

MESTRADO
ECONOMETRIA APLICADA E PREVISÃO

TRABALHO FINAL DE MESTRADO
TRABALHO DE PROJETO ELABORADO PARA OBTENÇÃO
DO GRAU DE MESTRE

ANÁLISE COMPORTAMENTAL DA POPULAÇÃO
PORTUGUESA DURANTE UM CENÁRIO PANDÉMICO:
ABORDAGEM DE SÉRIES TEMPORAIS

TIAGO JORGE LOUREIRO PIRES

ORIENTAÇÃO:

PROF^o DOUTOR NUNO SOBREIRA

PROF^o NUNO SANTOS

OUTUBRO-2021

MESTRADO
ECONOMETRIA APLICADA E PREVISÃO

TRABALHO FINAL DE MESTRADO
TRABALHO DE PROJETO ELABORADO PARA OBTENÇÃO
DO GRAU DE MESTRE

ANÁLISE COMPORTAMENTAL DA POPULAÇÃO
PORTUGUESA DURANTE UM CENÁRIO PANDÉMICO:
ABORDAGEM DE SÉRIES TEMPORAIS

TIAGO JORGE LOUREIRO PIRES

ORIENTAÇÃO:

PROFº DOUTOR NUNO SOBREIRA

PROFº NUNO SANTOS

OUTUBRO-2021

*“Para ser grande, sê inteiro: nada
Teu exagera ou exclui.
Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.
Assim em cada lago a lua toda
Brilha, porque alta vive.”*

RICARDO REIS

RESUMO

Durante o período da atual pandemia Covid-19, a eficácia da comunicação dirigida à população tem sido um dos maiores desafios das autoridades políticas e sanitárias. Com base no modelo COM-B (*Capability, Opportunity, Motivation e Behavior*) proposto por Michie *et al.* (2011), sobejamente utilizado na área das ciências comportamentais, pretende-se perceber que papel teve a comunicação das autoridades na adoção de comportamentos preventivos por parte da população portuguesa. De forma complementar, será interessante perceber se a população respondeu de forma semelhante face aos mesmos estímulos ou se foi ajustando os seus comportamentos em função das especificidades de cada momento. Realça-se que a monitorização de quadros comportamentais ao longo do tempo não tem sido muito explorada nas publicações científicas existentes na área da Economia, razão pela qual este estudo gera uma contribuição inovadora nesta matéria.

Na base de dados recolhida, destaca-se, entre outras, o nível de confinamento diário da população, fornecido pelo Painel Mobilidade PSE¹, para o período compreendido entre 1 de março de 2020 a 31 de março de 2021. Empregou-se a metodologia dos modelos de Séries Temporais ADL com o objetivo de relacionar o confinamento diário com os determinantes preconizadores do modelo COM-B. Encontrou-se forte evidência a favor da eficácia de uma comunicação assente em medidas fortemente restritivas sobre a contenção dos contactos sociais por parte dos portugueses, como a suspensão da atividade letiva presencial. Por outro lado, um estilo comunicacional baseado em medidas avulsas e pontuais, como a proibição de circular entre concelhos, terá produzido efeitos contraproducentes no confinamento da população. O cenário de fadiga pandémica da população, devidamente testado, também obteve suporte estatístico. As conclusões dos testes de Bai e Perron (2003) apontam para a existência de um conjunto de momentos em que a resposta da população face aos determinantes do modelo COM-B se alterou estruturalmente.

PALAVRAS-CHAVE: Covid-19; Comunicação; Modelo COM-B; Séries Temporais; Teste de Quebras de Estrutura.

¹PSE - Produtos e Serviços de Estatística é uma empresa especialista em Data Science e Pesquisa Avançada, responsável pelo Painel de Mobilidade que serviu de base para o presente estudo. Para mais informações consulte www.pse.pt/mobilidade.

ABSTRACT

During the period of the current Covid-19 pandemic, the effectiveness of communication aimed at the population has been one of the greatest challenges for political and health authorities. Based on the COM-B (Capability, Opportunity, Motivation and Behavior) model proposed by Michie *et al.* (2011), widely used in behavioral sciences, it's important to understand what role the communication of authorities played in the adoption of preventive behaviors by the Portuguese population. Furthermore, it will be interesting to check if the population responded in a similar way to the same stimuli or if they were adjusting their behavior according to the specifics of each moment. It should be noted that the monitoring of behavioral conditions over time has not been explored in Economic Literature, which is why this study provides an innovative contribution in this area.

From a collected database, it should be highlighted, among others, the level of daily confinement of the population, provided by the PSE² Mobility Panel, for the period from March 1, 2020 to March 31, 2021. The methodology of ADL Time Series models was used to relate the daily confinement with the determinants of the COM-B model. Strong evidence was found in favor of the effectiveness of a communication based on strongly restrictive measures in the containment of social contacts by the Portuguese population, such as the interruption of school activities in a presential way. In other hand, a communication style based in single and isolated policies, such as the prohibition to circulate between municipalities, will have produced counterproductive effects in the confinement of the population. The population pandemic fatigue scenario, duly tested, also received statistical support. The conclusions of the Bai and Perron (2003) tests lead to the existence of a set of moments in which the population's response to the determinants of the COM-B model structurally changed.

KEYWORDS: Covid-19; Communication; COM-B Model; Time Series; Multiple Structural Changes Test.

² PSE - Produtos e Serviços de Estatística is a company expert in Data Science and Advanced Research, responsible for the Mobility Panel that served as the basis for this study. For more information see www.pse.pt/mobilidade.

AGRADECIMENTOS

Orgulho e gratidão. São estas as únicas palavras que me ocorrem quando tento descrever o que sinto perante o culminar de uma das mais bonitas e gratificantes etapas da minha vida. Cumprir um trajeto difícil, mas meritório. Um trajeto do qual me orgulho indescritivelmente. Orgulho pelo elevadíssimo nível de superação que atingi, mas sobretudo, orgulho por ter sido sempre fiel a mim mesmo e aos meus valores e princípios. No entanto, expressar um enorme sentimento de gratidão pois trilhar este caminho apenas foi possível com o indispensável auxílio, força e encorajamento de várias pessoas a quem, mais uma vez, agradeço e dedico este Projeto.

Em primeiro lugar, aos professores que me orientaram ao longo desta jornada.

Ao Professor Doutor Nuno Sobreira pelas lúcidas sugestões e observações que contribuíram para o enriquecimento deste trabalho.

Ao Professor Nuno Santos, por toda a disponibilidade, apoio e colaboração incansáveis desde o primeiro ao último dia. Um obrigado sobretudo por toda a aprendizagem transmitida que extravasa o âmbito académico. Uma referência no que concerne ao profissionalismo, energia e vivacidade com que encara todos os desafios.

Aos meus avós, ao meu pai, à minha mãe, ao meu irmão e à Mikas por representarem tudo o que melhor tenho na vida.

À Cristiana, pelo amor e carinho demonstrados diariamente, pela força que me transmite quando dela mais preciso, por ser o meu maior apoio em tudo.

Ao ISEG por ser a melhor faculdade do mundo. Ao corpo docente do mestrado por pautar sempre pelo rigor e excelência no processo pedagógico. Aos meus colegas por proporcionarem um ambiente de entreajuda.

À Produtos e Serviços de Estatística (PSE) pela pronta disponibilização dos materiais necessários à elaboração do trabalho.

Por fim, um agradecimento a todos os meus amigos de sempre por personificarem o verdadeiro significado de amizade.

Índice

1. Introdução.....	1
2. Revisão da Literatura.....	3
2.1. Modelo <i>COM-B</i>	4
3. Metodologia.....	6
3.1. Análise de Componentes Principais	6
3.2. Abordagem de Séries Temporais.....	9
3.2.1. Análise de Especificação	11
3.2.1.1. Testes de Autocorrelação dos erros	11
3.2.1.2. Teste RESET	13
3.2.2. Teste de Quebra de Estrutura de Bai-Perron	13
4. Análise Exploratória dos Dados	15
4.1. Caracterização dos Dados.....	15
4.2. Obtenção das Componentes da Comunicação.....	18
4.3. Análise Descritiva das variáveis.....	21
5. Modelação e Análise de Resultados	22
5.1. Especificação e estimação do modelo	22
5.2. Interpretação dos resultados	25
5.3. Teste às Alterações Comportamentais.....	29
6. Conclusões, Limitações e Investigação Futura	33
Referências Bibliográficas.....	36
Anexo A – Figuras.....	40
Anexo B – Tabelas	47

Índice de Figuras

Figura I. Diagrama do modelo COM-B.....	4
Figura II. Ilustração da utilização do Modelo COM-B pela Escola Nacional de Saúde Pública nas reuniões no Infarmed.....	6
Figura III. Estrutura conceptual do modelo COM-B aplicada ao caso em estudo	18
Figura IV. Representação Gráfica dos pontos de quebra estimados da série do Confinamento Diário (sazonalidade ajustada).....	33
Figura V. Série do Confinamento Diário: % da População confinada.....	40
Figura VI. Série do Número de Novos Casos Reportados de Covid-19 e do DEI - Daily Economic Indicator.....	40
Figura VII. Representação Gráfica das Componentes obtidas e retidas pela ACP	41
Figura VIII. Estimação do modelo ADL(7,7)	41
Figura IX. Teste LM à autocorrelação dos erros de 1ª ordem ao modelo ADL(7,7)	42
Figura X. Teste LM à autocorrelação até à ordem 7 dos erros do modelo ADL(7,7)....	43
Figura XI. Teste RESET ao modelo ADL(7,7).....	44
Figura XII. Teste de hipóteses: $H_0: \beta C1t - 1 * 2020 = \beta C1t - 1 * 2021$ vs. $H_1: \beta C1t - 1 * 2020 > \beta C1t - 1 * 2021$	44
Figura XIII. Teste de hipóteses: $H_0: \beta C1t - 1 * 2020 = 2\beta C1t - 1 * 2021$ vs. $H_1: \beta C1t - 1 * 2020 \neq 2\beta C1t - 1 * 2021$	45
Figura XIV. Teste <i>double maximum</i> de Bai e Perron.....	45
Figura XV. Teste Sequencial de Bai e Perron.....	46

Índice de Tabelas

Tabela I. Caracterização das variáveis que personificam as dimensões Comportamento, Motivação e Oportunidade utilizadas no modelo	47
Tabela II. Caracterização das variáveis que personificam a dimensão Capacidade utilizadas no modelo	47
Tabela III. Descrição das variáveis da base de dados relativa à Comunicação.....	48
Tabela IV. Estatísticas de diagnóstico da ACP	50
Tabela V. Variância Total Explicada pelas Componentes Principais	50
Tabela VI. <i>Loadings</i> ou pesos de cada variável em cada componente.....	51
Tabela VII. Matriz de correlações de Pearson das variáveis do modelo.....	52

ABREVIACES

ACP – Anlise de Componentes Principais

ADL – *Autoregressive Distributed Lag*

ARIMA – *Autoregressive Integrated Moving Average*

BG – Breusch-Godfrey

COM-B – *Capability, Opportunity Motivation and Behavior*

DEI – *Daily Economic Indicator*

DGS – Direo Geral da Sade

FDL – *Finite Distributed Lag*

HAC – *Heteroskedastic and autocorrelation consistent*

HIV – *Human Immunodeficiency Virus*

LM – *Lagrange Multiplier*

OLS – *Ordinary Least Squares*

OMS – Organizao Mundial de Sade

PSE – Produtos e Servios de Estatstica

SEM - *Structural Equations Model*

SQR – Soma do Quadrado dos Resduos

1. INTRODUÇÃO

Em Portugal, a cobertura mediática da doença provocada pela infeção do coronavírus SARS-CoV-2 foi diferente de todas as outras doenças infecciosas, nomeadamente a do HIV/SIDA, segundo Cabrera *et al.* (2020). Enquanto a identificação da primeira como valor-notícia é imediata e antecede a sua chegada a Portugal, na segunda existe um longo tempo para que adquira noticiabilidade.

A eficácia da comunicação durante um período pandémico pode ser determinante para o controlo da disseminação da própria doença (Gupta & Reddy B, 2020). Neste sentido, a gestão da pandemia covid-19, não somente em termos sanitários, mas também em termos comunicacionais e políticos, tem sido dos maiores desafios dos governantes e das autoridades de saúde de todos os países. Desde o início do período pandémico, que a população portuguesa tem sido recetora de inúmeras mensagens - por exemplo, as conferências de imprensa diárias da Direção Geral da Saúde (DGS), as diversas declarações do primeiro-ministro – de múltiplas naturezas – por exemplo, proíbe-se a circulação entre concelhos e simultaneamente permite-se as celebrações do 1º de Maio, em 2020.

Dada a ausência de variáveis que definam expressamente os estímulos comunicacionais emitidos pelas autoridades políticas e sanitárias, foram recolhidas dezenas de variáveis binárias identificadoras das medidas adotadas pelas mesmas autoridades e das intervenções públicas mais mediáticas das mesmas. A partir deste *dataset* recolhido pelo autor, após uma extensa e criteriosa compilação de várias fontes, recorreu-se à Análise de Componentes Principais a fim de encontrar métricas que sintetizassem os principais estilos de comunicação e que perfilassem uma dimensão quantitativa do fenómeno.

Neste contexto, adotando uma abordagem macroeconómica, a partir do Painel de Mobilidade PSE pretende-se estudar as alterações comportamentais da população portuguesa ao longo do período de pandemia Covid-19. Com base num quadro conceptual amplamente utilizado na área das ciências comportamentais para avaliar este tipo de questões, o modelo COM-B (*Capability, Opportunity, Motivation and Behavior*), pretende-se estimar os impactos dos diferentes estímulos comunicacionais sobre os comportamentos preventivos dos portugueses, como a literatura prevê noutros quadros.

De forma complementar, propõe-se também testar eventuais diferenças de resposta da população face aos mesmos estímulos em períodos distintos, isto é, será que a propensão a confinar da população terá sido igual nos dois períodos de *lockdown*? Ou será que o cenário de “fadiga pandémica” terá contribuído para uma postura de descontração na adoção de comportamentos? Os testes de quebra de estrutura propostos por Bai e Perron serão uma preciosa ferramenta para a deteção de momentos-chave em que a população foi levada a ajustar os seus comportamentos em virtude das diversas conjunturas sanitárias, económicas e normativas.

A análise comportamental levada a cabo a partir do modelo COM-B permitirá modelizar três eixos que, segundo o mesmo, explicarão conjuntamente os níveis de confinamento dos portugueses ao longo da pandemia, são estes a Oportunidade, Motivação e Capacidade da população portuguesa. A vertente Oportunidade procurará enquadrar as tomadas de decisão dos portugueses com o ambiente e condições socioeconómicas com que a mesma se deparou; paralelamente, a vertente Motivação terá como papel medir a perceção ao risco de contágio da doença Covid-19 na população como fator explicativo da adoção de comportamentos preventivos dos portugueses; por fim, a vertente Capacidade servirá para avaliar o papel da Comunicação e da ação política, por intermédio das medidas restritivas impostas pelo Governo.

Como se irá aprofundar mais adiante, o acompanhamento de quadros comportamentais ao longo do tempo é uma lacuna nas publicações científicas existentes na área das Ciências Económicas, o que torna extremamente desafiante o desenvolvimento deste Projeto. Pelo exposto, e dado que se trata de uma temática extremamente atual e pertinente, as conclusões desta aplicação podem ser revestidas de grande relevância.

Este trabalho encontra-se organizado em quatro secções. A secção 2 em que se apresenta uma revisão da literatura onde serão enunciados estudos que vaticinam relações de causalidade da comunicação para determinados comportamentos dos agentes e, sobretudo, será explicitado o *framework* conceptual que os estudos de análise comportamental avançam para responder a questões como as que esta aplicação coloca. A secção 3 onde serão descritos os procedimentos metodológicos adotados para tratar o presente objeto de estudo. A secção 4 onde se procede a uma análise exploratória dos

dados e respetiva caracterização. Finalmente, a secção 5 em que se aplicará os métodos econométricos para responder às questões propostas.

2. REVISÃO DA LITERATURA

Se é verdade que os *media* contribuem positivamente para uma sociedade mais informada e consciencializada, sendo um dos principais veículos de difusão de informação, também é verdade que a comunicação pode alterar os estilos de vida e os comportamentos pessoais. Diferentes estudos têm evidenciado relações de causalidade da comunicação para determinados comportamentos dos agentes: em Engelberg & Parsons (2011) é encontrada forte evidência estatística a favor do impacto da cobertura dos *media* na probabilidade e magnitude de *trading* nas 500 empresas do Índice norte-americano *Dow Jones*; em Chemli *et al.* (2020) é analisada a influência da cobertura dos órgãos de comunicação social em matéria de situação sanitária Covid-19 na propensão à atividade turística internacional, de onde se conclui que esta última exhibe grande sensibilidade face à natureza da informação veiculada.

No domínio da saúde, as campanhas de sensibilização e de informação nos órgãos de comunicação social têm sido frequentemente utilizadas como veículo condutor de uma mensagem de prevenção, centradas na redução de comportamentos de risco face a uma diversidade de patologias (Wakefield *et al.*, 2010). Os resultados dessas intervenções comunicacionais perante a exposição da população a surtos de doenças infecciosas têm sido devidamente estudados. No caso do vírus da SIDA, Bertrand *et al.* (2006) e Bekalu & Eggermont (2014) evidenciam a utilidade de uma comunicação massificada baseada na informação e no conhecimento sobre a redução do comportamento sexual de alto risco. Relativamente à exposição à Gripe A, provocada pelo vírus H1N1, segundo Zhang *et al.* (2015) e Oh *et al.* (2015) a probabilidade de adoção de medidas preventivas, como a lavagem frequente das mãos e o distanciamento social, é positivamente afetada por uma pedagógica cobertura informativa.

A doença infecciosa mais recente e atual com que todo o mundo se depara à data da elaboração deste trabalho, é a denominada Covid-19, provocada pela infeção do Coronavírus SARS-CoV-2 (Síndrome Respiratória Aguda Grave – Coronavírus-2). A produção científica já existente destaca o importante papel das intervenções precoces da Organização Mundial de Saúde no contributo para a diminuição de comportamentos de

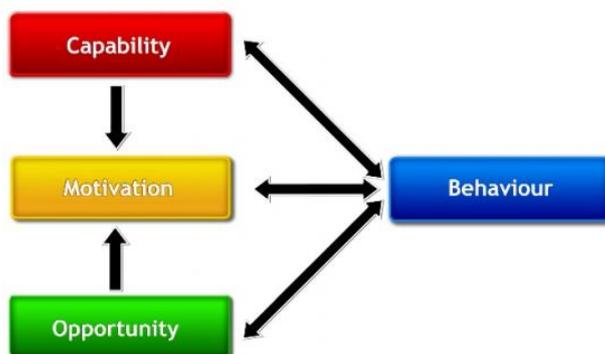
risco (Anwar *et al.*, 2020). Os autores alertam ainda para a importância de uma correta percepção das comunicações de saúde pública como ferramenta essencial ao combate da disseminação do covid-19 e de outros futuros surtos.

2.1. Modelo *COM-B*

O principal objetivo desta aplicação empírica é consolidar evidência sobre os impactos da comunicação nos comportamentos dos portugueses desde o início da pandemia covid-19 em Portugal, ou seja, analisar alterações do foro comportamental. Para isso, é necessário um enquadramento teórico de um quadro conceptual frequentemente utilizado na literatura para debater questões de atividade comportamental.

O modelo conceptual *COM-B* (*Capability, Opportunity, Motivation e Behavior*) desenvolvido em Michie *et al.* (2011) procura explicar a adoção de determinado comportamento com base num sistema de interações envolvendo três fatores: Capacidade, Oportunidade e Motivação. A Capacidade refere-se à capacidade física e psicológica que o indivíduo exibe para desenvolver a ação; a Oportunidade descreve o ambiente externo, as condições que estimulam determinada ação; e a Motivação que se prende com fatores internos ao individuo, fatores de resposta emocional. O resultado final desta interação, o Comportamento, define a tomada de decisão por parte do agente. Assim, segundo este modelo, as três primeiras componentes identificam e explicam conjuntamente o Comportamento individual. A Figura I ilustra este *framework* teórico.

Figura I. Diagrama do modelo COM-B



Fonte: Michie *et al.* (2011)

Este quadro conceptual tem sido amplamente utilizado em aplicações de ciências comportamentais, nomeadamente no campo da Psicologia, onde se pretende explorar os fatores associados aos comportamentos de *self-care* dos indivíduos, tipicamente expostos

a alguma doença ou patologia clínica. Para estudar os fatores associados aos comportamentos de cuidado pessoal em pacientes com insuficiência crónica, Zou *et al.* (2017) recolheu informação de 321 indivíduos com diagnóstico da referida patologia sobre a literacia na saúde, conhecimento, *status* socioeconómico, *self-care* entre outras variáveis de natureza sociodemográfica. Dada a impossibilidade de observar diretamente as dimensões Capacidade, Motivação, Oportunidade e Comportamento, os autores partiram para uma abordagem de Modelização de Equações Estruturais (ou *SEM – Structural Equations Model*), de modo que pudessem ser endogeneizadas como variáveis latentes.

Abordagem semelhante seguiram Willmott *et al.* (2021) cujo trabalho pretendeu identificar os fatores que mais e menos propiciam hábitos de vida saudáveis entre os jovens adultos. Através de uma amostra de 1040 indivíduos em que se recolheram variáveis sociodemográficas de cada participante, testou-se a aplicação do modelo COM-B como instrumento de planeamento e intervenção para a prevenção do ganho de peso em jovens adultos, recorrendo novamente à modelação de um sistema de equações estruturais.

Existem ainda outros estudos que optam por fazer referência a este modelo como um *framework* de partida para as suas análises, à semelhança dos dois trabalhos acima referidos, mas cujo desenvolvimento empírico se esgota numa dimensão qualitativa, como é o caso de McDonagh *et al.* (2018) e Jatau *et al.* (2019). Em ambas as obras, a aplicabilidade do modelo COM-B deve-se sobretudo ao fornecimento de uma adequada estrutura de compreensão do fenómeno em causa, no sentido em que permite dissecar teoricamente as dimensões (Comunicação, Oportunidade e Motivação) das respetivas temáticas. Enquanto no primeiro, o objetivo centrou-se nos condicionalismos da testagem voluntária da doença sexualmente transmissível clamídia, no segundo a atenção incidiu-se nas potenciais estratégias para melhorar o diagnóstico precoce da fibrilação auricular, a arritmia crónica mais frequente.

Importa frisar que este é o quadro comportamental utilizado pela Escola Nacional de Saúde Pública no âmbito das investigações relativas às perceções sociais sobre a Covid-19 em Portugal (Nunes, 2021). Nomeadamente, nas reuniões de acompanhamento da situação epidemiológica em Portugal, que reúne especialistas em Epidemiologia e em

Saúde Pública no Infarmed, é precisamente com base neste modelo que a discussão e o debate são promovidos. A Figura II ilustra a utilização deste *framework* na reunião do Infarmed de 23 de março de 2021 por parte dos técnicos da Escola Nacional de Saúde Pública.

Figura II. Ilustração da utilização do Modelo COM-B pela Escola Nacional de Saúde Pública nas reuniões no Infarmed



Fonte: [CARLA-NUNES.pdf \(sns.gov.pt\)](#)

Como se teve a oportunidade de constatar, todos estes estudos existentes na literatura apoiam-se numa perspetiva seccional, baseando-se numa amostra de indivíduos, ignorando a dimensão temporal associada ao fenómeno comportamental. Em muitos casos seria certamente interessante perceber qual a evolução das sensibilidades dos diferentes determinantes da ação individual ao longo do tempo. A monitorização de quadros comportamentais numa base temporal é assim uma lacuna patente na literatura da área, aspeto que torna este trabalho completamente diferenciador.

3. METODOLOGIA

Esta secção destina-se à apresentação de todo o processo metodológico que será utilizado nos pontos seguintes. Na primeira subsecção será explicitado o processo analítico que permitirá obter as componentes associadas aos estímulos comunicacionais, que consistirá um passo prévio da modelação econométrica, ao passo que na segunda subsecção será abordada toda a teoria econométrica que servirá de suporte da análise levada a cabo para responder às questões propostas nesta aplicação.

3.1. Análise de Componentes Principais

Face à ausência de variáveis que possam definir de forma concreta os estímulos comunicacionais e devido à sua complexidade, torna-se complicado definir um par de

variáveis que sejam representativas de todo este conceito. Estes podem advir de múltiplas fontes: seja uma imposição legal ditada pelo Governo, seja uma conferência de Imprensa da Direção Geral da Saúde, entre outras; e podem ter cargas percecionais diferentes: seja uma norma que limite a circulação das pessoas aos fins de semana, seja uma progressiva reabertura do comércio, entre outras.

Estas dimensões, não sendo diretamente observadas, necessitam de ser tratadas por abordagens analíticas que tenham em linha de conta a natureza latente destas variáveis. A metodologia que emprega abordagens onde se modelizam variáveis latentes como passo crucial das investigações empíricas encontra-se perfeitamente identificada na literatura. Neste sentido, sublinham-se os trabalhos de Zou *et al.* (2017) e Willmott *et al.* (2021) utilizando o método de Equações Estruturais (*SEM*) como parte fundamental das respetivas análises; e de Kevin *et al.* (1986) e Russell (2002) utilizando a Análise Fatorial Exploratória (*AFE*) como instrumento de modelização das respetivas variáveis latentes. De forma bastante genérica, o *SEM* procura estabelecer relações entre variáveis observadas e variáveis latentes (não observadas), no quadro dos modelos lineares, sendo tipicamente usado nas Ciências Sociais e Humanas para validar um determinado quadro teórico estabelecido *a priori*. Paralelamente, a *AFE* representa também outra ferramenta analítica que visa identificar os fatores latentes que explicam as correlações observadas nas variáveis originais.

É neste contexto que se optou pela Análise de Componentes Principais (*ACP*), uma técnica de análise exploratória multivariada cujo propósito passa por transformar um conjunto de variáveis correlacionadas num conjunto menor de variáveis independentes, combinações lineares das variáveis originais, designadas por componentes principais (variáveis latentes ou não observadas). Esta abordagem permitirá obter “índices” ou indicadores que resumem a informação disponível nas variáveis originais. Por exemplo, o conhecido índice PSI-20 é um caso de uma combinação linear das cotações da bolsa de valores das empresas com maior volume de negócios. A utilização desta ferramenta analítica exhibe evidentes vantagens do ponto de vista prático, nomeadamente em aplicações em que o analista dispõe de inúmeras variáveis e pretende reduzir a sua dimensão com a menor perda de informação possível. Voltando ao caso do PSI-20, é claramente mais fácil para o analista avaliar a evolução do mercado através de um único indicador do que com 20 variáveis que descrevam a cotação de 20 empresas.

Esta abordagem foi proposta por Pearson (1901) onde de um ponto de vista teórico as componentes principais populacionais são combinações lineares de p variáveis populacionais originais correlacionadas, sendo o intuito da análise formar p combinações lineares independentes:

$$\begin{aligned}\xi_1 &= \gamma_{11}X_1 + \gamma_{12}X_2 + \cdots + \gamma_{1p}X_p \\ \xi_2 &= \gamma_{21}X_1 + \gamma_{22}X_2 + \cdots + \gamma_{2p}X_p \\ &\vdots \\ \xi_p &= \gamma_{p1}X_1 + \gamma_{p2}X_2 + \cdots + \gamma_{pp}X_p\end{aligned}\tag{1}$$

Onde X_1, X_2, \dots, X_p são as p variáveis populacionais originais; $\xi_1, \xi_2, \dots, \xi_p$ são as p componentes principais e γ_{ij} é o peso da variável j na componente principal i . A estimação dos pesos γ_{ij} é realizada de modo que sejam satisfeitas as seguintes condições:

1. A primeira componente principal (ξ_1) explica a maior proporção total das variâncias originais;
2. A componente seguinte explica a maior proporção da variância não explicada pela primeira componente, ou seja, matematicamente formaliza-se: $\gamma_{i1}\gamma_{j1} + \gamma_{i2}\gamma_{j2} + \cdots + \gamma_{ip}\gamma_{jp} = 0$ ($i \neq j; i = 1, \dots, p; j = 1, \dots, p$);
3. A escala das novas componentes deve permanecer fixa de modo a manter constante a variância total, ou seja, matematicamente formaliza-se: $\gamma_{i1}^2 + \gamma_{i2}^2 + \cdots + \gamma_{ip}^2 = 1$ ($i = 1, \dots, p$).

Seja \mathbf{X} um vetor aleatório de dimensão p (em que p é o número de variáveis originais), a matriz de variâncias-covariâncias de \mathbf{X} é $\Sigma_x = E(\mathbf{X}\mathbf{X}')$. Seja $\boldsymbol{\gamma}' = [\gamma_1, \gamma_2, \dots, \gamma_p]$ o vetor transposto do vetor dos pesos e $\boldsymbol{\xi}' = [\xi_1, \xi_2, \dots, \xi_p]$ o vetor transposto do vetor das componentes principais. Então, cada componente principal em (1) pode ser apresentada matricialmente da seguinte forma:

$$\boldsymbol{\xi} = \boldsymbol{\gamma}'\mathbf{X}\tag{2}$$

A notação matricial (2) é particularmente importante para se deduzir o problema de otimização em causa. A variância de $\boldsymbol{\xi}$ pode ser apresentada:

$$\text{Var}(\boldsymbol{\xi}) = E(\boldsymbol{\xi}\boldsymbol{\xi}') = E(\boldsymbol{\gamma}'\mathbf{X}(\boldsymbol{\gamma}'\mathbf{X})') = E(\boldsymbol{\gamma}'\mathbf{X}\mathbf{X}'\boldsymbol{\gamma}) = \boldsymbol{\gamma}'E(\mathbf{X}\mathbf{X}')\boldsymbol{\gamma} = \boldsymbol{\gamma}'\Sigma_x\boldsymbol{\gamma}\tag{3}$$

O problema de otimização resume-se a encontrar o vetor dos pesos γ tal que $\gamma' \Sigma_x \gamma$ seja máxima sob a restrição $\gamma \gamma' = 1$. A solução deste problema obtém-se por recurso às derivadas parciais e ao Multiplicador de Lagrange para:

$$(\Sigma - \lambda I)\gamma = \mathbf{0} \text{ tal que } \gamma \gamma' = 1. \quad (4)$$

Onde λ é o Multiplicador de Lagrange, I é a matriz identidade e $\mathbf{0}$ é o vetor nulo. Para que este sistema de equações possua uma solução não trivial, o determinante de $\Sigma - \lambda I$ deve ser nulo, ou seja, $|\Sigma - \lambda I| = 0$. As p raízes desta equação ($\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_p$) designam-se por valores próprios. Após calculada a primeira componente, procede-se para a segunda através de um processo iterativo semelhante, mas com a restrição adicional da ortogonalidade entre as componentes ($Cov(\xi_1, \xi_2) = 0$). O processo de estimação das componentes principais é levado a cabo com o desconhecimento da matriz de variâncias-covariâncias Σ_x , pelo que terá que se substituir o referido parâmetro pelo seu estimador $\hat{\Sigma}_x$.

3.2. Abordagem de Séries Temporais

Como foi referido no término da Revisão da Literatura, para além de abordar uma temática extremamente atual e pertinente, em termos empíricos este trabalho diferencia-se dos demais no campo das alterações comportamentais, na medida em que introduz a dimensão temporal ao objeto em estudo. Assim, neste capítulo, a abordagem metodológica adotada recairá no tratamento econométrico de Séries Temporais.

Na Econometria o tratamento temporal dos fenómenos económicos difere substancialmente do tratamento levado a cabo a partir de unidades seccionais, começando logo por uma hipótese básica que sustenta todo o processo de estimação e inferência estatística dos últimos. Enquanto em análises microeconómicas a hipótese da aleatoriedade da amostra é absolutamente crucial e elementar, tipicamente definindo as observações como *iid* (independentes e identicamente distribuídas), quando se estudam variáveis indexadas pelo tempo essa é uma premissa impossível de ser verificada.

Uma dada série temporal de interesse Y_t é definida como um processo estocástico, sendo que o que é observado é somente uma concretização/realização desse mesmo processo estocástico, isto é $\{Y_t\}$ (William W. S., 2005). A aleatoriedade subjacente aos dados temporais é refletida na incerteza dos seus valores futuros. Por exemplo, alguém

sabe como irá fechar a bolsa hoje? As observações provenientes da concretização de um processo estocástico não só não são independentes, como na maioria das séries económicas e financeiras, exibem vincados padrões de autocorrelação. Então, conclui-se que a introdução da dimensão temporal na análise induz uma certa dependência temporal entre as observações que urge ser corretamente modelizada.

A literatura das Séries Temporais distingue, grosso modo, duas classes de modelos quanto à finalidade da análise. Se o fim único for a obtenção de um modelo preditivo e autoexplicativo, então o investigador será direcionado para os Modelos Univariados, onde se encontram os conhecidos modelos ARIMA (*Autoregressive Integrated Moving Average*). Por outro lado, se o investigador estiver interessado em analisar as relações de causalidade existentes entre diferentes variáveis, na maioria das vezes suportadas pela teoria económica, então um caminho a seguir poderá passar por modelos como os ADL (*Autoregressive Distributed Lag*) ou FDL (*Finite Distributed Lag*) que fazem depender a variável de interesse de um conjunto de variáveis explicativas (e da própria variável dependente desfasada no caso do modelo ADL).

Uma vez que a finalidade do presente Projeto assenta na estimação dos impactos da comunicação e de outras dimensões apontadas pelo Modelo conceptual COM-B sobre os comportamentos da população portuguesa, adotou-se a segunda classe de modelos, mais concretamente o modelo ADL. De forma genérica, sem perda de generalidade, seja Y_t a variável dependente com p desfasamentos e X_t a variável explicativa com q desfasamentos, o modelo ADL (p,q) pode ser escrito (Stock & Watson, 2007):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots + \beta_p Y_{t-p} + \delta_0 X_t + \delta_1 X_{t-1} + \delta_q X_{t-q} + u_t, \quad (5)$$

Onde $\beta_0, \beta_1, \dots, \beta_p, \delta_0, \dots, \delta_q$ são os parâmetros desconhecidos e u_t o termo de erro, com $E(u_t | Y_{t-1}, \dots, Y_{t-p}, X_t, X_{t-1}, \dots, X_{t-q}) = 0$. Note que o modelo FDL é um caso particular do ADL em que $\beta_1 = \dots = \beta_p = 0$. Os modelos ADL têm a vantagem relativamente aos FDL de preservar informação histórica da própria variável dependente e utilizá-la no modelo de regressão. A expressão em (5) é facilmente generalizada para o caso em que se introduz mais do que uma variável explicativa exógena, sendo q o comprimento do respetivo desfasamento mais longo.

No que diz respeito às hipóteses que sustentam a estimação pelo método dos Mínimos Quadrados (OLS) no contexto das suas propriedades assintóticas, isto é para grandes amostras, destacam-se três pela diferenciação face ao caso seccional e até mesmo relativamente ao caso de séries temporais em pequenas amostras, (MacKinnon & Davidson, 2003). A primeira tem que ver com a necessidade da estacionaridade em distribuição do processo estocástico conjunto $\{Y_t, X_t\}$. Esta é a versão para dados temporais da hipótese *iid* dos micro-dados. A segunda exige que os regressores sejam contemporaneamente exógenos, no sentido em que sejam ortogonais ao erro contemporâneo. Todavia, esta hipótese permite a existência de correlação do erro com regressores futuros. A verificação destas duas hipóteses é condição necessária para a consistência do estimador OLS. Por fim, a terceira hipótese, que juntamente com a satisfação das duas anteriores, possibilita a realização de inferência estatística (testes de hipóteses, intervalos de confiança, etc) é a de ausência de autocorrelação dos erros juntamente com a ausência de correlação dos mesmos com os regressores contemporâneos e passados.

3.2.1. Análise de Especificação

Na literatura de Séries Temporais designa-se análise de especificação à análise da verificação das hipóteses do modelo de regressão linear. Neste domínio, os testes privilegiados foram os testes de autocorrelação dos erros, para tentar garantir que a dinâmica do modelo é suficiente e o teste RESET, relacionado com a forma funcional. Saliente-se que dar-se-á mais ênfase aos primeiros por colocarem em causa uma das hipóteses mais sensíveis e cruciais em aplicações com dados temporais (a hipótese de ausência de autocorrelação dos erros).

3.2.1.1. Testes de Autocorrelação dos erros

A hipótese de ausência de autocorrelação dos erros, apesar de não colocar em causa a consistência do estimador OLS, é uma hipótese fundamental para empregar as estatísticas t e F no processo de inferência (Wooldridge, 2006). A ideia que está por detrás desta premissa é a de que um modelo corretamente especificado não deve apresentar autocorrelação dos erros, no sentido em que tudo o que é sistemático, ou seja, que tem algum padrão, deve ser incluído na componente principal do modelo.

A estatística de Durbin-Watson foi, durante largos anos, a única disponível para testar autocorrelação no modelo linear. Contudo, este teste apresenta duas limitações: em primeiro lugar, o modelo assumido para os erros é um AR(1) estacionário, por isso, apenas pode capturar autocorrelação de 1ª ordem, que diga-se, é a mais importante; e em segundo, exige a hercúlea hipótese de exogeneidade estrita dos regressores, algo que não se verifica, por exemplo, nos modelos ADL uma vez que incorporam a variável dependente desfasada. A menção a este teste terá apenas um carácter informal ao longo da exposição.

O teste estatístico a que se dará maior relevância e que será apresentado com maior formalidade, é o teste Breush-Godfrey ou na versão LM. As vantagens deste relativamente ao anterior é que não tem as limitações deste último, ou seja, permite a validade do procedimento estatístico mesmo na presença de regressores não estritamente exógenos e concede a hipótese de testar a ausência de autocorrelação dos erros em ordens superiores a 1. A formulação da hipótese alternativa do teste possibilita ter neste quadro uma interpretação bastante interessante:

H_0 : Ausência de autocorrelação até à ordem p vs.

H_1 : $u_t \sim ARMA(m, q), \forall m, q: m + q = p$

Se o teste incidir, por exemplo, na autocorrelação de 2ª ordem ($p = 2$), então sob H_1 os erros podem evoluir segundo um processo $AR(2)$, $MA(2)$ ou $ARMA(1,1)$. Uma rejeição de H_0 poderá ser um sintoma de uma especificação dinâmica pobre do modelo, ou seja, com grande plausibilidade o mesmo não será dinamicamente completo.

O procedimento de teste pode ser de forma simplificada dividido em três momentos: o primeiro, que consiste na regressão de Y_t sobre os k regressores x_1, \dots, x_k e de onde são obtidos os resíduos \hat{u}_t ; no segundo momento aplica-se uma regressão auxiliar de \hat{u}_t sobre os k regressores, x_1, \dots, x_k , e os próprios resíduos desfasados até à ordem que se pretender testar a autocorrelação, $\hat{u}_{t-1}, \dots, \hat{u}_{t-p}$.

$$\hat{u}_t = \mathbf{x}'_t \alpha + \phi_1 \hat{u}_{t-1} + \dots + \phi_p \hat{u}_{t-p} + \text{erro}_t, \quad (6)$$

Por fim, trata-se de testar:

$$H_0 : \phi_1 = \dots = \phi_p = 0 \text{ vs. } H_1 : \exists \phi_j \neq 0 (j = 1, \dots, p),$$

A estatística de teste é derivada a partir do Multiplicador de Lagrange:

$$LM = BG(p) = TR_{\hat{u}}^2 \xrightarrow{a} \chi_{(p)}^2, \text{ sob } H_0,$$

Onde T é o número de observações e $R_{\hat{u}}^2$ o R-quadrado da regressão auxiliar.

3.2.1.2. Teste RESET

O teste RESET, proposto por Ramsey (1969), visa testar a correta especificação da forma funcional do modelo linear estimado pelo OLS. A sua utilização também se estende aos modelos com dados seccionais. Dado um modelo original genérico expresso na forma matricial $Y = X\beta + U$, a equação de teste pode ser apresentada da seguinte maneira:

$$Y = X\beta + \gamma_1(X\hat{\beta})^2 + \dots + \gamma_{k-1}(X\hat{\beta})^k + \text{erro}, \quad (7)$$

Onde $X\hat{\beta} = \hat{Y}$ e representam os valores ajustados da estimação OLS do modelo original. As hipóteses a testar são: $H_0: \gamma_1 = \dots = \gamma_{k-1} = 0$ vs. $H_1: \exists \gamma_i \neq 0$ ($i = 1, \dots, k - 1$). Um simples teste F permite realizar este ensaio estatístico. Tipicamente, um valor de $k=2$ costuma ser suficiente para a deteção de alguma má especificação, caso ela exista verdadeiramente. A ideia com a introdução do quadrado dos valores ajustados é a de que estes podem configurar aproximações de variáveis omissas no modelo original. A rejeição de H_0 revelará indícios de que provavelmente o modelo precisará de ver a sua forma funcional ajustada, introduzindo interações entre variáveis, elevar algumas ao quadrado, etc, ou poderá indicar a ausência no modelo de outros regressores estatisticamente relevantes para explicar Y , algo que provoca enviesamento no estimador OLS (Hayashi, 2000).

3.2.2. Teste de Quebra de Estrutura de Bai-Perron

Com o intuito de testar estatisticamente a diferença de comportamentos estrutural da população portuguesa ao longo do período da amostra, recorreu-se aos testes sugeridos por Bai & Perron (1998). Estes testes permitem testar múltiplos pontos de quebras de estrutura, ocorrendo em datas desconhecidas, no modelo de regressão linear estimado pelo OLS. Considerando o modelo de regressão linear com m pontos de quebra, ou seja, $m+1$ regimes:

$$y_t = x_t'\beta + z_t'\delta_t + u_t, (t = T_{j-1} + 1, \dots, T_j), (j = 1, \dots, m + 1). \quad (8)$$

onde, por convenção, $T_0 = 0$ e $T_{m+1} = T$. Neste contexto, y_t representa a variável dependente observada; x_t o vetor ($p \times 1$) de variáveis cujos coeficientes β são invariantes entre os regimes; z_t o vetor ($q \times 1$) de variáveis cujos coeficientes δ_t são específicos de cada regime; e u_t o habitual termo de erro. O objetivo passa por estimar os coeficientes de regressão, juntamente com os m pontos de quebra (T_1, \dots, T_m) que são tratados como desconhecidos, a partir das T observações disponíveis da concretização do processo estocástico $\{y_t, x_t, z_t\}$. A grande diferença face ao modelo puro de quebra de estrutura é a presença de um conjunto de regressores cujos coeficientes se admite serem constantes entre os diferentes regimes.

Um aspeto relevante que favorece o recurso ao teste de Bai-Perron tem que ver com o facto de não ser necessário dispor-se de nenhuma indicação prévia sobre a(s) data(s) de mudança de regime, ao contrário do que é exigido na estatística de Chow. Se por um lado, a escolha arbitrária da data de quebra conduz a testes com potência reduzida, ou seja, mesmo que exista diferença de regimes ela dificilmente é detetada; por outro, uma prévia análise dos dados para tentar “adivinhar” o momento da quebra contamina a dimensão efetiva dos testes, que acaba por ser superior à dimensão nominal (tipicamente fixada nos 5%). Deste modo, o desconhecimento dos m pontos de quebra (T_1, \dots, T_m) não devidamente acautelado pode afetar seriamente as propriedades dos testes realizados, sendo que o teste proposto por Bai-Perron ao endogeneizar explicitamente a estimação desses pontos permite ultrapassar os problemas acima referidos.

Verificando-se as condições descritas em Bai & Perron (1998), as hipóteses a testar podem ser especificadas de formas diferentes, sendo a distribuição da estatística de teste de cada uma das especificações adaptada a cada caso. O primeiro caso consiste em testar a inexistência de pontos de quebra de estrutura contra um determinado número fixo de pontos de mudança de regime, ou seja: $H_0 : m = 0$ vs. $H_1 : m = k$ ($k > 0$) com k fixo. A grande desvantagem consiste na exigência da especificação de k sob H_1 e por isso não será considerada.

Alternativamente, poderá ser interessante não especificar na hipótese alternativa um número particular de quebras. Os autores introduziram dois novos testes que permitem testar a inexistência de quebras de estrutura na série contra um número desconhecido de quebras dado um limite superior M . O primeiro é uma versão com um peso igual definido

como $UDmaxF_T(M, q) = \max_{1 \leq m \leq M} F_T(\hat{\lambda}_1, \dots, \hat{\lambda}_m; q)$, onde $\hat{\lambda}_j = \frac{\hat{T}_j}{T}$ ($j = 1, \dots, m$) são as estimativas dos pontos de quebra obtidos através da minimização global da soma dos quadrados dos resíduos. O segundo teste aplica pesos aos testes individuais de modo que os valores-p marginais sejam iguais entre os valores de m e é designada por $WDmaxF_T(M, q)$.

Finalmente, a última especificação proposta pelos autores consiste em testar sequencialmente a existência de l quebras estruturais contra $l+1$, isto é, $H_0 : m = l$ vs. $H_1 : m = l + 1, \forall l \geq 0$. Este teste, definido por $supF_T(l + 1|l)$, começa por testar sob a hipótese nula a inexistência de nenhum ponto de quebra contra a hipótese alternativa de uma única quebra em toda amostra, sendo que no caso da rejeição de H_0 a mesma é dividida em duas subamostras repartidas na data de quebra estimada. O teste rejeita a hipótese nula se o valor mínimo da soma dos quadrados dos resíduos (SQR) do modelo com $l+1$ quebras for suficientemente mais pequena que a mesma soma do modelo com l quebras. O procedimento é repetido até à não rejeição de H_0 , situação na qual nenhuma das subamostras apresenta evidência da ocorrência de quebras de estrutura. Os valores críticos de cada especificação de teste podem ser consultados em Bai & Perron (1998).

4. ANÁLISE EXPLORATÓRIA DOS DADOS

4.1. Caracterização dos Dados

Como fonte primária de caracterização comportamental da população portuguesa encontra-se o Painel de Mobilidade PSE³, um painel pioneiro a nível mundial por ter a capacidade de rastrear a mobilidade da população em contínuo. Tratando-se de um painel representativo da população, medindo a mobilidade todos os dias, todas as horas, ao longo de todo o ano, possibilita a identificação de métricas que se podem revelar determinantes para o presente estudo, tais como a identificação de hábitos e padrões comportamentais da população portuguesa.

³ Através do Painel de Mobilidade PSE, a consultora PSE estuda a mobilidade dos portugueses em contínuo através de monitorização de localização e meios de deslocação via aplicação móvel de um painel de 4992 indivíduos representativos do Universo em estudo. Este estudo, que implica uma APP instalada nos telemóveis da amostra participante, apresenta um elevado rigor porque regista a localização via GPS. Deste modo, monitoriza-se a deslocação real da população 24 horas por dia. Para um Universo de 6.996.113 indivíduos residentes nas regiões estudadas a margem de erro imputável ao estudo é de 1.62% para um intervalo de confiança a 95%.

A partir da informação fornecida pelo Painel de Mobilidade PSE é permitido observar a resposta comportamental que os portugueses adotaram ao longo do período de pandemia. Assim, como variável de interesse deste estudo foi considerado o nível de confinamento diário dos portugueses, medido pela percentagem de pessoas que permanecem confinadas no lar, durante o período compreendido entre 1 de março de 2020 e 31 de março de 2021, o que perfaz um total de 396 observações.

A Figura V do Anexo A, que ilustra a respetiva série temporal, permite identificar visualmente os dois momentos de *lockdown* do país, o primeiro entre março e abril de 2020 e o segundo entre janeiro e fevereiro de 2021. De um modo geral, no primeiro *lockdown* verifica-se uma maior adesão do dever de confinamento por parte da população, com um valor máximo de aproximadamente 80% dos portugueses a permanecer confinados no domicílio, enquanto o segundo registou um valor máximo de 73%.

Adicionalmente, também é possível verificar que mesmo em períodos de relativa normalidade de circulação, com níveis de restrições relativamente baixos, como nos meses de julho e agosto de 2020, os valores de confinamento não se fixam abaixo de determinado intervalo, nomeadamente, entre os 25% e os 30%. Ou seja, parece existir um *stock* de população que tende diariamente a não sair de casa e que, em média, corresponde a um valor entre os 25% e os 30% da população. É necessário conhecer a estrutura demográfica da população portuguesa que, alertam diversos especialistas Azevedo *et al.* (2011), está ano após ano cada vez mais envelhecida, para obter um melhor enquadramento desta factualidade. Segundo as estimativas do INE para 2020, Portugal apresenta cerca de 22,4% da população com 65 ou mais anos. Estes elementos ajudam a compreender o fenómeno da permanência diária da população no seu lar.

Recuperando o quadro conceptual do modelo COM-B (Figura I) importa identificar as variáveis que, segundo o mesmo, possam justificar a tomada de ação por parte da população portuguesa. No que diz respeito à dimensão Motivação, esta será incorporada através da série diária do número de novos casos de covid-19 reportados pela Direção-Geral de Saúde⁴, na medida em que a adoção de comportamentos por parte da população será também resultado da perceção ao risco de contração da doença que a mesma exibiu nos diversos momentos da pandemia. A propensão da população a confinar deverá ou não

⁴ Fonte: <https://ourworldindata.org/coronavirus>

crescer com o aumento do número de novos casos reportados da doença? Através da representação gráfica da série que consta da Figura VI do Anexo A observa-se aquilo que foi a evolução de incidência da doença covid-19 em Portugal, de onde se destacam visualmente três vagas, cada uma com um nível crescente de incidência face à anterior: a primeira nos meses de março e abril de 2020, pouco expressiva; a segunda nos meses de novembro e dezembro do mesmo ano, com um máximo de cerca de 8000 casos diários; e a terceira nos meses de janeiro e fevereiro de 2021, onde se atingiu o recorde de 16000 casos num só dia.

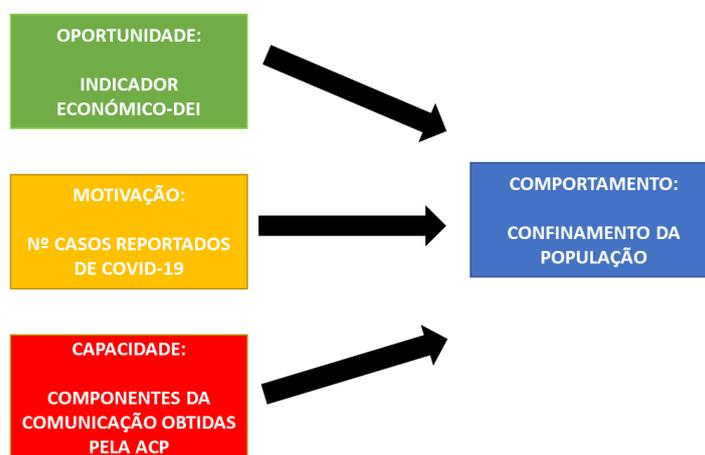
Relativamente à dimensão Oportunidade, a mesma será incluída através do *Daily Economic Indicator* (DEI), um indicador produzido e criado pelo Banco de Portugal que, nesta conjectura, tem o papel de configurar um *proxy* do nível diário da atividade económica no país, (Lourenço & Rua, 2020). O intuito é o de fornecer uma descrição do ambiente socioeconómico em que a população se encontrava mergulhada, percebendo se as condições socioeconómicas podem desempenhar um papel importante neste *framework*. Este indicador composto cobre diversas dimensões relacionadas com a atividade económica, tais como tráfego rodoviário de veículos comerciais pesados nas autoestradas, consumo eletricidade e gás natural, carga e correio desembarcados nos aeroportos nacionais, entre outras. Através da representação gráfica da série que consta da Figura VI do Anexo A verifica-se claramente o impacto que teve principalmente o primeiro grande confinamento de março e abril de 2020 na brusca redução da atividade económica, bem como a sua recuperação nos meses seguintes, sem nunca atingir taxas de crescimento registadas nos períodos pré-pandemia.

Com o intuito de atenuar as irregularidades típicas das séries de alta frequência decidiu-se utilizar no modelo uma média móvel centrada de 7 dias do DEI, também disponibilizada pelo Banco de Portugal ao invés da série original. A ilustração gráfica desta série pode também ser encontrada na Figura VI do Anexo A. A caracterização das variáveis utilizadas no modelo que personificam as dimensões de Comportamento, Motivação e Oportunidade pode ser sintetizada na Tabela I do Anexo B.

Por fim, será a dimensão Capacidade que irá preconizar os estímulos comunicacionais emitidos pelas autoridades, através da recolha de um largo conjunto de variáveis que qualificam diariamente as diferentes medidas e políticas adotadas. As

medidas são encaradas como instrumentos de comunicação das autoridades públicas, no sentido em que a adoção de uma política pretende transmitir implicitamente determinada mensagem para a população. Por exemplo, quando o Governo decretou a suspensão da atividade letiva presencial e uma série de outras medidas fortemente restritivas estava a pretender transmitir a mensagem para as pessoas permanecerem confinadas no seu domicílio. Com base nas variáveis recolhidas pelo autor e apresentadas na próxima secção será possível, através da Análise de Componentes Principais, identificar diferentes estímulos comunicacionais frequentes ao longo do período em análise. A Figura III representa graficamente a estrutura conceptual do modelo COM-B aplicado ao caso em estudo.

Figura III. Estrutura conceptual do modelo COM-B aplicada ao caso em estudo



Fonte: Adaptado de Michie *et al.* (2011)

4.2. Obtenção das Componentes da Comunicação

A procura de informação caracterizadora da comunicação incidu sobre a identificação de medidas adotadas pelas autoridades políticas e sanitárias, bem como sobre a sinalização das intervenções públicas mais mediáticas das mesmas, através da categorização de variáveis binárias ao longo do tempo.

O processo de *data collection* exigiu uma pesquisa exaustiva e minuciosa entre as múltiplas fontes disponíveis com o intuito de endogeneizar o máximo de informação possível, nomeadamente medidas aplicadas e intervenções públicas relevantes. Foi necessário visitar atentamente os mais de 70 comunicados das reuniões do Conselho de

Ministros⁵, de onde constavam as diversas medidas decretadas e respetivos períodos de vigência; procurar todas as datas das inúmeras intervenções do Presidente da República⁶ e de mais de uma centena de conferências de imprensa da DGS e ainda outras informações relevantes no site do ministério da Saúde. Todo este complexo processo de recolha culminou numa extensa e robusta base de dados com um total de trinta variáveis, como consta da Tabela III do Anexo B.

Encontram-se identificadas ao longo do tempo, nas respetivas variáveis, grande parte das medidas restritivas impostas pelos diferentes órgãos de soberania, principalmente pelo Governo, como por exemplo, a suspensão da atividade presencial (*X_suspensao_atividade_letiva_presencial*) e a obrigatoriedade do teletrabalho (*X_obrigatoriedade_teletrabalho*). Em termos de declarações formais ao país por parte das entidades políticas e sanitárias, dataram-se as comunicações do primeiro-ministro (*X_declaracoes_formais_governo*), as comunicações do Presidente da República (*X_declaracoes_formais_PR*), as conferências de imprensa da DGS (*X_conferencias_imprensa_DGS*) e as reuniões do Infarmed (*X_reunioes_Infarmed*). De forma complementar, e por se tratar de momentos com grande relevância mediática, também se coletaram alguns anúncios de realização/cancelamento de eventos desportivos ou culturais em Portugal, como é o caso do anúncio da realização da fase final da Liga dos Campeões em Portugal (*X_anuncio_Liga_dos_Campeoes*) e do cancelamento da organização dos Santos Populares (*X_cancelamento_Santos_Populares*).

A abordagem da Análise de Componentes Principais (ACP) cuja teoria e vantagens foram explicitadas na secção 3.1. aplicada às variáveis desta extensa base de dados tem como principal objetivo a redução da sua dimensão com a menor perda de informação possível. A utilização deste método analítico em variáveis com codificação binária como é o caso do presente estudo tem sido prática recorrente, existindo grande suporte na literatura, principalmente quando os investigadores não dispõem da totalidade das variáveis necessárias para a realização do seu trabalho metodológico. Neste sentido, destacam-se as investigações na área das Políticas Públicas de Houweling *et al.* (2003), Gwatkin *et al.* (2000), Vyas & Kumaranayake (2006), Mckenzie (2004) e Minujin & Bang (2002), em que este instrumento tem sido decisivo na construção de indicadores de

⁵ <https://www.portugal.gov.pt/pt/gc22/governo/comunicados-do-conselho-de-ministros>

⁶ <https://www.presidencia.pt/atualidade/noticias>

status socioeconómico a partir de variáveis com codificação binária, tipicamente elaboradas através de questões de resposta afirmativa ou negativa (por exemplo, se o indivíduo possui carro próprio ou não; se possui televisão ou não; etc).

Com o objetivo de analisar econometricamente o efeito da situação económica familiar sobre o desempenho escolar das crianças na Índia, em Filmer & Pritchett (2001) a metodologia adotada é em tudo semelhante à perfilada neste trabalho, no sentido em que o recurso inicial à ACP precedeu a modelização econométrica. Para obter um *proxy* da riqueza familiar, os autores utilizam um conjunto de variáveis *dummy* obtidas por questionário às famílias, por forma a estimar um índice de situação económica através da ACP. Posteriormente, este índice é incorporado no Modelo de Regressão Linear estimado por OLS.

Através da Tabela IV do Anexo B é possível verificar que as duas estatísticas que a literatura prevê de diagnóstico de ajustamento da ACP aos dados confirmam uma boa adequação. Em primeiro lugar, a estatística de Kaiser-Meyer-Olkin (KMO), que representa um índice variável entre 0 e 1, avalia a adequabilidade da ACP às variáveis em questão. O valor obtido foi de 0,597, superior aos 0,5 propostos por Hair *et al.* (1987) como limiar mínimo abaixo do qual o método deixa de produzir resultados satisfatórios. Em segundo lugar, o teste de esfericidade de Bartlett testa a hipótese nula de que as variáveis não estão correlacionadas na população, portanto o desejável será um valor-p que rejeite a hipótese nula. O valor-p produzido pelo respetivo teste sendo igual a zero permite rejeitar confortavelmente a hipótese nula do teste, pelo que existem fortes indícios estatísticos de que o ajustamento do método tenha sido estatisticamente adequado.

Por defeito, o software SPSS Statistics apresenta os resultados das componentes cujo valor próprio é superior a 1, como se pode verificar na Tabela V do Anexo B. Pode-se concluir que com 10 componentes é possível explicar 82,16% da variância das variáveis originais. No entanto, uma vez que existem muitas variáveis, a regra de reter componentes com valor próprio superior a 1 nem sempre se revelou vantajosa, provocando uma análise com excesso de componentes. Por isso, optou-se por conservar somente as componentes cujo valor próprio é superior a 2, ou seja, apenas as 4 primeiras calculadas, que explicam mais de 50% da variância da totalidade das variáveis originais (55,8%).

Analisando a Tabela VI do Anexo B, que detalha através dos *loadings* (ou pesos) a composição de cada uma das 4 componentes extraídas, seria interessante adotar uma designação sugestiva para cada uma. A primeira componente apresenta três variáveis que representam medidas fortemente restritivas da mobilidade dos portugueses com um peso muito elevado, acima de 0,8, que são: *X_encerramento_Centros_Comerciais*, *X_suspensao_atividade_letiva_presencial* e *X_encerramento_restaurantes_apenas_take_away*, pelo que será nomeada *C1_Medidas_agressivas*. A segunda componente apresenta três variáveis com um peso superior a 0,7 que representam medidas com um nível de restrição não tão elevado como as mencionadas na primeira componente, no entanto representam políticas preventivas no que diz respeito ao controlo da disseminação da doença: *X_encerramento_supermercados_20h_dia_util_17h_fds*, *X_proibicao_circular_entre_concelhos* e *X_obrigatoriedade_uso_mascaras*, pelo que será nomeada *C2_Medidas_Preventivas*. A terceira componente também apresenta duas variáveis com um peso elevado que podem ser enquadradas como medidas preventivas (*X_encerramento_comercio_retalho_22h* e *X_proibicao_circular_23h_5h_dia_util*), no entanto a variável *X_discriminacao_medidas_por_risco_concelho* com um peso de 0,859 atribui a esta componente uma dimensão territorial associada à aplicação das medidas restritivas que a anterior não tinha, pelo que será nomeada de *C3_Medidas_Territoriais*. Finalmente, a quarta componente exhibe como variáveis com maiores *loadings* (acima de 0,85) as seguintes: *X_Estado_Calamidade_só_19_freguesias_AM_Lisboa* e *X_encerramento_supermercados_AML_22h*. Esta componente sintetiza evidentemente as medidas restritivas direcionadas exclusivamente para a Área Metropolitana de Lisboa que foram aplicadas nos meses de julho e agosto de 2020, pelo que será designada por *C4_Medidas_AML*. A Tabela II do Anexo B caracteriza as quatro componentes que irão personificar a dimensão Capacidade no modelo a estimar. A Figura VII do Anexo A ilustra a representação gráfica destas componentes.

4.3. Análise Descritiva das variáveis

Uma vez apresentadas todas as variáveis que irão compor o modelo e antes de proceder à respetiva estimação, importa conhecer melhor como é que as mesmas se relacionam com a variável dependente. A análise da matriz de correlações é uma técnica de análise utilizada numa fase inicial das aplicações empíricas, em que o investigador

pretende obter *insights* adicionais, principalmente sobre o sinal e magnitude dos coeficientes de correlação das variáveis em estudo.

Através do cálculo da matriz de correlações de Pearson (Tabela VII do Anexo B) conclui-se que, excetuando apenas o coeficiente da variável *C3_Medidas_Territoriais*, todas as variáveis independentes exibem uma correlação significativa com a variável de interesse *confinamento* e ao nível de 1%. Em termos de magnitude dos coeficientes de correlação destaca-se o da variável *C1_Medidas_agressivas* com um valor de 0,725, indiciando assim uma forte e positiva medida de associação linear entre as políticas de restrições mais ferozes e o confinamento da população. Com o mesmo perfil também se encontra a variável *C2_Medidas_Preventivas*, apresentando no entanto um coeficiente mais baixo de 0,436. No que diz respeito à variável que representa o número de novos casos, o sinal do coeficiente estimado é positivo, fazendo antever aqui um eventual efeito positivo da mesma sobre a variável *confinamento*.

A interpretação dos resultados da matriz de correlações deve ser sempre efetuada com cautela e com caráter meramente exploratório, uma vez que quer os sinais, quer a magnitude dos coeficientes de cada variável são calculados de forma isolada, não contabilizando eventuais efeitos das restantes variáveis sobre as próprias estimativas. A fim de uma análise estatisticamente mais completa, robusta e detalhada utilizar-se-ão os métodos econométricos como veículo condutor do desenvolvimento e das conclusões desta aplicação.

5. MODELAÇÃO E ANÁLISE DE RESULTADOS

5.1. Especificação e estimação do modelo

A modelação de séries temporais de alta frequência exige que a especificação do modelo incorpore alguns efeitos específicos de dados desta natureza, como é o caso dos padrões sazonais e dos efeitos de calendário, que se espera que tenham influência na dinâmica da série. A endogeneização destes elementos é fundamental para evitar potenciais relações espúrias e isolar os efeitos das variáveis explicativas exógenas sobre a variável de interesse. Sendo a variável dependente deste estudo de periodicidade diária, é verosímil admitir-se a existência de determinados comportamentos específicos de cada dia da semana. Neste sentido, definiram-se variáveis *dummy* representativas de cada dia da semana, em que assumem o valor 1 caso seja referente ao respetivo dia da semana e 0

caso contrário, tais como: *Segunda, Terça, Quarta, Quinta, Sexta e Sábado*, adotando o domingo como grupo base.

Adicionalmente, através da observação empírica concluiu-se que poderia ser vantajoso introduzir também uma variável *dummy* que identificasse os dias feriados por serem momentos em que, genericamente, a mobilidade das pessoas é mais reduzida, designando-a por *Feriado*. Dado que é expectável que o efeito dos dias feriados na escolha de não sair de casa por parte dos portugueses seja diferente consoante os mesmos calharem ao dia útil ou ao fim de semana, foi adicionada a variável anterior com uma interação, isto é: *Feriado*Fim_Semana*, sendo que *Fim_Semana* assume o valor 1 nos dias fins de semana e 0 caso contrário.

Como já foi referido, o período amostral principia a 1 de março de 2020, ainda antes da notificação dos primeiros casos de covid-19 em Portugal, que ocorreu dia 3 de março. A verdade é que a partir desse dia o número de casos começou a subir. Subida essa que foi acompanhada de uma crescente cobertura mediática promovida à chegada do então novo coronavírus a Portugal. O ponto alto terá sido a declaração da doença como pandemia pela Organização Mundial de Saúde (OMS) a 11 de março, sendo que em Portugal o primeiro grande momento de tensão mediática terá sido o anúncio do encerramento de todas as atividades letivas pelo Primeiro-Ministro, imediatamente no dia seguinte (12 de março). Os dados mostram que a partir de dia 12 até à declaração do primeiro Estado de Emergência por parte do Presidente da República, que ocorreu apenas no dia 18 de março, os portugueses já contraíam voluntariamente os seus níveis de mobilidade, antecipando, ainda que de forma suave, as próprias restrições que a declaração do Estado de Emergência viria a impor. Assim, definiu-se a variável *Antecipação* que assume o valor 1 entre os dias 13 e 20 de março de 2020, período que antecedeu a entrada em vigor das medidas decretadas pelo governo, e 0 nos restantes dias, com o intuito de endogeneizar este fenómeno de antecipação de comportamentos por parte da população.

Recorrendo ao software econométrico EViews, estimou-se o seguinte modelo ADL(7,7) com as restrições $\lambda_2 = \dots = \lambda_6 = 0 \wedge \phi_3 = \dots = \phi_6 = 0$:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_7 Y_{t-7} + \theta D_t + \beta C_{t-1} + \lambda_1 X_{t-1} + \lambda_7 X_{t-7} + u_t, t = 1, \dots, 389 \quad (9)$$

onde $Y_t = \text{confinamento}_t$; D_t representa o conjunto das variáveis determinísticas, $D_t = (\text{segunda}_t, \dots, \text{sabado}_t, \text{feriado}_t, \text{feriado}_t \times \text{fim_semana}_t, \text{antecipacao}_t)$; θ é o respetivo vetor de coeficientes a estimar (1×9); C_{t-1} representa as variáveis explicativas exógenas das componentes, ou índices, obtidas pela ACP desfasadas 1 período, isto é, $C_{t-1} = (C1_Medidas_Agressivas_{t-1}, \dots, C4_Medidas_AML_{t-1})$, sendo β o respetivo vetor (1×4) de coeficientes a estimar; X_{t-1} representa as variáveis explicativas exógenas referentes à incidência da doença e ao indicador económico, $X_{t-1} = (X_novos_casos_rep_{t-1}, X_dei_ma_{t-1})$, sendo que λ_1 é o respetivo vetor (1×2) dos coeficientes a estimar; X_{t-7} representa somente a variável $X_dei_ma_{t-7}$ e λ_7 o respetivo coeficiente desconhecido; u_t é o habitual termo de erro assumindo a condição $u_t \sim iid(0, \sigma^2)$.

Deste modo, a equação (9) incorpora como regressores, para além das variáveis determinísticas, a própria variável dependente desfasada 1, 2 e 7 períodos, as quatro componentes principais obtidas e as variáveis X_dei_ma e $X_novos_casos_rep$ desfasadas 1 período e ainda X_dei_ma desfasada 7 períodos. Por se tratar de dados diários torna-se pouco verosímil a existência de algum efeito contemporâneo dos regressores sobre a variável dependente. Por exemplo, dado um aumento do número de novos casos hoje, não é expectável que as pessoas reduzam os seus contactos sociais no mesmo período. O mesmo se pode conjecturar relativamente aos estímulos comunicacionais e à atividade económica. Por esta razão decidiu-se não incluir nenhum regressor contemporâneo. A inclusão dos dois regressores que representam a variável dependente desfasada 1 e 2 períodos perspetiva-se relevante uma vez que fornecerá ao modelo uma componente dinâmica essencial para eliminar a evidência estatística de autocorrelação nos erros. Finalmente, os desfasamentos de 7 períodos quer na variável dependente, quer na variável X_dei_ma foram incorporados por forma a captar a existência de uma eventual sazonalidade estocástica existente na variável dependente que, a existir, não é caracterizada apenas pelas *dummies* determinísticas.

O método de estimação OLS foi aplicado à equação (9) assumindo a exogeneidade contemporânea dos regressores não-determinísticos, de modo a empregar a teoria assintótica pelo que, dada a robusta dimensão amostral deste exercício, a margem para a sua aplicação é confortável. Os resultados da estimação, que constam da Figura VIII do

Anexo A, demonstram a significância estatística individual das variáveis explicativas empregando os níveis habituais. Poder-se-ia especular que devido à estimação dos 20 parâmetros do modelo o estimador OLS pudesse vir ineficiente, com falta de precisão, o que implicaria uma redução da potência dos testes de significância estatística individual. Ora, não só apenas as variáveis *confinamento*_{t-2}, *sabado*_t e *C4_Medidas_Territoriais*_{t-1} não exibem significância ao nível de 5%, como todos os outros regressores são significativos a 5 e a 2%.

No que diz respeito à análise de especificação do modelo, ainda que de forma informal, a estatística de Durbin-Watson tendo produzido um valor de 1,915 está relativamente perto do valor em que se admite não existir evidência de autocorrelação de 1ª ordem, que é igual a 2. Aplicando o teste LM testando sob H_0 a ausência de autocorrelação de 1ª ordem dos erros (Figura IX do Anexo A), obteve-se um valor-p de 0,2519, suficiente para não se rejeitar a nula aos níveis de significância habituais, pelo que não existe evidência estatística de autocorrelação dos erros de 1ª ordem. Com o intuito de despistar a existência de algum padrão de autocorrelação sazonal, testou-se ainda sob H_0 a ausência de autocorrelação dos erros até à ordem 7. A estatística de Breusch-Godfrey, BG(7), produziu um valor-p igual a 0,1618, pelo que não se rejeita novamente a nula do teste a 5% (Figura X do Anexo A), logo não existem indícios estatísticos da presença de autocorrelação até à ordem 7. O suporte evidenciado pela estatística BG(7) é de que o padrão sazonal da variável dependente se encontra corretamente explicado pela componente principal do modelo.

Para encerrar o capítulo dos testes de especificação, procedeu-se à realização do teste RESET introduzindo na equação a estimar o termo \hat{Y}^2 e testando a sua significância estatística individual. A partir da Figura XI do Anexo A constata-se que o valor-p do teste sendo igual a 0,065 permite não rejeitar H_0 a 5%, logo não existe evidência a favor da significância do termo incluído. A evidência suportada pelos testes estatísticos efetuados encaminha para a correta especificação do modelo, principalmente no que diz respeito à sua especificação dinâmica, e por isso, toda a inferência estatística considerar-se-á válida.

5.2. Interpretação dos resultados

Reunidas as condições estatísticas o objetivo é simples e claro: através de um detalhado enquadramento do modelo COM-B, determinar se a comunicação impactou o

comportamento dos portugueses, nomeadamente, na contração dos seus contactos sociais, diferenciando os diferentes estímulos comunicacionais.

A performance da aplicação do modelo COM-B a dados temporais revelou-se extremamente bem-sucedida uma vez que reproduz integralmente o principal resultado do referido modelo que vaticina que a adoção de comportamentos é resultado da interação entre as vertentes Oportunidade, Capacidade e Motivação. Ora, a significância estatística individual das variáveis que personificam estas dimensões, nomeadamente, as componentes principais incluídas na regressão, a variável da incidência da doença e o indicador económico, espelha precisamente o poder explicativo das mesmas sobre a decisão comportamental dos portugueses.

No que diz respeito à variável $X_{novos_casos_rep_{t-1}}$ (significativa a 5% e a 1%), que traduz a perceção ao risco de contração da doença, sendo o sinal do coeficiente estimado positivo implica um aumento dos níveis de confinamento da população à medida que o número de novos casos reportados também aumenta. Este resultado é consistente com a literatura que sugere os indivíduos irão aumentar o distanciamento social, de acordo com o crescimento do número de novos casos (Klein *et al.*, 2007). A estimativa obtida para esta variável indica que, em média, por cada 10 000 casos de covid-19 reportados (valor médio da terceira vaga da pandemia em Portugal) a percentagem de pessoas confinadas no dia seguinte aumenta aproximadamente 5 pontos percentuais, o que representa cerca de 496 mil portugueses a confinar, *ceteris paribus*. A aversão ao risco de contágio constitui assim um determinante da redução das interações sociais da população que, ao se sentir mais exposta ao vírus, decide adotar comportamentos mais preventivos, como a permanência no seu domicílio.

Relativamente à questão da descrição do ambiente socioeconómico como determinante do nível de confinamento salienta-se o papel do regressor $X_{dei_ma_{t-1}}$ (significativo a 5%) como elemento preconizador da dimensão Oportunidade. O sinal negativo da sua estimativa sugere, tudo o resto constante, uma redução do número de pessoas confinadas face a um crescimento da atividade económica registado no dia anterior, isto é, a existência de um contexto económico mais favorável proporcionará à população outra predisposição para o consumo e para a vida ativa que certamente contextos de recessões económicas não motivam.

A significância estatística individual ao nível de 5% dos regressores $C1_Medidas_Agressivas_{t-1}$, $C2_Medidas_Preventivas_{t-1}$ e $C4_Medidas_AML_{t-1}$ sugere relações de causalidade entre os diferentes estímulos comunicacionais e o comportamento da população. Por se tratarem de combinações lineares, não se torna direta a interpretação dos coeficientes β_j ($j = 1, \dots, 4$), pelo que se dará maior relevância ao seu sinal e significância. No caso da primeira variável, o impacto em termos comportamentais fornecido pelo sinal positivo da estimativa do seu coeficiente vai ao encontro do esperado, ou seja, uma comunicação assente em medidas fortemente restritivas e agressivas induz à população uma atitude de recolhimento no seu lar. Existe forte evidência estatística de que a mensagem transmitida a partir da aplicação de medidas fortemente restritivas, como as do encerramento do setor da restauração ($X_encerramento_restaurantes_apenas_take_away$), da atividade letiva presencial ($X_suspensão_atividade_letiva_presencial$), e dos Centros Comerciais ($X_encerramento_Centros_Comerciais$) impostas pelo Governo, obteve o efeito desejado sobre o comportamento da população portuguesa, que elevou os seus níveis de confinamento.

Em oposição, uma comunicação levada a cabo a partir de medidas declaradamente preventivas parece não ter obtido o efeito desejado entre os portugueses. O sinal negativo da estimativa de $C2_Medidas_Preventivas_{t-1}$ vaticina inclusive uma redução do confinamento diário da população provocada pelo mesmo estímulo. Assim, medidas decretadas pelo Governo como a redução do horário de encerramento dos supermercados ($X_encerramento_supermercados_20h_dia_util_17h_fds$) e a inibição de circulação entre concelhos ($X_proibição_circular_entre_concelhos$) despertaram nos portugueses, ainda que de forma não contemporânea, um comportamento contrário ao que seria esperado: um aumento de circulação. Políticas desta natureza, principalmente no caso da proibição de circulação inter-concelhia aplicada em períodos cirúrgicos em que se antevê padrões de elevada mobilidade como feriados e pontes, tendem a fomentar na população comportamentos que visam contornar as restrições, na medida em que esta se aproveita dos dias anteriores e posteriores para realizar as suas deslocações, provocando então a referida redução do confinamento nesses momentos.

Face ao exposto, torna-se pertinente realçar a dicotómica resposta dos portugueses face a estes dois conjuntos de medidas: se por um lado medidas de carácter mais agressivo

e autoritário são bem correspondidas pela população, que as acata; por outro, medidas de carácter predominantemente pontual e avulso parecem não surtir o efeito desejado.

No que concerne à variável explicativa $C4_Medidas_AML_{t-1}$, o sinal positivo da sua estimativa precifica o desejado impacto da comunicação direcionada para os residentes da Área Metropolitana de Lisboa nos meses do verão de 2020 sobre os comportamentos da população, quando as medidas mais restritivas eram quase exclusivamente aplicadas a este território, nomeadamente o decretar do Estado de Calamidade para as 19 Freguesias da AML ($X_Estado_Calamidade_só_19_Freguesias_AML$) e o encerramento às 22 horas dos supermercados na AML ($X_encerramento_supermercados_AML_22h$). A evidência estatística corrobora, para este caso em específico, o sucesso das medidas aplicadas no sentido da adoção de comportamentos preventivos por parte da população.

De forma semelhante ao que foi detalhado para o caso do regressor $C2_Medidas_Preventivas_{t-1}$, cujo impacto estimado na variável dependente foi o oposto do que seria esperado pela natureza da variável, no que respeita ao regressor $C3_Medidas_Territoriais_{t-1}$ ao nível de significância de 5% pode-se afirmar que não se encontrou relevância estatística para explicar variações de $confinamento_t$. Deste modo, salienta-se a ineficácia de um estilo comunicacional baseado em medidas preventivas diferenciadas consoante o risco de incidência de cada concelho, nomeadamente a imposição do confinamento noturno nos dias úteis ($X_proibicao_circular_23h_5h_dia_util$) e de horários de encerramento reduzidos para as 22 horas no comércio a retalho ($X_encerramento_comercio_retalho_22h$).

Em resposta ao objetivo proposto no início desta secção, conclui-se que existe forte evidência estatística de diferentes impactos de estímulos comunicacionais na adoção de comportamentos preventivos por parte da população portuguesa. Estes impactos podem ser decompostos em três categorias em função da sua eficácia na contenção dos contactos sociais: Impactos Positivos, Impactos Negativos e Impactos Neutros. De entre os Impactos Positivos destacam-se claramente as medidas de cariz mais agressivo, como a suspensão da atividade letiva presencial e o encerramento do setor da restauração; relativamente aos Impactos Negativos evidenciam-se a inibição de circulação entre concelhos e a redução do horário de encerramento dos supermercados; finalmente, em

termos de Impactos Neutros surgem a imposição do confinamento noturno e o encerramento antecipado do comércio a retalho para as 22 horas.

No entanto, importa perceber se a resposta comportamental foi constante durante o período considerado ou se por outro lado, em determinados momentos, a reação dos portugueses foi diferente face a estímulos idênticos.

5.3. Teste às Alterações Comportamentais

O principal propósito desta secção consiste em testar alguns tipos de alterações de resposta comportamental dos portugueses durante o período considerado e será dividida em dois momentos: o primeiro, onde se testará exclusivamente diferenças de resposta face à vertente da comunicação; e o segundo, onde de uma forma mais generalizada, se realizará o teste de quebra de estrutura de Bai-Perron, tendo em vista a deteção de momentos de quebra na série.

Reteve-se da secção anterior que um estilo comunicacional mais agressivo, apoiado em medidas de restrições mais fortes, constituiu um importante determinante da contenção dos contactos sociais da população. Uma vez que foi essencialmente nos dois períodos de confinamento generalizado, entre março e abril de 2020 e entre janeiro e março de 2021, que este estilo se protagonizou, seria interessante perceber se os portugueses reagiram de igual forma face ao mesmo registo de comunicação nesses dois períodos. Com efeito, adicionou-se ao modelo duas variáveis *dummy* de interação a $C1_Medidas_Agressivas_{t-1}$ identificando os anos de 2020 e 2021. Pretende-se testar estatisticamente a hipótese simples de igualdade dos respetivos coeficientes contra a hipótese composta de que o coeficiente referente ao ano de 2020 é maior do que o de 2021, ou seja, $H_0: \beta_{C1_{t-1}*2020} = \beta_{C1_{t-1}*2021}$ contra $H_1: \beta_{C1_{t-1}*2020} > \beta_{C1_{t-1}*2021}$, sendo a região crítica do teste definida por $RC_{5\%} = \{t: t_{obs} > 1,645\}$.

A partir da Figura XII do Anexo A verifica-se que o valor observado da estatística t é igual a 2,179, então como $t_{obs} = 2,179 > 1,645$ rejeita-se a hipótese nula de igualdade dos coeficientes ao nível de significância de 5%, existindo evidência estatística a favor da hipótese de que o coeficiente do ano de 2020 é superior ao de 2021. Este resultado não será surpreendente se tivermos em linha de conta os vários alertas que os especialistas da área da Saúde têm vindo a fazer relativamente ao sentimento de “fadiga pandémica” cada vez mais presente na população. Segundo a OMS, a fadiga da pandemia

refere-se a um sentimento de sobrecarga e de cansaço, por obedecermos a restrições e alterações na nossa vida, estimando que pelo menos 60% da população mundial já padece desta patologia.

A evidência suporta a ideia de que os portugueses responderam mais afirmativamente aos pedidos de permanência na sua habitação durante o *lockdown* de 2020 do que durante o segundo, em 2021, onde o sentimento de relaxamento da população face às medidas adotadas pelo governo aparenta ter sido uma realidade. Mas qual a grandeza desta diferença de comportamentos? Terá sido esta tão significativa que justifique o cenário de “fadiga pandémica” vaticinado pelos especialistas? Para o efeito, testou-se a hipótese de que a propensão a confinar em função do quadro mais restritivo de medidas em 2020 terá sido o dobro da propensão de 2021, isto é, $H_0: \beta_{C1_{t-1}*2020} = 2\beta_{C1_{t-1}*2021}$ contra $H_1: \beta_{C1_{t-1}*2020} \neq 2\beta_{C1_{t-1}*2021}$. Recorrendo à Figura XIII do Anexo A, verifica-se que o valor-p do teste é de 0,243, permitindo não rejeitar H_0 aos níveis de significância habituais, o que quer dizer que estatisticamente não se rejeita a hipótese de que a referida propensão terá sido em 2020 o dobro da registada em 2021. Esta grande discrepância no que diz respeito à adesão da população portuguesa às medidas restritivas entre os dois períodos pode ser consequência do desgaste emocional sentido pela mesma, devido à constante observância de medidas e restrições completamente desenquadradas com o seu habitual estilo de vida pré-pandemia.

A abordagem anterior testa somente mudanças comportamentais associadas a um determinante em particular, a variável *C1_Medidas_Agressivas*. A metodologia desenvolvida por Bai e Perron permite testar múltiplos pontos de quebra de estrutura da série, tendo a possibilidade de fazer variar os coeficientes dos regressores entre os diferentes regimes. Esta abordagem surge com o objetivo de analisar de forma mais completa e detalhada eventuais alterações de resposta dos portugueses, no enquadramento global do modelo COM-B.

No contexto dos testes de Bai-Perron, pretendeu-se especificar um modelo bem mais parcimonioso do que o apresentado na equação (9), apenas com os regressores que personificam os determinantes do modelo COM-B, até para deixar graus de liberdade suficientes para o procedimento do teste estimar todos os parâmetros, incluindo número e pontos de quebra. Com efeito, devido à elevada sazonalidade presente na variável

dependente *confinamento*, expurgaram-se os respetivos efeitos de sazonalidade através da decomposição STL (*Seasonal and Trend decomposition using Loess*), um método de ajustamento sazonal que decompõe a série em três componentes, Sazonal, Tendência e Resíduo, assumindo uma relação aditiva entre as mesmas. Consequentemente, o modelo estimado que servirá de base para a aplicação dos testes pode ser expresso da seguinte forma:

$$Y_t^{sa} = \alpha_t + \beta_t C_{t-1} + \lambda_t X_{t-1} + e_t, t = 1, \dots, 395 \quad (10)$$

onde Y_t^{sa} representa a variável *confinamento*_{*t*} ajustada sazonalmente, sendo que as matrizes C_{t-1} e X_{t-1} representam exatamente as mesmas variáveis; e e_t é o habitual termo de erro. Note que, neste quadro, a colocação dos índices t nos coeficientes α, β e λ permite a sua variação entre os regimes definidos pelos pontos de quebra. É permitida autocorrelação dos erros desde que não seja incluída como regressor a variável dependente desfasada (ver *Assumption A4* em Bai & Perron (1998)).

A estratégia adotada para determinar o número de quebras, como sugerido em Bai & Perron (2003), foi a seguinte: numa primeira instância, a partir dos testes $UDmaxF_T(M, q)$ ou $WDmaxF_T(M, q)$ aferir a existência de pelo menos um momento de alteração de estrutura; depois, se o teste apontar para a presença de pelo menos um desses momentos, então o número de pontos de quebra e as respetivas datas deverão ser estimados pelo procedimento sequencial, $supF_T(l + 1|l)$.

Procedeu-se à estimação prévia do modelo da equação (10) acautelando eventuais questões de autocorrelação com o estimador HAC para a matriz de covariâncias OLS. Estabeleceu-se um máximo de 5 pontos de quebra ($m=5$) e utilizou-se um parâmetro de *trimming* igual a 0,15 ($\varepsilon = 0,15$) para um nível de significância de 5%. Permitiu-se que os coeficientes dos seis regressores pudessem variar entre regimes e ainda a existência de distribuições heterogêneas dos erros entre segmentos. Os resultados são apresentados nas Figuras XIII e XIV do Anexo A.

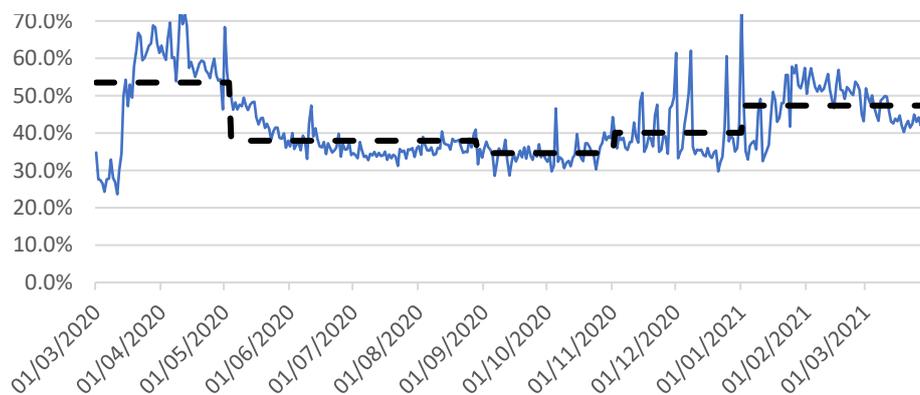
Os resultados dos testes *double maximum* permitem rejeitar a hipótese nula de inexistência de quebras a favor da hipótese de um número desconhecido de quebras dado um limite superior de 5, tendo sido estimados pelas estatísticas $UDmaxF_T(5, q)$ e $WDmaxF_T(5, q)$ os seguintes números de quebras $\hat{m} = 1$ e $\hat{m} = 5$, respetivamente. Determinada a evidência a favor da existência de pelo menos um momento de mudança

estrutural de comportamentos da população portuguesa, os testes $supF_T(l+1|l)$ são importantes para a estimação do número e respetivas datas de quebra. Os mesmos foram estatisticamente significativos para $l=0,1,2,3$, sendo que a não rejeição de H_0 verificou-se quando $H_0: l = 4$. Consequentemente, o número de pontos de quebra e as respetivas datas estimadas pelo teste sequencial foi $\hat{m} = 4$; $\hat{T}_1 = 04/05/2020$, $\hat{T}_2 = 29/08/2020$, $\hat{T}_3 = 02/11/2020$ e $\hat{T}_4 = 02/01/2021$.

O primeiro desses momentos ($\hat{T}_1 = 04/05/2020$) prender-se-á com o início do primeiro processo de desconfinamento, depois de um longo período com níveis elevadíssimos de confinamento da população; o segundo ($\hat{T}_2 = 29/08/2020$) terá como base uma primeira tentativa de conciliação e de adaptação, após o período de férias, entre o regresso a uma vida ativa com relativa normalidade, sobretudo com a suspensão da obrigatoriedade do teletrabalho e com o início da atividade escolar presencial, e um cenário de pandemia ainda existente com as restrições que isso acarreta; o terceiro ($\hat{T}_3 = 02/11/2020$) terá tido como principal impulsionador a segunda vaga da pandemia, aliada a um quadro normativo intermédio de restrições; por fim, o quarto momento ($\hat{T}_4 = 02/01/2021$) coincide com o escalar de casos confirmados da doença e com a entrada na terceira e mais alarmante vaga da pandemia até ao momento no país, no que diz respeito aos indicadores de mortes e doentes em cuidados intensivos provocados pela Covid-19.

A Figura IV permite analisar graficamente os pontos de quebra estimados na série do Confinamento Diário com sazonalidade ajustada. É por demais evidente que as alterações estruturais da resposta dos portugueses desde o início do primeiro processo de desconfinamento (inícios de maio de 2020) até ao início da segunda vaga da pandemia (inícios de novembro de 2020) foram no sentido de uma maior libertação e de tentativa de regresso à “sua normalidade”, comprovada pela redução do nível dos regimes. A partir daí e até ao final da amostra, o ajustamento comportamental foi no sentido de uma maior contração dos comportamentos de mobilidade, motivado sobretudo pela escalada de casos do início de 2021 e pela entrada num novo período de fortes restrições.

Figura IV. Representação Gráfica dos pontos de quebra estimados da série do Confinamento Diário (sazonalidade ajustada)



Fonte: Elaboração Própria.

Em suma, os resultados fornecidos pelos testes de Bai e Perron permitem confirmar a ideia de que ao longo dos 13 meses em análise os comportamentos dos portugueses foram alvo de algumas alterações de estrutura no quadro do modelo COM-B. A evidência estatística direciona para a existência de quatro momentos em que, em função das circunstâncias específicas de cada um, a população foi levada a ajustar as suas tomadas de decisão.

6. CONCLUSÕES, LIMITAÇÕES E INVESTIGAÇÃO FUTURA

O presente trabalho procura fornecer um contributo relevante na investigação que se centra na questão comportamental dos indivíduos, na medida em que aborda a temática explorando-a numa dimensão temporal, algo que não tem sido frequentemente observado no campo das ciências comportamentais.

A incidência reportada da doença e o indicador da atividade económica revelaram-se fatores preponderantes na dinâmica da adoção de comportamentos preventivos da população no decurso da pandemia em Portugal. Enquanto a primeira teve um papel disseminador dos contactos sociais em virtude da maior perceção ao risco provocada no seio da população, um contexto económico menos motivador despoletou nos indivíduos uma menor predisposição para a vida ativa e para a livre circulação, o que implicou uma diminuição da mobilidade dos cidadãos.

No que concerne à dimensão enfoque e que personifica um dos objetivos deste estudo, foi possível distinguir os vários impactos da comunicação do Governo nos comportamentos dos portugueses ao longo do período de gestão da pandemia. Uma

comunicação assente em fortes medidas restritivas, tais como a suspensão da atividade letiva presencial e o encerramento dos centros comerciais e da restauração, parece ter sido eficaz na transmissão da mensagem pretendida.

No entanto, existiram algumas tomadas de decisão por parte do Governo que estiveram longe de ser bem-sucedidas. Neste sentido, destaca-se a ineficácia da adoção de medidas como a proibição de circulação entre concelhos e a redução do horário de encerramento dos supermercados para as 20 horas nos dias úteis e 17 horas nos fins de semana, que não terá sido corretamente percecionado pela população, respondendo contrariamente às solicitações das autoridades. As restrições ajustadas à heterogeneidade da situação sanitária de cada concelho, como a imposição do recolhimento noturno e a redução do horário de encerramento do comércio a retalho para as 22 hora, não produziram efeitos significativos.

A monitorização diária do comportamento dos portugueses permitiu concluir que a propensão a confinar em função dos apelos mais veementes das autoridades durante o primeiro confinamento, em 2020, foi sobejamente superior relativamente ao segundo confinamento, em 2021. Daqui decorre que o desgaste provocado por todo o contexto pandémico terá resultado numa adesão com menos afinco aos pedidos de permanência no domicílio do primeiro para o segundo *lockdown*. Este resultado é consistente com o cenário de fadiga pandémica traçado por diversos especialistas e organizações de saúde, que têm alertado para uma eventual displicência no que à adoção de comportamentos de proteção individual diz respeito com o avançar da pandemia.

Nunca a sociedade como a conhecemos tinha experienciado viver num contexto de uma pandemia disseminada à escala mundial, pelo que as populações de todo o mundo têm se adaptado diariamente aos diversos constrangimentos daí decorrentes. No caso português, este estudo permitiu concluir que a resposta comportamental esteve longe de ser uniforme ao longo do tempo. A evidência reunida corrobora a ideia de que a população portuguesa foi ajustando os seus comportamentos em conformidade com a evolução da conjuntura sanitária, económica e normativa. Em particular, os testes de Bai e Perron identificaram quatro momentos-chave que justificaram uma mudança estrutural de comportamentos dos portugueses.

Algumas limitações enfrentadas pelo presente estudo prendem-se, essencialmente, pela dificuldade de obtenção de outro tipo de variáveis relacionadas com a dimensão da comunicação. A presença exclusiva de variáveis de natureza qualitativa que mensuram os estímulos comunicacionais encurtou o leque de possibilidades de interpretação deste determinante. Seria interessante a inclusão de variáveis que atribuíssem a este fenómeno uma dimensão mais quantitativa, como por exemplo, o número diário de notícias veiculadas na comunicação social relacionadas com a doença, que complementaria esta vertente com outra fonte de informação muito relevante no decurso desta pandemia: os órgãos de comunicação social, principalmente os noticiários das televisões.

Finalmente, em termos de investigação futura destaco o papel que a Econometria pode assumir como ferramenta cada vez mais importante para complementar as análises comportamentais tradicionais na medida em que, como se conclui deste trabalho, a riqueza de informação extraída pode não somente auxiliar na tomada de decisões por parte das entidades de saúde competentes, como ser caracterizadora das dinâmicas comportamentais de uma população ao longo do tempo. Uma abordagem que possibilitaria uma melhor compreensão do fenómeno, seria analisar eventuais diferenças de resposta entre os vários segmentos populacionais, nomeadamente, em função do género, da idade e do local de residência. Tendo em conta o quadro conceptual e a metodologia adotados, os mesmos poderão ser aplicados às respostas comportamentais dos diferentes subgrupos.

REFERÊNCIAS BIBLIOGRÁFICAS

- Anwar, A., Malik, M., Raees, V., & Anwar, A. (2020). Role of Mass Media and Public Health Communications in the COVID-19 Pandemic. *Cureus*, 12(9). <https://doi.org/10.7759/cureus.10453>
- Azevedo, A. B., Gomes, C. S., Mendes, M. F., Baptista, M. I., & Cabral, M. V. (2011). *Demográficas E Envelhecimen O*.
- Bai, J., & Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66(1), 47. <https://doi.org/10.2307/2998540>
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22. <https://doi.org/10.1002/jae.659>
- Bekalu, M. A., & Eggermont, S. (2014). Media use and HIV/AIDS knowledge: A knowledge gap perspective. *Health Promotion International*, 29(4), 739–750. <https://doi.org/10.1093/heapro/dat030>
- Bertrand, J. T., O'Reilly, K., Denison, J., Anhang, R., & Sweat, M. (2006). Systematic review of the effectiveness of mass communication programs to change HIV/AIDS-related behaviors in developing countries. *Health Education Research*, 21(4), 567–597. <https://doi.org/10.1093/her/cyl036>
- Cabrera, A., Martins, C., & Cunha, I. F. (2020). A cobertura televisiva da pandemia de Covid-19 em Portugal. *Media & Jornalismo*, 20(37), 185–204. https://doi.org/10.14195/2183-5462_37_10
- Chemli, S., Toanoglou, M., & Valeri, M. (2020). Current Issues in Tourism The impact of Covid-19 media coverage on tourist ' s awareness for future travelling. *Current Issues in Tourism*, 0(0), 1–8. <https://doi.org/10.1080/13683500.2020.1846502>
- Engelberg, J. E., & Parsons, C. A. (2011). The Causal Impact of Media in Financial Markets. *Journal of Finance*, 66(1), 67–97. <https://doi.org/10.1111/j.1540-6261.2010.01626.x>
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data - Or tears: An application to educational enrollments in states of India. *Demography*,

- 38(1), 115–132. <https://doi.org/10.2307/3088292>
- Gupta, A., & Reddy B, V. (2020). Importance of effective communication during COVID-19 infodemic. *Journal of Family Medicine and Primary Care*, 6(2), 169–170. <https://doi.org/10.4103/jfmmpc.jfmmpc>
- Gwatkin, D. R., Rutstein, S., Johnson, K., Suliman, E., Wagstaff, A., & Amouzou, A. (2000). Socio-economic differences in health, nutrition, and population within developing countries: an overview. *Nigerian journal of clinical practice*, 10(4), 272–282. <https://doi.org/10.1001/jama.298.16.1943>
- Hair, J., Anderson, R. E., & Tatham, R. L. (1987). *Multivariate Data Analysis with Readings*. Macmillan Publishing Company.
- Hayashi, F. (2000). *Econometrics*. Princeton University Press. Section, 1, 60–69.
- Houweling, T. A. J., Kunst, A. E., & Mackenbach, J. P. (2003). Measuring health inequality among children in developing countries: Does the choice of the indicator of economic status matter? *International Journal for Equity in Health*, 2, 1–12. <https://doi.org/10.1186/1475-9276-2-1>
- Jatau, A. I., Peterson, G. M., Bereznicki, L., Dwan, C., Black, J. A., Bezabhe, W. M., & Wimmer, B. C. (2019). *Applying the Capability , Opportunity , and Motivation Behaviour Model (COM-B) to Guide the Development of Interventions to Improve Early Detection of Atrial Fibrillation*. <https://doi.org/10.1177/1179546819885134>
- Kevin, F. J., Maccallum, R. C., & Tait, M. (1986). the Application of Exploratory Factor Analysis in Applied Psychology: a Critical Review and Analysis. *Personnel Psychology*, 39(2), 291–314. <https://doi.org/10.1111/j.1744-6570.1986.tb00583.x>
- Klein, E., Laxminarayan, R., L. Smith, D., & A. Gilligan, C. (2007). Economic incentives and mathematical models of disease. *Environment and Development Economics*, 12, 707–732.
- Lourenço, N., & Rua, A. (2020). *The Daily Economic Indicator: tracking economic activity daily during the lockdown* (Vol. 100). <https://doi.org/10.1016/j.econmod.2021.105500>
- MacKinnon, J. G., & Davidson, R. (2003). *Econometric Theory and Methods* (Vol. 34).

Oxford University Press.

- McDonagh, L. K., Saunders, J. M., Cassell, J., Curtis, T., Bastaki, H., Hartney, T., & Rait, G. (2018). Application of the COM-B model to barriers and facilitators to chlamydia testing in general practice for young people and primary care practitioners: A systematic review. *Implementation Science*, *13*(1). <https://doi.org/10.1186/s13012-018-0821-y>
- Mckenzie, D. J. (2004). *Measuring Inequality with Asset Indicators forthcoming, Journal of Population Economics*.
- Michie, S., M.van Stralen, M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *121*(7), 929–930. <https://doi.org/10.1001/archderm.1985.01660070119033>
- Minujin, A., & Bang, J. H. (2002). Indicadores de inequidad social. Acerca del uso del «índice de bienes» para la distribución de los hogares. *Instituto de Desarrollo Económico Y Social*, *42*(165), 129–146.
- Nunes, C. (2021). Situação Epidemiológica em Portugal. *Percepções Sociais sobre a Covid-19*.
- Oh, S. H., Paek, H.-J., & Hove, T. (2015). Cognitive and emotional dimensions of perceived risk characteristics, genre-specific media effects, and risk perceptions: the case of H1N1 influenza in South Korea. *Asian Journal of Communication*.
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572. <https://doi.org/10.1080/14786440109462720>
- Ramsey. (1969). Tests for specification errors in classical least-squares regression analysis. *Journal of the Royal Statistical Society*.
- Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, *28*(12), 1629–1646. <https://doi.org/10.1177/014616702237645>
- Stock, J. H., & Watson, M. W. (2007). Introduction to Econometrics. Em *Pearson*

Addison-Wesley (2nd ed). <https://doi.org/10.12737/7786>

Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, 21(6), 459–468. <https://doi.org/10.1093/heapol/czl029>

Wakefield, M. A., Loken, B., & Hornik, R. C. (2010). Use of mass media campaigns to change health behaviour. *The Lancet*, 376(9748), 1261–1271. [https://doi.org/10.1016/S0140-6736\(10\)60809-4](https://doi.org/10.1016/S0140-6736(10)60809-4)

William W. S., W. (2005). *Time Series Analysis-Univariate and Multivariate Methods* (2nd Editio).

Willmott, T. J., Pang, B., & Rundle-Thiele, S. (2021). Capability, opportunity, and motivation: an across contexts empirical examination of the COM-B model. *BMC Public Health*, 21(1), 1–17. <https://doi.org/10.1186/s12889-021-11019-w>

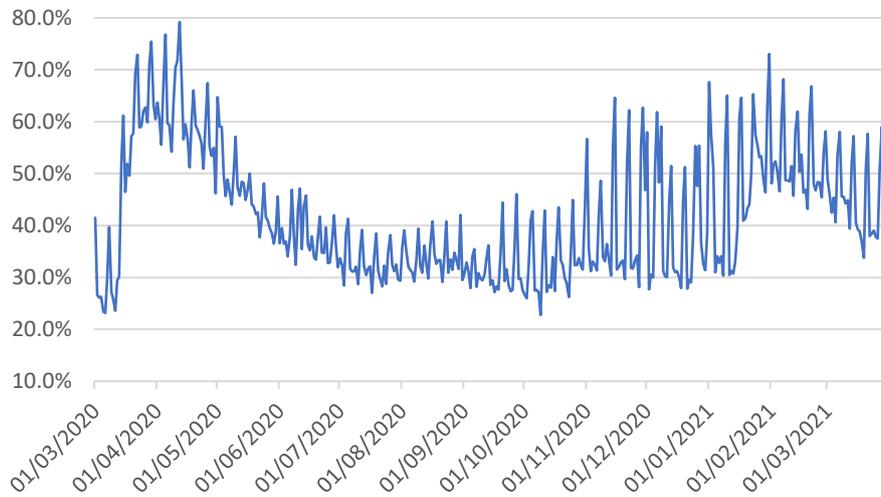
Wooldridge, J. M. (2006). *Introductory Econometrics - A modern approach* (2nd Editio, Vol. 101, Número 476). <https://doi.org/10.1198/jasa.2006.s154>

Zhang, L., Kong, Y., & Chang, H. (2015). Media Use and Health Behavior in H1N1 Flu Crisis: The Mediating Role of Perceived Knowledge and Fear. *Atlantic Journal of Communication*.

Zou, H., Chen, Y., Fang, W., Zhang, Y., & Fan, X. (2017). Identification of factors associated with self-care behaviors using the COM-B model in patients with chronic heart failure. *European Journal of Cardiovascular Nursing*, 16(6), 530–538. <https://doi.org/10.1177/1474515117695722>

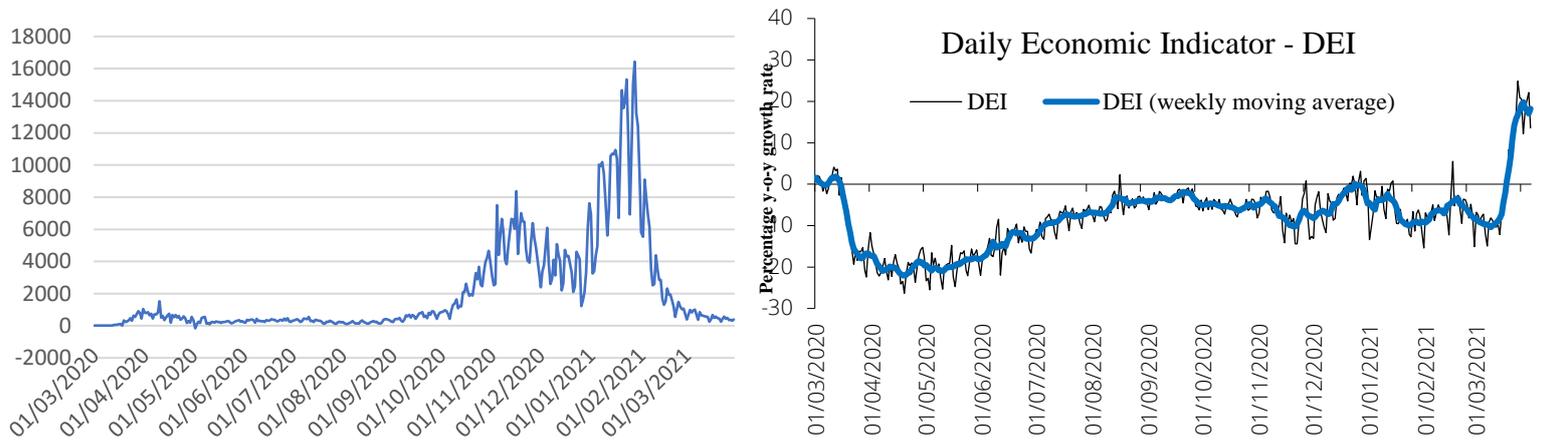
ANEXO A – FIGURAS

Figura V. Série do Confinamento Diário: % da População confinada



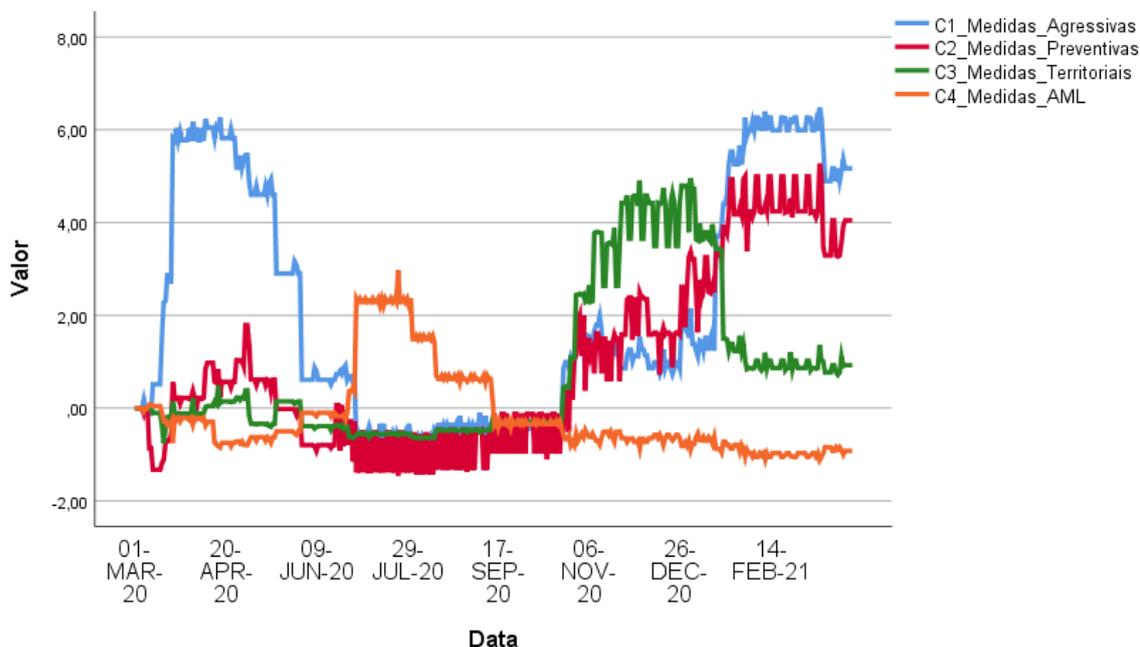
Fonte: Painel Mobilidade PSE

Figura VI. Série do Número de Novos Casos Reportados de Covid-19 e do DEI - Daily Economic Indicator



Fonte: Casos Covid-19 - <https://ourworldindata.org/coronavirus>
Dei – Lourenço & Rua, 2020

Figura VII. Representação Gráfica das Componentes obtidas e retidas pela ACP



Fonte: Elaboração própria. Software: SPSS Statistics

Figura VIII. Estimação do modelo ADL(7,7)

Dependent Variable: CONFINAMENTO
 Method: Least Squares
 Sample (adjusted): 3/08/2020 3/31/2021
 Included observations: 389 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.189396	0.019408	9.758597	0.0000
CONFINAMENTO(-1)	0.333927	0.040882	8.168051	0.0000
CONFINAMENTO(-2)	-0.067050	0.039781	-1.685484	0.0927
CONFINAMENTO(-7)	0.332518	0.036878	9.016796	0.0000
SEGUNDA	-0.103078	0.009760	-10.56118	0.0000
TERCA	-0.069012	0.011796	-5.850326	0.0000
QUARTA	-0.070030	0.009396	-7.453258	0.0000
QUINTA	-0.069291	0.009503	-7.291712	0.0000
SEXTA	-0.088375	0.009721	-9.090640	0.0000
SABADO	-0.010468	0.008584	-1.219480	0.2234
FERIADO	0.161568	0.012569	12.85468	0.0000
FERIADO*FIM_SEMANA	-0.125952	0.023294	-5.407168	0.0000
ANTECIPACAO	0.110629	0.017921	6.173121	0.0000
C1_MEDIDAS_AGRESSIVAS(-1)	0.023290	0.002878	8.091495	0.0000
C2_MEDIDAS_PREVENTIVAS(-1)	-0.016583	0.002953	-5.614809	0.0000
C3_MEDIDAS_TERRITORIAIS(-1)	0.003769	0.001927	1.956284	0.0512
C4_MEDIDAS_AML(-1)	0.007895	0.003202	2.465205	0.0141
X_NOVOS_CASOS_REP(-1)	4.96E-06	9.23E-07	5.377193	0.0000
X_DEI_MA(-1)	-0.001587	0.000655	-2.424477	0.0158
X_DEI_MA(-7)	0.003273	0.000693	4.724474	0.0000
R-squared	0.906584	Mean dependent var	0.422224	
Adjusted R-squared	0.901774	S.D. dependent var	0.121126	
S.E. of regression	0.037962	Akaike info criterion	-3.654415	
Sum squared resid	0.531772	Schwarz criterion	-3.450632	
Log likelihood	730.7836	Hannan-Quinn criter.	-3.573626	
F-statistic	188.4773	Durbin-Watson stat	1.915549	
Prob(F-statistic)	0.000000			

Fonte: Elaboração própria. Software: EViews.

Figura IX. Teste LM à autocorrelação dos erros de 1ª ordem ao modelo ADL(7,7)

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.316682	Prob. F(1,368)	0.2519
Obs*R-squared	1.386857	Prob. Chi-Square(1)	0.2389

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Sample: 3/08/2020 3/31/2021

Included observations: 389

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.009378	0.021051	0.445484	0.6562
CONFINAMENTO(-1)	-0.049086	0.059159	-0.829722	0.4072
CONFINAMENTO(-2)	0.025012	0.045346	0.551578	0.5816
CONFINAMENTO(-7)	0.004986	0.037117	0.134335	0.8932
SEGUNDA	0.001088	0.009802	0.110989	0.9117
TERCA	-0.006897	0.013235	-0.521108	0.6026
QUARTA	-0.003749	0.009944	-0.376970	0.7064
QUINTA	-0.004126	0.010157	-0.406286	0.6848
SEXTA	-0.003745	0.010251	-0.365344	0.7151
SABADO	-0.004796	0.009544	-0.502460	0.6156
FERIADO	-0.000470	0.012570	-0.037419	0.9702
FERIADO*FIM_SEMANA	-0.001317	0.023312	-0.056491	0.9550
ANTECIPACAO	0.002262	0.018022	0.125527	0.9002
C1_MEDIDAS_AGRESSIVAS(-1)	0.000763	0.002953	0.258248	0.7964
C2_MEDIDAS_PREVENTIVAS(-1)	-0.000303	0.002964	-0.102108	0.9187
C3_MEDIDAS_TERRITORIAIS(-1)	0.000174	0.001932	0.090281	0.9281
C4_MEDIDAS_AML(-1)	0.000191	0.003205	0.059604	0.9525
X_NOVOS_CASOS_REP(-1)	9.81E-08	9.27E-07	0.105877	0.9157
X_DEI_MA(-1)	-9.64E-05	0.000660	-0.146063	0.8840
X_DEI_MA(-7)	7.68E-05	0.000696	0.110455	0.9121
RESID(-1)	0.087854	0.076564	1.147468	0.2519
R-squared	0.003565	Mean dependent var	1.37E-16	
Adjusted R-squared	-0.050589	S.D. dependent var	0.037021	
S.E. of regression	0.037946	Akaike info criterion	-3.652845	
Sum squared resid	0.529876	Schwarz criterion	-3.438873	
Log likelihood	731.4783	Hannan-Quinn criter.	-3.568017	
F-statistic	0.065834	Durbin-Watson stat	1.990455	
Prob(F-statistic)	1.000000			

Fonte: Elaboração própria. Software: EViews.

Figura X. Teste LM à autocorrelação até à ordem 7 dos erros do modelo ADL(7,7)

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.512046	Prob. F(7,362)	0.1618
Obs*R-squared	11.05066	Prob. Chi-Square(7)	0.1364

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Sample: 3/08/2020 3/31/2021

Included observations: 389

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.008958	0.023797	0.376452	0.7068
CONFINAMENTO(-1)	-0.023343	0.062073	-0.376062	0.7071
CONFINAMENTO(-2)	0.078203	0.057698	1.355384	0.1761
CONFINAMENTO(-7)	-0.051992	0.047646	-1.091202	0.2759
SEGUNDA	-0.013202	0.011157	-1.183318	0.2375
TERCA	-0.021205	0.014726	-1.440035	0.1507
QUARTA	-0.011459	0.010501	-1.091253	0.2759
QUINTA	-0.010989	0.010636	-1.033228	0.3022
SEXTA	-0.010645	0.010807	-0.984974	0.3253
SABADO	-0.006106	0.009608	-0.635514	0.5255
FERIADO	-0.000239	0.012785	-0.018683	0.9851
FERIADO*FIM_SEMANA	-2.82E-05	0.023319	-0.001207	0.9990
ANTECIPACAO	-0.001904	0.018070	-0.105366	0.9161
C1_MEDIDAS_AGRESSIVAS(-1)	-0.000273	0.003243	-0.084174	0.9330
C2_MEDIDAS_PREVENTIVAS(-1)	0.000195	0.003008	0.064724	0.9484
C3_MEDIDAS_TERRITORIAIS(-1)	-3.71E-05	0.001979	-0.018723	0.9851
C4_MEDIDAS_AML(-1)	-0.000170	0.003210	-0.052880	0.9579
X_NOVOS_CASOS_REP(-1)	-4.70E-08	9.31E-07	-0.050519	0.9597
X_DEI_MA(-1)	0.000239	0.000675	0.354508	0.7232
X_DEI_MA(-7)	-0.000329	0.000730	-0.450811	0.6524
RESID(-1)	0.066422	0.078244	0.848907	0.3965
RESID(-2)	-0.111953	0.068183	-1.641958	0.1015
RESID(-3)	-0.052618	0.057779	-0.910689	0.3631
RESID(-4)	-0.075443	0.053702	-1.404857	0.1609
RESID(-5)	0.006272	0.053933	0.116301	0.9075
RESID(-6)	0.023179	0.053578	0.432616	0.6656
RESID(-7)	0.126255	0.068121	1.853376	0.0646
R-squared	0.028408	Mean dependent var	1.37E-16	
Adjusted R-squared	-0.041375	S.D. dependent var	0.037021	
S.E. of regression	0.037779	Akaike info criterion	-3.647244	
Sum squared resid	0.516666	Schwarz criterion	-3.372137	
Log likelihood	736.3890	Hannan-Quinn criter.	-3.538179	
F-statistic	0.407089	Durbin-Watson stat	2.003739	
Prob(F-statistic)	0.996236			

Fonte: Elaboração própria. Software: EViews.

Figura XI. Teste RESET ao modelo ADL(7,7)

Ramsey RESET Test
Equation: UNTITLED
Specification: CONFINAMENTO C CONFINAMENTO(-1)
CONFINAMENTO(-2) CONFINAMENTO(-7) SEGUNDA TERCA
QUARTA QUINTA SEXTA SABADO FERIADO FERIADO
*FIM_SEMANA ANTECIPACAO C1_MEDIDAS_AGRESSIVAS(-1)
C2_MEDIDAS_PREVENTIVAS(-1) C3_MEDIDAS_TERRITORIAIS
(-1) C4_MEDIDAS_AML(-1) X_NOVOS_CASOS_REP(-1)
X_DEI_MA(-1) X_DEI_MA(-7)
Omitted Variables: Squares of fitted values

	Value	df	Probability
t-statistic	1.849476	368	0.0652
F-statistic	3.420561	(1, 368)	0.0652
Likelihood ratio	3.599055	1	0.0578

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	0.004897	1	0.004897
Restricted SSR	0.531772	369	0.001441
Unrestricted SSR	0.526875	368	0.001432

LR test summary:

	Value
Restricted LogL	730.7836
Unrestricted LogL	732.5832

Unrestricted Test Equation:
Dependent Variable: CONFINAMENTO
Method: Least Squares
Sample: 3/08/2020 3/31/2021
Included observations: 389

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.202470	0.020596	9.830629	0.0000
CONFINAMENTO(-1)	0.239411	0.065361	3.662913	0.0003
CONFINAMENTO(-2)	-0.051893	0.040489	-1.281657	0.2006
CONFINAMENTO(-7)	0.234338	0.064569	3.629255	0.0003
SEGUNDA	-0.076146	0.017513	-4.348101	0.0000
TERCA	-0.051886	0.014966	-3.466949	0.0006
QUARTA	-0.053326	0.013011	-4.098605	0.0001
QUINTA	-0.052470	0.013131	-3.995808	0.0001
SEXTA	-0.068720	0.014381	-4.778425	0.0000
SABADO	-0.006838	0.008778	-0.778930	0.4366
FERIADO	0.119407	0.026012	4.590444	0.0000
FERIADO*FIM_SEMANA	-0.098562	0.027539	-3.579047	0.0004
ANTECIPACAO	0.087069	0.021940	3.968600	0.0001
C1_MEDIDAS_AGRESSIVAS(-1)	0.017539	0.004231	4.145703	0.0000
C2_MEDIDAS_PREVENTIVAS(-1)	-0.011428	0.004054	-2.818950	0.0051
C3_MEDIDAS_TERRITORIAIS(-1)	0.002957	0.001970	1.500977	0.1342
C4_MEDIDAS_AML(-1)	0.005540	0.003437	1.611986	0.1076
X_NOVOS_CASOS_REP(-1)	3.58E-06	1.19E-06	3.014348	0.0026
X_DEI_MA(-1)	-0.001171	0.000690	-1.697005	0.0906
X_DEI_MA(-7)	0.002233	0.000891	2.507062	0.0126
FITTED^2	0.283706	0.153398	1.849476	0.0652

R-squared	0.907444	Mean dependent var	0.422224
Adjusted R-squared	0.902414	S.D. dependent var	0.121126
S.E. of regression	0.037838	Akaike info criterion	-3.658526
Sum squared resid	0.526875	Schwarz criterion	-3.444553
Log likelihood	732.5832	Hannan-Quinn criter.	-3.573697
F-statistic	180.3990	Durbin-Watson stat	1.909870
Prob(F-statistic)	0.000000		

Fonte: Elaboração própria. Software: EViews.

Figura XII. Teste de hipóteses: $H_0: \beta_{C1_{t-1} * 2020} = \beta_{C1_{t-1} * 2021}$ vs. $H_1: \beta_{C1_{t-1} * 2020} > \beta_{C1_{t-1} * 2021}$

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
t-statistic	2.179012	368	0.0300
F-statistic	4.748093	(1, 368)	0.0300
Chi-square	4.748093	1	0.0293

Null Hypothesis: C(14)=C(15)
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(14) - C(15)	0.007734	0.003549

Restrictions are linear in coefficients.

Fonte: Elaboração própria. Software: EViews.

Figura XIII. Teste de hipóteses: $H_0: \beta_{C1_{t-1}*2020} = 2\beta_{C1_{t-1}*2021}$ vs. $H_1: \beta_{C1_{t-1}*2020} \neq 2\beta_{C1_{t-1}*2021}$

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
t-statistic	-1.168575	368	0.2433
F-statistic	1.365567	(1, 368)	0.2433
Chi-square	1.365567	1	0.2426

Null Hypothesis: C(14)=2*C(15)
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(14) - 2*C(15)	-0.008573	0.007336

Restrictions are linear in coefficients.

Fonte: Elaboração própria. Software: EViews.

Figura XIV. Teste *double maximum* de Bai e Perron

Multiple breakpoint tests
Bai-Perron tests of 1 to M globally determined breaks
Sample: 3/01/2020 3/31/2021
Included observations: 395
Breaking variables: C C1_MEDIDAS_AGRESSIVAS(-1)
C2_MEDIDAS_PREVENTIVAS(-1) C3_MEDIDAS_TERRITORIAIS(-1)
C4_MEDIDAS_AML(-1) X_NOVOS_CASOS_REP(-1)
X_DEI_MA(-1)
Break test options: Trimming 0.15, Max. breaks 5, Sig. level 0.05
Test statistics employ HAC covariances (Bartlett kernel, Newey-West automatic bandwidth, NW automatic lag length)
Allow heterogeneous error distributions across breaks

Sequential F-statistic determined breaks:	5
Significant F-statistic largest breaks:	5
UDmax determined breaks:	1
WDmax determined breaks:	5

Breaks	F-statistic	Scaled F-statistic	Weighted F-statistic	Critical Value
1 *	13.80234	96.61638	96.61638	21.87
2 *	11.99018	83.93128	96.71113	18.98
3 *	9.007008	63.04906	80.02803	17.23
4 *	9.436682	66.05678	92.90429	15.55
5 *	9.794228	68.55959	111.8954	13.40

UDMax statistic*	96.61638	UDMax critical value**	22.04
WDMax statistic*	111.8954	WDMax critical value**	23.81

* Significant at the 0.05 level.
** Bai-Perron (Econometric Journal, 2003) critical values.

Estimated break dates:
1: 5/04/2020
2: 5/02/2020, 1/03/2021
3: 5/04/2020, 10/31/2020, 1/02/2021
4: 5/04/2020, 9/05/2020, 11/03/2020, 1/02/2021
5: 5/04/2020, 7/02/2020, 9/05/2020, 11/03/2020, 1/02/2021

Fonte: Elaboração própria. Software: EViews.

Figura XV. Teste Sequencial de Bai e Perron

Multiple breakpoint tests
 Bai-Perron tests of L+1 vs. L sequentially determined breaks
 Sample: 3/01/2020 3/31/2021
 Included observations: 395
 Breaking variables: C C1_MEDIDAS_AGRESSIVAS(-1)
 C2_MEDIDAS_PREVENTIVAS(-1) C3_MEDIDAS_TERRI
 TORIAIS(-1) C4_MEDIDAS_AML(-1)
 X_NOVOS_CASOS_REP(-1) X_DEI_MA(-1)
 Break test options: Trimming 0.15, Max. breaks 5, Sig. level
 0.05
 Test statistics employ HAC covariances (Bartlett kernel,
 Newey-West automatic bandwidth, NW automatic lag
 length)
 Allow heterogeneous error distributions across breaks

Sequential F-statistic determined breaks: 4

Break Test	F-statistic	Scaled F-statistic	Critical Value**
0 vs. 1 *	13.80234	96.61638	21.87
1 vs. 2 *	4.531054	31.71738	24.17
2 vs. 3 *	4.503965	31.52776	25.13
3 vs. 4 *	4.222602	29.55822	26.03
4 vs. 5	0.000000	0.000000	26.65

* Significant at the 0.05 level.

** Bai-Perron (Econometric Journal, 2003) critical values.

Break dates:

	Sequential	Repartition
1	5/04/2020	5/04/2020
2	1/03/2021	8/29/2020
3	10/31/2020	11/02/2020
4	8/29/2020	1/02/2021

Fonte: Elaboração própria. Software: EViews.

ANEXO B – TABELAS

Tabela I. Caracterização das variáveis que personificam as dimensões Comportamento, Motivação e Oportunidade utilizadas no modelo

Variável	Descrição	Dimensão do Modelo COM-B
confinamento	Percentagem da população que permanece confinada em cada dia	Comportamento
X_novos_casos_rep	Número de novos casos reportados pela DGS em cada dia	Motivação
X_dei_ma	Média Móvel de 7 dias do DEI – Indicador Económico Diário	Oportunidade

Fonte: Elaboração própria.

Tabela II. Caracterização das variáveis que personificam a dimensão Capacidade utilizadas no modelo

Componente	Variável	Dimensão
Componente 1	C1_Medidas_agressivas	Capacidade
Componente 2	C2_Medidas_Preventivas	Capacidade
Componente 3	C3_Medidas_Territoriais	Capacidade
Componente 4	C4_Medidas_AML	Capacidade

Fonte: Elaboração Própria.

Tabela III. Descrição das variáveis da base de dados relativa à Comunicação

Variável	Descrição
X_encerramento_Centros_Comerciais	igual a 1 nos dias em que esteve em vigor o encerramento dos Centros Comerciais; 0 no caso contrário.
X_suspensao_atividade_letiva_presencial	igual a 1 nos dias em que esteve em vigor a suspensão da atividade letiva presencial; 0 no caso contrário.
X_encerramento_restaurantes_apenas_take_away	igual a 1 nos dias em que esteve em vigor o encerramento dos restaurantes em que a única modalidade disponível era o take away; 0 no caso contrário.
X_dever_geral_recolhimento	igual a 1 nos dias em que esteve em vigor a imposição do dever geral de recolhimento; 0 no caso contrário.
X_controlo_fronteras_terrestres	igual a 1 nos dias em que esteve em vigor o controlo de fronteiras terrestres; 0 no caso contrário.
X_teletrabalho_obrigatorio	igual a 1 nos dias em que esteve em vigor o teletrabalho obrigatório; 0 no caso contrário.
X_Transportes_sem_custos	igual a 1 nos dias em que os transportes públicos na AML eram gratuitos; 0 no caso contrário.
X_Estado_Emergência	igual a 1 nos dias em que esteve em vigor o Estado de Emergência; 0 no caso contrário.
X_fixação_intervalo_encerramento_supermercados_20h_23h	igual a 1 nos dias em que esteve em vigor o intervalo obrigatório de encerramento dos supermercados entre as 20h e as 23h; 0 no caso contrário.
X_Periodo_Vacinação	igual a 1 nos dias em que foram administradas vacinas contra a doença; 0 no caso contrário.
X_conferencias_imprensa_DGS	igual a 1 nos dias em que a DGS realizou conferências de imprensa; 0 no caso contrário.
X_encerramento_supermercados_20h_dia_util_17h_fds	igual a 1 nos dias em que os supermercados eram obrigados a fechar às 20h no dia útil e às 17h nos fins de semana; 0 no caso contrário.
X_proibicao_circular_entre_concelhos	igual a 1 nos dias em que estava em vigor a proibição de circular entre concelhos; 0 no caso contrário.
X_obrigatoriedade_uso_mascaras	igual a 1 nos dias em que esteve em vigor a obrigatoriedade do uso de máscaras na rua; 0 no caso contrário.
X_encerramento_comercio_retalho_22h	igual a 1 nos dias em que esteve em vigor o encerramento obrigatório do comércio a retalho até às 22h; 0 no caso contrário.
X_proibicao_circular_23h_5h_dia_util	igual a 1 nos dias em que esteve proibida a circulação de pessoas entre as 23h e as 5h em dias úteis; 0 no caso contrário.
X_discriminação_medidas_por_risco_concelho	igual a 1 nos dias em que as medidas adotadas pelo Governo eram

	diferenciadas pelo risco de incidência de cada concelho; 0 no caso contrário.
X_Estado_Calamidade_só_19_freguesias_AM_Lisboa	igual a 1 nos dias em que esteve em vigor o estado de Calamidade apenas em 19 freguesias da AM de Lisboa; 0 no caso contrário.
X_encerramento_supermercados_AML_22h	igual a 1 nos dias em que os supermercados eram obrigados a encerrar até às 22h apenas na AML; 0 no caso contrário.
X_anuncio_F1	igual a 1 no dia em que foi anunciada a realização da prova de Fórmula 1 em Portugal em 2020; 0 no caso contrário.
X_Ferias_escolares	igual a 1 no dia em que estavam a decorrer as férias escolares; 0 no caso contrário.
X_Estado_Contingência_apenas_AML	igual a 1 nos dias em que estava em vigor o estado de Contingência apenas na AM de Lisboa; 0 no caso contrário.
X_anuncio_Avante	igual a 1 no dia em que foi anunciada a realização do festival Avante em 2020; 0 no caso contrário
X_reunioes_infarmed	igual a 1 nos dias em que se realizaram reuniões no infarmed entre especialistas da área da Saúde e órgãos de soberania; 0 no caso contrário
X_declarações_formais_PR	igual a 1 nos dias em que existiram declarações formais por parte do Presidente da República; 0 no caso contrário
X_anuncio_medidas_verao	igual a 1 nos dias em que foram anunciadas as medidas para o verão de 2020 por parte do governo; 0 no caso contrário
X_anuncio_Liga_dos_campeões	igual a 1 no dia em que foi anunciada a realização da fase final da Liga dos Campeões em Portugal; 0 no caso contrário
X_declarações_formais_governo	igual a 1 no dia em que o Governo realizou declarações formais; 0 no caso contrário
X_celebrações_1_maio	igual a 1 no dia em que foi anunciada a realização das celebrações do 1º de Maio em 2020; 0 no caso contrário
X_cancelamento_santos_populares	igual a 1 no dia em que foi anunciado o cancelamento dos Santos Populares de 2020; 0 no caso contrário.

Fonte: Elaboração própria.

Tabela IV. Estatísticas de diagnóstico da ACP

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.597
Bartlett's Test of Sphericity	Approx. Chi-Square	1454.856
	df	435
	Sig.	.000

Fonte: Elaboração própria. Software: SPSS Statistics.

Tabela V. Variância Total Explicada pelas Componentes Principais

Component	Total Variance Explained								
	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
		% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.958	26.525	26.525	7.958	26.525	26.525	5.597	18.655	18.655
2	4.811	16.036	42.561	4.811	16.036	42.561	4.388	14.627	33.282
3	2.652	8.841	51.401	2.652	8.841	51.401	4.239	14.131	47.414
4	2.204	7.345	58.747	2.204	7.345	58.747	2.521	8.402	55.815
5	1.503	5.010	63.756	1.503	5.010	63.756	1.674	5.581	61.397
6	1.211	4.036	67.792	1.211	4.036	67.792	1.643	5.478	66.874
7	1.185	3.950	71.742	1.185	3.950	71.742	1.240	4.132	71.006
8	1.077	3.591	75.333	1.077	3.591	75.333	1.165	3.882	74.889
9	1.043	3.476	78.809	1.043	3.476	78.809	1.109	3.696	78.585
10	1.006	3.355	82.163	1.006	3.355	82.163	1.073	3.578	82.163
11	.827	2.755	84.919						
12	.707	2.357	87.276						
13	.552	1.841	89.117						
14	.522	1.739	90.856						
15	.452	1.505	92.361						
16	.377	1.256	93.617						
17	.363	1.210	94.827						
18	.268	.892	95.719						
19	.238	.793	96.512						
20	.221	.737	97.250						
21	.178	.594	97.843						
22	.150	.501	98.344						
23	.119	.395	98.739						
24	.104	.346	99.085						
25	.100	.334	99.419						
26	.065	.217	99.636						
27	.056	.186	99.822						
28	.034	.114	99.936						
29	.015	.051	99.986						
30	.004	.014	100.000						

Extraction Method: Principal Component Analysis.

Fonte: Elaboração própria. Software: SPSS Statistics.

Tabela VI. *Loadings* ou pesos de cada variável em cada componente

	Componente			
	C1_Medidas _Agressivas	C2_Medidas _Preventivas	C3_Medidas _Territoriais	C4_Medidas _AML
X_encerramento_Centros_Comerciais	.864	.305	-.229	-.081
X_suspensao_atividade_letiva_presencial	.848	.272	-.270	-.074
X_encerramento_restaurantes_apenas_take_away	.848	.363	-.213	-.051
X_dever_geral_recolhimento	.757	.163	.236	-.126
X_controlo_fronteras_terrestres	.735	.065	-.365	-.176
X_teletrabalho_obrigatorio	.670	.310	.526	-.194
X_Transportes_sem_custos	.637	-.468	-.082	-.032
X_Estado_Emergência	.584	.416	.564	-.090
X_encerramento_supermercados_20h_dia_util_17h_fds	.342	.787	-.139	-.002
X_cancelamento_santos_populares	.281	-.230	-.016	-.024
X_proibicao_circular_entre_concelhos	.276	.758	.158	-.082
X_reunioes_infarmed	.263	-.022	-.053	-.045
X_celebrações_1_maio	.259	-.209	-.023	-.022
X_Periodo_Vacinação	.249	.866	.142	-.043
X_declarações_formais_PR	.216	.274	.343	-.069
X_declarações_formais_governo	.189	-.123	-.017	-.076
X_anuncio_medidas_verao	.152	-.131	-.048	-.043
X_obrigatoriedade_uso_mascaras	.093	.691	.611	-.133
X_anuncio_F1	-.025	-.067	-.022	.614
X_Ferias_escolares	-.047	-.346	-.258	.536
X_anuncio_Avante	-.066	-.043	-.028	-.130
X_anuncio_Liga_dos_campeões	-.098	-.061	-.085	-.081
X_conferencias_imprensa_DGS	-.124	-.861	-.023	.076
X_proibicao_circular_23h_5h_dia_util	-.131	-.047	.928	-.069
X_discriminação_medidas_por_risco_concelho	-.177	.135	.859	-.066
X_encerramento_supermercados_22h	-.181	.077	.933	-.096
X_Estado_Calamidade_só_19_freguesias_AM_Lisboa	-.212	-.078	-.124	.890
X_encerramento_supermercados_AML_22h	-.268	-.099	-.161	.858
X_Estado_Contingência_apenas_AML	-.312	-.120	-.207	.089
X_fixação_intervalo_encerramento_supermercados_20h_23h	-.435	-.121	-.295	-.350

Extraction Method: Principal Component Analysis.
Rotation Method: Quartimax with Kaiser Normalization. ^a

a. Rotation converged in 7 iterations.

Fonte: Elaboração própria. Software: SPSS Statistics

Tabela VII. Matriz de correlações de Pearson das variáveis do modelo

Correlações		Confinament o
C1_Medidas_Agressivas	Correlação de Pearson	,725**
	Sig. (2 extremidades)	,000
	N	396
C2_Medidas_Preventivas	Correlação de Pearson	,436**
	Sig. (2 extremidades)	,000
	N	396
C3_Medidas_Territoriais	Correlação de Pearson	,062
	Sig. (2 extremidades)	,218
	N	396
C4_Medidas_AML	Correlação de Pearson	-,401**
	Sig. (2 extremidades)	,000
	N	396
X_dei_ma	Correlação de Pearson	-,392**
	Sig. (2 extremidades)	,000
	N	396
X_novos_casos_rep	Correlação de Pearson	,183**
	Sig. (2 extremidades)	,000
	N	396

** A correlação é significativa no nível 0,01 (2 extremidades).

Fonte: Elaboração própria. Software: SPSS Statistics