

**MESTRADO EM**  
**MÉTODOS QUANTITATIVOS PARA A DECISÃO**  
**ECONÓMICA E EMPRESARIAL**

**TRABALHO FINAL DE MESTRADO**  
**RELATÓRIO DE ESTÁGIO**

**DETERMINANTES PARA A ADOÇÃO DE TECNOLOGIAS**  
**DE INFORMAÇÃO EM PORTUGAL**

**RICARDO JORGE AZEVEDO PIRES**

**ORIENTAÇÃO:**

**PROFESSORA DOUTORA ESMERALDA RAMALHO**  
**DOUTORA ANA CHUMBAU**

**AGOSTO - 2021**

## RESUMO

Dada a importância das tecnologias de informação e comunicação, este trabalho tem como principal contributo o estudo dos determinantes da probabilidade da adoção de tecnologias de informação, cujo tema tem sido abordado na literatura ao nível empresarial, mas que revela alguma carência em Portugal, sendo que se destaca o facto de terem sido utilizados dados de painel para Portugal, para o período recente de 2010-2018. Os dados utilizados foram obtidos através do Inquérito à Utilização de Tecnologias da Informação e da Comunicação nas Empresas (IUTICE), realizado pelo INE anualmente, e dos Quadros de Pessoal, cujos dados são da responsabilidade do Gabinete de Estratégia e Planeamento (GEP) do Ministério do Trabalho, Solidariedade e Segurança Social (MTSSS). Através de modelos com variáveis de interesse binárias para dados de painel, seguindo uma modelação a duas partes, foram encontrados diversos determinantes para a posse de *website* e para a realização de comércio *online*, sendo que estes determinantes não são necessariamente os mesmos. Os resultados sugerem que características referentes aos recursos humanos, tais como a idade, o género e as habilitações, e referentes à dimensão da empresa e ao setor de atividade económica em que a empresa se encontra inserida têm impacto na adoção das tecnologias de informação por parte das empresas.

CÓDIGOS JEL: C23; C55; M15; O30.

PALAVRAS-CHAVE: Tecnologias de Informação e Comunicação; Dados de Painel; Logit; Probit; Comércio *Online*.

## ABSTRACT

Given the importance of information and communication technology, the main contribution of this work lies in the study of the determinants of the probability of information technology adoption, which is a theme that has been investigated in the literature at business level but represents some shortage of research in Portugal. Panel data for Portugal was used for the period 2010-2018. Specifically, the data were obtained through the Information and Knowledge Society Business Survey (IUTICE), annually done by National Institute of Statistics (INE), and through Staff Establishment Plan, provided by the Strategy and Planning Office (GEP) from the Ministry of Labour, Solidarity and Social Security (MTSSS). Using panel data binary models, following a two-part modelling approach, several determinants were found for both website and e-commerce adoption, in a situation where these drivers are not necessarily the same. Findings suggest that human resources' characteristics, such as age, gender and qualifications, and the firm's dimension and sector of economic activity have an impact on the information technology adoption by firms.

JEL CODES: C23; C55; M15; O30.

KEY WORDS: Information and Communication Technology; Panel Data; Logit; Probit; e-Commerce.

#### AGRADECIMENTOS

Gostaria de agradecer ao Instituto Nacional de Estatística (INE), pela oportunidade de estágio e pela disponibilidade. Agradecer também às minhas orientadoras, Prof. Dr.<sup>a</sup> Esmeralda Ramalho e Dr.<sup>a</sup> Ana Chumbau, pelas suas sugestões importantes e por todo o auxílio prestado ao longo deste processo.

Ao ISEG, por ter sido a minha casa ao longo destes 5 anos e que, sem dúvida, me ajudou a crescer tanto a nível pessoal, como académico e profissional. Independentemente do que o futuro me reserve, terei sempre o orgulho de levar o ISEG ao peito.

Um agradecimento também à minha família, que sempre me possibilitou com tudo, motivou e sempre acreditou em mim, e aos meus amigos, que sempre estiveram ao meu lado, pois sem eles tudo seria mais difícil.

Em especial, agradecer aos meus pais, por serem a minha inspiração, por todos os esforços que fizeram por mim, por todo o apoio incondicional e que, desde sempre, tudo fizeram para que eu conseguisse atingir os meus objetivos tanto pessoais como profissionais.

Por fim, dedico ao meu avô Francisco, por quem tenho o orgulho de considerar um modelo para a minha vida e sei o quanto ele gostaria de estar presente para celebrar este marco comigo.

## ÍNDICE

1.	INTRODUÇÃO.....	1
2.	REVISÃO DE LITERATURA .....	2
3.	METODOLOGIA.....	8
3.1.	MODELOS COM VARIÁVEL DEPENDENTE BINÁRIA PARA DADOS EM PAINEL ....	9
3.2.	MODELO LOGIT DE EFEITOS FIXOS.....	10
3.3.	MODELOS DE EFEITOS ALEATÓRIOS.....	11
3.4.	TESTE DE HAUSMAN.....	13
4.	DADOS .....	14
4.1.	PREPARAÇÃO DOS DADOS.....	14
4.2.	ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS .....	16
5.	RESULTADOS .....	23
5.1.	MODELOS ESTIMADOS.....	24
5.1.1.	Modelos com Website como Variável Binária.....	24
5.1.2.	Modelos com Comércio Online como Variável Binária .....	26
5.2.	ESTIMATIVAS DOS EFEITOS PARCIAIS MÉDIOS .....	29
6.	CONCLUSÕES E SUGESTÕES PARA INVESTIGAÇÕES FUTURAS .....	33
	REFERÊNCIAS BIBLIOGRÁFICAS .....	36
	ANEXO .....	38

## LISTA DE TABELAS

TABELA I - MÉDIA DAS VARIÁVEIS EXPLICATIVAS CONFORME AS VARIÁVEIS DE INTERESSE .....	20
TABELA II - ESTIMATIVAS DOS MODELOS PARA A VARIÁVEL <i>WEBSITE</i> .....	24
TABELA III - ESTIMATIVAS DOS MODELOS PARA A VARIÁVEL <i>COMÉRCIO ONLINE</i> .....	26
TABELA IV - EFEITOS PARCIAIS MÉDIOS.....	29
TABELA A2 - VARIÁVEIS EM ESTUDO.....	38
TABELA A3 - ESTATÍSTICAS DESCRITIVAS DAS VARIÁVEIS EXPLICATIVAS.....	40
TABELA A4 - FREQUÊNCIAS E PERCENTAGENS DAS VARIÁVEIS DE INTERESSE NO TOTAL DE OBSERVAÇÕES.....	41
TABELA A5 - CLASSIFICAÇÃO DOS SETORES DE ATIVIDADE ECONÓMICA DAS EMPRESAS .....	41
TABELA A6 – CLASSIFICAÇÃO DA DIMENSÃO DAS EMPRESAS .....	42
TABELA A9 - NÚMERO DE EMPRESAS CONSIDERADAS POR ANO .....	43
TABELA A10 - NÚMERO DE EMPRESAS CONSIDERADAS POR SETOR DE ATIVIDADE .....	43
TABELA A11 - NÚMERO DE EMPRESAS CONSIDERADAS POR DIMENSÃO .....	43

## LISTA DE FIGURAS

FIGURA 1 - ETAPAS DA MODELAÇÃO A DUAS PARTES. ....	9
FIGURA 2 - PROPORÇÕES DE EMPRESAS POR VARIÁVEIS DE INTERESSE POR ANO. ....	16
FIGURA 3 - PROPORÇÕES DE EMPRESAS POR VARIÁVEIS DE INTERESSE POR SETOR DE ATIVIDADE ECONÓMICA.....	18
FIGURA 4 - PROPORÇÕES DE EMPRESAS POR VARIÁVEIS DE INTERESSE POR DIMENSÃO DA EMPRESA, TENDO EM CONTA O NÚMERO DE PESSOAS AO SERVIÇO. ....	19
FIGURA A1 - FLUXOGRAMA DA JUNÇÃO DAS BASES DE DADOS.....	38
FIGURA A7 - OUTPUT DO COMANDO XTDESCRIBE ANTES DE ELIMINAR OBSERVAÇÕES. 42	
FIGURA A8 - OUTPUT DO COMANDO XTDESCRIBE APÓS ELIMINAR OBSERVAÇÕES.....	42

## 1. INTRODUÇÃO

Segundo o INE (2020), em Portugal aproximadamente 97% das empresas com 10 ou mais pessoas ao serviço e 42.8% das pessoas em serviço utilizam computador com ligação à Internet para fins profissionais, em 2020. No mesmo ano, cerca de 61.1% das empresas referem ter *websites* próprios ou do grupo económico a que pertencem, registando um crescimento de 2.6% relativamente ao ano transato. Muitas das empresas com website disponibilizam a descrição e o preço dos seus produtos ou serviços, referindo ainda os perfis das redes sociais da empresa. Outro fator a ter em conta é que, aproximadamente, 22.9% das empresas empregam especialistas em tecnologias de informação e comunicação. Para além disto, cada vez mais o comércio eletrónico ganha peso na atualidade dos negócios, sendo que em 2019, as vendas de bens e serviços através deste meio representaram aproximadamente 20% do total do volume de negócios das empresas com 10 ou mais pessoas ao serviço, registando assim também um crescimento percentual relativamente a anos anteriores.

O presente relatório enquadra-se no contexto da realização do Trabalho de Final de Mestrado relativo ao Mestrado em Métodos Quantitativos para a Decisão Económica e Empresarial, do Instituto Superior de Economia e Gestão (ISEG). Este relatório vem assim no seguimento de um estágio curricular realizado no Instituto Nacional de Estatística (INE), no qual o principal objetivo foi encontrar os determinantes, em especial relativos aos recursos humanos, para a adoção das tecnologias de informação e comunicação (TIC) por parte das empresas portuguesas.

Este estudo tem como base algumas investigações passadas sobre o tema. Uma primeira, em que Bharadwaj (2000) defende que empresas que apresentam uma elevada capacidade nas TIC tendem a apresentar performances e desempenhos financeiros superiores. Num segundo estudo, Galve-Górriz & Castel (2010) apresentam resultados que sugerem que as TIC estão relacionadas com níveis mais altos de qualificação e formação específica dos trabalhadores. Já num terceiro estudo a partir do qual a investigação deste relatório foi motivada, Oliveira & Martins (2009) destacam vários determinantes que afetam a posse de website e a adoção do comércio *online* nos contextos tecnológico, organizacional e ambiental em que as empresas se encontram. Todas estas

investigações sugerem para estudos futuros a utilização de uma metodologia com dados longitudinais, como a utilização de dados em painel.

De forma a aplicar esta metodologia com dados longitudinais, para esta investigação foram utilizados dados do Inquérito à Utilização de Tecnologias da Informação e da Comunicação nas Empresas (IUTICE), realizado pelo INE anualmente, e dos Quadros de Pessoal, cujos dados são da responsabilidade do Gabinete de Estratégia e Planeamento (GEP) do Ministério do Trabalho, Solidariedade e Segurança Social (MTSSS). Estas duas fontes de dados foram integradas numa única base de dados com dados relativos ao período de 2010 a 2018.

Consideraram-se então modelos logit e probit destinados a dados de painel para modelar primeiramente a posse de *website* como variável de interesse binária e, em seguida e para as empresas que possuem *website*, modelar a realização de comércio *online* como variável de interesse binária, tendo em conta diversas características relativas às tecnologias, aos recursos humanos e às próprias características intrínsecas das empresas. O objetivo principal é apurar que determinantes se revelam estatisticamente significativos para explicar a probabilidade condicional de uma empresa portuguesa possuir *website* próprio ou do grupo e, em seguida e para as empresas que já possuem *website*, explicar a probabilidade condicional de uma empresa portuguesa realizar comércio *online* e ainda estimar os efeitos parciais desses determinantes sobre a probabilidade de interesse.

O restante relatório encontra-se organizado de acordo com a seguinte explicação. No Capítulo 2 é apresentada a revisão de literatura, onde se encontram presentes diversos estudos relativos à utilização das TIC em contexto empresarial. No Capítulo 3 encontram-se apresentados os modelos aplicados na estimação. No Capítulo 4 encontram-se detalhados os dados utilizados para esta investigação. Relativamente ao Capítulo 5, são apresentados e interpretados os resultados a destacar deste estudo. Por fim, no Capítulo 6 são apresentadas as conclusões retiradas da investigação, as limitações encontradas no seu decorrer e ainda algumas sugestões para investigações futuras nesta área.

## 2. REVISÃO DE LITERATURA

As TIC de uma organização são um conceito que pode ter diversas definições. Segundo o Instituto Nacional de Estatística (2020), podem ser incluídas no campo da ciência da computação e da sua utilização prática que tenta classificar, conservar e

disseminar a informação pelas organizações. Em termos práticos, podem ser definidas como a aplicação de sistemas de informação e de conhecimentos nos negócios e na aprendizagem. São os aparelhos de *hardware* e de *software* que formam a estrutura eletrónica de apoio à lógica da informação.

De acordo com Bharadwaj (2000), é perceptível que o conjunto de TIC de uma empresa é um recurso fulcral para obter vantagens competitivas no longo prazo. Isto deve-se, essencialmente, ao facto deste conjunto de infraestruturas permitir às entidades a criação de processos multifuncionais e ainda promover a redução de tempos de ciclos, como por exemplo de reposição de *stocks*, entre outros. É por estes motivos que as TIC determinam a liberdade que uma empresa dispõe para os seus planos de negócios. Estas tecnologias podem ser adquiridas ou produzidas pelas próprias organizações, porém a sua criação exige não só engenho e conhecimento tecnológico, mas também tempo, que é um fator fulcral e determinante para as organizações conseguirem atingir vantagens competitivas em relação aos seus concorrentes diretos.

No que diz respeito aos recursos humanos de uma organização, Bharadwaj (2000) ainda refere que estes podem ser entendidos não só como o conjunto dos seus funcionários, ou seja, o capital humano empregue numa determinada empresa, mas também como o treino, a experiência, as relações e as perceções dos seus empregados. Existem alguns fatores essenciais para o sucesso dos recursos humanos nas TIC, entre eles, as aptidões técnicas dos utilizadores, como por exemplo os conhecimentos e habilidades na programação e análise de dados. Outros fatores importantes para o sucesso da utilização destas tecnologias são a capacidade de gestão para um nível tecnológico, que compreende a gestão e manutenção dos sistemas de informação, a coordenação e interação com os utilizadores do sistema e ainda uma boa capacidade de liderança.

De acordo com a Comissão Europeia (2020b), Portugal encontra-se classificado no 19º lugar de um ranking relativo ao Índice de Digitalidade da Economia e da Sociedade (DESI) de 2020, de entre os 28 países pertencentes à União Europeia. Este índice é um índice composto que resume indicadores como a conectividade, o capital humano, a utilização de serviços Internet, a integração das tecnologias digitais e os serviços públicos digitais. Para além disto, o DESI acompanha a evolução dos Estados-Membros da União Europeia na competitividade digital. Apesar da baixa classificação no ranking, Portugal

tem registado progressos na dimensão do capital humano ao longo dos últimos anos, devido essencialmente a uma melhoria no nível básico de competência digitais e a um crescimento na percentagem de licenciados em TIC no país.

Para um nível estritamente empresarial, a Comissão Europeia (2020a), desenvolveu um Índice de Intensidade Digital (DII), que mede a utilização de diferentes tecnologias digitais em empresas nos diferentes países, sendo que a pontuação do DII varia entre 0 e 12, consoante o número de tecnologias digitais utilizadas. Alguns indicadores para este índice são: a utilização de medidas de segurança para as tecnologias de informação e comunicação; ter a velocidade máxima de download de internet mais rápida de pelo menos 30 Mb/s; a utilização de um *software* de *Enterprise Resource Planning* para partilhar informação; a utilização de redes sociais; a utilização de um *software* de *Customer Relationship Management*; ter mais de 50% de funcionários que usem computadores e internet; ter vendas online (pelo menos 1% do volume de negócios).

No que diz respeito à gestão dos recursos humanos, a introdução das TIC nas organizações desempenhou um papel determinante na alteração do modelo de gestão de recursos humanos tradicional, uma vez que este se revelou ineficiente (Zheng, 2017; Gardner et al, 2003). A inovação e implementação destas tecnologias oferece formas para gerir os negócios que são radicalmente diferentes relativamente às que se utilizavam no passado. E quando estas tecnologias são corretamente implementadas, então regista-se um aumento da eficiência e da competitividade e uma redução dos custos das empresas

Os recursos humanos das organizações também têm vindo a sofrer alterações profundas ao longo do tempo devido essencialmente a um rápido e constante crescimento da utilização das TIC nas empresas. Uma das alterações que os recursos humanos sofreram foi um maior envolvimento em funções e decisões estratégicas mais específicas, uma vez que as tecnologias de informação são capazes de processar grandes quantidades de informação, libertando assim os recursos humanos de tarefas administrativas repetitivas e de rotina (Gardner et al, 2003; Haines & Lafleur, 2008). Outra alteração que os recursos humanos sofreram com a utilização deste tipo de tecnologias foi um aumento da exigência do nível de qualificações (Arvanitis & Hollenstein, 2001; Bartel et al, 2007).

Com a chegada da chamada “era digital”, as empresas investiram bastante em TIC com o objetivo de melhorar a sua performance e os seus resultados. Algumas empresas

conseguiram atingir os objetivos que pretendiam, enquanto outras empresas, apesar dos elevados investimentos neste campo, não conseguiram tirar os benefícios desejados das tecnologias adquiridas. De acordo com Ravichandran & Lertwongsatien (2005), estas diferenças podem ser explicadas através de uma teoria baseada nos recursos, na qual os recursos intangíveis e as capacidades funcionais dos sistemas de informação são fatores críticos para o sucesso da implementação e utilização das tecnologias de informação nas empresas.

Segundo Powell & Dent-Micallef (1997), as TIC, por si só, têm um grande poder na definição da produtividade, porém quando incorretamente utilizadas, não atingem o potencial de rendimento que as organizações pretendem delas. Ou seja, a tecnologia por si só não é suficiente. É assim necessário que existam recursos humanos que complementem este poder da tecnologia, de modo a atingir os níveis de performance pretendidos. Para além disto, é necessário um fator fulcral para o sucesso. Este fator é o tempo para que os recursos humanos assimilem e tenham um melhor conhecimento relativamente às novas TIC que vão surgindo.

Assim, é necessário aprofundar o conhecimento das várias tecnologias deste tipo e ter uma clara noção das características fundamentais das empresas, como por exemplo a importância da dimensão e do setor de atividade económica, bem como das características dos recursos humanos, como por exemplo as suas qualificações e habilitações literárias (Bayo-Moriones & Lera-López, 2007).

Townsend & Bennett (2003) referem que recursos humanos qualificados, bem formados e motivados e uma gestão de recursos humanos de qualidade, quando são auxiliados por novas e mais sofisticadas tecnologias de informação que dispõem de ferramentas que possibilitam as organizações para atingirem os seus objetivos, são uma vantagem competitiva fundamental.

Tornatsky et al (1990) criaram uma estrutura teórica TOE (tecnologia-organização-ambiente), que serviu de base para diversas investigações nesta área, e segundo esta estrutura, existem três fatores que podem influenciar a adoção das TIC. O primeiro fator é o contexto tecnológico da empresa, que inclui as tecnologias internas e externas consideradas relevantes para a empresa. Outro fator é o contexto organizacional da empresa que se refere às suas características e recursos. Por fim, o terceiro fator é o

contexto ambiental da empresa, que engloba a dimensão, tipo e estrutura da indústria em que a empresa se encontra.

Bharadwaj (2000), que tem uma visão baseada nos recursos de tecnologias de informação, defende que as empresas se diferenciam e adquirem vantagens competitivas através de todos os seus recursos de TIC, isto é, das infraestruturas, dos recursos humanos e da sua capacidade para aproveitar os benefícios intangíveis das tecnologias de informação. Assim, desenvolve-se o conceito das TIC como uma capacidade organizacional. Como hipótese para este estudo, defende-se que as empresas que apresentam sucesso na criação de uma capacidade nas TIC tendem a registar um desempenho a nível financeiro superior, quer por aumento das suas receitas, quer por diminuição dos seus custos. No que diz respeito às empresas que investem em tecnologias de informação sem desenvolverem uma forte e estável capacidade nas tecnologias vão estar em desvantagem competitiva. A metodologia aplicada neste estudo foi um método de comparação de amostras emparelhadas, de modo a analisar empiricamente a relação entre o desempenho e a capacidade nas tecnologias de informação. Assim, pode afirmar-se que as empresas que apresentam uma elevada capacidade nas TIC tendem a apresentar performances e desempenhos financeiros superiores. Neste artigo a causalidade é ao contrário do que se pretende investigar nesta tese, que se centra no estudo dos determinantes do uso das TIC ao nível empresarial. De qualquer modo, também no artigo em análise se reconhecia a necessidade de para estudos futuros se basearem numa perspetiva com dados longitudinais para tentar entender como as empresas conseguem desenvolver a sua capacidade nas tecnologias ao longo tempo, de modo a determinar quais são os recursos das tecnologias de informação e comunicação mais importantes para alcançar vantagens competitivas.

Uma outra visão também foi estudada e contribuiu para avanços nesta área, tendo sido utilizada uma metodologia diferente da anterior. Num estudo desenvolvido por Galve-Górriz & Castel (2010) são analisadas as principais características da relação entre as TIC e os recursos humanos nas empresas espanholas. O objetivo deste estudo foi analisar as diferenças entre empresas que investem em TIC quando comparadas com empresas que não fazem investimentos neste tipo de tecnologias. A amostra analisada diz respeito ao ano de 2002 e contém um número de observações relativo a 1269 organizações, sendo que de entre estas empresas, 738 investem em TIC. Assim sendo, os resultados deste

estudo sugerem que as TIC estão relacionadas com níveis mais altos de qualificação e formação específica dos trabalhadores. Para além disto, também se destacou o facto das empresas que investem nestas tecnologias oferecerem níveis de salários médios mais elevados e investirem mais nas áreas de investigação e desenvolvimento. Assim este artigo, apesar de abordar a causalidade no sentido contrário ao que irá ser pretendido para o nosso estudo, é de grande relevo e importância para os gestores das organizações, uma vez que a partir das conclusões retiradas acerca das relações existentes entre as TIC e os recursos humanos, é possível otimizar várias decisões no que dizem respeito aos investimentos nos departamentos de tecnologia, investigação e recursos humanos. Como sugestões para trabalhos futuros nesta área, é sugerido novamente que deve ser abordada uma metodologia que utilize dados longitudinais, de modo a permitir determinar e explicar as relações causais entre características dos recursos humanos e as TIC.

O artigo encontrado na literatura que mais se assemelha à abordagem que se pretende fazer nesta tese pertence a Oliveira & Martins (2009) e aplica-se a Portugal, com o objetivo de explicar os determinantes do uso das TIC. Neste estudo é utilizada a estrutura teórica TOE com o objetivo de analisar os fatores que influenciam os níveis de adoção de *Internet Business Solutions* (IBS) nas empresas portuguesas. Especificamente, foram considerados diferentes níveis para: os determinantes tecnológicos, como por exemplo níveis de propensão à integração de novas tecnologias; para os determinantes organizacionais, tendo como exemplo os níveis de benefícios percebíveis da correspondência eletrónica; e para os determinantes ambientais, por exemplo o nível de pressão sentida por uma empresa provocada pelos seus concorrentes de setor de atividade económica. A amostra utilizada neste estudo foi obtida através do INE, mais especificamente através do IUTICE de 2006, e engloba observações para 2 626 empresas com mais de 9 empregados, sendo esta amostra representativa do setor privado (excluindo o setor financeiro) em janeiro de 2006, em Portugal. A metodologia utilizada foram modelos ordenados probit em duas etapas, que permitem dois tipos de decisões, nos quais são examinados os efeitos de diferentes fatores da estrutura teórica TOE. O processo de modelação foi dividido em duas etapas uma vez que a adoção de IBS (etapa 2) só se verifica nas empresas que possuem um *website* próprio (etapa 1). Após a estimação dos modelos, os resultados indicam que existem oito determinantes relevantes para a posse de *website* e cinco para a adoção de comércio *online*. Os determinantes encontrados para

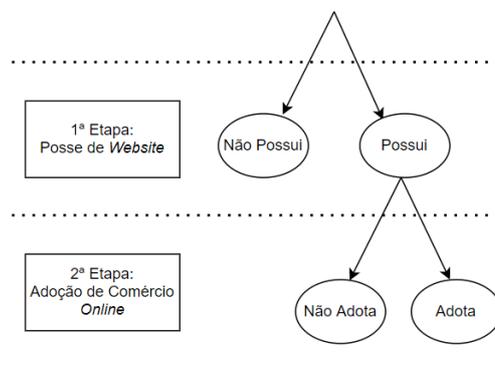
a posse de website são: a propensão para aceitar as tecnologias e a segurança das aplicações, para o nível tecnológico; a dimensão da empresa, a formação em tecnologias de informação (TI), o acesso a sistemas de TI da empresa, as normas da internet e do *e-mail* e os benefícios perceptíveis da correspondência eletrónica, para o nível organizacional; pressão competitiva nos *websites*, para o nível ambiental. Relativamente aos determinantes que influenciam positivamente a adoção do comércio *online*, estes são: a integração tecnológica, no contexto tecnológico; os benefícios perceptíveis da correspondência eletrónica e o acesso a sistemas de TI da empresa, no contexto organizacional; pressão competitiva no comércio *online*, no contexto ambiental. No contexto organizacional, destaca-se o facto de alguns bens e serviços não serem suscetíveis a serem vendidos *online*, sendo este o maior entrave à adoção do comércio *online*. Como sugestão para investigações futuras na área, os autores sugerem a investigação com dados em painel. Dada esta sugestão, este artigo serve como base e inspiração para a investigação apresentada nesta tese. Uma vez que os dados utilizados nesta tese são diferentes, de natureza de painel, a diferença recai na decisão do segundo passo ser binária e não ordinal (sobretudo porque se pretende considerar a possibilidade de testar efeitos aleatórios contra fixos no painel, o que só é possível fazer no âmbito dos modelos Logit, via teste de Hausman, como se pode verificar no Capítulo 3. Assim, utiliza-se uma modelação a duas partes, permitindo a separação da decisão, considerando modelos binários em ambas as partes.

### 3. METODOLOGIA

Os dados disponibilizados pelo INE contêm três variáveis que refletem a adoção das TIC: internet, *website* e comércio *online*. Todas estas variáveis são de natureza binária e, potencialmente, poderiam ter sido consideradas como variáveis de interesse. Acontece que com a estatística descritiva desenvolvida no capítulo seguinte se verificou que a primeira variável, sobretudo nos últimos anos do painel, era 1 para quase todas as empresas da amostra, o que levou a que não fosse considerada como variável dependente na modelação. No que respeita às outras duas variáveis, suspeita-se de uma decisão a duas partes, pois a empresa numa primeira fase decide ter ou não um *website* e, numa segunda fase, no caso de possuir *website*, avança ou não para o comércio *online*. Assim, nesta tese é adotada uma metodologia quantitativa a duas partes, baseada numa análise econométrica do tipo da utilizada por Oliveira & Martins (2009), mas utilizando uma

variável binária na segunda fase. Esta modelação a duas partes pode ser observada na Figura 1.

FIGURA 1 - ETAPAS DA MODELAÇÃO A DUAS PARTES.



Dado que a proporção de empresas que utilizam Internet é bastante elevada ao longo dos anos e que esta variável é decerto determinante da posse de *website* e da adoção do comércio *online*, a variável Internet foi considerada como variável potencialmente explicativa (embora a sua pouca variabilidade levasse desde logo a suspeitar que poderia não se revelar estatisticamente significativa). Assim, o principal objetivo da metodologia aplicada é estimar a probabilidade de uma empresa possuir *website* e, em seguida, de entre as empresas que possuem *website*, estimar a probabilidade de realizar comércio online.

O estudo de variáveis dependentes discretas e a sua utilização com dados de painel tende a complicar a estimação de modelos econométricos. Este facto pode ser justificado por no caso dos dados de painel não se poder habitualmente afirmar que diferentes observações para a mesma unidade amostral são independentes. As correlações existentes entre os termos de erro ao longo do tempo complicam as funções de verosimilhança dos modelos e, desta forma, complicam também a estimação dos mesmos (Verbeek, 2004).

### 3.1. MODELOS COM VARIÁVEL DEPENDENTE BINÁRIA PARA DADOS EM PAINEL

Os modelos com variável de interesse binária são habitualmente escritos em termos de uma variável latente, apresentando-se da seguinte forma:

$$(1) \quad y_{it}^* = x_{it}'\beta + \alpha_i + \varepsilon_{it},$$

onde  $x_{it}'$  é o vetor das variáveis explicativas de dimensão  $K$ , que não inclui termo constante,  $\beta$  é o vetor dos coeficientes das variáveis explicativas que se pretende estimar,

$\alpha_i$  é o termo dos efeitos individuais, que não variam com o tempo nem são observáveis, e  $\varepsilon_{it}$  é o termo erro idiossincrático. Da equação (1) observamos  $y_{it} = 1$  se  $y_{it}^* > 0$  e  $y_{it} = 0$  caso contrário.

A presença de  $\alpha_i$  complica a estimação, quer no caso em que estes efeitos individuais sejam tratados como parâmetros desconhecidos fixos ou como termos de erro aleatórios. Para ultrapassar estes problemas pode ser utilizado o método da máxima verosimilhança condicional, onde se considera uma função de verosimilhança condicional para um conjunto de estatísticas  $t_i$  que são suficientes para a variável aleatória  $\alpha_i$ . Isto significa que condicionar a contribuição da verosimilhança individual em  $t_i$  faz com que esta não dependa mais de  $\alpha_i$ , mas continue a depender de  $\beta$ . A função de densidade conjunta de  $y_{i1}, \dots, y_{iT}$  pode ser escrita como  $f(y_{i1}, \dots, y_{iT} | \alpha_i, \beta)$ , que depende de  $\beta$  e  $\alpha_i$ . Se existir uma estatística suficiente  $t_i$ , isto significa que existe uma variável observada  $t_i$  em que  $f(y_{i1}, \dots, y_{iT} | \alpha_i, \beta) = f(y_{i1}, \dots, y_{iT} | t_i, \beta)$ , sendo que assim esta função não depende de  $\alpha_i$ . Este facto tem como consequência podermos maximizar a função de verosimilhança condicional sobre  $f(y_{i1}, \dots, y_{iT} | t_i, \beta)$ , de modo a obter um estimador consistente para  $\beta$ . Em alternativa, a literatura considera a possibilidade de se assumir uma distribuição para  $\alpha_i$ . Esta abordagem é especialmente popular no âmbito do modelo probit, assumindo-se a distribuição normal para a heterogeneidade individual, mas também pode ser facilmente empregue com o modelo logit.

Assim, irão ser utilizados três modelos de variáveis dependentes binárias para dados de painel, sendo eles o modelo logit de efeitos fixos, o modelo logit de efeitos aleatórios e o modelo probit de efeitos aleatórios. Para escolher entre efeitos aleatórios e efeitos fixos, utiliza-se o teste de Hausman com o modelo logit (com o modelo probit, tal não é possível, pois não existe na literatura um modelo probit de efeitos fixos).

### 3.2. MODELO LOGIT DE EFEITOS FIXOS

Neste modelo, assume-se que  $\alpha_i$  e  $x_{it}$  podem estar correlacionados. Assume-se ainda que  $t_i = \bar{y}_i$  é uma estatística suficiente para  $\alpha_i$  e assim é possível a estimação consistente através da função de verosimilhança condicional. Neste modelo, destaca-se o facto da distribuição condicional de  $y_{i1}, \dots, y_{iT}$  ser degenerada se  $t_i = 0$  ou se  $t_i = 1$  e, por conseguinte, os indivíduos que apresentam estas características não contribuem para a verosimilhança condicional, sendo removidos da estimação. Ou seja, apenas os

indivíduos que mudam de estado pelo menos uma vez são relevantes para estimar  $\beta$ . Considerando o caso em que  $T = 2$  e onde os 2 possíveis resultados são (0,1) e (1,0), a probabilidade condicional do primeiro resultado do modelo logit de efeitos fixos pode ser escrito da seguinte forma:

$$(2) \quad P\{(0,1)|t_i = 1, \alpha_i, \beta\} = \frac{P\{(0,1)|\alpha_i, \beta\}}{P\{(0,1)|\alpha_i, \beta\} + P\{(1,0)|\alpha_i, \beta\}}.$$

Sabendo que

$$(3) \quad P\{(0,1)|\alpha_i, \beta\} = P\{y_{i1} = 0|\alpha_i, \beta\} P\{y_{i2} = 1|\alpha_i, \beta\},$$

em que

$$(4) \quad P\{y_{i2} = 1|\alpha_i, \beta\} = \frac{e^{\{\alpha_i + x'_{i2}\beta\}}}{1 + e^{\{\alpha_i + x'_{i2}\beta\}}},$$

então a probabilidade condicional é escrita como:

$$(5) \quad P\{(0,1)|t_i = 1, \alpha_i, \beta\} = \frac{e^{\{(x_{i2} - x_{i1})'\beta\}}}{1 + e^{\{(x_{i2} - x_{i1})'\beta\}}}.$$

Para o segundo resultado:

$$(6) \quad P\{(1,0)|t_i = 1, \alpha_i, \beta\} = \frac{1}{1 + e^{\{(x_{i2} - x_{i1})'\beta\}}}.$$

A partir das equações (5) e (6) é possível verificar que as probabilidades condicionais não dependem do termo dos efeitos individuais  $\alpha_i$  e por este motivo podemos estimar o modelo de efeitos fixos logit para  $T = 2$  utilizando um logit standard com  $x_{i2} - x_{i1}$  como variáveis explicativas e a mudança que ocorre em  $y_{it}$  como evento endógeno. Para o caso do nosso estudo, em que  $T$  é maior, torna-se um pouco mais complicado derivar as probabilidades condicionais necessárias, mas continua a ser uma extensão intuitiva desta ilustração.

### 3.3. MODELOS DE EFEITOS ALEATÓRIOS

Se puder ser assumido que  $\alpha_i$  é independente das variáveis explicativas em  $x_{it}$ , então o modelo de efeitos aleatórios é o mais apropriado para a modelação. A derivação deste modelo é mais facilmente explicada no contexto dos modelos Probit. Assim, após uma

abordagem geral, adequada para modelos logit e probit, a descrição centra-se no modelo probit.

A variável latente do modelo pode ser especificada como:

$$(7) \quad y_{it}^* = x_{it}'\beta + u_{it} ,$$

em que  $y_{it} = 1$  se  $y_{it}^* > 0$  e  $y_{it} = 0$  se  $y_{it}^* \leq 0$ .

Nesta especificação  $u_{it}$  é o termo de erro com média 0 e variância 1, sendo independente de  $(x_{i1}, \dots, x_{iT})$ . Para ser possível estimar  $\beta$ , faz-se uma suposição sobre a distribuição conjunta de  $u_{i1}, \dots, u_{iT}$ . A contribuição de verosimilhança do indivíduo  $i$  é a probabilidade conjunta de observar  $T$  resultados  $y_{i1}, \dots, y_{iT}$ .

Se puder ser assumido que todos os termos de erro  $u_{it}$  são independentes, tem-se que:

$$(8) \quad f(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \beta) = \prod_t f(y_{it} | x_{it}, \beta) ,$$

que envolve  $T$  integrais de uma só dimensão. Assim, se podermos assumir que  $u_{it} = \alpha_i + \varepsilon_{it}$ , em que  $\varepsilon_{it}$  é independente ao longo do tempo (e individualmente), então a probabilidade conjunta pode ser escrita da seguinte forma:

$$(9) \quad \begin{aligned} f(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \beta) &= \int_{-\infty}^{\infty} f(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \alpha_i, \beta) f(\alpha_i) d\alpha_i \\ &= \int_{-\infty}^{\infty} [\prod_t f(y_{it} | x_{it}, \alpha_i, \beta)] f(\alpha_i) d\alpha_i . \end{aligned}$$

Esta formulação necessita de integração numérica apenas para uma dimensão, permitindo que os termos de erro sejam correlacionados ao longo dos diferentes períodos, ainda que de forma restritiva.

Concretizando, assume-se que a distribuição conjunta de  $u_{i1}, \dots, u_{iT}$  é normal com média 0 e variância 1 e que  $cov\{u_{it}, u_{is}\} = \sigma_\alpha^2$ , para  $s \neq t$ . Sendo que isto é o mesmo que assumir que  $\alpha_i$  é i.i.d. com média 0 e variância  $\sigma_\alpha^2$  e que  $\varepsilon_{it}$  é i.i.d. com média 0 e variância  $(1 - \sigma_\alpha^2)$ . De modo a normalizar as variâncias dos erros, tipicamente assume-se que a variância do erro num determinado período seja igual 1.

Para o modelo probit de efeitos aleatórios, as expressões da função de probabilidade são dadas por:

$$(10) \quad f(y_{it}|x_{it}, \alpha_i, \beta) = \Phi\left(\frac{x'_{it}\beta + \alpha_i}{\sqrt{1 - \sigma_\alpha^2}}\right) \quad \text{se } y_{it} = 1,$$

$$f(y_{it}|x_{it}, \alpha_i, \beta) = 1 - \Phi\left(\frac{x'_{it}\beta + \alpha_i}{\sqrt{1 - \sigma_\alpha^2}}\right) \quad \text{se } y_{it} = 0,$$

onde  $\Phi$  representa a função da distribuição normal *standard*. Relativamente à função de densidade de  $\alpha_i$ , esta é dada por:

$$(11) \quad f(\alpha_i) = \left(\frac{1}{\sqrt{2\pi\sigma_\alpha^2}}\right) e^{\left\{-\frac{1}{2}\frac{\alpha_i^2}{\sigma_\alpha^2}\right\}}.$$

Esta adaptação para o modelo logit de efeitos aleatórios é simples uma vez que as suposições da distribuição para  $\alpha_i$  também são facilmente empregues com os modelos logit.

#### 3.4. TESTE DE HAUSMAN

De modo a testar se os efeitos são fixos ou aleatórios e, por conseguinte, averiguar qual dos dois tipos de estimador anteriormente apresentados é o mais eficiente, é utilizado o teste de Hausman. Neste teste são consideradas as seguintes hipóteses:

$$(12) \quad H_0: E(\alpha_i|x_{it}) = 0;$$

$$H_1: E(\alpha_i|x_{it}) \neq 0.$$

Sendo que a estatística de teste é dada por:

$$(13) \quad H = (\hat{\beta}_{EF} - \hat{\beta}_{EA})' [\hat{V}(\hat{\beta}_{EF}) - \hat{V}(\hat{\beta}_{EA})]^{-1} (\hat{\beta}_{EF} - \hat{\beta}_{EA}) \sim X_K^2,$$

onde  $\hat{\beta}_{EF}$  é o vetor dos coeficientes do estimador de efeitos fixos,  $\hat{\beta}_{EA}$  é o vetor dos coeficientes do estimador de efeitos aleatórios,  $\hat{V}(\hat{\beta}_{EF})$  é o estimador da matriz de covariâncias do estimador de efeitos fixos e  $\hat{V}(\hat{\beta}_{EA})$  é o estimador da matriz de covariâncias do estimador de efeitos aleatórios. Sob  $H_0$  ambos os estimadores de efeitos fixos e de efeitos aleatórios são consistentes, mas apenas o último é eficiente. Se  $H_0$  for rejeitada, deve-se utilizar o estimador de efeitos fixos, uma vez que de entre os dois é o único consistente. No caso de  $H_0$  não ser rejeitada, então utiliza-se o estimador de efeitos aleatórios.

#### 4. DADOS

Neste capítulo numa primeira fase será descrito o processo de junção das diferentes bases de dados utilizadas, os respetivos dados perdidos e as fontes utilizadas. Posteriormente, os dados serão examinados com pormenor, através de uma análise descritiva. Para o tratamento de dados recorreu-se ao *software* R. O código R criado foi entregue ao INE, para posterior atualização da base de dados conjunta, caso o INE pretenda.

##### 4.1. PREPARAÇÃO DOS DADOS

Os dados foram obtidos através do INE, no âmbito da Lei do SEN – artigo 6º, que prevê que os dados estatísticos individuais sobre pessoas singulares e coletivas podem ser cedidos para fins científicos sob forma anonimizada, a investigadores credenciados. Uma das fontes de dados, é o IUTICE, que é realizado anualmente em Portugal pelo INE no âmbito de legislação comunitária, sendo que a metodologia é harmonizada e os resultados são comparáveis a nível europeu. É realizado por amostragem, sendo que a base de amostragem é estratificada por classes de pessoas ao serviço (dimensão), classes de volumes de negócios e por classes de atividade económica. Esta operação estatística visa contribuir para um conhecimento mais aprofundado sobre a utilização de TIC nas empresas, assim como possibilitar o estudo da importância e da prioridade que poderá revestir o recurso a este tipo de tecnologias em termos da competitividade no conjunto da economia. Desde o início da realização deste inquérito, e dado que se trata de uma operação estatística que se foca em aspetos em permanente mutação, o inquérito tem sofrido alterações todos os anos, com a introdução, remoção e alteração de alguns módulos e variáveis. No sentido de obter informações relativas à maior ou menor utilização de tecnologias pelas empresas, no âmbito deste estudo foram selecionadas as seguintes variáveis consideradas de interesse: a utilização de Internet, a posse de *website* próprio ou do grupo e a realização comércio *online*, através do seu *website* ou através de apps.

A outra fonte de dados utilizada são os dados dos Quadros de Pessoal. Esta informação, da responsabilidade do GEP, do MTSSS, é uma operação do tipo recenseamento, que resulta do aproveitamento administrativo de normas legais que estabelecem a obrigatoriedade da entrega do Quadro de Pessoal por parte de todas as

entidades com trabalhadores ao seu serviço, com exceção da administração central, regional e local e os institutos públicos e a empregadores de trabalho doméstico. Esta informação tem como período de referência o mês de outubro de cada ano e âmbito geográfico o país. O objetivo desta operação é disponibilizar informação sobre estrutura empresarial, emprego, duração de trabalho, remunerações e regulamentação coletiva de trabalho. No âmbito deste trabalho foi utilizada a informação dos Quadros de pessoal por trabalhador, a qual permitiu obter informações, tais como a proporção por escalão de idade do pessoal ao serviço da empresa, a proporção de pessoas ao serviço por sexo e ainda as proporções por tipo de habilitação literária do pessoal ao serviço das empresas.

Com o objetivo de relacionar a intensidade das empresas no domínio tecnológico com as qualificações dos recursos humanos e as características das empresas, para o período de 2010 a 2018, as informações destas duas fontes de dados foram integradas numa única base de dados. Primeiramente e antes do processo de junção das informações presentes nas duas fontes: IUTICE e Quadros de Pessoal, dado que as informações presentes nas bases de dados dos Quadros de Pessoal diziam respeito a cada trabalhador individualmente, procedeu-se a um agrupamento por empresa das observações através do número de identificação fictício das empresas, que é único para cada uma das diferentes empresas. De facto, a unidade amostral considerada neste estudo será a empresa. Como consequência deste facto, este procedimento envolveu também a criação de variáveis de quantidades e proporções das características para o estudo. Resultante deste processo e apenas com base nos Quadros de Pessoal, obtiveram-se observações relativas a aproximadamente 290 000 empresas por cada ano de estudo. De forma a trabalhar com dados em painel, agregaram-se estas observações referentes às diferentes empresas ao longo dos anos numa só base de dados e deste procedimento resultaram 4 796 415 observações relativas a 532 935 empresas distintas. Em seguida, procedeu-se ao tratamento das bases de dados do IUTICE, que correspondem a amostras anuais e que contêm observações relativas a aproximadamente 6 000 empresas distintas por ano. Durante o tratamento transformaram-se algumas variáveis, devido às alterações que ocorreram no inquérito ao longo dos anos. Após este procedimento, realizou-se o processo de agregação das bases de dados anuais numa só, onde se juntaram as observações relativas às empresas presentes no IUTICE ao longo dos 9 anos do estudo e, deste modo, resultaram 307 566 observações relativas a 34 174 empresas distintas. Na

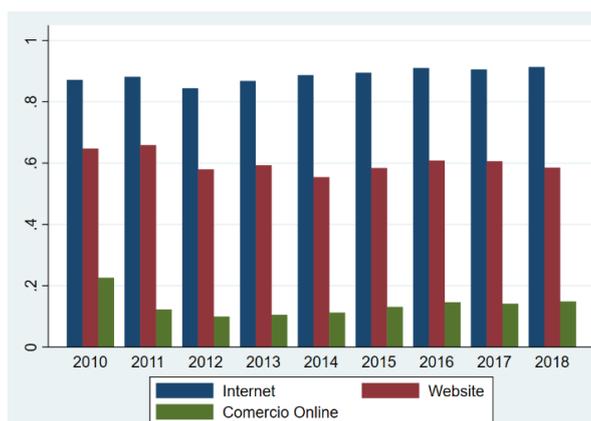
base de dados final, após juntar os dados resultantes das bases de dados do IUTICE e dos Quadros de Pessoal e de reter apenas as observações relativas às empresas comuns em ambas as bases de dados através da utilização do número de identificação fictício das empresas e do ano como chave, resultaram um total de 279 612 observações para o período de 2010 a 2018, que dizem respeito a 31 068 empresas. Ou seja, conseguiu-se manter cerca de 90.9% das empresas presentes nas amostras do IUTICE e aproximadamente 5.8% das empresas presentes na base de dados dos Quadros de Pessoal. Este processo de junção de bases de dados pode ser observado na Figura A1, presente no Anexo.

#### 4.2. ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS

Todas as variáveis que vão ser utilizadas na estimação dos modelos, juntamente com as suas respetivas descrições, codificações e tipo, estão representadas na Tabela A2, presente no Anexo. As estatísticas descritivas das variáveis explicativas que mais tarde serão apresentadas para a modelação econométrica também podem ser encontradas no Anexo, mais especificadamente na Tabela A3.

Relativamente às variáveis de interesse, que são binárias, tomando o valor 1 quando a característica em causa se verifica e 0 caso não se verifique, encontram-se apresentadas na Figura 2 as suas proporções ao longo do período de estudo.

FIGURA 2 - PROPORÇÕES DE EMPRESAS POR VARIÁVEIS DE INTERESSE POR ANO.



Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

Através da análise desta representação, verifica-se que a proporção de empresas em estudo que utilizam Internet registou um mínimo em 2012, ano a partir do qual se verificou um crescimento desta proporção, sendo que esta registou um máximo no ano de

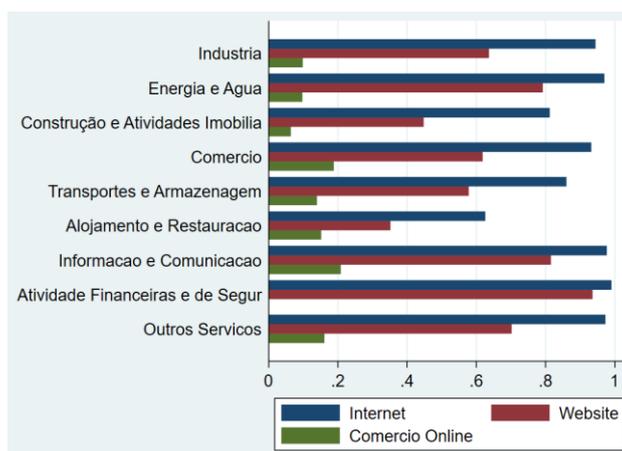
2018, no qual 91.3% das empresas em estudo utilizavam Internet no seu funcionamento. Relativamente à proporção de empresas com *website* próprio, é possível observar que esta proporção aumenta e diminui ao longo dos 9 anos, mas é sempre superior a 0.55, sendo que atinge um máximo em 2011 (0.659) e um mínimo em 2014 (0.554). No que diz respeito às empresas que realizam comércio *online*, retira-se que em 2010 a proporção regista um máximo de 0.226, sendo que esta sofre uma queda abrupta até 2012, onde regista um mínimo (0.099). Este facto pode ser explicado pelas alterações da pergunta referente ao comércio *online* do IUTICE, uma vez que em 2010 esta pergunta se refere ao comércio *online* efetuado através de todas as redes eletrónicas, sendo que a partir de 2011 a pergunta já só se refere ao comércio *online* efetuado exclusivamente através do *website* ou *apps*. A partir de 2012 esta proporção cresce continuamente até 2016, diminuindo em 2017 e voltando a subir em 2018.

Para uma análise mais detalhada destas variáveis e da sua variação ao longo do tempo, é possível retirar da Tabela A4, presente no Anexo, que, 88.84% do total das observações dizem respeito a empresas que utilizam Internet. 84.60% das empresas (26283 das 31068 empresas) utilizaram Internet em pelo menos 1 dos anos. De entre essas 26283 empresas, 99.26% utilizaram Internet sempre, ao longo dos 9 anos observados. Relativamente à variável *Website*, 59.94% das observações totais dizem respeito a empresas que possuem *website* próprio. 47.78% das empresas (14844 das 31068 empresas) possuíram *website* próprio em pelo menos 1 dos anos. De entre essas 14844 empresas, 96.93% possuíram *website* próprio sempre, ao longo dos 9 anos. No que diz respeito à variável *ComOnline*, 13.60% do total de observações dizem respeito a empresas que realizam comércio *online*. 13.08% das empresas (4065 das 31068 empresas) realizaram comércio *online* em pelo menos 1 dos anos e 77.69% dessas 4065 empresas realizaram sempre comércio *online*, durante todo o período de 2010 a 2018.

De forma a analisar as variáveis de interesse tendo em conta o setor de atividade económica desenvolvido pela empresa (atividade principal) e a dimensão da empresa, em termos de pessoal ao serviço, podemos observar as Figuras 3 e 4, respetivamente. Os critérios de classificação das variáveis utilizadas nas Figuras 3 e 4 anteriores, podem ser observados com detalhe no Anexo, na Tabela A5 e na Tabela A6, respetivamente.

A Figura 3 apresenta a informação por setores de atividade económica, que correspondem a agregações da Secção da Classificação das Atividades Económicas Portuguesas, Revisão 3 (CAE Rev. 3).

FIGURA 3 - PROPORÇÕES DE EMPRESAS POR VARIÁVEIS DE INTERESSE POR SETOR DE ATIVIDADE ECONÓMICA.

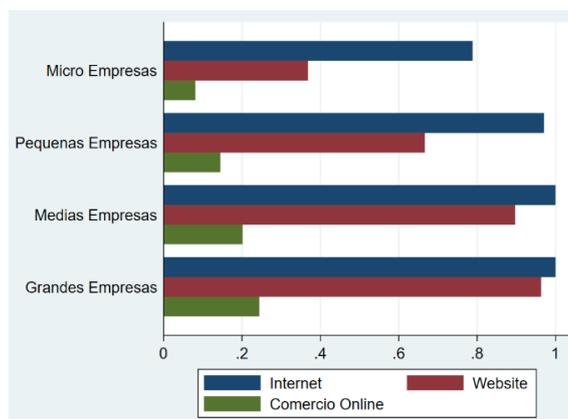


Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

A partir da Figura 3, é possível retirar a ideia de que os setores de atividades financeiras e de seguros, da informação e comunicação e da energia e água são os que apresentam uma maior proporção de empresas que utilizam Internet e que possuem *website* próprio. Por outro lado, os setores de alojamento e restauração e de construção e atividades imobiliárias são os que apresentam uma menor proporção de empresas que utilizam os serviços de Internet e que possuem *website* próprio ou do grupo. Relativamente ao comércio *online*, o setor de informação e comunicação e o setor do comércio são os que apresentam uma maior proporção de empresas que realizam vendas de produtos ou serviços através do seu *website* ou de *apps*, enquanto o setor de atividade financeira e de seguros é o que apresenta uma menor proporção relativamente a esta variável.

No caso da Figura 4, apresenta informação por dimensão da empresa. Para classificar as empresas por dimensão foi utilizada a variável relativa ao número de pessoas ao serviço.

FIGURA 4 - PROPORÇÕES DE EMPRESAS POR VARIÁVEIS DE INTERESSE POR DIMENSÃO DA EMPRESA, TENDO EM CONTA O NÚMERO DE PESSOAS AO SERVIÇO.



Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

Ao observar a Figura 4, verifica-se que tanto a utilização de Internet, como a posse de *website* próprio e como a realização de comércio *online* tendem a crescer com a dimensão da empresa, isto é, as empresas com mais de 250 pessoas ao serviço (grandes empresas) são as que apresentam maiores proporções das variáveis de interesse e as empresas com menos de 10 pessoas ao serviço (microempresas) são as que apresentam menores proporções.

Dado que a maioria das variáveis referentes às proporções relativas às habilitações literárias na desagregação original apresentam valores bastante reduzidos, estas variáveis foram agrupadas em 3 variáveis, sendo uma das variáveis a proporção de trabalhadores com habilitações equivalentes ao ensino primário (PICiclo), uma segunda variável a proporção de trabalhadores com habilitações equivalentes ao ensino secundário (PEnsSec) e, por fim, uma variável relativa à proporção de trabalhadores com habilitações equivalentes ao ensino superior (PEnsSup).

Com o objetivo de analisar a relação entre as variáveis explicativas e as variáveis de interesse, calcularam-se as médias das variáveis explicativas conforme o valor de cada uma das diferentes variáveis de interesse. Estes cálculos encontram-se representados na Tabela I, sendo que através da análise desta tabela, existem alguns fatores que merecem destaque.

TABELA I<sup>1</sup>

MÉDIA DAS VARIÁVEIS EXPLICATIVAS CONFORME AS VARIÁVEIS DE INTERESSE

	Internet		Website		Comércio Online	
	0	1	0	1	0	1
N	4.21	169.391	19.662	234.679	131.318	284.468
Pmulheres	0.44	0.389	0.411	0.383	0.388	0.431
PIdade30	0.126	0.181	0.149	0.192	0.17	0.209
PIdade3154	0.631	0.685	0.661	0.691	0.68	0.68
PCursoSup	0	0.0001	0	0.0001	0.0001	0.0001
PInf1Ciclo	0.007	0.004	0.005	0.004	0.004	0.003
PEnsBas	0.776	0.486	0.658	0.427	0.533	0.405
PEnsPS	0.005	0.008	0.007	0.008	0.008	0.009
PBach	0.008	0.026	0.016	0.03	0.023	0.03
PLic	0.045	0.176	0.091	0.208	0.157	0.198
PMest	0.004	0.018	0.007	0.022	0.016	0.019
PDout	0.001	0.002	0.001	0.002	0.002	0.002
PICiclo	0.783	0.49	0.663	0.43	0.537	0.408
PEnsSec	0.154	0.28	0.216	0.299	0.257	0.334
PEnsSup	0.062	0.23	0.121	0.27	0.206	0.258

Fonte: INE, Inquérito à utilização de TIC nas empresas e GEP, Quadros de pessoal, cálculos do autor.

Ao analisar a Tabela I, em relação à utilização de Internet, retira-se que o número médio de pessoas ao serviço é bastante superior nas empresas em estudo que utilizam (169.391) quando comparado com as empresas que não utilizam (4.21). Para além disso, verifica-se que proporção de mulheres é inferior nas empresas que utilizam Internet

<sup>1</sup> No cálculo das proporções, os valores apresentados resultam da divisão do número de trabalhadores que apresentam determinada característica pelo número total de trabalhadores da empresa. Por exemplo, os valores para Pmulheres resultam da divisão do número de mulheres pelo número total de trabalhadores da empresa.

(0.389) quando comparada com as empresas que não utilizam (0.44). Relativamente às proporções por escalão de idade do pessoal ao serviço da empresa, é possível observar que estas são superiores nas empresas que utilizam Internet e com pessoas ao serviço com idades entre 31 e 54 anos, o que indicia que o pessoal ao serviço tende a ser mais jovem nas empresas que utilizam Internet. No que diz respeito às habilitações literárias do pessoal ao serviço, os cálculos sugerem que as proporções de trabalhadores com qualificações mais baixas, equivalentes ao ensino básico, são inferiores no caso das empresas que têm acesso à Internet. Verifica-se que a proporção de trabalhadores com habilitações equivalentes ao ensino secundário é superior no caso das empresas que têm acesso. Ainda é possível retirar a ideia de que as proporções de trabalhadores com habilitações mais altas, equivalentes ao bacharelato, licenciatura, mestrado e doutoramento são superiores no caso das empresas que utilizam Internet quando comparadas com as que não utilizam.

Relativamente à variável de interesse que diz respeito ao *website*, verifica-se que o número médio de pessoas ao serviço é bastante superior nas empresas que possuem (234.679) quando comparado com as empresas que não possuem (19.662). Relativamente à proporção de mulheres, e tal como no caso da Internet, observa-se que é ligeiramente inferior nas empresas em estudo que possuem (0.383) quando comparada com as empresas que não possuem (0.411). Relativamente às idades, verifica-se que as proporções dos escalões de idades mais jovens são superiores nas empresas que possuem *website*. Já no que diz respeito às habilitações literárias, observou-se que a proporção de trabalhadores com habilitações inferiores, isto é, inferiores ao primeiro ciclo e equivalentes ao ensino básico, verifica-se inferior no caso das empresas que possuem *website* e que as proporções de trabalhadores com habilitações equivalentes à licenciatura, mestrado e doutoramento é superior no caso das empresas que possuem.

No caso da variável relativa ao comércio *online*, tal como nos casos anteriores, verifica-se um número médio de pessoas ao serviço superior no caso das empresas em estudo que realizam comércio *online* através de *websites* ou *apps* (284.486), quando comparado com as empresas não realizam (131.318). Relativamente à proporção de mulheres empregadas nas empresas, verifica-se que, contrariamente aos casos das outras variáveis de interesse, no caso das empresas que realizam comércio *online* esta proporção é superior (0.431) quando comparada com as empresas que não realizam (0.388). Quanto

às habilitações, a situação é idêntica à observada anteriormente, sendo que se verifica nas empresas que realizam comércio *online* um decréscimo nas proporções equivalentes a habilitações inferiores e um aumento nas proporções relativas a habilitações superiores.

De modo a proceder ao levantamento do padrão da falta de dados observou-se o output do comando *xtdescribe* no Stata. De acordo com este output representado na Figura A7, no Anexo, retira-se a ideia de que não há falta de dados para nenhum dos anos em estudo, o que levaria a que o painel fosse equilibrado. Contudo existem algumas observações que contêm apenas dados relativos às variáveis dos quadros de pessoal, não existindo alguns dados relativos às variáveis de interesse das tecnologias, isto porque nestes casos essas empresas não foram selecionadas para fazerem parte das amostras do IUTICE em alguns dos anos em análise. Também se verifica o inverso, sendo que existem algumas observações que apenas contêm dados relativos às variáveis das tecnologias de informação, dado que os quadros de pessoal incluem apenas empresas com pelo menos um trabalhador remunerado.

Dado que para a modelação irão ser utilizadas apenas as observações que contêm dados completos para todas as variáveis, deixam-se cair todas as observações incompletas. Por outro lado, devido à alteração que ocorreu na pergunta do IUTICE relativa ao comércio online, como explicado anteriormente, a modelação irá incidir sobre o período de 2012 a 2018, tendo-se eliminado da base inicial os anos 2010 e 2011. Assim, ao observar o padrão do output do *xtdescribe* representado na Figura A8, no Anexo, já é possível verificar que existem dados em falta para alguns anos. Deste modo e como seria de esperar obtemos um painel não equilibrado, com um total de 23 636 empresas distintas para efetuar a modelação econométrica no período de 2012 a 2018.

De forma a analisar a distribuição das empresas que irão ser utilizadas na modelação, encontram-se representados no Anexo, na Tabela A9 na Tabela A10 e na Tabela A11, o número de empresas por ano, por setor de atividade e por dimensão, respetivamente. Verifica-se que nos anos mais recentes há um maior número de empresas na amostra, sendo que o setor de atividade económica que tem maior peso no total de observações é o da Indústria. Já relativamente à dimensão das empresas, as microempresas são as que estão mais presentes na amostra.

## 5. RESULTADOS

Esta secção tem como objetivo apresentar e interpretar os resultados mais relevantes da aplicação da metodologia descrita no capítulo 3 os dados apresentados no capítulo 4. Encontra-se dividida em duas partes, sendo a primeira dedicada à análise dos resultados dos diferentes modelos considerados e a segunda dedicada à análise das estimativas para os efeitos parciais médios, sendo que os efeitos marginais foram calculados a partir do comando automático do Stata *margins*, o que significa que  $\alpha_i$  foi considerado como zero.

Para a estimação dos modelos apresentados foi utilizado o *software* Stata, no qual foram testados vários conjuntos de regressores, com a inclusão e a exclusão de algumas variáveis. Para os modelos com *website* como variável binária são incluídas todas as observações referidas no fim do capítulo 4, mas dado que a abordagem utilizada foram modelos a duas partes, então para os modelos com comércio *online* como variável de interesse irão ser utilizadas apenas as observações das empresas que possuem *website* próprio ou do grupo. Note-se ainda que, das variáveis explicativas disponíveis, numa primeira fase a variável *Internet* foi incluída na modelação. Contudo esta não se verificou estatisticamente significativa em nenhum dos modelos, o que certamente se deve à sua pouca variabilidade, especialmente nos últimos anos da amostra, sendo que os modelos propostos foram estimados sem a sua inclusão. No caso dos modelos com *website* como variável binária, as variáveis *dummy* relativas aos setores de atividade económica e à dimensão das empresas foram incluídas numa primeira fase, mas dado que se verificou a sua insignificância estatística estas foram removidas dos modelos. No caso dos modelos com comércio *online* como variável de interesse binária, a variável *dummy* relativa às empresas pertencentes ao setor das Atividades Financeiras e de Seguros foi removida uma vez que através da análise descritiva dos dados se apurou que nenhuma destas empresas realiza comércio *online*. Foram ainda incluídas variáveis *dummy* temporais para todos os modelos, de modo a estimar os efeitos temporais na posse de *website* próprio ou do grupo e na adoção do comércio *online* comparativamente ao primeiro ano do estudo (2012). Note-se que em ambas as partes do modelo são considerados modelos logit de efeitos aleatórios e de efeitos fixos numa primeira fase. Após a realização do teste de Hausman, caso se conclua pela existência de efeitos aleatórios, considera-se ainda o modelo probit de efeitos aleatórios.

## 5.1.MODELOS ESTIMADOS

Nesta secção, apresentam-se vários resultados para os modelos estimados, nomeadamente, estimativas dos coeficientes de regressão e respetivos desvios-padrão, testes de significância global e o teste de Hausman para o modelo logit de efeitos fixos e de efeitos aleatórios.

## 5.1.1. MODELOS COM WEBSITE COMO VARIÁVEL BINÁRIA

Para a variável *website*, vejam-se os resultados na Tabela II. De acordo com o teste de Hausman, que pode ser observado na Tabela II, rejeita-se a hipótese nula na qual os efeitos são aleatórios, para um nível de significância de 5%. Dado que esta hipótese não pode ser assumida, conclui-se que os efeitos são fixos e considera-se o modelo logit de efeitos fixos como modelo final.

TABELA II  
ESTIMATIVAS DOS MODELOS PARA A VARIÁVEL *WEBSITE*

	Logit de Efeitos Fixos	Logit de Efeitos Aleatórios
Pmulheres	0.522 (0.636)	-3.492*** (0.487)
PIdade30	2.197*** (0.713)	7.757*** (1.161)
PIdade3154	1.197** (0.59)	3.946*** (0.95)
PEnsSec	-0.346 (0.635)	12.292*** (2.008)
PEnsSup	-0.396 (0.866)	16.13*** (2.653)
2013	0.215 (0.179)	0.086 (0.151)
2014	-0,542*** (0.191)	-0.853*** (0.211)
2015	0.309 (0.206)	0.139 (0.203)
2016	0.605*** (0.208)	0.477* (0.201)
2017	0.624*** (0.217)	0.38* (0.21)

2018	0.777*** (0.237)	0.264 (0.216)
constante	-	-9.013*** (1.097)
Significância Global	0.00	0.00
Significância Tempo	0.00	0.00
N	2 112	36 233
Teste de Hausman	0.00	

Notas:

1. Os desvios-padrão estão apresentados entre parênteses e foram estimados com Bootstrap.
2. Níveis de significância: \*10%, \*\*5%, \*\*\*1%.
3. Para os testes de significância conjunta e para o teste de Hausman, apresentam-se os valores-p.

Através da observação da Tabela II, os resultados sugerem que para um nível de significância de 5%, apenas as variáveis relativas às proporções de idades dos trabalhadores e as variáveis *dummy* temporais, exceto as relativas a 2013 e a 2015, são estatisticamente significativas individualmente, ou seja, apenas alguns determinantes contribuem significativamente para a explicação da posse de *website* por parte das empresas. Verifica-se também que as variáveis explicativas são conjuntamente significativas. Note-se que caso se tivesse obtido um resultado no teste de Hausman que suportasse o uso dos modelos de efeitos aleatórios, no âmbito do logit de efeitos aleatórios existiria um maior número de variáveis explicativas estatisticamente significativas, sendo que apenas a posse de as *dummies* temporais relativas a 2013, 2015 e 2018 não seriam consideradas estatisticamente significativas para um nível de significância de 10%.

Por outro lado, dado que o sinal dos coeficientes apresentados indica se uma variável influencia positivamente (sinal positivo) ou negativamente (sinal negativo) a probabilidade de uma empresa possuir *website* próprio ou do grupo, retira-se desta análise que os efeitos das variáveis estatisticamente significativas não se opõem ao que seria esperado para cada. Verifica-se assim que os aumentos das proporções de trabalhadores mais jovens contribuem positivamente para esta probabilidade. Relativamente às *dummies* temporais, observa-se que o ano de 2014 contribui negativamente, enquanto os

anos de 2016, 2017 e 2018 contribuem positivamente para a probabilidade, *ceteris paribus* e quando comparado com 2012.

### 5.1.2. MODELOS COM COMÉRCIO ONLINE COMO VARIÁVEL BINÁRIA

No que se refere à variável comércio *online*, os resultados apresentam-se na Tabela III.

TABELA III  
ESTIMATIVAS DOS MODELOS PARA A VARIÁVEL COMÉRCIO *ONLINE*

	Logit de Efeitos Fixos	Logit de Efeitos Aleatórios	Probit de Efeitos Aleatórios
Pmulheres	0.911 (0.803)	1.266*** (0.22)	0.725*** (0.116)
PIdade30	0.286 (0.935)	0.694* (0.36)	0.404* (0.238)
PIdade3154	0.598 (0.77)	0.215 (0.289)	0.13 (0.198)
PEnsSec	-0.179 (0.872)	1.047*** (0.255)	0.594*** (0.147)
PEnsSup	-0.694 (1.061)	0.077 (0.247)	0.037 (0.117)
Pequenas Empresas	-0,239 (0.319)	0.166 (0.121)	0.09 (0.071)
Médias Empresas	-0.046 (0.537)	0.588*** (0.147)	0.322*** (0.075)
Grandes Empresas	-0.19 (0.643)	0.712*** (0.183)	0.386*** (0.116)
Energia e Água	13.054* (7.719)	0.176 (0.368)	0.106 (0.227)
Construção e Atividades Imobiliárias	1.202 (0.00)	0.497** (0.201)	0.296*** (0.113)
Comércio	0.565 (6.118)	1.762*** (0.165)	1.034*** (0.097)
Transportes e Armazenagem	Omitido	1.212*** (0.259)	0.715*** (0.186)

Alojamento e Restauração	12.786 (0.00)	3.673*** (0.276)	2.082*** (0.134)
Informação e Comunicação	1.65 (13.545)	1.995*** (0.231)	1.167*** (0.115)
Outros Serviços	1.705 (11.405)	1.14*** (0.187)	0.681*** (0.115)
2013	0.24* (0.134)	0.254* (0.139)	0.121* (0.068)
2014	0.679*** (0.171)	0.628*** (0.153)	0.328*** (0.082)
2015	0.952*** (0.191)	0.93*** (0.161)	0.498*** (0.084)
2016	1.146*** (0.192)	1.124*** (0.152)	0.605*** (0.089)
2017	1.069*** (0.196)	1.004*** (0.157)	0.537*** (0.091)
2018	1.29*** (0.202)	1.264*** (0.157)	0.692*** (0.087)
constante	-	-7.077*** (0.412)	-3.954*** (0.259)
Significância Global	0.00	0.00	0.00
Significância Dimensão	0.84	0.00	0.00
Significância Setor	0.111	0.00	0.00
Significância Tempo	0.00	0.00	0.00
N	3 106	21 271	21 271
Teste de Hausman		0.202	

Notas:

1. Os desvios-padrão estão apresentados entre parênteses e foram estimados com Bootstrap.
2. Níveis de significância: \*10%, \*\*5%, \*\*\*1%.
3. Para os testes de significância conjunta e para o teste de Hausman, apresentam-se os valores-p.

De acordo com o resultado do teste de Hausman, que pode ser observado na Tabela III, neste caso já não se rejeita a hipótese nula na qual os efeitos são aleatórios, para um

nível de significância de 5%. Assim, admite-se que os efeitos são aleatórios e então tanto os modelos de efeitos aleatórios como os modelos de efeitos fixos são consistentes, porém os modelos de efeitos aleatórios são mais eficientes. Assim sendo, os modelos de efeitos aleatórios serão considerados como os mais apropriados para a modelação. Dado que podemos supor que os efeitos são aleatórios, também se pode considerar o modelo probit, que até é mais conhecido do que o logit neste contexto, e, por conseguinte, a interpretação também irá considerar este modelo. De qualquer modo, os resultados em termos de significância e direção dos efeitos parciais são praticamente coincidentes entre o logit e o probit e, para além disso, a magnitude dos efeitos parciais é também muito similar nestes dois modelos, como pode ser observado mais à frente na Tabela IV.

Os resultados do modelo probit apresentados na Tabela III sugerem que para um nível de significância de 1%, quase todas as variáveis explicativas consideradas são estatisticamente significativas individualmente, ou seja, existem diversos determinantes que contribuem significativamente para a adoção de comércio *online* por parte das empresas. De facto, apenas seis não se revelaram significativas para este nível de significância, sendo que a proporção de trabalhadores com idade inferior a 30 anos e a *time-dummy* para 2013 apenas são significativas a 10%. Verifica-se também que as variáveis explicativas são conjuntamente significativas. Note-se que caso não se tivesse obtido um resultado no teste de Hausman que suportasse o uso dos modelos de efeitos aleatórios, no âmbito do logit de efeitos fixos apenas a pertença ao sector de Energia e Água e as *dummies* temporais seriam consideradas estatisticamente significativas.

Por outro lado, através do sinal dos coeficientes apresentados retira-se desta análise que os efeitos das variáveis estatisticamente significativas não se opõem ao que seria esperado para cada variável. Verifica-se que aumentos na proporção de trabalhadores com idades mais jovens, na proporção de trabalhadores com habilitações equivalentes ao secundário e na proporção de mulheres têm impactos positivos na probabilidade de uma empresa adotar comércio *online*. O facto de uma empresa pertencer ao grupo das médias ou grandes empresas também influencia positivamente esta probabilidade, relativamente às microempresas. Quando comparado com o setor da Indústria, verifica-se que o facto de uma empresa pertencer a qualquer um dos setores considerados tem um impacto positivo na probabilidade de realizar comércio *online*, exceto o setor da energia e água, cujo efeito não difere estatisticamente em relação ao setor Indústria. Já no que diz respeito

às *dummies* temporais, verifica-se que todos os anos influenciam positivamente esta probabilidade, relativamente ao ano de 2012.

## 5.2. ESTIMATIVAS DOS EFEITOS PARCIAIS MÉDIOS

De modo a avaliar a magnitude dos efeitos que cada variável tem sobre a probabilidade uma empresa possuir *website* próprio ou do grupo e de uma empresa realizar comércio *online* foram estimados os efeitos parciais médios (EPM). Na Tabela IV podem ser observados os EPM das variáveis explicativas para os modelos selecionados na secção anterior.

TABELA IV  
EFEITOS PARCIAIS MÉDIOS

	<i>Website</i>	<i>Comércio Online</i>	
	Logit de Efeitos Fixos	Logit de Efeitos Aleatórios	Probit de Efeitos Aleatórios
Pmulheres	0.077 (0.094)	0.077*** (0.011)	0.081*** (0.012)
PIdade30	0.322*** (0.081)	0.042** (0.021)	0.045* (0.023)
PIdade3154	0.176*** (0.061)	0.013 (0.019)	0.015 (0.021)
PEnsSec	-0.051 (0.099)	0.0364*** (0.013)	0.067*** (0.014)
PEnsSup	-0.058 (0.134)	0.005 (0.012)	0.004 (0.014)
Pequenas Empresas	-	0.01 (0.007)	0.01 (0.008)
Médias Empresas	-	0.036*** (0.008)	0.036*** (0.009)
Grandes Empresas	-	0.044*** (0.011)	0.043*** (0.012)
Energia e Água	-	0.009 (0.014)	0.009 (0.016)
Construção e Atividades Imobiliárias	-	0.025*** (0.009)	0.027** (0.011)

Comércio	-	0.1*** (0.009)	0.11*** (0.01)
Transportes e Armazenagem	-	0.065*** (0.015)	0.072*** (0.018)
Alojamento e Restauração	-	0.258*** (0.017)	0.262*** (0.015)
Informação e Comunicação	-	0.117*** (0.014)	0.128*** (0.015)
Outros Serviços	-	0.06*** (0.009)	0.068*** (0.011)
2013	0.035 (0.031)	0.014** (0.007)	0.012* (0.007)
2014	-0.107** (0.044)	0.035*** (0.008)	0.034*** (0.008)
2015	0.049 (0.036)	0.053*** (0.008)	0.053*** (0.008)
2016	0.089** (0.04)	0.066*** (0.008)	0.065*** (0.008)
2017	0.091** (0.041)	0.058*** (0.008)	0.057*** (0.008)
2018	0.108** (0.045)	0.075*** (0.009)	0.075*** (0.009)

Notas:

- Os desvios padrão estão apresentados entre parênteses. O estimador para a variância utilizado foi o Bootstrap, no caso do *Website*, e o Huber-White Sandwich, no caso do comércio *online*.
- Níveis de significância: \*10%, \*\*5%, \*\*\*1%.

Para o modelo logit de efeitos fixos com *website* como variável binária, existem poucos determinantes estatisticamente significativos para a probabilidade de uma empresa possuir *website* próprio ou do grupo. De entre as características dos recursos humanos, apenas as proporções de idades influenciam positivamente esta probabilidade, sendo que, *ceteris paribus*, o aumento de 0.1 da proporção de trabalhadores com idades iguais ou inferiores a 30 anos provoca um acréscimo de 0.032 na probabilidade de posse de *website*, sendo este acréscimo superior ao que se verifica com um aumento de igual quantidade na proporção de trabalhadores com idades compreendidas entre os 31 e os 54 anos, o qual é de 0.018. Ou seja, existe uma relação que determina que quanto mais jovens

forem os trabalhadores de uma empresa, maior será a probabilidade de essa empresa possuir *website*.

Relativamente às variáveis *dummy* temporais e tal como esperado, os anos mais recentes (especialmente a partir de 2016) são aqueles que têm um maior impacto na probabilidade de posse de *website*, comparativamente a 2012 e mantendo tudo o resto constante. De notar que o ano de 2014 apresentou um impacto negativo na probabilidade de as empresas possuírem *website*, quando comparado com o ano de 2012 e mantendo tudo o resto constante. Este facto está de acordo com a análise descritiva dos dados. Já os anos 2013 e 2015 parecem não ter efeitos estatisticamente significativos nesta probabilidade relativamente a 2012.

Em geral, para os modelos com comércio *online* como variável de interesse, os EPM de ambos os modelos são muito semelhantes tanto em significância como em termos das magnitudes, como esperado. Assim a interpretação foca-se no modelo probit.

Observa-se que com respeito aos determinantes relativos aos recursos humanos, *ceteris paribus*, o aumento em 0.1 da proporção de mulheres ao serviço numa empresa provoca o aumento da probabilidade de essa mesma empresa realizar comércio *online* em 0.0081. De notar ainda que o aumento de 0.1 da proporção de trabalhadores com habilitações equivalentes ao ensino secundário e da proporção de trabalhadores com idades iguais ou inferiores a 30 anos, *ceteris paribus*, também provocam aumentos dessa probabilidade, com impactos de intensidades aproximadamente iguais a 0.007 e 0.005, respetivamente. Ou seja, consegue-se estabelecer uma relação entre o aumento das habilitações dos recursos humanos e uma maior utilização e aproveitamento das TIC nas empresas, conforme indicado na literatura. Como seria de esperar, também nestes modelos se estabelece uma relação entre empregados mais jovens com um maior aproveitamento das tecnologias por parte das empresas, fator que pode ser explicado pela maior adaptabilidade que o pessoal ao serviço mais jovem apresenta geralmente relativamente aos trabalhadores com idades mais avançadas.

Relativamente à dimensão das empresas, *ceteris paribus* e comparativamente às microempresas, é observável que as médias e grandes empresas têm uma maior probabilidade de adotar o comércio *online* (com aumentos na probabilidade iguais a 0.036 e 0.043, respetivamente). Este resultado seria de esperar uma vez que as médias e grandes

empresas possuem mais recursos tanto financeiros como humanos que permitem uma integração mais fácil e rápida das TIC nos seus serviços. No caso da comparação entre as pequenas e a microempresas, este efeito já não é estatisticamente significativo.

Já no que diz respeito aos setores de atividade económica, mantendo tudo o resto constante e comparativamente ao setor da Indústria, verifica-se que a pertença a qualquer um dos setores considerados gera acréscimos na probabilidade de realizar comércio *online*, exceto o setor da energia e água, cujo efeito não difere estatisticamente em relação ao setor Indústria. De entre os setores com efeitos estatisticamente significativos, é perceptível que o setor do Alojamento e Restauração é o que apresenta um efeito com maior magnitude na probabilidade de as empresas adotarem comércio *online*, com um EPM igual a 0.262, *ceteris paribus* e comparativamente à Indústria. Por outro lado, verifica-se que o setor da Construção e Atividades Imobiliárias é o que apresenta um impacto mais reduzido nessa probabilidade (0.027), isto quando comparado com o setor da Indústria e mantendo tudo o resto constante.

As variáveis *dummy* temporais permitem retirar a ideia de que, tal como esperado, os anos mais recentes são aqueles que têm um maior impacto na probabilidade da adoção do comércio *online*, comparativamente a 2012 e mantendo tudo o resto constante. A partir de 2015 os efeitos parciais médios são superiores a 0.05 pontos de probabilidade, enquanto em 2013 e 2014 os efeitos são mais reduzidos, ainda que estatisticamente significativos. Este facto pode ser explicado pela evolução das TIC, na sua globalidade, que levou as empresas a entenderem os benefícios inerentes à sua utilização e integração. Para além disso, muitas empresas devem ter pressentido a pressão competitiva no comércio *online* por parte das empresas concorrentes, que levou a que muitas das empresas apostassem em desenvolver e melhorar os seus serviços em tecnologias, de modo a conseguirem acompanhar ou até mesmo ultrapassar as restantes.

A terminar, destaca-se a diferença nos efeitos parciais individualmente significativos nas duas variáveis de interesse. A probabilidade de efetuar comércio *online* reage em menor magnitude às variações ao nível das proporções de grupos etários nas empresas, quando comparada com a probabilidade de possuir *website*, sendo que a diferença de impacto é cerca de 10 vezes superior nesta última probabilidade, quando comparada com a primeira. De notar a diferença na direção do impacto do ano de 2014, quando comparado com 2012.

De facto, em 2014, relativamente a 2012, enquanto a probabilidade de ter um *website* diminuiu, a probabilidade de realizar comércio *online* aumentou. Para os anos de 2016, 2017 e 2018 os aumentos de probabilidade foram maiores na posse de *website* do que na realização de comércio *online*. Ainda assim, apesar da diminuição em 2014, o ritmo de crescimento da probabilidade de posse de *website* foi superior à da probabilidade de efetuar comércio *online*.

## 6. CONCLUSÕES E SUGESTÕES PARA INVESTIGAÇÕES FUTURAS

O presente trabalho fornece algumas informações importantes no que diz respeito à adoção e utilização das TIC nas empresas portuguesas. Foram estimados modelos com variável de interesse binária para dados de painel com o objetivo de encontrar determinantes que têm impacto na probabilidade condicional de uma empresa portuguesa possuir *website* próprio ou do grupo e na probabilidade condicional de uma empresa portuguesa realizar comércio *online* (através do *website* ou de *apps*).

Os resultados sugerem que existe evidência estatisticamente significativa para diversas características, nos contextos organizacional e competitivo em que as empresas se encontram, que têm impacto na probabilidade de uma empresa portuguesa possuir *website* próprio e na probabilidade de uma empresa portuguesa adotar comércio *online*.

No que diz respeito à posse de *website*, verifica-se que, de entre as diferentes características dos recursos humanos testadas, apenas a idade das pessoas ao serviço tem impacto na probabilidade de uma empresa possuir *website* próprio. Deste facto retira-se que quanto mais jovens forem as pessoas ao serviço das empresas então maior a probabilidade de essas mesmas empresas possuírem *website*. Para esta probabilidade, verifica-se que tanto a dimensão como o setor económico não têm impacto estatisticamente significativo.

Em relação à realização de comércio *online*, os resultados já sugerem a existência de um conjunto mais variado de fatores que influenciam a probabilidade de uma empresa portuguesa adotar o comércio *online*. Quanto aos recursos humanos, destaca-se que o aumento da proporção de mulheres ao serviço das empresas nas empresas tem um efeito positivo nesta probabilidade. Para além disso e tal como verificado para a posse de *website*, o facto de uma empresa possuir uma maior proporção de trabalhadores mais jovens também influencia positivamente a sua probabilidade de realizar comércio *online*.

Relativamente aos níveis de educação dos recursos humanos e quando comparados com habilitações inferiores (equivalentes ao ensino básico), empresas na qual se encontram ao serviço trabalhadores com habilitações um pouco mais elevadas também têm uma maior probabilidade de realizar vendas através de *websites* e *apps*. De notar que não se observou a necessidade de as qualificações dos recursos humanos serem muito elevadas (iguais ou superiores à licenciatura). Para esta probabilidade e contrariamente ao que se verificou para a probabilidade da posse de *website*, a dimensão e o setor económico que as empresas se encontram também têm impacto. Relativamente à dimensão, observou-se que o facto de uma empresa ser média ou grande tem efeitos positivos para esta probabilidade. Já no contexto dos setores de atividade económica, estes têm impactos diferentes para a probabilidade de realizar comércio *online*, sendo que o setor do alojamento e restauração é aquele que tem um maior efeito positivo para esta probabilidade, isto quando comparado com o setor da indústria.

Para ambas as situações e no contexto tecnológico, apenas se incluiu a variável Internet, que por si só não representa um grande fator de diferenciação tecnológica uma vez que nos dias de hoje praticamente todas as empresas utilizam serviços de Internet. Por este motivo no contexto tecnológico não foram encontrados determinantes para nenhuma das probabilidades, sendo este facto uma limitação deste estudo.

Em ambos os casos, verificou-se que os últimos anos observados tiveram efeitos positivos superiores em ambas as probabilidades, o que pode ser explicado pela evolução das TIC em si como também pelo aumento da sua utilização a nível empresarial. E como se tem observado nos dias que correm, ainda para mais com a situação da pandemia provocada pelo COVID-19 que veio a alterar todo o funcionamento das nossas vidas, é muito provável que a sua utilização a nível empresarial seja cada vez maior e que as próprias tecnologias sejam cada vez ainda mais avançadas, permitindo às empresas atingir os seus objetivos de uma forma continuamente mais eficaz e eficiente. Por este motivo este estudo é de grande importância para as empresas portuguesas, fornecendo indicações importantes no que diz respeito às características dos recursos humanos e das empresas em si.

De modo a aumentar a probabilidade das empresas portuguesas aumentarem o nível de adoção das tecnologias de informação e comunicação, este estudo fornece assim

resultados que incentivam a contratação de candidatos mais jovens e de mais mulheres para os cargos disponíveis, sendo que este último incentivo ajudaria também na evolução da igualdade de género em Portugal.

Para investigações futuras sobre o tema abordado no presente trabalho, deixo como sugestão ao nível da modelação, o explorar de modelos sequenciais para dados binários de painel, que incorporem uma possível correlação entre a decisão do primeiro passo, de possuir ou não website, e a decisão do segundo passo, de realizar ou não comércio *online*. Por outro lado, em termos de análise de outros determinantes das variáveis de interesse, seria interessante explorar mais aprofundadamente o contexto tecnológico, dadas as inovações constantes que se verificam nas tecnologias. Podem ser incluídas diversas variáveis como a utilização de robôs industriais ou de serviço, a análise de *Big Data* ou o facto de uma empresa possuir ou não especialistas em TIC. Seria também interessante, ao nível dos recursos humanos, utilizar informações relativas aos donos das empresas, como por exemplo a idade, as habilitações literárias ou o género.

## REFERÊNCIAS BIBLIOGRÁFICAS

Arvanitis, S. & Hollenstein, H. (2001). The Determinants of the Adoption of Advanced Manufacturing Technology. *Economics of Innovation and New Technology* 10 (5), 377-414.

Bartel, A., Ichniowski, C. & Shaw, K. (2007). How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills. *The Quarterly Journal of Economics* 122 (4), 1721-1758.

Bayo-Moriones, A. & Lera-López, F. (2007). A firm-level analysis of determinants of ICT adoption in Spain. *Technovation* 27 (6-7), 352-366.

Bharadwaj, A. (2000). A Resource-Based Perspective on Information Technology Capability and Firm Performance: An Empirical Investigation. *MIS Quarterly* 24 (1), 169-196.

Comissão Europeia (2020a). *Integração de Tecnologia Digital por Empresas* [Em linha]. Disponível em: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=67076](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=67076) [Acesso em: 2021/03/1].

Comissão Europeia (2020b). *O Índice de Economia e Sociedade Digital (DESI)* [Em linha]. Disponível em: [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=67086](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=67086) [Acesso em: 2021/03/1].

Gabinete de Estratégia e Planeamento (2005). *Documento Metodológico (Quadros de Pessoal)* [Em linha]. Disponível em: <https://smi.ine.pt/UploadFile/Download/380> [Acesso em: 2021/04/19].

Galve-Górriz, C. & Castel, A. (2010). The relationship between human resources and information and communication technologies: Spanish firm-level evidence. *Journal of Theoretical and Applied Electronic Commerce Research* 5 (1), 11-24.

Gardner, S., Lepak, D. & Bartol, K. (2003). Virtual HR: The impact of information technology on the human resource professional. *Journal of Vocational Behavior* 63 (2), 159-179.

Haines, V. & Lafleur, G. (2008). Information Technology Usage and Human Resource Roles and Effectiveness. *Human Resource Management* 47 (3), 525-540.

Instituto Nacional de Estatística (2020). *Sociedade da Informação e do Conhecimento - Inquérito à Utilização de Tecnologias da Informação e da Comunicação nas Empresas* [Em linha]. Disponível em: [https://www.ine.pt/ngt\\_server/attachfileu.jsp?look\\_parentBoui=465221840&att\\_display=n&att\\_download=y](https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=465221840&att_display=n&att_download=y) [Acesso em: 2021/01/29].

Oliveira, T. & Martins, M.F.O. (2009). Determinants of information technology adoption in Portugal. In: *ICETE 2009 - International Joint Conference on e-Business and Telecommunications*. pp. 264-270, ICETE 2009 - International Joint Conference on e-Business and Telecommunications, Milan, Italy, 7 de junho de 2009.

Powell, T. & Dent-Micallef, A. (1997). Information Technology as Competitive Advantage: The Role of Human, Business, and Technology Resources. *Strategic Management Journal* 18 (5), 375-405.

Ravichandran, T. & Lertwongsatien, C. (2005). Effect of Information Systems Resources and Capabilities on Firm Performance: A Resource-Based Perspective. *Journal of Management Information Systems* 21 (4), 237-276.

Tornatsky, L., Chakrabarti, A.K. & Fleischer, M. (1990). *The Process of Technology Innovation*, Lexington, Mass.: Lexington Books.

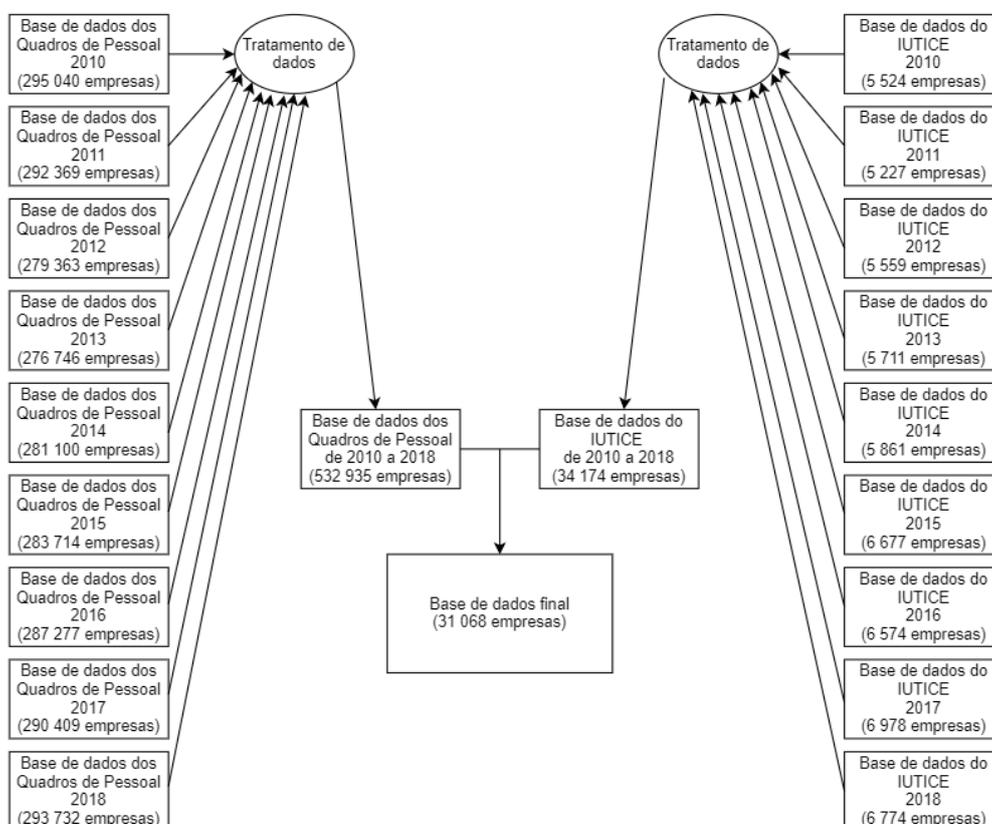
Townsend, A. & Bennett, J. (2003). Human Resources and Information Technology. *Journal of Labor Research* 24 (3), 361-363.

Verbeek, M. (2004). *A Guide to Modern Econometrics*, 2ª Ed. Chichester, Inglaterra: John Wiley & Sons.

Zheng, X. (2017). Research on the Application of Information Technology in Human Resource Management. *2017 International Conference on Smart City and Systems Engineering (ICSCSE)*, 85-88.

## ANEXO

FIGURA A1 - FLUXOGRAMA DA JUNÇÃO DAS BASES DE DADOS.



Fonte: INE, Inquérito à utilização de TIC nas empresas e GEP, Quadros de pessoal.

TABELA A2

## VARIÁVEIS EM ESTUDO

Variável	Descrição	Codificação	Tipo
NPC_FIC	Número de identificação da empresa (anonimizado)		Numérica
Ano	Ano de referência dos dados		Numérica
Internet	Acesso à Internet	0 = Não utiliza 1 = Utiliza	Binária
Website	Possui website próprio ou do grupo	0 = Não possui 1 = Possui	Binária
ComOnline	Realizou comércio online através de website ou apps (recebeu encomendas de bens/serviços)	0 = Não realiza 1 = Realiza	Binária

Pmulheres	Proporção de mulheres no total de trabalhadores		Numérica
PIdade30	Proporção de trabalhadores com idade igual ou inferior a 30 anos		Numérica
PIdade3154	Proporção de trabalhadores com idade entre os 31 e os 54 anos		Numérica
PCursoSup	Proporção de trabalhadores com habilitações equivalentes ao curso técnico superior		Numérica
PInf1Ciclo	Proporção de trabalhadores com habilitações equivalentes ao 1º ciclo		Numérica
PEnsBas	Proporção de trabalhadores com habilitações equivalentes ao ensino básico		Numérica
PEnsPS	Proporção de trabalhadores com habilitações equivalentes ao ensino pós-superior		Numérica
PBach	Proporção de trabalhadores com habilitações equivalentes ao bacharelato		Numérica
PLic	Proporção de trabalhadores com habilitações equivalentes à licenciatura		Numérica
PMest	Proporção de trabalhadores com habilitações equivalentes ao mestrado		Numérica
PDout	Proporção de trabalhadores com habilitações equivalentes ao doutoramento		Numérica
Setor	Setor de atividade económica da empresa	1 = Indústria 2 = Energia e Água 3 = Construção e Atividades Imobiliárias 4 = Comércio 5 = Transportes e Armazenagem 6 = Alojamento e Restauração 7 = Informação e Comunicação 8 = Atividades Financeiras e de Seguros 9 = Outros Serviços	Conjunto de Variáveis Binárias

dim	Escalão do pessoal ao serviço	1 = Microempresas 2 = Pequenas Empresas 3 = Médias Empresas 4 = Grandes Empresas	Conjunto de Variáveis Binárias
PIciclo	Proporção de trabalhadores com habilitações equivalentes ao 1º ciclo		Numérica
PEnsSec	Proporção de trabalhadores com habilitações equivalentes ao ensino secundário		Numérica
PEnsSup	Proporção de trabalhadores com habilitações equivalentes ao ensino superior		Numérica

Fonte: INE, Inquérito à utilização de TIC nas empresas e GEP, Quadros de pessoal, cálculos do autor.

TABELA A3  
ESTATÍSTICAS DESCRITIVAS DAS VARIÁVEIS EXPLICATIVAS

Variável Explicativa	Média	Desvio Padrão
Pmulheres	0.397	0.314
PIdade30	0.173	0.206
PIdade3154	0.672	0.255
PCursoSup	0.0001	0.007
PInf1Ciclo	0.005	0.037
PEnsBas	0.571	0.364
PEnsPS	0.008	0.057
P Bach	0.021	0.083
PLic	0.135	0.237
PMest	0.012	0.066
PDout	0.002	0.025
Indústria	0.273	0.446
Energia e Água	0.022	0.148
Construção e Atividades Imobiliárias	0.116	0.32
Comércio	0.207	0.405

Transportes e Armazenagem	0.048	0.213
Alojamento e Restauração	0.128	0.334
Informação e Comunicação	0.073	0.261
Atividades Financeiras e de Seguros	0.012	0.108
Outros Serviços	0.121	0.326
Microempresas	0.603	0.489
Pequenas Empresas	0.263	0.44
Médias Empresas	0.104	0.306
Grandes Empresas	0.029	0.169
P1Ciclo	0.576	0.365
PEnsSec	0.247	0.275
PEnsSup	0.177	0.276

Fonte: GEP, Quadros de Pessoal, cálculos do autor.

TABELA A4

FREQUÊNCIAS E PERCENTAGENS DAS VARIÁVEIS DE INTERESSE NO TOTAL DE  
OBSERVAÇÕES

Variável de Interesse	Valor da variável	Overall		Between		Within
		Freq.	Percentagem	Freq.	Percentagem	Percentagem
Internet	0	5733	11.16	5186	16.69	95.99
	1	45615	88.84	26283	84.6	99.26
Website	0	20568	40.06	17238	55.48	96.76
	1	30780	59.94	14844	47.78	96.93
Comércio Online	0	44366	86.4	28524	91.81	97.85
	1	6982	13.6	304	13.08	77.69

Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

TABELA A5

CLASSIFICAÇÃO DOS SETORES DE ATIVIDADE ECONÓMICA DAS EMPRESAS

Setor de Atividade Económica	Secção da CAE Rev. 3
Indústria	B + C
Energia e Água	D + E

Construção e Atividades Imobiliárias	F + L
Comércio	G
Transportes e Armazenagem	H
Alojamento e Restauração	I
Informação e Comunicação	J
Atividades Financeiras e de Seguros	K
Outros Serviços	M + N + S

Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

TABELA A6

CLASSIFICAÇÃO DA DIMENSÃO DAS EMPRESAS

Classe de Dimensão	Número de pessoas ao serviço
Microempresa	0 a 9
Pequena Empresa	10 a 49
Média Empresa	50 a 249
Grande Empresa	250 ou mais

Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

FIGURA A7 - OUTPUT DO COMANDO XTDESCRIBE ANTES DE ELIMINAR OBSERVAÇÕES.

```
. xtdescribe
NPC_FIC: 5.005e+08, 5.010e+08, ..., 5.032e+08      n =      31068
Ano: 2010, 2011, ..., 2018                        T =          9
Delta(Ano) = 1 unit
Span(Ano) = 9 periods
(NPC_FIC*Ano uniquely identifies each observation)

Distribution of T_i:  min    5%   25%   50%   75%   95%   max
                   9      9      9      9      9      9      9

   Freq.  Percent  Cum. | Pattern
-----|-----
  31068   100.00 100.00 | 111111111
   31068   100.00      | XXXXXXXXX
```

FIGURA A8 - OUTPUT DO COMANDO XTDESCRIBE APÓS ELIMINAR OBSERVAÇÕES.

```
. xtdescribe
NPC_FIC: 5.010e+08, 5.010e+08, ..., 5.032e+08      n =      23636
Ano: 2012, 2013, ..., 2018                        T =          7
Delta(Ano) = 1 unit
Span(Ano) = 7 periods
(NPC_FIC*Ano uniquely identifies each observation)

Distribution of T_i:  min    5%   25%   50%   75%   95%   max
                   1      1      1      1      1      5      7

   Freq.  Percent  Cum. | Pattern
-----|-----
   3024   12.79   12.79 | .....1.
   2945   12.46   25.25 | .....1
   2753   11.65   36.90 | ...1...
   2716   11.49   48.39 | ....1..
   2383   10.08   58.47 | ..1....
   2205    9.33   67.80 | 1.....
   2180    9.22   77.03 | .1.....
    680    2.88   79.90 | 1111111
    282    1.19   81.10 | .....11
   4468   18.90  100.00 | (other patterns)
  23636   100.00      | XXXXXXX
```

TABELA A9

## NÚMERO DE EMPRESAS CONSIDERADAS POR ANO

Ano	Frequência
2012	4 423
2013	4 637
2014	4 907
2015	5 507
2016	5 430
2017	5 765
2018	5 564

Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

TABELA A10

## NÚMERO DE EMPRESAS CONSIDERADAS POR SETOR DE ATIVIDADE

Setor de atividade económica	Frequência
Indústria	10 328
Energia e Água	872
Construção e Atividades Imobiliárias	3 685
Comércio	7 461
Transportes e Armazenagem	1 737
Alojamento e Restauração	5 149
Informação e Comunicação	2 672
Atividades Financeiras e de Seguros	227
Outros Serviços	4 102

Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.

TABELA A11

## NÚMERO DE EMPRESAS CONSIDERADAS POR DIMENSÃO

Dimensão	Frequência
Microempresas	17 449
Pequenas Empresas	8 727
Médias Empresas	6 439
Grandes Empresas	3 618

Fonte: INE, Inquérito à utilização de TIC nas empresas, cálculos do autor.