internet”, Paris: OECD.

Patel, A., 2010, “The Survival Of The Newspaper In The Digital Age Of Communication”, New York University.

Pauwels, K., Dans, E., 2001, “Internet marketing the news: leveraging brand equity from marketplace to marketspace”, Brand Management 8(4), 303-314.

Pew Research Center’s Project For Excellence In Journalism, 2012, “The State Of The News Media 2012”, Annual Report On American Journalism.

Pino, A., 2010, “Did The Internet Kill Magazines And Newspapers?”, Harvard College.

PWC Global Entertainment & Media Outlook 2013-2017.

Rathmann, T., 2002, “Supplement or substitution? The relationship between reading a local print newspaper and the use of its online version”, Communications, Volume 27, Issue 4, Pages 485–498, DOI: 10.1515/comm.2002.004.

Reddaway, 1963, “The economics of newspapers”, Economic Journal, 73, 201-218.

Schrøder, K. C., and Kobbarnagel, Christian, 2010, "Towards a typology of cross-media news consumption: a qualitative-quantitative synthesis", *Northern Lights*, vol. 8, pp. 115-138.

Schrøder e Larsen, 2010, "The shifting cross-media news landscape: Challenges for news producers", *Journalism Studies* 11 (4), 524-534.

Seamans, R. e Zhu, F., 2011, "Technology Shocks In Multi-Sided Markets: The Impact Of Craigslist On Local Newspapers", Copenhagen Business School.

Simon, D., 2004, "The Effect of a magazine's digital content on its print circulation: cannibalization or complementarity?", Cornell University working paper.

Sims, C., 1980, "Macroeconomics And Reality". *Econometrica*, Vol. 48 (1), 1-48.

Sims, Stock e Watson, 1990, "Inference in linear time series models with some unit roots", *Econometrica*, Vol. 58, No. 1, 113-144.

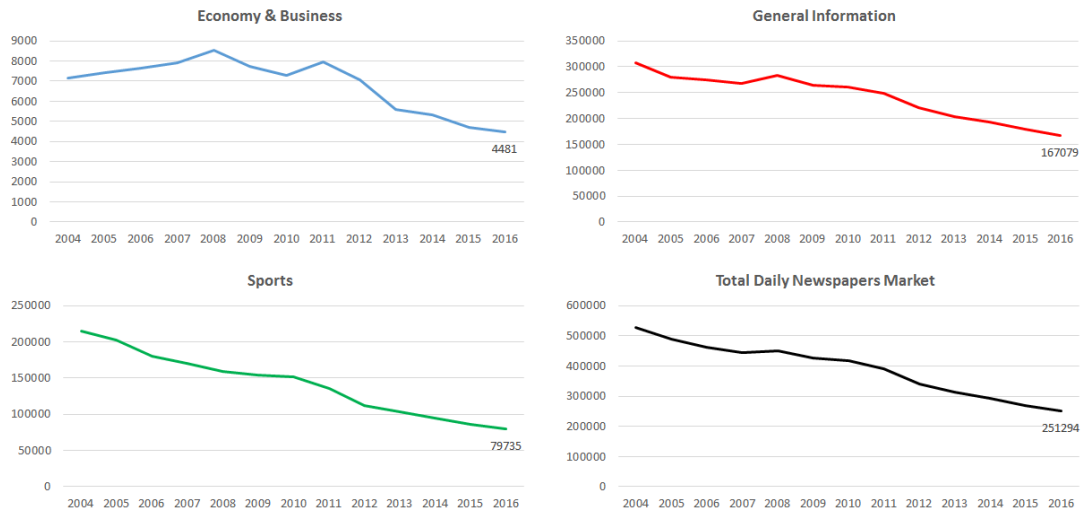
Slot e Munniks de Jongh Luchsinger, 2011, "To read or not to read , Een onderzoek naar de meest gebruikte Nederlandse nieuwsmedia" , INM (2011).

Wan-Ifra & NAA (2014). *World Press Trends 2014*.

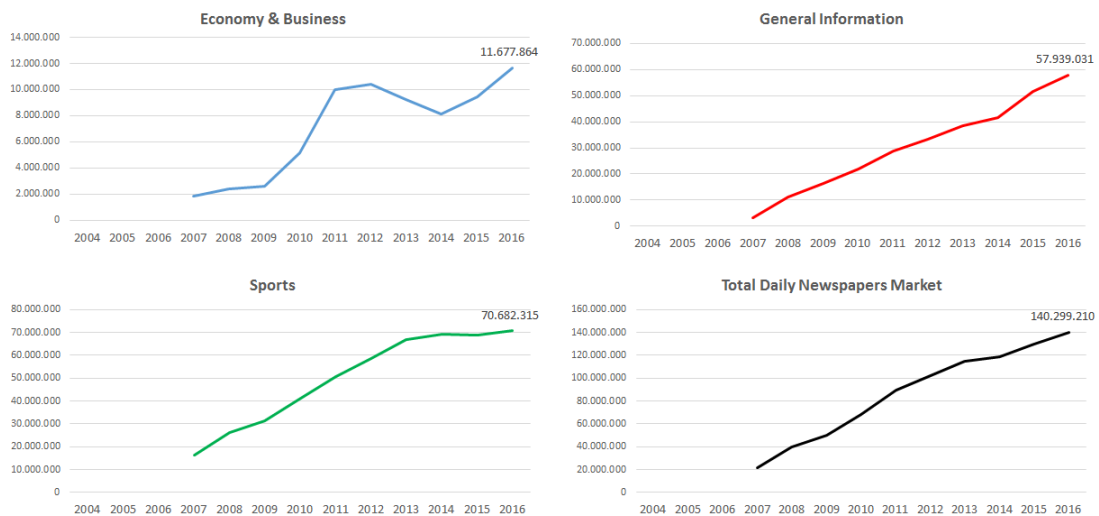
Wan-Ifra & NAA (2014). *World Press Trends 2015*.

## Apêndice I – Gráficos e Figuras

**Figura 3: Vendas médias diárias (2004- Fev/2016)** <sup>33</sup>



**Figura 4: Visitas médias mensais** <sup>34</sup>



<sup>33</sup> Fonte: APCT, 1º Bimestre de 2016; Sociedade Vicra Desportiva SA

<sup>34</sup> Fonte: Marktest - Ranking netScope de tráfego web, Junho 2016 ((Visitas: número de visitas recebidas no site, ou seja, o número total de sessões dos internautas, com um intervalo de tempo entre duas sessões para um mesmo internauta de 30 minutos mínimo (norma IFABC): se um visitante efectua duas visitas no seu site num intervalo inferior a 30 minutos, apenas uma visita é considerada))

Figura 5: Evolução do logaritmo das vendas em banca<sup>35</sup>

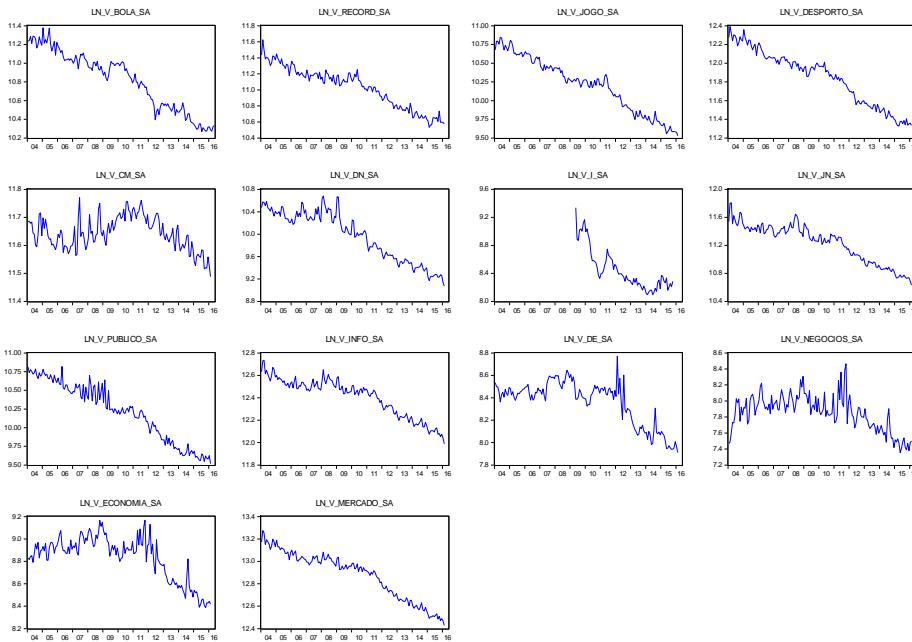
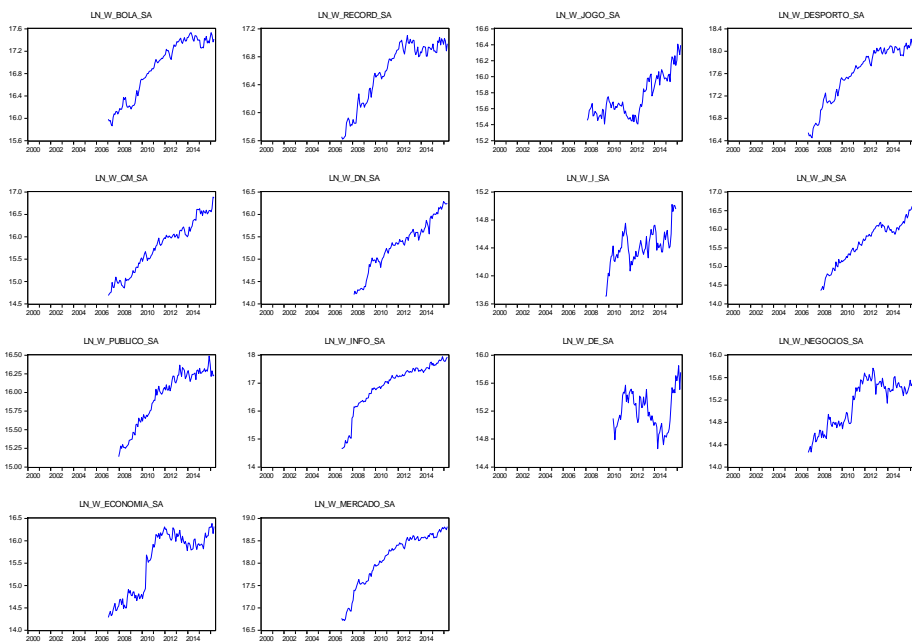


Figura 6: Evolução do logaritmo das visitas aos *websites*



<sup>35</sup> Nota: as variáveis estão identificadas com a estrutura LN\_Y\_PUBLICAÇÃO\_SA. Por exemplo, tal como LN\_V\_BOLA\_SA se refere ao logaritmo das Vendas (V) com ajustamento sazonal (SA) da publicação “A Bola”, LN\_W\_BOLA\_SA refere-se ao logaritmo das visitas ao *website* (W) também ajustadas sazonalmente da mesma publicação.

## Apêndice II – Testes de Raízes Unitárias e Cointegração

Antes da estimação dos VAR, testámos a existência de raízes unitárias nas taxas de crescimento. Numa aproximação *GTS t-sig*, começamos definindo  $k_{max}$  usando a habitual regra  $l_{12}$  para definir a ordem máxima de desfasamento onde:

$$l_{12}: k_{max} = \left[ 12 * \left( \frac{T}{100} \right)^{1/4} \right] = \left[ 12 * \left( \frac{90}{100} \right)^{1/4} \right] \cong 12 \quad (9)$$

Utilizando testes ADF, a equação inicial de testes será para cada variável:

$$\Delta y_t = \alpha + \phi y_{t-1} + \sum_{j=1}^{12} \gamma_j \Delta y_{t-j} + v_t \quad (10)$$

Onde no caso,  $y_t = [\Delta v_t \ \Delta w_t]'$ .

**Tabela VII – Testes ADF para as taxas de crescimento mensais**

	$y_t$	$\hat{k}$	$\hat{\phi}$	ADF			$\hat{\gamma}_k/se(\hat{\gamma}_k)$	BG(2)	
				$\hat{\phi}/se(\hat{\phi})$ (ADF)	p-value	Critical Value 1%		N*R <sup>2</sup>	p-value
Vendas $\Delta v_t$	A Bola	1	-1.62220	-13.07609	0	-3.472534	3.570897	2.970128	0.2265
	Record	2	-1.79713	-11.50629	0	-3.464827	1.832825	4.961014	0.0837
	O Jogo	1	-1.38639	-13.06646	0	-3.464643	3.369710	0.299949	0.8607
	Segmento Desportivo	1	-1.35007	-12.55305	0	-3.464643	2.697612	0.887443	0.6416
	Diário Económico	3	-2.05303	-11.87869	0	-3.465014	4.635377	0.893255	0.6398
	Jornal de Negócios	2	-1.90492	-12.54282	0	-3.464827	3.926533	1.906953	0.3854
	Segmento Economia e Negócios	3	-2.2806	-11.77138	0	-3.465014	2.354386	0.412171	0.8138
	Correio da Manhã	3	-1.79944	-9.526503	0	-3.465014	0.700527	4.196602	0.1227
	Diário de Notícias	3	-1.83173	-11.44483	0	-3.465014	3.971830	0.268917	0.8742
	I	0	-1.08144	-11.22085	0	-3.519050	-	0.928201	0.6287
	Jornal de Notícias	3	-1.71301	-9.827493	0	-3.465014	3.358333	3.586087	0.1665
	Público	1	-2.31038	-21.87022	0	-3.464643	8.965540	3.714959	0.1561
	Segmento Informação Geral	3	-1.88409	-10.29901	0	-3.465014	2.752899	0.771277	0.6800
Total Market	1	-1.56727	-14.95322	0	-3.464643	4.650374	1.672587	0.4333	
Visitas aos Websites $\Delta w_t$	A Bola	0	-0.97512	-10.14605	0	-3.490772		1.873433	0.3919
	Record	1	-1.19120	-8.790728	0	-3.491345	2.005874	3.457283	0.1775
	O Jogo	1	-1.50918	-10.04203	0	-3.499910	3.038387	0.226994	0.8927
	Segmento Desportivo	0	-0.98765	-10.27585	0	-3.490772		3.542254	0.1701

$y_t$	$\hat{k}$	$\hat{\phi}$	ADF			$\hat{\gamma}_k/se(\hat{\gamma}_k)$	BG(2)	
			$\hat{\phi}/se(\hat{\phi})$ (ADF)	p-value	Critical Value 1%		N*R <sup>2</sup>	p-value
Diário Económico	0	-1.15799	-9.528960	0	-3.527045		3.838594	0.1467
Jornal de Negócios	1	-1.39873	-9.924786	0	-3.491345	2.617457	5.272737	0.0716
Segmento Economia e Negócios	0	-1.13091	-11.79830	0	-3.490772		4.080681	0.1300
Correio da Manhã	0	-1.16632	-12.27687	0	-3.490772		3.370424	0.1854
Diário de Notícias	0	-1.15736	-11.45767	0	-3.499167		4.846925	0.0886
I	0	-1.04141	-8.766764	0	-3.522887		2.200302	0.3328
Jornal de Notícias	2	-1.21971	-5.485704	0	-3.500669	-2.008156	2.334026	0.3113
Público	1	-1.58937	-10.27292	0	-3.498439	2.740495	5.345428	0.0691
Segmento Informação Geral	0	-1.10336	-11.53501	0	-3.490772		0.276899	0.8707
Total Market	0	-0.99632	-10.37814	0	-3.490772		0.579234	0.7486

Dadas as equações de teste, as hipóteses de teste são:

$$H_0: \phi = 0 \text{ Vs. } H_1: \phi < 0.$$

Mesmo para um nível de significância de 1%, rejeitamos claramente a hipótese nula que as taxas de crescimento contêm uma raíz unitária. Como conclusão, aceitamos que as taxas de crescimento são I(0).

Apesar de não serem alvo do presente trabalho, os testes das variáveis em níveis encontram-se na tabela seguinte. Todavia, no caso as equações iniciais de testes assumem a seguinte estrutura:

$$\Delta y_t = \alpha + \phi y_{t-1} + \sum_{j=1}^{12} \gamma_j \Delta y_{t-j} + \beta t + v_t \tag{11}$$

Onde, temos agora,  $y_t = [v_t \ w_t]'$ .

**Tabela VIII – Testes ADF para os logaritmos das vendas e visitas**

	$y_t$	$\hat{k}$	$\hat{\phi}$	ADF			$\hat{\gamma}_k/se(\hat{\gamma}_k)$	BG(2)	
				$\hat{\phi}/se(\hat{\phi})$ (ADF)	p-value	Critical Value 1%		N*R <sup>2</sup>	p-value
Vendas $\Delta v_t$	A Bola	10	-0.16783	-2.476273	0.3395	-4.021254	-1.974786	0.139880	0.9324
	Record	10	-0.10060	-1.541086	0.8119	-4.008987	-2.379346	1.554301	0.4597
	O Jogo	2	-0.03007	-1.446776	0.8440	-4.006824	-3.334024	0.709098	0.7015
	Segmento Desportivo	2	-0.02715	-1.461665	0.8393	-4.006824	-2.687963	0.675508	0.7134
	Diário Económico	4	-0.11453	-2.472022	0.3417	-4.007347	-4.135427	0.888481	0.6413
	Jornal de Negócios	3	-0.15879	-3.098922	0.1096	-4.007084	-3.283398	0.493352	0.7814
	Segmento Economia e Negócios	4	-0.10453	-2.404154	0.3762	-4.007347	-2.140816	0.641284	0.7257
	Correio da Manhã	5	-0.06964	-1.898894	0.6512	-4.007613	-1.795398	1.163774	0.5588
	Diário de Notícias	4	-0.14741	-2.868513	0.1752	-4.007347	-3.114781	1.474461	0.4784
	I	0	-0.19528	-3.267803	0.0794	-4.081666		0.104024	0.9493
	Jornal de Notícias	4	-0.06380	-1.823023	0.6898	-4.007347	-3.206735	3.850907	0.1458
	Público	2	-0.10664	-2.330357	0.4151	-4.006824	-8.351157	3.878392	0.1438
	Segmento Informação Geral	4	-0.04046	-1.079141	0.9288	-4.007347	-2.699625	1.177480	0.5550
	Total Market	2	-0.02272	-1.032673	0.9359	-4.006824	-4.720361	1.865913	0.3934
Visitas aos Websites $\Delta w_t$	A Bola	0	-0.04485	-1.328157	0.8757	-4.042819		1.556984	0.4591
	Record	0	-0.06484	-1.886511	0.6549	-4.042819		4.010081	0.1347
	O Jogo	2	-0.07694	-1.366563	0.8646	-4.056461	-2.691949	0.269584	0.8739
	Segmento Desportivo	0	-0.06239	-1.846790	0.6751	-4.042819		3.366285	0.1858
	Diário Económico	0	-0.10186	-1.693752	0.7439	-4.092547		1.326822	0.5151
	Jornal de Negócios	2	-0.07156	-1.583494	0.7933	-4.044415	-2.347277	4.742388	0.0934
	Segmento Economia e Negócios	0	-0.06193	-1.862377	0.6672	-4.042819		4.413055	0.1101
	Correio da Manhã	12	-0.23423	-2.097398	0.5406	-4.053392	-1.757548	3.341567	0.1881
	Diário de Notícias	0	-0.17910	-3.154242	0.0999	-4.054393		2.556227	0.2786
	I	0	-0.23308	-3.265283	0.0802	-4.086877		0.529981	0.7672
	Jornal de Notícias	2	-0.09153	-2.003111	0.5920	-4.056461	-3.011894	5.262476	0.0720
	Público	4	-0.00488	-0.093674	0.9943	-4.056461	-2.620422	0.089730	0.9561
	Segmento Informação Geral	0	-0.06601	-2.627757	0.2691	-4.042819		2.869683	0.2382
	Total Market	0	-0.04605	-1.849079	0.6740	-4.042819		1.228408	0.5411

Como esperado, as séries em níveis mostram-se ser I(1) mas estarão cointegradas?

**Tabela IX – Testes de Cointegração de Johansen Cointegration para as variáveis em níveis <sup>36</sup>**

		Number of Coint. Equations	Trace Test		Maximum Test	
			Trace Statistic	Prob. <sup>37</sup>	Max. Eigen. Statistic	Prob. <sup>38</sup>
Bola	Constant in the coint equation and test VAR	None	11.66734	0.1736	8.628116	0.3184
		At most 1	3.039221	0.0813	3.039221	0.0813
	Constant and trend in the coint equation	None	12.43344	0.7811	9.229404	0.7007
		At most 1	3.204041	0.8514	3.204041	0.8514
Record	Constant in the coint equation and test VAR	None	9.410713	0.3288	8.026955	0.3760
		At most 1	1.383758	0.2395	1.383758	0.2395
	Constant and trend in the coint equation	None	17.14018	0.4046	11.19130	0.4938
		At most 1	5.948881	0.4668	5.948881	0.4668
Jogo	Constant in the coint equation and test VAR	None	10.35330	0.2545	10.30806	0.1925
		At most 1	0.045241	0.8315	0.045241	0.8315
	Constant and trend in the coint equation	None	18.21852	0.3294	14.86379	0.2011
		At most 1	3.354728	0.8320	3.354728	0.8320
Segmento Desportivo	Constant in the coint equation and test VAR	None	9.940280	0.2853	8.224629	0.3563
		At most 1	1.715651	0.1903	1.715651	0.1903
	Constant and trend in the coint equation	None	12.23246	0.7957	8.308478	0.7930
		At most 1	3.923987	0.7531	3.923987	0.7531
Diário Económico	Constant in the coint equation and test VAR	None	1.739347	0.9976	1.547811	0.9976
		At most 1	0.191536	0.6616	0.191536	0.6616
	Constant and trend in the coint equation	None	10.18722	0.9168	8.834355	0.7414
		At most 1	1.352863	0.9946	1.352863	0.9946
Jornal de Negócios	Constant in the coint equation and test VAR	None	8.531672	0.4104	6.497177	0.5503
		At most 1	2.034495	0.1538	2.034495	0.1538
	Constant and trend in the coint equation	None	22.74708	0.1167	17.51659	0.0916
		At most 1	5.230490	0.5637	5.230490	0.5637
Segmento Economia e Negócios	Constant in the coint equation and test VAR	None	4.437534	0.8652	4.247209	0.8324
		At most 1	0.190325	0.6626	0.190325	0.6626
	Constant and trend in the coint equation	None	18.87980	0.2880	16.95615	0.1089
		At most 1	1.923646	0.9731	1.923646	0.9731
Correio da Manhã	Constant in the coint equation and test VAR	None	8.629481	0.4008	7.985067	0.3802
		At most 1	0.644415	0.4221	0.644415	0.4221
	Constant and trend in the coint equation	None	20.70877	0.1921	13.47919	0.2909
		At most 1	7.229588	0.3212	7.229588	0.3212
Diário de Notícias	Constant in the coint equation and test VAR	None	11.33984	0.1914	10.15619	0.2019
		At most 1	1.183654	0.2766	1.183654	0.2766
	Constant and trend in the coint equation	None	22.41762	0.1269	17.37418	0.0957
		At most 1	5.043443	0.5902	5.043443	0.5902
I	Constant in the coint equation and test VAR	None	19.05488	0.0139	12.89539	0.0813
		At most 1	6.159489	0.0131	6.159489	0.0131
	Constant and trend in the coint equation	None	22.37588	0.1282	16.19508	0.1371
		At most 1	6.180801	0.4377	6.180801	0.4377
Público	Constant in the coint equation and test VAR	None	8.603748	0.4033	4.747834	0.7733
		At most 1	3.855914	0.0496	3.855914	0.0496
	Constant and trend in the coint equation	None	11.84065	0.8229	7.874169	0.8326
		At most 1	3.966484	0.7470	3.966484	0.7470
Segmento Informação Geral	Constant in the coint equation and test VAR	None	14.88047	0.0617	14.66616	0.0432
		At most 1	0.214314	0.6434	0.214314	0.6434
	Constant and trend in the coint equation	None	27.05837	0.0355	14.67973	0.2116
		At most 1	12.37865	0.0527	12.37865	0.0527
Total Market	Constant in the coint equation and test VAR	None	14.81894	0.0630	13.65993	0.0621
		At most 1	1.159005	0.2817	1.159005	0.2817
	Constant and trend in the coint equation	None	23.18891	0.1041	14.32848	0.2329
		At most 1	8.860438	0.1891	8.860438	0.1891

<sup>36</sup> Lag 4 para variáveis endógenas diferenciadas

<sup>37</sup> MacKinnon-Haug-Michelis (1999) p-values

<sup>38</sup> MacKinnon-Haug-Michelis (1999) p-values

A tabela anterior mostra os testes do Traço e do Máximo de Johansen, em ambos os casos permitindo uma tendência determinística. Podemos verificar que os testes indicam que globalmente não rejeitamos a hipótese nula de ausência de relações de cointegração a um nível de 5%. A única publicação que é exceção é o caso do “I” (note-se que mesmo neste caso, apenas o teste do Traço rejeita a hipótese de ausência de cointegração e quando a equação de cointegração não inclui o termo de tendência). Uma suspeita que será negligenciada uma vez que estamos a lidar com o título mais recente, e como tal, o que de menos observações dispõe (sendo de alguma forma excessivo derivar relações de longo-prazo com uma janela temporal tão restrita).

Em síntese, não verificamos evidências de relações de cointegração (individualmente) entre as visitas e as vendas.

### Testes em Painel

**Tabela X – Testes de Raízes Unitárias para as taxas de crescimento das vendas e visitas**

<i>Estatística (Prob.)</i> <sup>39</sup>	$\Delta$ Sales	$\Delta$ Website	Sales	Website
Levin, Lin & Chu t	-7.09324 (0.00)	-13.0135 (0.00)	-0.49378 (0.3107)	-3.10415 (0.0010)
Im, Pesaran and Shin W-stat	-21.5801 (0.00)	-18.1267 (0.00)	1.74889 (0.9598)	0.16712 (0.5664)
ADF - Fisher Chi-square	381.434 (0.00)	312.492 (0.00)	11.7657 (0.9239)	16.7012 (0.6723)
PP - Fisher Chi-square	339.944 (0.00)	413.555 (0.00)	19.4895 (0.4902)	21.7941 (0.3518)

Como vemos na tabela anterior, podemos rejeitar a hipótese nula da existência de raízes unitárias nas taxas de crescimento para os níveis de significância habituais (não rejeitando para as séries em nível, identificando que são I(1)).

De idêntica forma, testando a cointegração em painel:

<sup>39</sup> As probabilidades para os testes de Fisher são calculados usando uma distribuição assintótica de qui-quadrado. Os restantes testes assumem uma distribuição assintótica normal.

**Tabela XI – Testes de Cointegração em Painel**

Test	Assumption	Number of Coint. Equations	Trace Test		Maximum Test	
			Trace Statistic	Prob.	Max. Eigen. Statistic	Prob.
<b>Johansen, Fisher</b>	<b>No deterministic trend</b>	0	31.46	0.05	20.00	0.46
			<b>t Statistic</b>	<b>Prob.</b>		
<b>Kao</b>	<b>No deterministic trend</b>		0.260593	0.40		
			<b>Statistic</b>	<b>Prob.</b>	<b>Weighted Statistic</b>	<b>Prob.</b>
<b>Pedroni<sup>40</sup></b>	<b>No deterministic trend</b>	Panel v-Statistic	0.560553	0.2876	0.433628	0.3323
		Panel rho-Statistic	-7.808217	0.0000	-7.211771	0.0000
		Panel PP-Statistic	-5.172442	0.0000	-4.851331	0.0000
		Panel ADF-Statistic	-0.115481	0.4540	0.265677	0.6048
		Group rho-Statistic	-6.149380	0.0000		
		Group PP-Statistic	-4.814345	0.0000		
		Group ADF-Statistic	1.489243	0.9318		

Como é possível identificar na tabela anterior, os testes de cointegração para as variáveis em níveis não são concordantes: se por um lado os testes de Johansen-Fisher e Kao, bem como os testes de 5 das 11 estatísticas de Pedroni (estatísticas v e ADF), não rejeitam a ausência de causalidade, por outro lado as estatísticas rho e PP rejeitam a ausência de causalidade ao nível habitual de 0.05.

<sup>40</sup> Panel Statistics (within dimension); Group Statistics (between dimension)

## Apêndice III – Sumário dos VAR Independentes

Tabela XII – Síntese dos VAR Independentes

VAR	Ordem VAR	Estabilidade	Serial Correlation LM Tests -Lag3 (Prob)	Heteroskedasticity Tests: No Cross Terms (only levels and squares) (Prob)	Normality Tests Orthogonalization: Cholesky (Lutkepohl) Joint (Prob)	Variáveis Exógenas
A Bola	2	No root lies outside the unit circle. VAR satisfies the stability condition.	0.5203	0.1539	0.0015	X_CONSUM_CONF_IND; X_GAP_BENFICA_PORTO; X_GOOGLE_BOLA; X_GOOGLE_RECORD; X_LPV; X_SMARTPHONES.
Record	3	“	0.8454	0.5101	0.1690	X_CONSUM_CONF_IND; X_GAP_SPORTING_BENFICA; X_SMARTPHONES.
O Jogo	2	“	0.5423	0.1095	0.1960	X_CONSUM_CONF_IND; X_GAP_BENFICA_PORTO; X_TOTAL_INTERNET.
Segmento Desportivo	2	“	0.2501	0.3909	0.0000	X_CONSUM_CONF_IND; X_GAP_BENFICA_PORTO; X_GOOGLE_BOLA; X_GOOGLE_JOGO; X_GOOGLE_RECORD; X_GOOGLE_SPORT; X_SMARTPHONES.
Diário Económico	2	“	0.9547	0.9917	0.0000	X_BANIF; X_BES; X_CONSTCOURTDECISION; X_LPV; X_SMARTPHONES; X_TABLET; X_TERROR; X_TOTAL_INTERNET.
Jornal de Negócios	6	“	0.3495	0.7793	0.6319	X_BANIF; X_BES; X_CONSUM_CONF_IND; X_DEBTCRISIS; X_ELECTIONS; X_FOREIGNAID; X_SOCRATES; X_TROIKACRISIS.
Economy and Business Segment	2	“	0.4020	0.9991	0.000	X_BES; X_CONSUM_CONF_IND; X_ELECTIONS; X_FOREIGNAID; X_IPI_CONSUMPTION; X_TROIKACRISIS.
Correio da Manhã	2	“	0.6904	0.5601	0.1749	X_CONSUM_CONF_IND; X_CPCOURTSENT; X_ELECTIONS; X_EUSEBIO; X_GOOGLE_CM; X_GOOGLE_I; X_GOOGLE_JNOT; X_IPI_CONSUMPTION; X_MINISTERGRADUATION; X_TOTAL_INTERNET.
Diário de Notícias	12	“	0.9236	0.0625	0.0043	X_DEBTCRISIS; X_ELECTIONS; X_GOOGLE_DN; X_GOOGLE_INFO; X_IPI_CONSUMPTION; X_SMARTPHONES.
I	2	“	0.5303	0.3058	0.0474	X_DEBTCRISIS; X_ELECTIONS; X_FOREIGNAID; X_TOTAL_INTERNET; X_TROIKACRISIS.
Jornal de Notícias	2	“	0.2644	0.1600	0.0000	X_ELECTIONS; X_GOOGLE_CM; X_GOOGLE_I; X_GOOGLE_JNOT; X_GOOGLE_PUBLICO; X_TOTAL_INTERNET.
Público	2	“	0.7850	0.1269	0.4998	X_BES; X_CONSUM_CONF_IND; X_LPV.
Segmento Informação Geral	12	“	0.9586	0.2889	0.000	X_BANIF; X_BES; X_DEBTCRISIS; X_ELECTIONS; X_GOOGLE_CM; X_GOOGLE_JNOT; X_INF; X_IPI_CONSUMPTION; X_MINISTERGRADUATION; X_SOCRATES; X_TABLET; X_TERROR; X_TOTAL_INTERNET.
Total Market	12	“	0.5165	0.4733	0.5962	X_BANIF; X_BES; X_DEBTCRISIS; X_ELECTIONS; X_FOREIGNAID; X_IPI_CONSUMPTION; X_LPV; X_SECRETSERVICES; X_SMARTPHONES; X_TERROR; X_TOTAL_INTERNET; X_TROIKACRISIS.

## Apêndice IV – Glossário de Variáveis

**Tabela XIII–Lista de variáveis significativas nos diferentes modelos estimados**

Variável	Descrição
X_BANIF	Variável dummy que identifica manchetes envolvendo a crise do BANIF (Dez 2015)
X_BES	Variável dummy para o caso BES (Agosto 2014)
X_BUDGET	Variável dummy que identifica manchetes envolvendo orçamento rectificativo (Março 2012)
X_CONSTCOURTDECISION	Variável dummy com decisões do tribunal constitucional (Julho 2012)
X_CONSUM_CONF_IND	Índice de Confiança do Consumidor
X_CPCOURTSENT	Variável dummy com a decisão judicial do caso Casa Pia (Fevereiro 2013)
X_DEBTCRISIS	Variável dummy para a crise da dívida e a crise portuguesa (Janeiro 2011, Setembro 2011, Outubro 2011)
X_ECON_SENT_IND	Indicador de Sentimento Económico
X_ELECTIONS	Variável dummy que identifica eleições legislativas (Setembro 2009, Junho 2011)
X_EUSEBIO	Variável dummy que identifica a morte de Eusebio da Silva Ferreira (Janeiro 2012)
X_FIXED_INTERNET	Número total de clientes com acesso a internet fixa
X_FOREIGNAID	Variável dummy que identifica o pedido de assistência financeira ao Estado Português (Abril 2011)
X_GAP_BENFICA_PORTO	Gap pontual entre SLB e FCP
X_GAP_SPORTING_BENFICA	Gap pontual entre SCP e SLB
X_GOOGLE_BOLA	Pesquisas Google "A Bola"
X_GOOGLE_CM	Pesquisas Google "Correio da Manhã"
X_GOOGLE_DN	Pesquisas Google "Diário de Notícias"
X_GOOGLE_ECON	Pesquisas Google jornais economia e negócios
X_GOOGLE_I	Pesquisas Google "I"
X_GOOGLE_INFO	Pesquisas Google jornais informação geral
X_GOOGLE_JNOT	Pesquisas Google "Jornal de Notícias"
X_GOOGLE_JOGO	Pesquisas Google "O Jogo"
X_GOOGLE_PUBLICO	Pesquisas Google "Público"
X_GOOGLE_RECORD	Pesquisas Google "Record"
X_GOOGLE_SPORT	Pesquisas Google jornais desportivos
X_INF	Série mensal da taxa de variação homóloga do IPC geral
X_IPI_CONSUMPTION	Índice de Produção Industrial de bens de consumo
X_LPV	Logaritmo natural do número de pontos de venda
X_MANIFFACEBOOK	Variável dummy que identifica manchetes envolvendo manifestação organizada via facebook (Set 2012)
X_MINISTERGRADUATION	Variável dummy que identifica manchetes sobre o caso da licenciatura de ministro português (Outubro 2012)
X_SECRETSERVICES	Variável dummy que identifica manchetes sobre a controvérsia do caso "secretas" (Maio 2012)
X_SMARTPHONES	Variável dummy para a penetração de smartphones após 2009
X_SOCRATES	Variável dummy que identifica manchetes envolvendo a operação marquês (Nov 2014, Out 2015)
X_TABLET	Variável dummy que identifica uma penetração de Tablets > 10% (Jan 2014)
X_TERROR	Variável dummy que identifica manchetes envolvendo atentados terroristas (Jan 2015, Nov 2015)
X_TOTAL_INTERNET	Número total de clientes com acesso a internet fixa ou móvel
X_TROIKACRISIS	Variável dummy que identifica manchetes envolvendo a crise portuguesa e a Troika (Setembro-Novembro 2011)
X_UNEMP	Série mensal da variação homóloga da taxa de desemprego

