

SPECIAL ISSUE: REGULATORY PROBLEMS
AND SOLUTIONS IN SERVICES OF GENERAL
ECONOMIC INTEREST

ANDREA GUERRINI / RITA MARTINS
Introduction

JOSÉ AMADO DA SILVA / EDUARDO CARDADDEIRO
Is There Still a Role for Independent Regulators?

VÍTOR MARQUES
Novas Perspetivas para a Regulação do Setor
Elétrico

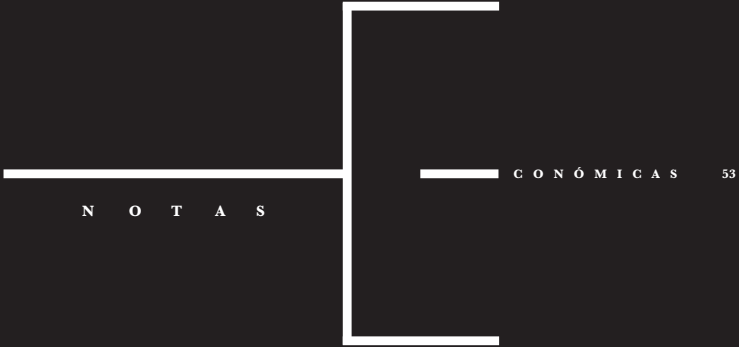
JOÃO CONFRARIA
Défice Tarifário, Rendas Excessivas e Privatizações

BLANDINA OLIVEIRA / ADELINO FORTUNATO
The Effect of Intermittent Renewable Energy
Generation on Electricity Prices: A Literature Survey

RITA MARTINS / MICAELA ANTUNES / PATRÍCIA SILVA /
ADELINO FORTUNATO
Tarifa Social de Energia: Génese, Incidência e Lições

DANIEL MURTA
Autonomous Vehicles and Public Transportation

SÉRGIO HORA LOPES
Aggregation in the Portuguese Water Industry: The
Case of Retail Service Operators



| | | |
|-----------|-----------|----|
| N O T A S | CONÓMICAS | 53 |
|-----------|-----------|----|

Informações aos autores (Submission guidelines)

<https://impactum-journals.uc.pt/notaseconomicas/about/submissions>

Publicação resumida e indexada por (Indexing and Abstracting)

EconLit
RePEc
REDIB
latindex

Design

Imprensa da Universidade de Coimbra

Execução gráfica

BOOKPAPER

Tiragem 200 exemplares

Propriedade

Faculdade de Economia
Universidade de Coimbra
Av. Dias da Silva, 165
3004-512 Coimbra

Apoios (Sponsors)

Esta publicação é financiada por fundos nacionais através da
FCT – Fundação para a Ciência e a Tecnologia, I.P.,
no âmbito do projeto UIDB/05037/2020

FCT Fundação
para a Ciência
e a Tecnologia

Número avulso

Portugal 7,5 €
Outros países 15 €

Condições de assinatura bienal

Portugal: 20 €
Outros países: 40 €

ISSN 0872-4733

ISSN Digital 2183-203X

DOI https://doi.org/10.14195/2183-203X_53

Depósito Legal 64932/93

Editors

Paulino Teixeira (Universidade de Coimbra, Portugal)
Tiago Sequeira (Universidade de Coimbra, Portugal)

Co-Editors

Mark Roberts (Pennsylvania State University, U.S.A.)
Lars Ehlers (Université de Montréal, Canada)
Markus Jäntti (Stockholm University, Sweden)
Alberto Bucci (Università degli Studi di Milano, Italy)
Manuel Gómez (Universidade da Coruña, Spain)
Constantin Chilarescu (Université de Lille, France)
Francisco Veiga (Universidade do Minho, Portugal)
Joana Pais (Universidade de Lisboa, Portugal)
José Varejão (Universidade do Porto, Portugal),
Pedro Brinca (Universidade Nova de Lisboa, Portugal)

Associate Editors

Guido Cozzi (University of St. Gallen, Switzerland)
Ludger Schuknecht (OCDE, Paris)
Qihong Liu (University of Oklahoma, U.S.A.)
Vasco Gabriel (University of Surrey, U.K.)
António Afonso (Universidade de Lisboa, Portugal)
Anabela Carneiro (Universidade do Porto, Portugal)
Esmeralda Ramalho (Universidade de Lisboa, Portugal)
Francisco Louçã (Universidade de Lisboa, Portugal)
Joaquim Ramalho (Instituto Universitário de Lisboa, Portugal)
Jorge Braga de Macedo (Universidade Nova de Lisboa, Portugal)
José Rui Figueira (Universidade de Lisboa, Portugal)
Miguel Portela (Universidade do Minho, Portugal)
Óscar Afonso (Universidade do Porto, Portugal)
Rosa Branca Esteves (Universidade do Minho, Portugal)
Rui Baptista (Universidade de Lisboa, Portugal)

Contents

Special Issue: Regulatory Problems and Solutions in Services of General Economic Interest

Andrea Guerrini / Rita Martins

Introduction • 7

José Amado da Silva / Eduardo Cardadeiro

Is There Still a Role for Independent Regulators? • 11

Vítor Marques

Novas Perspetivas para a Regulação do Setor Elétrico • 31

João Confraria

Défice Tarifário, Rendas Excessivas e Privatizações • 49

Blandina Oliveira / Adelino Fortunato

The Effect of Intermittent Renewable Energy Generation on Electricity
Prices: A Literature Survey • 73

Rita Martins / Micaela Antunes / Patrícia Silva / Adelino Fortunato

Tarifa Social de Energia: Génese, Incidência e Lições • 85

Daniel Murta

Autonomous Vehicles and Public Transportation • 103

Sérgio Hora Lopes

Aggregation in the Portuguese Water Industry: The Case of
Retail Service Operators • 123

Introduction

Introdução

Andrea Guerrini
Rita Martins

According to EU law there are different approaches in the services of general economic interest's organization, namely in the electricity, gas, water and wastewater sectors.

In the energy sector, since 2009, the “third energy package” rules the presence of an independent market watchdogs at country level, and the institution of an Agency for the cooperation of Energy Regulators (ACER). This Agency was established to promote the cooperation among national regulators and solve cross boarder issues. In electricity and gas sector, article 36 of the Directive 2009/72/EC (now article 58 of the Directive 2019/944/EU and article 41 of the Directive 2009/73/EC rules “General objectives of National Regulatory Authority” (NRA). The regulatory authority shall take all reasonable measures in pursuit of the following objectives: the promotion of competitive, secure and environmentally sustainable internal market in electricity and natural gas; the elimination of constraints in electricity and natural gas trade among member states; the creation of an efficient, reliant, adequate, secure and consumer oriented electric and natural gas system; the elimination of barriers to new market entrants; the creation of incentive based model for promoting efficiency through an effective competition, and ensure customer protection; finally the achievement of high standards of universal and public service in electricity supply and natural gas sector.

After the 2009 Directives, NRAs should be able to take decision in national energy sectors and to work with a full independence from any other public or private interests, political or economic clout.

According to OECD (2017)¹, the protection of regulatory Authorities from undue influence can be achieved by monitoring five main dimensions: i) Role clarity and responsibility (clarity of role and relations in law; outreach to governmental entities; strategic foresight function); ii) Transparency and accountability (performance reporting; transparent engagement; feedback and appeals; code of ethics); iii) Financial independence (funding sources; identification of needs; multi-year budget; autonomous management; external and internal audit); iv) Leadership (nomination; appointment; mandates; conflict of interest; exit process); v) Staff behaviour (recruitment; incentives; salary scales; employment restrictions).

¹ OECD (2017). *Creating a Culture of Independence. Practical Guidance against Undue Influence*. April. Paris.

First of all, the legislation should clearly define aims and duties of the NRA, assigning a power consistent with them; then the relationship between NRA and public actors should be defined by law, to avoid any powers overlap and conflict of interests (role clarity). The NRA should account its work, in terms of target achieved, output and outcomes; then, specific procedures should be set in order to assure a transparent stakeholder engagement (transparency and accountability). The achievement of an effective independence depends on clear role for budget assignment and full autonomy of NRA in spending its budget (financial independence). The board must be nominated and appointed according to transparent and accountable procedures, aiming to assure the greatest degree of members' legitimation; also the exiting process has to be developed under specific rules, as the cooling-off periods (independence of leadership). Finally, the internal organization procedures to attract and motivate staff should be based on a transparent performance evaluation system, not only based to monetary incentives, and linked to the system of strategic objectives set by the board plan; the integrity and freedom of action of staff member could be achieved also with employment restrictions concerning the interaction with regulated operators, and the exiting process (staff behaviour).

Despite this common set of rules at EU level and the mentioned clear definition of independence, European energy NRA often operate with different methods and procedures, and they not always achieve a full degree of independence².

Differently from the energy sector, there isn't a common governance framework in water and wastewater sector (WWS) at European level. According to the objectives of environmental protection, existing EU law mainly regulates technical requirements and defines the proper managerial approaches to reduce risks of water contamination and pollutions caused by effluent discharges.

This implies a wide heterogeneity on how WWS is regulated in Europe. Several countries of central and northern Europe (e.g. Germany, Austria, the Netherlands, Norway and Finland) have a locally based model, with a power of regulation exerted by local authorities within a general framework set at national level and concerning common rules and objectives for the sector. In the rest of European countries there are two different models: i) ministerial administrative bodies (Greece, Belgium Flanders and Spain) or regulation by contracts (France); ii) independent/autonomous regulators, i.e. mono sectorial regulation on water sector or environment water and waste (UK, Portugal); competition authority (Denmark, Estonia); NRAs on energy (Albania, Bulgaria, Republic of Ireland, Leetonia, Lithuania, Malta, Italy, Hungary).

WAREG, the European association of water regulators, gathers this last group of regulators, with the aim to: monitor the evolution of the sector in each country; exchange best practices in terms of regulation among members; support members and non-members states to design their own system of rules; promote water regulation in EU and national institutions. The umbrella of WAREG³ shows a wide variety of organizational approach to the regulation of WWS, including utilities regulators, ministerial administrative bodies,

² CEER (2016). Safeguarding the independence of regulators Insights from Europe's energy regulators on powers, resources, independence, accountability and transparency. December. Brussels.

³ WAREG (2015). Institutional regulatory frameworks. A comparative assessment. Available at <https://www.wareg.org>

and competition authorities; the association gathers regulators with a different degree of independence and scope of action, joining ministerial agencies and full independent regulators, pure water NRAs and multisectorial authorities.

In order to investigate, justify and support this variety of solutions in water regulation, WAREG refers to the so called “contingency approach”. The contingency approach is followed to design a new structure process, adapting its features to the external context (input) and to the targets which should be achieved. According to contingency theory of organizational science, the Water regulation in a country depends on specific exogenous variables like cultural/economic environment (economic growth, degree of market liberalization, presence of private investors, attention to water rights and to the environment) and its WWS (size of utilities, ownership, number of operators). WWS governance could impact in expected outcomes (investments growth rate; trend of tariff; variation of quality indexes; access to water; firms ‘failures; citizens ‘perception of the sector). Following this theory, WAREG investigates how regulators can improve or better design their organization through this scientific method.

The Covid 19 pandemic has reinforced the relevance of guaranteeing the fulfilment of obligations of public service, including universality, affordability and transparency, which are mandatory in services of general economic interest. These requirements are critical to avoid exclusion and to allow the adoption of alternative forms of work and education that also prevents the dissemination of the virus in the community. In addition to the immediate measures imposed by governments and regulators, such as non-disconnection safeguard or deferred payment, the pandemic situation and the subsequent economics effects will certainly require the discussion concerning new challenges for economic regulation.

This special issue of *Notas Económicas* on “Regulatory problems and solutions in services of general economic interest” is designed to be a tribute to Professor Adelino Fortunato.

The economic regulation of utilities is one of the research and teaching areas of Adelino Fortunato, Professor of Economics and researcher at the Center for Business and Economics Research (CeBER) at the University of Coimbra. The articles included in this special issue represent contributions from authors who share common interests on the topic of economic regulation.

In the article “Is there still a role for independent regulators?” José Amado da Silva and Eduardo Cardadeiro address the legitimacy of independent regulators in the democratic organization of the State. They discuss legitimacy and its role inherently to independence and in accordance with the requirements of legitimacy and accountability. The authors define the conditions for regulators to be appropriately integrated in State institutions suggesting a triangle to support independent regulators, whose apexes are Transparency and Participation, Adequate Resources, and Appeal and Prompt Judicial Response.

Focusing on sector-specific regulation, Vitor Marques presents the evolution of electricity regulation in Portugal, since the beginning of the century in his article “Perspectives for the electricity sector regulation in the context of energy transition”. By identifying the main technological and organizational changes or environmental requirements as well as taking into account the increased scrutiny on regulators’ action that has taken place in recent years, the author presents key drivers and challenges for the regulation of the electricity sector in Portugal in the European context.

Still in the context of the electricity sector, in the article “Tariff deficit, excessive rents and privatizations”, João Confraria provides a reflexion on the welfare impacts from several governments’ policy options of in favour of a tariff deficit. Additionally, the author addresses the rationale for the excessive rents, specifically in terms of efficiency or in terms of their potential to increase privatization revenues from strategic segments of the Portuguese energy sector. This is an important contribution to the public and scientific debate in Portugal regarding excessive rents.

The article by Blandina Oliveira and Adelino Fortunato, “Effect of intermittent renewable energy’s generation on electricity prices: A Literature Survey”, is a contribution focused on the apparent contradiction between the decrease on generation costs from renewable energy, on the one hand, and the increase of retail electricity price, on the other hand. Based on the literature review the authors found evidence for the merit-order effect, meaning that an increase in intermittent sources generation would reduce the spot electricity market price. Concerning the scarcer studies applied to retail market, no consensus was found about the sign of the impact of the support of renewable energy sources over retail electricity price. This approach, underlining the trade-off between the need to contribute to the promotion of environmental goals and ensure the affordability of electricity services, exemplifies the complexity inherent to the regulation of the electricity sector.

Rita Martins, Micaela Antunes, Patrícia Pereira da Silva and Adelino Fortunato, deal precisely with affordability in the energy sector in the article “Energy social tariff: origin, prevalence and lessons”. The authors address the rationale and impact of social tariffs for energy services, to subsequently derive recommendations regarding their extension to other utilities such as water or telecommunications. The authors also emphasize the relevance of social tariffs in the context of the social and economic crisis brought about by the Covid-19 pandemic to ensure that no one is left behind due to unaffordable essential services.

More oriented to environmental concerns, Daniel Murta’s study “Autonomous Vehicles and Public Transportation” contributes to the ongoing economic debate on the future of transportation. It provides an analysis on the broad opportunities and challenges for welfare from autonomous vehicles. Focusing at the same time on potential contradictions such as discouragement from the use of public transportation, which might have negative effects on congestion and pollutant emissions, it reinforces the need for support and for the regulation of public transportation.

The last article deals with sector organization in the Portuguese water industry, one of the components of its structural regulation. Sérgio Hora Lopes discusses “The aggregations in the Portuguese water industry – the case of retail service operators” following the literature justification for integrations. The focus is on the aggregations which occurred in the last decade in Portugal. Contrarily to most integration processes which took place in the late 1980’s until the middle of 2000/2010, the most recent are mainly aggregations of retail operators. Moreover, the most recent integrations correspond to aggregations of smaller operators in terms of the population served and the area covered. According to the author, the different contexts under which aggregations occur set equally different conditions for their success.

José Amado da Silva
Universidade Autónoma de Lisboa
jamado@autonoma.pt
orcid.org/0000-0002-5449-9308

Eduardo Cardadeiro
Universidade Autónoma de Lisboa, CICEE
Economic and Business Sciences Department
ecardadeiro@autonoma.pt
orcid.org/0000-0001-8819-3326

DOI: https://doi.org/10.14195/2183-203X_53_2

Is There Still a Role for Independent Regulators?

Haverá (ainda) um Papel para Reguladores Independentes?

José Amado da Silva
Eduardo Cardadeiro

Received for publication: July 20, 2020

Revision accepted for publication: April 9, 2021

ABSTRACT

Although regulators have been around for three decades, their Independence is still not fully understood as an administrative institution in the governance of modern states. Neoclassic approaches support their existence, in addition to valuable contributions made from the perspective of transaction costs, institutional economics and political economy. To improve the trade-off between independence and accountability of regulators, several conditions for the appointment and functioning of regulators have been put forward, but this remains a hot topic. The authors discuss the integration of regulators in the democratic organization of the State, address their legitimacy, and discuss the powers and duties inherent to independence, consistent with the requirements of legitimacy and accountability. Finally, they define the requirements for regulators to be appropriately integrated in State institutions and suggest a “triangle” to support independent regulators.

Keywords: Economic regulation; independent regulators; regulators’ accountability; regulators’ legitimacy.

JEL Classification: L51; K23; H10.

RESUMO

Embora proliferando há três décadas, a independência dos reguladores ainda não é completamente compreendida como uma instituição administrativa na governação dos Estados modernos. Tem fundamentação na teoria neoclássica, mas a economia institucional e dos custos de transação, bem como as perspetivas da economia política também têm dado contributos relevantes. Para melhorar o trade-off entre independência e responsabilização, têm sido propostas várias condições na nomeação e funcionamento dos reguladores, mas o debate mantém-se aceso. Os autores debatem a integração dos reguladores na organização do Estado, a sua legitimação, e discutem as competências e deveres compatíveis com essa legitimação e com a sua responsabilização. Finalmente, definem condições necessárias à

adequada integração dos reguladores nas instituições do Estado e propõem um “triângulo” para suportar a regulação independente.

Palavras-chave: Regulação económica; reguladores independentes; entidades reguladoras; organização do Estado.

1. THE ROLE OF INDEPENDENT REGULATORS IN STATE ADMINISTRATIVE ORGANIZATION

Independent economic regulation dates from the 19th century but become particularly important in the 1980s with the proliferation of regulators – or independent administrative authorities (IAA), as they are called in State administration – in democratic societies and developed or developing economies (Rosanvallon, 2008; Tucker, 2018).

The fundamentals of neoclassical economic theory, resorting to a whole theory on incentives, imperfect information, and to the reasons behind market failures, but also some developments of the behavioural economics, such as intertemporal inconsistency, have been used to explain the need for these IAA as agents of economic policy, as a means of overcoming issues related with commitment and sub-investment, information asymmetry, competitive neutrality in relationship-specific assets' contexts, with long lifecycles and strong economies of scale and scope – basically, in the public utilities (Amstrong and Sappington, 2007; Bawn, 1995; Edwards and Waverman, 2006; Gilbert and Newbery, 1988, 1994; Laffont and Tirole, 1986; Majone, 1996; Waverman and Koutroumpis, 2011). The independence of these IAA has been duly grounded to face the risks of capturing and influencing the different stakeholders, including those of regulated companies and political power (Dal Bo, 2006; Stigler, 1971).

However, the existence of the IAA with the power to decide key aspects of markets that provide essential services and have a high impact on the well-being of the population, without that power deriving from traditional sovereign bodies in representative democracies, has raised the issue of their legitimacy in state administration. A common approach in the literature has been to present these independent regulators as an institutional response from the State to deal with the State's own failures, deriving from the limitations of economic policy, assuming that the delegation of power in technocratic bodies¹ allows that better decisions are taken than those by “political” bodies (Bénassy-Quéré et al., 2019), which is in accordance with the theory on transaction costs, according to which the state's own institutional arrangements meet the needs to respond more and more efficiently to social contexts in constant change (Williamson, 1985; 1998), as well as with the new institutional economy, according to which only after institutions have been correctly designed and the property rights have been defined is it possible to discuss the governance of the said institutions and analyse the short-term decision of economic agents (Williamson, 2000).

When confronting technocratic decisions – by the IAA – and political decisions – by executive bodies democratically elected – some authors have focused on the conditions in which decisions made by the IAA are better than those taken by political bodies, mostly considering the nature of those decisions and the control mechanisms of said decisions. Several elements contribute to this line separating which decisions should be made by which institution, namely, the technicality of the decision, the stability of social preferences

¹ The idea of the regulator as a technocrat is associated with the creation of this figure, namely in the economic field, to overcome market failures. Considering the link with overcoming State failures and the consequent integration in the democratic organization of the State, the characteristics required for regulators go beyond their technocratic ability, and include, at least, their independence from a hierarchy structure dependent on political power, since it is from the combination of these two features that one might derive the superiority of certain decision making.

related with regulated activity, the ability that voters have of seeing the results, the risk of intertemporal inconsistency of decisions, the impact on intra and intergeneration surplus distribution, and the need to consider incompatible goals or the involvement of pressure groups (Alesina and Tabellini, 2007; Maskin and Tirole, 2004; Tucker, 2018).

Obviously, these criteria do not always allow for a clear separation of which decision should be made by which institution. However, not only do they help to set a rationale for the allocation, but they also allow for the definition of intermediary institutional solutions, as when political institutions define the objectives and delegate decisions to attain those objectives on technocratic institutions with operational independence. This is what has mostly happened with the IAA in more developed economies (Bénassy-Quéré et al., 2019).

In any case, the existence of a rationale to accept more efficient decisions by IAA in specific circumstances may provide the right behavioural incentives and resolve the issue of operational optimization in public governance, though it does not resolve the (never ending) tension between independence and accountability in the legitimation of IAA, which appears almost as a “foreign body” in the administrative organization of a democratic State built on the principles of the Social Contract and the separation of powers inherited from 19th century Enlightenment. This never quite understood “foreign body” has been increasingly questioned, both due to people’s decrease in confidence in governing bodies and to the pressure resulting from poor economic results over more than a decade in a large number of economies in which this model has been in place (Tirole, 2016; Tucker, 2018). This makes the questioning on legitimacy particularly relevant in order to make it compatible (as best as possible) with its inherent independence, especially since, after the global financial crisis at the end of the last decade, we now find ourselves in a new crisis, this time caused by the Covid-19 pandemic.

2. THE (DEMOCRATIC) LEGITIMACY OF INDEPENDENT REGULATORS

Given the fast change in social and economic conditions, hindered or sped by the current pandemic, to question whether economic regulation makes any sense requires a prior answer to the wider issue of regulation in general and its legitimacy in the context of a country’s political governance. Moreover, it requires a prior answer to the issue of the democratic legitimacy of its existence and statutes, in the two aforementioned contexts.

We must recognise that, in the field of political science, these matters have not been deeply analysed, the issue of their legitimacy and their need or merit or their action remaining open for discussion. Additionally, many of these discussions often occur in tense environments due to specific situations, ignoring or belittling the debate on the referred topics.

Interestingly, in an important publication, Laffont and Tirole (2000), when discussing in depth the fast development of telecommunications, almost at the end of their book (p. 279), ask: “Toward the demise of regulation?”

In a paper published in “Finance & Development”, by the International Monetary Fund, Frieder (2020, p. 5) acknowledges (only now?) that “The COVID19 pandemic strikingly illustrates the intersection of politics, economics and other considerations”.

Stating that the “Political Economy is about how politics affects economy and the economy affects politics” (p. 6) (thus showing the deep link between the two, the play on adjectives and nouns reinforcing the conceptual blending), the author recalls that Adam Smith, David Ricardo and John Stuart Mill defined themselves as “political economists”, history, namely 20th century history, being responsible for the separation of the two fields.

Frieder (2020) advocates that, in the past 50 years, political economy has become increasingly important, both in economic and in political science, but his statement that “politicians can be thought of as analogues to firms, with voters as consumers or governments as monopoly providers of goods and services to constituent customers” (p. 7) shows that he still considers the idea that the economic model is prevalent (though only at an initial stage) and that political principles are subordinate to economic reasoning. This adds to the need for an analysis of the legitimacy of regulation, particularly economic regulation, from a wider perspective of the democratic process and its governance.

This debate is framed by Tucker (2018) from a point of view of inevitable change of State organization, which Tirole (2016) calls “a new concept of State”, motivated by the answer to two questions: Do government structures work and help people to live good and happy lives, with realistic expectations and opportunities? In democratic countries, are governments in touch with and shaped by people’s motivations so that their representatives may limit, supervise and reform “non-elected governing bodies” (i.e. IAA)? The surprise evidenced by the author in regard to the fact that these societies show little interest in debating and setting principles for a new governing geometry reinforces the importance of this issue in the context of Political Economy and not simply of Economy.

Barak Obama² offers a similar perspective: “Our regulatory system must protect public health, welfare, safety and our environment while promoting economic growth, innovation, competitiveness, and job creation... It must identify and use the best, most innovative, and least burdensome tools for achieving regulatory ends”.

But is this not a demanding set of possibly conflicting goals when, in regulation, you forget the key financial issues, typical in a government plan? The regulatory system is then more and more viewed from a political perspective. What is its legitimacy in representative democracy?

Rosanvallon (2008), in a remarkable text on the IAA, historically describes how these bodies were a response to the loss of confidence in parties, namely in the USA, in relation to the decisions on the railroads at the end of the 19th century.

It is from this distrust of the democratic process based on parties, whose specific interests are opposed to the construction of the common good, that Rosanvallon refers to a double legitimacy, which explains the emergence of the IAA.

By advocating and explaining this legitimacy, in a chapter curiously named “Une révolution encore indéterminée”, he considers two types of institutions, the two first figures of this double legitimacy: the independent administrative authorities of “surveillance et de régulation” and the constitutional courts.

² Quoted in “The regulation of tomorrow: Rulemaking and enforcement in an era of exponential change”, Deloitte University Press, 2015.

He considers that the first benefit from a legitimacy of impartiality, due to the way they were created and are composed. He refers in particular that “Elles ont été soit des créations du pouvoir législatif soucieux de limiter et encadrer un pouvoir exécutif jugé très partisan, soit suscitées pour le pouvoir exécutif lui-même, prêt à se dépouiller d’un certain nombre de ses attributions pour restaurer une crédibilité affaiblie ou se décharger des tâches pour la gestion desquelles il n’estime pas disposer de compétences nécessaires” (p. 22).

The quote is purposefully long to evidence the way the IAA are formed and the reasons underlying the way that, something deriving from two sources: the Parliament, and due to the distrust of government’s partisan perspective (even when ruling with absolute majority?); and the government and its attempt to self-constrain, either because it strategically wants to regain legitimacy or because it acknowledges that, due to its structure, it does not have the adequate and sufficient competences to meet society’s wishes.

It is also evident that this double source for the creation of the IAA will eventually influence its composition while adding a new requirement to base the legitimacy of these authorities: their technical and scientific skills which, on their own, do not ensure their legitimacy, as we will see later in the text.

Curiously, the way Rosanvallon describes the legitimacy of the Constitutional Court appears to be an example of that potential distrust of Parliament towards the government, as it conveys the potential distrust of parliamentary decisions.

In our opinion, this distrust, if combined with the distrust of partisan priorities over society’s needs, has special importance in the case of absolute majorities.

Therefore, according to Rosanvallon, the function of the Constitutional Courts is to “encadrer la production législative en la soumettant à une contrainte renforcée de généralité par rapport à l’expression majoritaire. Leur proximité est liée au caractère réflexif de leur intervention” (p. 22).

The combination of the features of the two types of IAA leads to two essential features being required of them: impartiality and reflection, the latter being implicit in the technical and scientific skills required to declare that legitimacy.

There is yet another requirement that Rosanvallon adds to the two mentioned earlier: proximity (p. 268 onwards) which is the key requirement to overcome the issue of the “contrainte renforcée de la généralité par rapport à l’expression majoritaire”.

As Rosanvallon refers (p. 10), the majority principle became both a justification to represent the will of the people and a decision technique, blending both without considering their different character. This has led to a distortion of democracy.³ In fact, democracy in this general sense should express society in general. The blending of the two aforementioned principles leads to the majority becoming confused with the total, the more so that a part stood for the whole and the elections stood for the whole mandate.

Proximity meets that need to pay attention to each and every one, an attention to specificity, a policy of presence rather than representation and a democracy of interaction (p. 265).

It is obvious that a deep discussion of this issue is not possible in this paper. Noteworthy is the fact that the IAA and, therefore, Regulators may find their legitimacy in considering

³ This becomes a means to not give an advantage to anyone and place all voters in the same position so that the value of each member of the community is not questioned.

democracy a dynamic process, in constant change, in which all citizens may and should be active agents. That is democratic utopia, for sure, which irrevocably assumes an ethic reflecting a policy of attention (p. 280).

Important authors that should be mentioned here are (Rosanvallon, p. 282) Nussbaum (2006, 2007), who urges others to recognise the philosophical contribution of literature (and not just of technical and the so-called pure science), due to the complexity of characters and situations, and Wittgenstein (1965), who has constantly fought what he described as the “*désir constant de généralisation*” and “*mépris pour les cas particuliers*”.

Rosanvallon even considers that Wittgenstein, by inviting philosophers to base their research on firstly paying attention “*au sol raboucheux de l’ordinaire*” (an expression that philosophers have long discussed in the field of democratic ethics), creates a wider redefinition of democracy in terms of “conduct” rather than a simple structure, and thus pushing regulation, and in particular, economic regulation, to the field of Behavioral Economics, and, consequently, to the field of Ethics.

Rosanvallon considers, then, that the IAA’s legitimacy is grounded on three elements: impartiality, reflection, and proximity, which he describes in three separate chapters: the legitimacy of impartiality, the legitimacy of reflexivity and the legitimacy of proximity. We consider that the three “legitimacies”, in terms of a really democratic regulation, cannot be separated because there is no impartiality if there is no proximity (as some will be ignored) and if there is no reflection – the key to any technical and scientific decision.

These are the requirements (and they are neither few nor easy) that underlie the legitimacy of the existence and the action of an IAA. An IAA that does not comply to them, cannot justify its existence (Rosanvallon 2008, p. 3).

Rosanvallon even adds another requirement, a more operational one, which he calls legitimacy and the idea of procedural justice (p. 270), stating that “*la légitimité des agents publics est fonction des qualités de “justice procédurale” attachés à leur comportement*”, considering it “*l’entrelacement de l’impartialité et de la proximité*”.

It is important to realize that the claim goes beyond the formality of “delegated legitimacy”, a result of a formal action by a state authority, rather describing behavioural requirements, in line with Tucker’s (2018) ideas when he states that a regulator’s credibility,⁴ its active principle, depends on a wide public discussion and acceptance, and its legitimacy depends on the right balance with the three powers of the State⁵ through careful delegation, constant surveillance and public debate. He adds that credibility requires legitimacy.

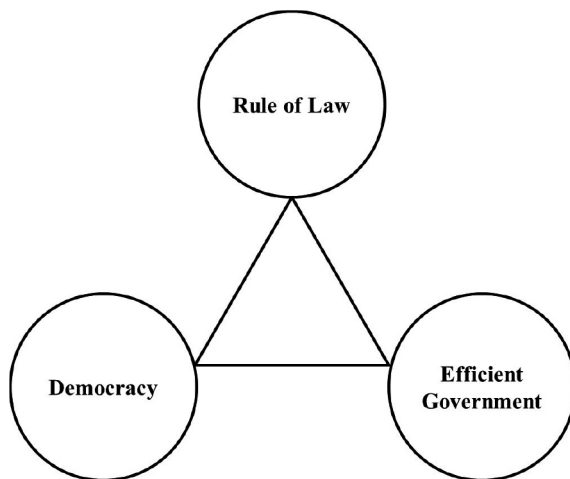
In the search for legitimacy in the framework and organization of the State, Tucker (2018) states that a modern constitutional state is based on two “triangles”: that of functions and values, whose apexes are Democracy, Rule of Law and efficient administration of Government (Figure 1); and the triangle of the institutional form, the separation of powers, whose apexes are the Judicial power, the Legislative power and the Executive power.

⁴ Here, Tucker refers to the national banks, but in a context that can be applied to other IAA and to independent regulators.

⁵ Legislative, judicial and executive branches.

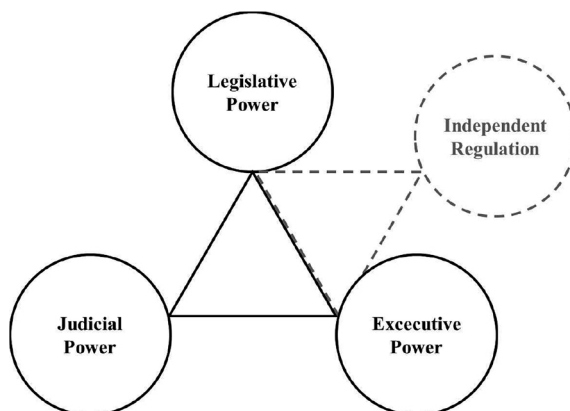
However, this balance may be jeopardized by the division of the executive power between elected representatives and non-elected technocrats, squaring this latter triangle into a rhombus (see Figure 2).

Figure 1: Triangle of Functions and Values



Source: Tucker (2018).

Figure 2: Triangle of Institutional Form



Source: Tucker (2018).

In the attempt to resolve the framing of this “foreign body” in the structure of the State, either to directly reinforce its legitimation, or to indirectly attempt to meet all the legitimacy requirements mentioned above, several proposals have been put forward and measures implemented, namely to the creation of IAAs, to the appointing its directors, and regarding the resources and IAA’s functioning, among others. However, almost all are just tools to support the independence of the IAA (OECD, 2016).

3. INDEPENDENCE: POWER OR DUTY?

We will now focus on the adjective “independent”, which is in the name of these Authorities and whose scope and interpretation seem to be at the core of the combination between these authorities and the institutions supporting modern democratic societies.

Having already established that, from the perspective of the dynamic construction of a democracy, regulation has a role in that construction, it is important to discuss whether the reasons listed for its existence and its usefulness might become reasons against its existence, if those requirements are not met or if its procedural construction is incorrect or insufficient.

We consider that the biggest danger lies in the interpretation and use (rather, abuse) of the concept of “independence”.

First of all, independence cannot mean anonymity, silence, or being above the organization of the State, since the actions of an IAA requires a constant interaction with state sovereign bodies, regulated economic agents, citizens (the ultimate beneficiaries of its actions) and all other stakeholders (OECD, 2016; Tirole, 2016). Moreover, this independence must be necessarily limited by its mandate, conflict of interest resolution mechanisms and the supervisory power of state sovereign bodies (Tirole, 2016).

Noteworthy is the reference to the speech given by Cathryn Ross, then Chief Executive of OFWAT, at the 2016 RPI Annual Westminster Conference, with the title “The purpose and functions of economic regulation”, in which she started by clarifying what is meant by “independent economic regulation”. It is important to point out that our concern with independence is not restricted to economic regulation but to all types of regulation, as evidenced in our rationale. Ross starts with an interesting phrase: “Independence: Yes, but it is not a vacuum”. It is true that she immediately focuses on independent economic regulation, considering that there lies a paradox: “Regulation exists because the sectors they regulate matter – not only to their customers, but also to our economy and society more widely”, adding that “The political salience of sectors we regulate mean it shouldn’t be surprising when politicians take a keen interest” (p. 2).

Both observations, though in different contexts, place regulation and its action in the field of politics, as they should, thus making the need for legitimacy important.

It also raises another issue, in the specific level of Economics: is economic regulation (and should it be?) a sectoral regulation? Or is it, in terms of a global response to society, which is also a stakeholder, an integrated regulation? And how can we separate it from governmental politics?

These issues should be discussed, and we will return to them later.

Let us then go back to the potential issue of the “distortion” of independence. Here, Ross fills that “vacuum” she referred to in the beginning when she recognised “our legitimacy underpins our independence...”, adding that “I think our legitimacy is closely linked to our accountability” and concluding with a careful “broad accountability”. That is, “accountability” is not only for stakeholders (and here there is already a generalization in view of “classical” regulation) of the regulated sector, but to society in general, being “judged” for all the consequences that derive from that regulation, at the risk of accountability in political terms and in terms of supervision, whose functions remain not entirely separate from regulation. In fact, this confusion is increased by the fact that there are several IAA with both functions.

Noteworthy is to ensure that “independence” is not misused. But how can we ensure that its compositional structure and its behaviour is compliant?

As Ross states, the regulator is accountable, not only before regulated companies and their consumers, but also before a wide range of stakeholders, investors, NGOs, and, this being a key, “how externalities, such as environment impacts are dealt with”.

This is obviously another challenge to clarify the possible or preferable scope of economic regulation adoption.

This concern with accountability is similar to Rosavallon’s question (p. 259): “*Qui gardera les gardiens?*”, a question everyone asks, especially in view of a misunderstood status of independence.

In any case, Tirole (2016) suggests that the legislative (parliamentary) qualified majority should have the power to suspend IAA leaders, not based on specific, politically sensitive decisions but based on the overall behaviour of these authorities. Without questioning the legitimacy of action of elected political bodies (mandated to govern the State), the issue lies in establishing the limit between a globally negative assessment of a regulator because it does not comply with the competences of its mandate, does not deliver the expected results to society, and a negative assessment that results from dissonance between what the political decision would be and the regulator’s technocratic decision.

It is important to realize that it is exactly because those decisions tend to be different in specific moments and considering (with due distancing) that the decision taken by the technocrats are more correct, that is at the basis of creating the IAA; therefore, the tension between political power and the IAA is natural and inevitable. Moreover, there tends to exist greater hostility from political power towards the IAA in certain stages of the electoral cycle, hostility that is easily followed by voters, given the referred limitations to assessing the regulator’s action (Tirole, 2016).

We acknowledge that the issue will remain unresolved and this is not the place to discuss it further. As Rosavallon said “*Elle n’a pas de solution logique et soulève les mêmes difficultés formelles que la notion d’auto-fondation*” (p. 259).

For now let us express agreement to Constant (1991): “*On ne peut donner une garantie à la garantie elle-même*”.⁶

⁶ Cap. XV, *Des garanties contre les abus du pouvoir préservateur* (Constant, 1991). See Rosavallon, p. 260.

This is an issue that derives from the binomial independence/accountability and that is difficult or impossible to resolve; a binomial which political science can and should always consider in its quest to improve the dynamic democratic process.

It is true that the several conditions listed in the literature to “balance” that binomial must include a transparency in its action in which all stakeholders and society in general see the action as being in conformity with those conditions, which requires ethically irreproachable behaviour from the institution and, as a consequence, will lead to society’s trust in it.

This trust, from our perspective, is a necessary condition for the legitimacy of the regulator’s action, which must be kept and conquered again at every new action.

However, if the regulator does not have that behaviour but does not break any law or regulation, is it possible to question the existence of an Independent Authority?

In a specific study on airport regulation in the United Kingdom, and in face of a group of tests he considered might question the quality of that regulation, though the regulator had not considered it as such, Littlechild (2018) stated: “This is not surprising: regulators commonly have difficulty in conceiving that their services might not be needed”. The author goes further, suggesting that this assessment should be extended to other regulated sectors to avoid over-regulation. Tucker (2018) also draws attention to the risk of over-regulation through super-powerful non-elected citizens.

Noteworthy here is to again mention Rosanvallon, who contributed to defining the scope of independence, which is referred in his texts as a requirement for impartiality in the regulator’s decision. And you cannot be impartial if you depend on one of the parties!

Rosanvallon sums up what being independent entails in a sentence (p. 149): “Les membres de tels organismes peuvent même considérer qu’ils ont un “devoir d’ingratitude”, ... pour être à la hauteur de leur tâche. C’est une situation aux antipodes de celle de l’élu...”.

The author draws the attention to impartiality, which is a quality and not a statute, it must be constantly built and validated (p. 152). He emphasizes that this validation is based not only on the IAA composition but also on its collegiality and, in particular on its decision-making processes.

From our own experience, the effect of collegiality, one primarily based on difference⁷ rather than homogeneity in skills and in background knowledge, is key to impartial, reflected decision-making.

On the other hand, it is crucial to resort to public consultation on regulatory decisions, previously submitting the draft decision and its grounds to public scrutiny in a document that describes “the probable decision”. This document that is made available for public consultation should be written after collection and analysis of data, collected in proximity, duly reflected by the IAA’s services that process them and approved (with impartiality) by the decision-making body. At a later stage, the final decision should be published with a report on the referred consultation and an analysis of all contributions.

It is important to ensure free participation of each individual and group, and the current communication means may improve that. The regulator should take all contributions into account for its final decision.

⁷ Rosanvallon (p. 149) refers that “La diversité corporative est souvent plus important qu’une simple compétence pour prendre une bonne décision”, and even claims that, as a consequence, the Independent Administrative Authorities have an advantage over the common ruler decision-maker.

Collegiality has a key role here, since, as Rosanvallon states, the instances of regulation are not what he calls “congrès d’ambassadeurs!”, but “reunions de personnes sans mandat”, in which members have equal voice and are recognised by the others as competent, thus creating a type of “collective intelligence” due to the rational conditions of a “délégation rationnelle” (p. 148).

The IAA decision-making process is therefore a critical element to its balanced independence, which may condition the regulator’s own existence due to it not being needed any more, as suggested by Littlechild (2018). This careful consideration of the extension of regulatory activity is evident in Khan’s (1981) words “If I were asked to offer one single piece of advice to would-be regulators, on the basis of my own experience, it is that as they perform their every single regulatory action they ask themselves: «Why am doing this? Is it really necessary? »”, which shows extreme care in not abusing the independence that is especially awarded to the IAA.

4. WHICH GOVERNANCE STRUCTURE FOR REGULATION?

For now, we will not discuss the issue of continuity for regulation institutions but rather focus on an issue prior to the current challenges to economic regulation, an issue raised by the high geographical and interdisciplinary inter-penetration of economic decisions as referred above by Ross. In fact, since economic decisions have relevant environmental and social consequences, and given the possible inequality of income and living conditions at national level and in comparison, with other countries also heavily influenced by the current financial (supervisory and regulatory) conditions, does it still make sense to have an autonomous economic regulation in other fields, for example, in the environmental and financial fields?

And does it make sense that there are environmental and financial regulations separate from economic regulation?

These are political economy issues, or economic policy issues, quoting Frieder (2020) once more, i.e., an issue regarding the position of regulatory bodies in the general context of the government. But which government: the national government or the European one (in the case of Portugal)?

In regard to this last question, and in theoretical terms, the principle of subsidiarity will allow for resolving the conflict, but the first issue remains.

This is an open question, even in the name given to regulatory authorities, whose name, in general, includes the word “Authority”. The concept of Authority implies⁸ “the power to make decisions/ the right to command and control other people/ quality (...) that someone has that causes other people to obey them”.

This description evidences the potential clash with the executive power, or even with the legislative power, showing the double legitimacy Rosanvallon mentions.

Hence, the issue is more an issue of State organization than that of “governance structure for regulation”, which cannot be discussed without first referring to that issue. That

⁸ Collins Cobuild, Collins Birmingham University English Language Dictionary, London, 1987, p. 84.

is a core issue in Political Science, evidencing that the issues on the support for regulation, supervision, among others, remain because they have not been sufficiently studied.

Though this is not a topic for our paper, we must point out that the basic conditions for economic regulation are relevant and must not be forgotten.

Considering that these will be discussed, and scenarios will be created that allow for adequate economic regulation framing, the discussion on the future of sectoral regulation implies a prior choice (also dynamically open to change) between specific sectoral autonomy and, alternatively, the integration in a more global institution with sectoral specificities but whose decisions are integrated decisions. Keeping the sectoral regulators independent in terms of sectors require appropriate coordinated decisions in areas that are clearly connected.

In this respect, it is especially important the role that Competition Authorities have, which, in the more classical perspective, is still almost exclusively *ex post*, making the relations with sectoral Authorities more difficult, especially when regulating and supervisory accountabilities intermingle, as was the case of the failed SONAE's takeover on Portugal Telecom.

Considering the characteristics that legitimize the Regulator, which we have already mentioned, it seems clear that the emergence of the behaviour of economic agents in the way markets function drives the debate to the regulatory process itself. In fact, it has become a compulsory reference for the governance model, namely when taking into consideration the reaction of economic agents to fast changes.

In fact, this optimization of the regulatory procedure will be key to better reconcile this "independent power" with political power, because it will make it more visible to society the possible merits of the existence of independent regulators.

The urgent debate between (sectoral) regulation *ex ante* and (competition) regulation *ex post* becomes especially relevant. There is a growing trend to migrate from the first to the second approach.

Its seminal (and still decisive) role in the field of economic regulation is clearly evidenced by Littlechild (2018) in the abstract of his paper: "Regulators are often required to assess the extent of competition in a market and to promote competition or a substitute of it...".

This statement implicitly reflects that competition is the final (sometimes the initial) objective of economic regulation, at least in most situations in which regulation is required. Obviously, this presupposes that there is no specific regulation in a market⁹ where there is competition.

Therefore, an efficient regulation, together with basic favourable conditions, may lead to the loss of the object of regulation and, consequently, determine that the regulator (because it is no longer necessary) is eliminated in that specific market or, more commonly, of several markets, given the restrictive (and controversial) definitions of market that regulators commonly use. This means that an efficient regulator may "commit suicide" as a body, which may lead to the issue of "distortion" referred to by Littlechild (2018), when he states that regulators tend to not recognise their uselessness.

⁹ It is increasingly harder to define a market, both in terms of geography and product, i.e., neglecting other basic conditions of supply and demand!

On the other hand, the Authority that decides whether there is competition is the Competition Authority, which, in this sense, assesses the activity of the sectoral Regulator. But where are its capacity and legitimacy to do that?

A statement by Littlechild is crucial here, as the author questions the model of perfect competition as a reference and adds that: “Interpreting competition as a rivalrous discovery process calls for lighter handed regulation”.

To consider competition as a dynamic process rather than an equilibrium, in which market agents continuously try to adapt to change, or even add more advantages through changes strategically introduced, may jeopardize the “classical” binary view of regulation: *ex ante* vs. *ex post*

We fall into the “soft regulation”, a field which has not only the role of behavioural economics underlying but also those of technological development and relevant aspects of social organization that the more classical paradigm of competition clearly forgets and that most university syllabus also neglect.

Noteworthy is this context is to emphasize that the dynamic perspective of the market is closely linked to the substitution of market concentration (both of supply and of the dangerous and sometimes neglected demand), a clearly structural variable, if seen from a statistical point of view, but whose development analysis may lead to a dynamic perspective, the valuing of entry barriers, with the seminal appearance of the concept of contestable market, and with it, the concept of potential competition. A contestable market is one where there are not barriers to entry and to exit, a key and rather forgotten feature (Baumol and Willig, 1981).

In this type of market, the structure (measured by concentration, for example) will not be an obstacle to competition, though here, for the concept to be effective, rather than perfect information, the required is information symmetry (or little asymmetry) so that new competitors enter the market (Baumol et al., 1988).

In terms of dynamics and of a closer link with the actual functioning of markets, the requirement that no entry and exit barriers exist can (and should) be replaced by the possibility of “low” barriers (meaning, that can be easily overcome), very similar to an earlier perspective, by Clark (1961), who proposed the concept of “workable competition”, which later led to the dynamic perspective on competition.

In fact, Clark (1963), though he advocated the dynamic vision, considers it still imperfect, since, as it consists in identifying the development trajectories which clearly do not tend to have static, full and definable equilibrium, it does not have to be limited to examining the discrepancies between the real values and static level, but impose (p. 225) a normative change. But his contribution goes much further (Silva, 1991) as he considers the static theory, used as an irrefutable norm, as a theory of irresponsibility, ignoring social responsibilities, because society is not merely the result of a mechanical sum of independent results. And he refers with astonishment the position of people who shows social commitment, leaving out this concern in economics, claiming “business is business”.

So, we can see that Littlechild’s position has historical grounds, almost a century old, introducing economy’s social accountability and the ethical obligation that economic agents must take on that accountability. The role of the behaviour of economic agents is made evident here, as well as the ethical principles guiding that behaviour. Isn’t it this paradigm that appears, either in the context of the UN Sustainable Development Goals (SDO) or

especially in a concept that is emerging and has gained relevance – ESG Environment, Social, Governance – as a reference to the functioning of economy and to company behaviour, and in particular, of their managers?

Why have the ideas of some authors not become relevant, authors with continuous, deep and differentiated approaches in the context of economy's historical development and companies' organization? Is this not the reason behind the current profound inequality in living conditions and income, perhaps even behind environmental degradation?

Yet would the world have otherwise reached this level of production and product and service diversity available today? But at what price?

And today, can the “regulatory system” that Barak Obama speaks about or the “oversight” Littlechild (2018) mentions overlook these issues?

Regulation should also be discussed in this context, setting aside assumptions that development eliminates, with not always desirable consequences, at least in the long run.

Laffont and Tirole (2000) do discuss the topic from the traditional “internal” logic of economic regulation, confronting regulation with anti-trust policy, and refer that, in dealing with this confrontation, they feel that “competition policy should be given a more prominent role in the overall process”.

And though they opt for complementariness rather than incompatibility of “anti-trust enforcement and regulation”, they recognise that “their coexistence may jeopardize the coherency of the oversight institution”.

This concern leads us to a different debate that does not question the existing types of economic regulation, but rather which type of institution should implement those actions, one (regulation) linked to ex ante decisions and the other (anti-trust) linked to ex post decisions.

Though there is the concern to further the conditions and governance of good regulation (in the general sense) in the field they call “industry”, it is clear that the question does not include any concern with the democratic legitimacy of regulation but rather with the best way to exercise the regulation.

The subliminal question regarding the relation between sectoral regulation and competition regulation is the argument that the former should disappear when markets are already competing. Overlooking the issue of finding the criteria to verify that existence, which marks the transition from an ex ante to an ex post regulation, you can even question, as others have hypothesized, that the rapid change in the basic conditions of supply, and even of demand, will lead to a dynamic that will create a market that is dynamic in terms of competition, thus making ex ante regulation not relevant, at least one with the characteristics and the more classical regulation procedures.

As mentioned in a document published by Delloite University Press (see footnote 2), in the section entitled “It's tough to be a regulator today”, namely on the argument that ex ante regulation may hinder innovation, they mention five trends to assess that discussion: an exponential technological change, the emergence of new business models, the decrease in entrance barriers in supply and demand that new technologies provide, what they call “The ‘ignore until large’ phenomenon” (which ignores the principle of incipience) and the birth and death of negotiating ecosystems.

5. THE TRIANGLE AT THE BASIS OF INDEPENDENT REGULATION AND SOME FINAL REMARKS

Since we have already discussed the issue of legitimacy, the implications of independence and the changes in the scope of (economic) regulation, it is time to focus on the conditions that may support independent regulators.

As made evident, there are conditions linked to the appointment of regulators - among personalities acknowledged for their competence, involving more than one state authority, imposing periods of mourning, individual, long and lagged mandates - to how regulators work - requiring collegiality, transparent and participated decisions, types of relation with political power and other authorities, compliance with the inevitable principle of legality within the mandate appointed by the state authorities - to conditions for operational, administrative and financial autonomy - requiring autonomous funding from the State Budget, the ability to select human resources with the adequate skills and general management authority - to the clarity in regards to the mandate - defining clearly the objectives of the policy, the degree of freedom of decision and the separation between regulatory (independent) decisions and assistance to political power - and lastly, to accountability - before the citizens, the economic agents of the regulated field, the executive power and the legislative power (OECD, 2016).

In our opinion, all of these conditions may be grouped into two main ones:

Firstly, the “Transparency and Participation” condition regarding their decisions, which is a counterpart of the independence, since there is no control mechanism that citizens can use in the democratic delegation process, which exists in the case of elected political power. This condition, considered in its general sense, encompasses all the duty of traceability in the decision-making process and in compliance with administrative procedures, taken to a much higher level than what is in place for elected political decision-makers, but which in these cases, needs not be as demanding because there is the referred control mechanism of the decisions of the mentioned political decision-makers.

Secondly, the condition “Adequate Resources”, so that the decisions made by the regulators may be made with competence and impartiality; this includes material, financial and human resources, as well as all conditions for selecting, recruiting, appointing and retaining human resources, a key condition for the activity of the IAA. This is an indispensable condition for independence, as well as for recognition and credibility before all stakeholders, also a condition for the legitimization of these bodies.

However, even when these (1) functional and (2) institutional conditions are met, you cannot exclude the possibility that the regulator behaves in a deviant manner, which leads us back to the question: who controls the regulators?

It is true that the conditions for transparency and participation already allow for a decentralized control by all stakeholders, as well as, together with the adequate resource conditions, for decreasing the risk of that type of behaviour. However, it is also true that separation between the regulator’s abusive decision and its legitimate decision is not always evident. This legitimate decision is different from the correct decision as viewed by the citizens and by political power, in particular when the political power behaves according to short-term interests, or even media-shaped public opinion, not having the courage to sacrifice political careers to take responsibility for unpopular decisions, as stated by Tirole (2016).

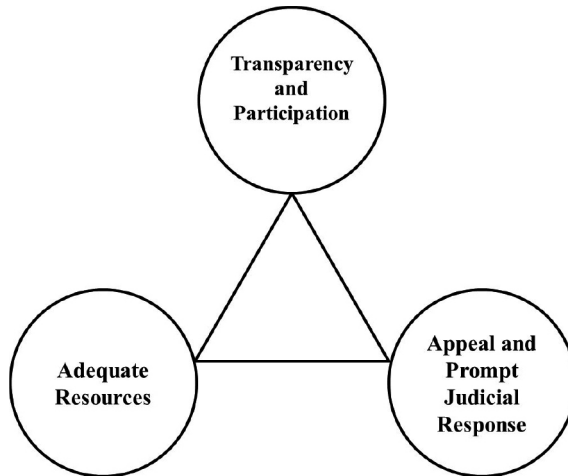
Thus, we believe it is crucial to add one additional condition to support for an independent regulator (more) compatible with the administrative structure of modern societies' democratic governments.

We remind the idea that a regulator cannot be someone anonymous, silent and above the system. We also note the systematic use of judges' independence requirement as one of the cornerstones of non-elect State power.

Thus, we consider that a third condition for "Appeal and Prompt Judicial Response", on the decisions of independent regulators whenever a stakeholder believes there has been deviant behaviour, together with transparency and participation, and adequate resources, would be a third apex of a triangle supporting independent regulation without compromising the triangle of the separation of powers, which emerged with the French revolution.

We suggest that a third triangle be added to the two proposed by Tucker (2018) which structured the thought on modern constitutional states – Democracy/Rule of Law/Efficient Governance (Figure1) and separation of powers Legislative Δ Judicial Δ Executive (Figure 2) – in order to avoid the squaring of the latter triangle forced by the proliferation of IAA in "modern States". This third triangle (Figure 3) is crucial to the support of these Authorities– Transparency and Participation Δ Adequate Resources Δ Appeal and Prompt Judicial Response.

Figure 3: Triangle of Independent Regulation



This Judicial response, of appeal, should not focus primarily on the substance of the regulator's decision, but above all on the compliance to the requirements for grounds, transparency, participation, procedure and pondering in the decision-making process. Obviously, this requires resorting to experts on the matters in question and to comparison with comparable decisions by other regulators. However, to consider the multiple and legitimate interests at play could lead the judge to becoming a political decider, thus distorting the reasons for including the IAA in State administrative organization.

Probably, this concern explains why legislators in several democratic states have created courts specialized in Regulation and Competition, which will be able to respond (more) promptly to (technically complex) appeals on the decisions by independent regulators.

Unfortunately, perhaps due to lack of understanding the role these courts have, their scope of action has not always been defined so as to encompass basic regulating decisions. This is the case of Portugal, in which the action of this court is restricted to penalty matters while the actual decisions on regulation are the responsibility of administrative courts.

We believe that the failure of this apex of the triangle compromises the whole structure of the IAA within the framework of State governance.

If we add failures in the other apexes, for example, failures due to the conditioning of resources imposed by political power and by IAA's deviant behaviour, and the hostile, populist and easy criticism by specific social actors often motivated by political agendas, namely political power itself, tarnishing the prestige and credibility of regulators - you can question if maintaining these Authorities in State organization still makes sense.

The flaw of any piece of the puzzle clouds the image and may even compromise the perception others have of the IAA.

We aim in this study to have contributed to raising awareness towards the need of continuing to analyse the grounds, critical thinking, and explanation of the IAA, following Tirole's suggestion in one of the sub-chapters of his 2016 book: "Explain, Don't Complain."¹⁰

¹⁰ English version, published in 2017, of the original words "Un peu de pédagogie..."

REFERENCES

- Alesina, A.; Tabellini, G. (2007) Bureaucrats or politicians? Part I: A single policy task. *American Economic Review*, 97(1), 169-79.
- Armstrong, M.; Sappington, D. (2007) Recent developments in the theory of regulation. In Armstrong, M.; Porter, R. (eds.), *Handbook of Industrial Organization* (Vol. III), Elsevier Science Publishers, 1557-1700.
- Baumol, W.; Willig, R. (1981) Fixed costs, sunk cost, entry barriers and sustainability of monopoly. *Quarterly Journal of Economics*, 95, 405-431.
- Baumol, W.; Panzar, J.; Willig, R. (1988) *Contestable Markets and the Theory of Industry Structure*. Harcourt College Pub.
- Bawn, K. (1995) Political control versus expertise: Congressional choices about administrative procedures. *American Political Science Review*, 89(1), 62-73.
- Bénassy-Quéré, A.; Couré, B.; Jacquet, P.; Pisani-Ferry, J. (2019) *Economic Policy: Theory and Practice*, 2nd Ed. Oxford University Press.
- Clark, J. (1936) *Preface to Social Economics: Economic Theory and Social Problems*. New York, Routledge.
- Clark, J. (1961) *Competition as a Dynamic Process*. Washington, The Brookings Institution.
- Constant, B. (1991) Fragments d'un Ouvrage Abandonné sur la Possibilité d'une Constitution Républicaine dans Grand Pays. Paris, Aubier.
- Dal Bo, E. (2006) Regulatory capture: A Review. *Oxford Review of Economic Policy*, 22(2), 203-225.
- Edwards, G.; L. Waverman (2006) The effects of public ownership and regulatory independence on regulatory outcomes. *Journal of Regulatory Economics*, 29(1), 23-67.
- Frieder, J. (2020) The Political Economy of Economic Policy. *Finance & Development*, June 2020, International Monetary Fund.
- Gilbert, R.J.; Newbery, D. (1988) Regulation games, CERP Discussion paper 267.
- Gilbert, R.J.; Newbery, D. (1994) The dynamic efficiency of regulatory constitutions. *Rand Journal*, 25(4), 538-554.
- Kahn, A. (1981) *The Economics of Regulation: Principles and Institutions*. Cambridge, MIT Press.
- Laffont, J.J.; Tirole, J. (1986) Using cost observation to regulate firms. *Journal of Political Economy*, 94, 614-641.
- Laffont, J. J.; Tirole, J. (2000) *Competition in Telecommunications*. Cambridge, MIT Press.
- Littlechild, S. (2018) Regulation and the nature of competition. *Journal of Air Transport Management*, 67, 211-223.
- Majone, G. (1996) Temporal consistency and policy credibility: Why democracies need non-majoritarian institutions, EUI Working paper RSC, 96/57, 1-14.
- Maskin, E.; Tirole, J. (2004) The Politician and the Judge: Accountability in government. *American Economic Review*, 94, 1034-54.
- Newbery, D. (2001) *Privatization, Restructuring and Regulation of Network Utilities*. MIT Press.
- Nussbaum, M. (2006) La fêlure dans le cristal : La Coupe d'or de James et la littérature comme philosophie morale. In Laugier, S. (Ed.), *Éthique, Littérature, Vie Humaine*, Paris, PUF, 19-51.
- Nussbaum, M. (2007) *Frontiers of Justice: Disability, Nationality, Species Membership*. Harvard University Press.
- OCDE (2016) *Being and Independent Regulator: The Governance of Regulators*. OECD Publishing.
- Rosanvallon, P. (2008) *La Légimité Démocratique: Impartialité, Réflexivité, Proximité*. Paris, Seuil.
- Ross, C. (2016) *The purposes and functions of economic regulation*, Ofwat, U.K.

- Silva, J.A. (1991) *Economia Industrial e Excesso de Capacidade*. Lisboa, Instituto de Novas Profissões.
- Stigler, G. J. (1971) The theory of economic regulation. *Bell Journal of Economics and Management Science*, 2(1), 3-21.
- Tirole, J. (2016) *Économie du Bien Commun*. Paris, Presses Universitaires de France.
- Tucker, P. (2018) *Unelected Power: The Quest for Legitimacy in Central Banking and the Regulatory State*. New Jersey, Princeton University Press.
- Waverman, L.; Koutroumpis, P. (2011) Benchmarking telecoms regulation: The telecommunications regulatory governance index (TRGI). *Telecommunications Policy*, 35(5), 450-468.
- Williamson, O. (1985) *The Economic Institutions of Capitalism*. New York, Free Press.
- Williamson, O. (1998) Transaction cost economics: How it works; where it is headed. *De Economist*, 146(1), 23-58.
- Williamson, O. (2000) The new institutional economics: Taking stock, looking ahead. *Journal of Economic Literature*, 38(3), 595-613.
- Wittgenstein, L. (1965) *Le Cahier Bleu et le Cahier Brun*. Paris, Gallimard.

Novas Perspetivas para a Regulação do Setor Elétrico

NEW PERSPECTIVES FOR THE REGULATION OF THE ELECTRICITY SECTOR

Vitor Marques

Received for publication: June 15, 2020

Revision accepted for publication: April 9, 2021

RESUMO

O presente texto apresenta a evolução da regulação do setor elétrico em Portugal desde o início do século XXI, quando muitos setores elétricos, entre os quais o nacional, ainda se encontravam no início do processo de liberalização. Este processo era movido pelo objetivo europeu de criação do Mercado Interno da Energia. Desde então, o contexto da regulação do setor elétrico na Europa e em Portugal evoluiu significativamente. Os avanços tecnológicos, as necessidades decorrentes das alterações climáticas e o maior escrutínio dos consumidores sobre as ações dos reguladores constituem motores da alteração drástica desse contexto. Ao discutir as principais alterações tecnológicas e organizativas que se verificaram nos últimos anos, este estudo apresenta, igualmente, um conjunto de desafios para a regulação do setor elétrico em Portugal no contexto europeu.

Palavras-chave: Regulação sectorial; monopólios naturais; transição energética; setor elétrico.

ABSTRACT

This study presents the development of energy regulation in Portugal, in particular, progress in the electricity sector since the beginning of the century when the electricity sector was still at the start of the process of liberalization, accompanied by new agents entering with private equity, along its value chain. This pattern, common to most European countries, was largely the result of the foundation of the Internal Energy Market. This process was designed to allow consumers to benefit from a more competitive market and from a better use of Europe's energy resources. The context of the electricity sector has evolved significantly ever since. The technological advances, the needs resulting from climate change and the increased scrutiny of regulators' activities, in particular by consumers, are not only drastic changes, but also main drivers of the new regulatory context. By presenting the main technological and organizational changes that have taken place in recent years, this study also highlights many challenges for the regulation of the electricity sector in Portugal in the European context.

Keywords: Sectoral regulation; natural monopolies; energy transition; electricity sector.

JEL Classification: L12; L51; L94.

Disclaimer: Os resultados e opiniões apresentados neste artigo são da inteira responsabilidade do autor e não devem estar de forma alguma associados às opiniões oficiais da ERSE, nem vinculam esta instituição.

1. INTRODUÇÃO

O presente texto é um exercício retrospectivo, focado na regulação da energia e, em particular, na regulação do setor elétrico. O ano de referência para este exercício é 2003, quando foi publicado “Poder de Mercado e Regulação nas Indústrias de Rede” (Marques, 2003), livro que apresentava um conjunto de conceitos económicos, que enquadram a regulação setorial e concorrencial, e abordava problemáticas que a regulação do setor elétrico enfrentava à data. Nesse ano, a maioria dos setores elétricos dos países europeus, tal como o nacional, encontrava-se numa fase embrionária do processo de liberalização, que se foi consolidando até à atualidade. Este processo é em grande parte fruto do objetivo europeu de criação do Mercado Interno da Energia (MIE), cujas bases tinham sido lançadas no final dos anos 90 do século XX. O MIE visa permitir aos consumidores europeus beneficiarem das vantagens de um mercado competitivo de grande dimensão, capaz de aproveitar os recursos endógenos existentes na Europa de forma eficaz.

Essa pretensão confrontava-se, na altura, com barreiras ao nível das infraestruturas existentes nos setores elétricos europeus, assim como ao nível das suas estruturas organizativas. No que dizia respeito a estas últimas, esses setores deveriam permitir a entrada de novos agentes em condições de igualdade face aos incumbentes nacionais. Deveriam igualmente ser criadas condições para a existência de mercados de índole nacional ou regional a montante e a jusante da cadeia de valor do setor elétrico que pudessem, posteriormente, ser fundidos com outros mercados da mesma natureza, com vista à criação de mercados alargados que englobassem vários países europeus. Finalmente, deveriam ser criados reguladores setoriais que, apesar de pertencerem à administração pública, fossem independentes do Estado concedente, o que lhes permitiria, por exemplo, não serem influenciados pelos ciclos eleitorais. Essas entidades distinguiam-se dos restantes órgãos do Estado não apenas devido à sua natureza independente, mas também pelos instrumentos que privilegiariam na prossecução dos seus objetivos que assentariam na teoria económica.¹

Nos últimos 18 anos, o contexto regulatório evoluiu significativamente, em grande parte devido aos avanços tecnológicos, às necessidades decorrentes das alterações climáticas e ao maior escrutínio dos consumidores sobre as ações dos reguladores. Ao apresentar os principais fatores que explicam a evolução do contexto regulatório ao longo desse período, o presente texto apresenta, igualmente, os novos desafios enfrentados pelo regulador independente quando procura assegurar que o setor elétrico contribua de forma sustentável para o desenvolvimento económico do país.

¹ A existência de entidades reguladoras independentes do poder executivo, externas aos poderes legislativo e judicial, não era, contudo, uma novidade em toda a Europa e noutros países fora deste continente. Refira-se, em particular, o caso dos países anglo-saxónicos, onde o Estado, tradicionalmente, não desempenha um papel centralizador no funcionamento da economia e onde o poder judicial baseia-se no *common law*.

2. NOVA ESTRUTURA DO SETOR ELÉTRICO

2.1. MUDANÇAS TECNOLÓGICAS

No início do século XXI, o setor elétrico caracterizava-se por ser um setor verticalmente integrado, com agentes com papéis claramente definidos ao longo da sua cadeia de valor. Nesse contexto, a energia elétrica, fornecida aos consumidores através de redes de distribuição e de transporte,² provinha de centros electroprodutores afastados dos locais de consumo, que produziam energia elétrica com base em energia fósseis ou em aproveitamentos hidroelétricos de grande dimensão. Refira-se que esta estrutura produtiva tinha impactes negativos nos ecossistemas porque a produção de energia elétrica com recurso a combustíveis fósseis emite grandes quantidades de gases com efeito de estufa (CO₂, N₂O, CH₄, etc.), enquanto os grandes aproveitamentos hídricos provocam erosão dos solos e, embora em muito menor escala do que a produção com recurso a combustíveis fósseis, também geram emissões de gases de efeito de estufa (Dai e Liu, 2013; Lima, 2008).

A liberalização dos setores elétricos, designadamente nas atividades de produção e de comercialização, acompanhada pela separação dessas atividades das atividades de redes que são monopólios naturais, o transporte e a distribuição de energia elétrica, provocou o fim das *utilities* verticalmente integradas.

Assistiu-se, entretanto, a uma forte diminuição dos custos de produção dos centros electroprodutores com fontes de energia renováveis, em particular da produção fotovoltaica, que permitiu um crescimento significativo da produção descentralizada de energia elétrica. Ao contrário da produção de energia elétrica convencional, a produção descentralizada, também chamada de produção distribuída, não é despachada³ centralmente pelo gestor do sistema e não é encaminhada pela rede de transporte, sendo, pelo contrário, diretamente injetada na rede de distribuição e, por isso, encontra-se de um modo geral mais perto dos centros de consumo. Esta mudança tecnológica levanta grandes desafios à gestão do sistema por parte dos operadores das redes de transporte e de distribuição.⁴ Um dos principais desafios que enfrenta estes operadores advém de a produção de energia elétrica distribuída recorrer, na maior parte das vezes, a fontes de energia renováveis, que são por natureza intermitentes e não despacháveis centralmente por dependerem de fatores não controláveis, como por exemplo a velocidade do vento e a intensidade da luz solar.

Acresce que neste novo contexto os fluxos de energia elétrica podem, igualmente, direccionar-se no sentido dos níveis de tensão mais elevados, isto é, da rede de distribuição para a rede de transporte, se as potências injetadas tiverem algum peso face ao consumo local

² O transporte de electricidade consiste grosso modo no encaminhamento da energia eléctrica desde os centros electroprodutores, em muita alta tensão (normalmente a partir de 110 kV), até às redes de distribuição, a partir das quais a electricidade será fornecida aos consumidores finais, com níveis de tensão que não ultrapassam 50 kV.

³ O despacho é a ordem dada aos centros electroprodutores pelo gestor do sistema eléctrico, atividade geralmente assumida pelo operador da rede de transporte, de injeção na rede de uma determinada potência durante um determinado período para fazer face à variação de procura de energia nesse período.

⁴ Entende-se por gestão dos sistemas elétricos as ações desenvolvidas pelos operadores das redes para garantirem um equilíbrio contínuo entre a produção e o consumo de energia elétrica, com níveis adequados de segurança, estabilidade e qualidade de serviço.

(Leiß, 2013), o que torna mais complexa a gestão dos setores elétricos, tanto a nível físico, como económico. Esta complexidade é ampliada com o grande aumento de consumidores que passaram a produzir energia elétrica para consumo próprio, ou até para injeção na rede de distribuição, devido à diminuição dos custos de produção⁵. No entanto, o progresso tecnológico também permitiu responder a estes novos desafios e suportar a reconfiguração da organização dos setores elétricos em torno de dois vetores: i) a maior capacidade de tratamento e armazenamento de dados decorrente das tecnologias de informação e ii) a maior eficiência do armazenamento de energia elétrica através de baterias.

2.1.1. REDES ELÉTRICAS INTELIGENTES

O primeiro vetor referido anteriormente está na base das chamadas “redes inteligentes”. Segundo a definição dada pelos reguladores europeus num documento de 2009 (ERGEG, 2009), as “redes inteligentes” integram, de modo eficiente em termos de custo, o comportamento e as ações de todos os utilizadores que lhes estão ligados – produtores, consumidores e utilizadores com ambas as funções – com o objetivo de assegurar um sistema energético economicamente eficiente e sustentável, com baixo nível de perdas e elevados níveis de qualidade de serviço, de segurança de abastecimento e de proteção.

As tecnologias que promovem a “inteligência” das redes situam-se ao longo da cadeia de valor desde os contadores inteligentes ao nível dos consumidores até, mais a montante, aos equipamentos e sistemas para recolha, transmissão e armazenamento de dados e, ainda, aos equipamentos de comando e controlo ao nível das infraestruturas de redes elétricas (Comissão Europeia, 2011). A integração de “inteligência” nas redes por parte dos operadores de redes europeus não é apenas fruto do acompanhamento da evolução tecnológica que se verifica no conjunto da economia, mas deve-se sobretudo à necessidade de responder ativamente aos novos desafios que a gestão dos fluxos de energia e económicos, principalmente nas redes de distribuição, enfrenta.

Tal como referido no ponto anterior, estes desafios decorrem em grande parte da penetração da produção a partir de fontes de energias renováveis, cuja disponibilidade é intermitente, bem como de novas formas de utilização da energia elétrica, designadamente com os veículos elétricos. Em Portugal, entre 2005 e 2015, a produção de energia elétrica a partir de fontes de energia renovável intermitente passou de cerca de 28% para mais de 52% da produção total de energia elétrica (Comissão Europeia, 2017). Estas alterações ao nível da produção e do consumo de energia elétrica geram novos fluxos de energia, com sentidos opostos aos que ocorriam até à data (dos níveis mais baixos de tensão para os níveis de tensão mais elevados, por exemplo), em novos locais das redes e em novos períodos do dia e, conseqüentemente, geram também novos picos de consumo em termos de horários e de localização.

A resposta tradicional a estas alterações nos fluxos energéticos e económicos passaria por um reforço dos investimentos nas redes de transporte e de distribuição de energia elétrica, bem como em centros electroprodutores com tecnologias convencionais, que servissem de

⁵ Para o caso português, ver, por exemplo, Smart grid: Uma Visão da Regulação (Esteves et al., 2017).

back-up à produção de energia elétrica com base em fontes de energia renováveis. Todavia, parte desses investimentos e, conseqüentemente do acréscimo de custos que lhes estão associados, pode ser evitada se os sistemas elétricos se tornarem mais flexíveis.

O conceito de flexibilidade dos setores elétricos corresponde à capacidade de estes responderem às flutuações da procura e da oferta, enquanto, em paralelo, mantêm-se fiáveis (CEER, 2017). Este conceito é capacitado pela existência de “redes inteligentes”. Assim, tanto os produtores como os consumidores podem contribuir para uma maior flexibilidade dos sistemas elétricos, desde que o sistema seja suficientemente “inteligente” para permitir que esses agentes adaptem o seu comportamento às suas necessidades.

O melhor conhecimento por parte dos operadores das redes do grau de saturação destas e dos perfis de consumos de energia elétrica, conjugado com um melhor conhecimento por parte dos consumidores da evolução dos seus consumos em cada momento, permite valorizar economicamente a flexibilidade das redes para o seu uso mais eficiente. Esta valorização pode concretizar-se através de tarifas dinâmicas de uso das redes ou através da venda em mercado de serviços de sistema associados à gestão da carga, isto é, à procura num determinado momento (Lavrijssen e Carrillo Parra, 2017).

Neste último caso, os consumidores podem vender estes serviços, agregando-os através de um agregador (Crampes e Waddams, 2017). A valorização desta flexibilidade requer a evolução, em paralelo, do quadro regulatório. Este requisito será desenvolvido mais adiante.

2.1.2. NOVAS FORMAS DE ARMAZENAMENTO DA ENERGIA

Além das tendências de evolução tecnológica nos sistemas elétricos, o armazenamento da energia elétrica com recurso a baterias constitui outra tendência que possibilita aos utilizadores das redes desempenharem novas funções, contribuindo para a garantia da flexibilidade dessas infraestruturas. O armazenamento de energia elétrica tem sido, desde há muitos anos, possível em grandes quantidades através da bombagem da água das albufeiras de aproveitamentos hidrológicos ou ainda de forma mais difusa, com recurso aos sistemas de armazenamento de água quentes para usos domésticos. Contudo, a tendência de forte diminuição dos custos com baterias que se tem verificado, que se perspetiva deva continuar (Hatzigiorgiou, et al., 2007; Esteves, et al., 2017; Irena, 2017), conjugada com a diminuição dos custos da produção fotovoltaica, comparativamente com o preço da energia elétrica, torna economicamente racional ao cliente doméstico adquirir baterias para armazenar a energia fotovoltaica produzida e consumi-la posteriormente, em lugar de adquiri-la.⁶

Esta nova realidade permite antecipar novas estruturas organizativas nas quais os sistemas de armazenamento permitirão aos consumidores dissociar os períodos de consumo, dos períodos de uso das redes e de aquisição de energia elétrica, possibilitando um menor recurso ao sistema elétrico nos períodos de ponta. Como exemplo do papel que poderão desempenhar novos agentes detentores de sistemas de armazenamento nestas estruturas

⁶ No que diz respeito à redução dos custos de produção de fotovoltaica, não se pode deixar de registar o caso excecional dos leilões de atribuição de capacidade de injeção na Rede Elétrica de Serviço Público de capacidade renovável para produção de energia fotovoltaica, que ocorreram em Portugal em 2019 e em 2020, e que atingiram preços de fecho extremamente baixos, impensáveis de puderem ser atingidos ainda há poucos anos.

organizativas, refira-se o dos proprietários de veículos elétricos que poderão contribuir para uma maior flexibilidade dos sistemas elétricos através das baterias dos seus automóveis.

2.2. NOVO PAPEL DOS CONSUMIDORES

Por toda a Europa, os setores elétricos nacionais têm assistido não apenas à democratização da produção de energia elétrica, como também ao aparecimento de novos serviços (serviços de gestão de energia, serviços de consultoria energética e serviços de comercialização de energia), facultados por novos agentes (consumidores produtores (prosumers), agregadores, cooperativas energéticas) sob diferentes configurações ou plataformas (mercados grossistas, retalhistas, contratos bilaterais de prestação de serviço, comunidades energéticas...). Em particular, o autoconsumo, isto é, o consumo de energia elétrica produzida pelos próprios consumidores tem-se generalizado, apoiado na diminuição dos custos de produção do solar fotovoltaico e do armazenamento de energia através de baterias. Com a diminuição em cerca de 80% dos custos de produção através de painéis fotovoltaicos nos últimos 10 anos (IRENA, 2018), a chamada paridade com a rede em termos de custo⁷ tem-se generalizado, principalmente para os consumidores domésticos (GfK, 2017).

Em algumas situações, os consumidores organizam-se coletivamente para produzirem energia elétrica que consomem, dando lugar às chamadas coletividades energéticas. Estas coletividades podem configurar sistemas elétricos parcial ou integralmente independentes da rede de distribuição de pública, chamados micro-redes. As micro-redes são redes de distribuição de energia elétrica em baixa tensão às quais estão ligados pequenos produtores de energia elétrica e que, por isso, têm um certo grau de autonomia relativamente à rede de distribuição principal em termos de fornecimento de energia.

Os desafios para a gestão das redes decorrentes da participação mais ativa de todos os consumidores podem ser superados com uma otimização dos fluxos energéticos gerados por esses agentes localmente, que possibilita, por exemplo, um menor nível de perdas de energia elétrica e uma maior resiliência da rede. Neste contexto, são reconhecidas potenciais mais-valias decorrentes da maior proximidade entre os locais de consumo e de produção (Chiradeja, 2005; Wang e Nehrir, 2004; Chiradeja e Ngaopitakkul, 2013).

Por sua vez, esta otimização subentende uma participação economicamente eficiente dos consumidores, com o desenvolvimento de novos comportamentos por parte dos utilizadores das redes de transporte e, principalmente, das redes de distribuição, tanto em termos de consumo e de gestão da carga, como de produção. Assim, perspectiva-se um “círculo virtuoso” de reorganização do setor elétrico movido pela evolução tecnológica: as novas tecnologias permitem a descarbonização do setor elétrico através de uma participação mais ativa de todos os agentes ao longo da cadeia de valor, facto que levanta novos desafios à gestão das redes que, por sua vez, apenas podem ser superados com o desenvolvimento de comportamentos economicamente eficientes desses mesmos agentes possibilitados pela evolução tecnológica.

⁷ A paridade com a rede (grid parity) corresponde ao ponto em que o autoconsumo é economicamente mais eficiente para o consumidor do que a aquisição de energia através da rede. No entanto, mesmo quando consomem energia elétrica que produzem, os consumidores não deixam de pagar os custos das redes a que estão ligados.

A organização do setor elétrico assente no recurso a combustíveis fósseis e em fluxos energéticos e financeiros que percorrem a cadeia de valor de montante para jusante vai perder relevância. Assiste-se, desta forma, à quebra de várias barreiras à entrada, tanto tecnológicas, como regulamentares e legislativas, em especial nas atividades de produção e de comercialização, mas também nas atividades de rede.⁸ Consequentemente, as *utilities* que se desenvolveram em torno do paradigma tradicional de organização do setor elétrico têm vindo a rever as suas estratégias de desenvolvimento.

2.3. NOVAS ESTRUTURAS EMPRESARIAIS

Os mercados de capitais têm assistido a uma perda de valor das principais *utilities*, principalmente na Europa ocidental (AT Kearney, 2017), que se destaca pelas medidas implementadas com vista à descarbonização da economia. Saliente-se que, após a ratificação do protocolo de Quioto em 2005, a preocupação da União Europeia em descarbonizar as suas economias se acentuou. Para responder a esta preocupação, vários instrumentos foram instituídos a nível nacional e europeu. A nível europeu, destaca-se o estabelecimento de um mercado de licenças de emissão de CO₂ (CELE, Comércio Europeu de Licenças de Emissão de CO₂), que visa internalizar os custos ambientais (associados à emissão de gases de efeito de estufa) no preço de um conjunto de bens ou serviços, dos quais se destaca a energia elétrica, tornando a produção de energia elétrica com recurso a combustíveis fósseis mais cara e, consequentemente, menos competitiva.

A perda do valor das principais *utilities* europeias verificou-se, principalmente, nas empresas que assentam o seu mix produtivo na utilização de combustíveis fósseis ou na energia nuclear. Registe-se, por exemplo, o caso de dois gigantes europeus, a francesa EDF francesa e a alemã EON, que viram o seu valor de mercado diminuir em mais de 80% entre 2008 e 2018 (AT Kearney, 2017).

Os principais conglomerados energéticos europeus têm procurado reagir e adaptar-se a este novo contexto. Assistiu-se a reestruturações desses grupos económicos, que procuraram separar as atividades associadas ao paradigma energético vigente até há pouco tempo, designadamente a produção de energia com base em combustíveis fósseis ou no nuclear, das atividades associadas à descarbonização da economia (produção a partir de fontes de energia renováveis, prestação de novas soluções para os agentes ou, ainda, investimentos em infraestruturas de rede). Esta dinâmica de reorganização empresarial foi patente em alguns países há já mais de 10 anos. Assim, foi criada, em 2004 em França, a EDF Energie Nouvelle, na qual a gigante francesa EDF integrou as suas atividades de produção de energia renovável. Esta empresa passou a denominar-se EDP Energie Renouvelables a partir de 2018. Em Portugal, em 2007, a EDP agrupou numa nova empresa sua subsidiária, a EDP Renováveis, toda a produção com origem em fontes renováveis, enquanto na Itália, em 2008, a gigante Enel criou a Enel Green Power onde agregou a sua produção renovável. Mais tarde, em 2016, surge a Innogy da RWE alemã, com as atividades associadas às energias

⁸ Registe-se que a legislação em Portugal acompanhou esta tendência (em especial com o Decreto-Lei n.º 162/2019, de 25 de outubro), concebendo as bases para a existência de comunidades de energia renováveis, baseadas no auto-consumo coletivo, e permitindo associar instalações de produção e de consumo de energia elétrica.

renováveis e às infraestruturas de rede. Nesse mesmo ano, a alemã EON optou por agrupar os negócios mais tradicionais numa nova empresa, deixando as energias associadas à nova dinâmica na EON.

Esta tendência observou-se juntamente com o realinhamento dos investimentos dessas *utilities* nas atividades associadas à descarbonização da economia. A aquisição de *start-ups* associadas aos novos serviços na área de energia, por parte de grandes grupos energéticos europeus, tem sido uma forma desses grupos ultrapassarem as suas dificuldades históricas quanto à inovação. Entre estes novos serviços na área de energia, destacam-se a integração de renováveis, a mobilidade elétrica, a eficiência energética, as *smart cities*, as plataformas digitais, etc.

Há que referir, contudo, que, apesar das mudanças assinaladas no setor elétrico decorrentes da evolução tecnológica, este setor destaca-se pela negativa, comparativamente com outros, pelo pouco peso da inovação no volume de negócios das suas empresas (Gates, 2015). Existe, por conseguinte, uma margem importante para que a dinâmica de reorganização do setor elétrico devido à evolução tecnológica ganhe ainda maior impulso.

2.4. TRANSIÇÃO ENERGÉTICA AINDA POUCO INCLUSIVA

A democratização da participação dos agentes no setor elétrico enfrenta ainda muitas barreiras tanto financeiras, como comportamentais ou ainda técnicas (Goulden et al., 2014). Os equipamentos que suportam a maior participação dos consumidores exigem esforços financeiros no momento do investimento, designadamente ao nível da produção, que podem ser dificilmente suportados por muitos consumidores (GfK, 2017).⁹

A maior “contestabilidade” (Baumol et al., 1982) dos mercados de produção e de comercialização de energia elétrica que advirá desta democratização da participação dos agentes poderá ter, na prática, ainda pouco efeito no melhor funcionamento desses mercados, atendendo às especificidades desta dinâmica tendencialmente limitada ao nível do autoconsumo ou centrada em comunidades energéticas. Assim, a diminuição de barreiras à entrada nas atividades destes novos agentes apenas beneficiará, num primeiro tempo, uma minoria de consumidores e a maioria dos consumidores deverá ainda ficar dependente nos próximos anos, em exclusivo, dos operadores das redes de distribuição. O aumento do autoconsumo gera um menor consumo de energia elétrica através da rede de distribuição e, consequentemente, uma menor base de diluição dos custos das infraestruturas de rede que são, predominantemente, fixos. Deste modo, os reguladores europeus deverão acompanhar os efeitos da pressão acrescida sobre as tarifas reguladas de uso das infraestruturas de rede que serão, principalmente, suportadas pelos consumidores que não terão a oportunidade ou a capacidade financeira para poderem ficar independentes das redes de distribuição.

⁹ Refira-se, no entanto, que no caso da gestão ativa da procura, designadamente através dos contadores inteligentes, os custos têm sido socializados por beneficiarem o conjunto da sociedade. A União Europeia pretendia que esses contadores atingissem 80% do total dos contadores até 2020, sempre que se verificam benefícios líquidos decorrentes desta substituição. No entanto, os últimos dados apontam para situações muito dispare entre países europeus (Comissão Europeia, 2019)

3. PERSPETIVAS PARA A REGULAÇÃO DO SETOR ELÉTRICO

Nos últimos 18 anos, a regulação dos setores elétricos na Europa procurou dar continuidade aos processos de transição de sistemas elétricos inicialmente nacionais, centralizados e, muitas vezes, detidos pelo Estado, para sistemas transnacionais, assentes na economia de mercado, que se desenvolvem em torno de infraestruturas de rede consideradas monopólios naturais.

A regulação setorial enquadrada pela regulamentação comunitária e, subsidiariamente, nacional, assentou em premissas económicas de promoção de eficiência. Neste contexto, a regulação do setor elétrico focou-se na definição de tarifas eficientes de uso das infraestruturas de rede, em termos de alocação de recursos, na dinamização de mercados retalhistas e grossistas concorrenciais, bem como na promoção da qualidade técnica e comercial dos serviços prestados.

Nas infraestruturas de rede foram principalmente aplicadas metodologias regulatórias focadas no controlo do nível de custos, ou inputs base, tanto sob a forma da regulação dita por incentivos do tipo *price-cap*, como sob a forma de uma regulação mais tradicional baseada em custos aceites ou *rate of return*. Em Portugal, a aplicação dessas metodologias por parte do regulador setorial, Entidade Reguladora dos Serviços Energéticos (ERSE), permitiu que o peso das tarifas de uso das infraestruturas de transporte e distribuição, cujos rendimentos são exclusivamente definidos pelo regulador, tenham passado entre 2003 e 2020 de cerca de 33% para 19% do valor médio da fatura do consumidor de energia elétrica. Noutra perspetiva, a componente dos rendimentos sujeita às regras da concorrência, nos mercados grossistas e retalhistas, ou dependente de subsídios ou taxas, definidos centralmente pelo Estado, passou de 67% para 81% do valor médio da fatura do consumidor final.

No que diz respeito aos mercados grossistas e retalhistas, para além de terem sido criadas regras para garantir a liquidez desses mercados e que todos os agentes ajam de forma racional e informada (pressupostos -base da existência de mercados concorrenciais), os reguladores europeus procuraram monitorizar de perto o funcionamento dos mercados.¹⁰ Atualmente em Portugal, a grande maioria da energia elétrica produzida tem os seus preços definidos no mercado grossista e o mercado retalhista, com mais de 20 comercializadores ativos, representa mais de 95% do consumo total da energia.

Ao nível da qualidade de serviço tem-se observado melhorias substanciais incentivadas por metodologias regulatórias focadas na apresentação de resultados, ou *output based* (CEER, 2018), designadamente para a promoção da qualidade de serviço. Em Portugal, a aplicação de um incentivo à promoção da qualidade de serviço por parte da ERSE contribuiu para que o TIEPI (Tempo de Interrupção Equivalente da Potência Instalada, medida da qualidade de serviço) tenha passado entre 2003 e 2018 de 222 minutos para menos de 50 minutos (ERSE, 2017).

Aparentemente e tomando por exemplo a situação nacional, a regulação setorial terá cumprido os seus propósitos. No entanto, o setor elétrico sofreu alterações significativas que obrigam a reequacionar os objetivos e as abordagens regulatórias. Assistiu-se a um incremento

¹⁰ Ao nível dos mercados grossistas destaca-se o REMIT (Regulation on Wholesale Energy Market Integrity and Transparency) de 2011 que é um Regulamento do Conselho e do Parlamento Europeu de outubro 2011 (Reg. n.º 1227/2011, de 25 de outubro) que visa promover a transparência e a estabilidade dos mercados europeus e evitar a manipulação do mercado, com base na recolha sistemática de informação sobre as transações efetuadas.

dos custos da energia elétrica, em particular para os consumidores domésticos (Price e Deller, 2015), o que tanto pode refletir um funcionamento desajustado dos mercados para este segmento de consumidores (Crampes et al., 2017), como a não existência de um mercado funcional. O caso português é paradigmático desta última situação: enquanto o peso dos custos das infraestruturas, regulados pela ERSE, diminuiu de forma sistemática na fatura de energia elétrica, o peso dos subsídios, em particular à produção, cresceu significativamente. Entre 2003 e 2020, o peso dos custos de interesse económico gerais, principalmente respeitantes a subsídios à produção, passou de cerca de 5% para 29% da fatura. Os preços praticados nos mercados grossistas, nos quais é transacionada a maioria da energia produzida em Portugal, não refletem a remuneração da energia produzida que é, em grande parte, subsidiada (ERSE, 2018). Estes subsídios levaram ao incremento da fatura dos consumidores de energia elétrica. Observou-se entre 2003 e 2020 um incremento acumulado dos preços médios das tarifas de referência de venda a clientes finais dos consumidores domésticos em Portugal de 6% em termos reais.¹¹

A estas exigências juntam-se, no caso português, as preocupações com a sustentabilidade económica do setor no longo prazo, designadamente com a diminuição da dívida tarifária, que corresponde em grande parte a montantes acumulados de subsídios à produção que ainda não foram refletidos nas tarifas de energia elétrica.

Observa-se, igualmente, que os países que foram mais longe na liberalização, como o Reino Unido, nem sempre apresentam consumidores mais satisfeitos ou com acesso a energia mais barata, comparativamente com outros países que optaram por processos mais faseados, como a França (ACER/CEER, 2019; Comissão Europeia, 2017).¹²

Esta tendência de crescimento da fatura de eletricidade poderá ser ampliada em alguns segmentos de consumidores, com as alterações da organização do setor elétrico que se perspetivam e que foram referidas no capítulo anterior. As alterações dos hábitos de consumo da eletricidade, passando esta energia a ser utilizada de forma mais intensa em períodos ou em locais que não eram habituais, justificam mais investimentos nas redes, em particular nas redes de distribuição. A necessidade de realizar os investimentos nas infraestruturas de rede surge num contexto em que o consumo da energia elétrica fornecida por essas infraestruturas estagna ou até diminui em resultado de mais eficiência energética e do aumento do autoconsumo. Esta tendência implicará, em primeiro lugar, o aumento do custo unitário das infraestruturas de rede reguladas, com o consequente impacto tanto para a capacidade financeira no consumo de energia elétrica (*affordability*) por parte dos consumidores domésticos (Mc Kinsey, 2018), como para a sustentabilidade económica dos sistemas.

A pressão nos custos das infraestruturas de rede ocorre, igualmente, devido a alterações a montante da cadeia de valor do setor elétrico, em resultado da liberalização dos mercados e da evolução tecnológica que induzem um maior peso da produção distribuída e o aparecimento de micro-produtores em larga escala.

¹¹ Em 2020, o preço médio estimado da tarifa transitória de venda a clientes finais foi igual a 0,1909 €/kWh (ERSE, 2019).

¹² Todavia, esta comparação deve ser interpretada com alguma cautela, visto que o baixo custo da produção de energia elétrica em França depende em grande parte de subsídios à produção nuclear que, não obstante não serem pagos pelos consumidores franceses nas tarifas de energia elétrica, são-no através dos seus impostos (ECOFYS, 2014).

Para além destes fatores que irão pressionar os custos das infraestruturas de rede, os reguladores deverão igualmente responder aos desafios relacionados com a proteção de dados e a cibersegurança, que surgirão com o armazenamento de informação associada aos consumidores e aos seus hábitos de consumo, cada vez mais valiosa para as empresas dentro e fora do setor (CEER, 2015; Sandys et al., 2017).

Porém, as alterações que se perspetivam são igualmente oportunidades que não podem ser desperdiçadas, porque permitirão, a prazo, diminuir os custos da energia consumida (que vão além dos custos das infraestruturas de rede), graças aos avanços verificados nas tecnologias de produção com base em fontes de energia renováveis, como permitirão aos consumidores tornarem-se agentes ativos do setor elétrico. O papel do regulador deverá ir além da proteção dos consumidores, facultando aos consumidores o suficiente *empowerment* para que se possam tornar agentes ativos neste novo contexto, através de quatro linhas de orientação, que são apresentadas de seguida.

3.1. MANTER A PRÁTICA REGULATÓRIA EM PROL DOS CONSUMIDORES

Tendo em conta as características técnicas e económicas do setor elétrico, a promoção da eficiência económica deverá continuar a ser um dos princípios orientadores da regulação sectorial. Contudo, para além de promover a eficiência económica, tanto ao nível dos monopólios naturais, como dos mercados, importa garantir que os consumidores possam capturar parte dos ganhos de eficiência obtidos. Ora, este objetivo passa por criar as condições para que os consumidores sejam devidamente informados sobre as condições de mercado, tanto em termos de preço, como de serviços (tendo em conta, por exemplo, o fornecimento de serviços integrados), bem como sobre os seus direitos. Em paralelo, deve-se garantir a manutenção da qualidade dos serviços prestados e o cumprimento dos padrões definidos na regulamentação em vigor e adaptar estes padrões às potencialidades tecnológicas das chamadas redes inteligentes.

3.2. CONTRIBUIR PARA UMA MAIOR INTEGRAÇÃO DO SETOR ELÉTRICO NACIONAL NO CONTEXTO EUROPEU

Os benefícios da implementação do Mercado Interno de Energia para os consumidores de energia elétrica são claros. O aproveitamento das vantagens competitivas de cada país ou região da Europa em termos de produção de energia elétrica e de uma maior segurança de abastecimento propiciada pela existência de uma gestão dos sistemas elétricos efetuada a um nível europeu, são dois exemplos dos benefícios expetáveis. O facto de, em termos energéticos, a Península Ibérica ser praticamente uma ilha face ao resto da Europa é um exemplo das consequências negativas da não efetivação do MIE.¹³

¹³ O sistema elétrico português e espanhol encontram-se bastante interligados. No entanto, a capacidade de interligação de Espanha e, conseqüentemente também de Portugal, com o resto da Europa é bastante deficitária. Assim, a capacidade de interligação entre Espanha e França representa menos de 3% da potência instalada de

Pelo seu carácter transnacional e devido às questões que levanta em termos de soberania, a implementação do MIE depende principalmente de iniciativas e vontades políticas. Assim, o papel dos reguladores europeus deverá ser o de facilitadores de processos, bem como de divulgadores dos benefícios do MIE junto dos vários agentes económicos. Os atrasos verificados até à data neste processo estão em grande parte associados à falta de investimentos nas interligações. Contudo, esses investimentos não poderão deixar de ser eficientes. Os reguladores setoriais deverão garantir que os custos destes investimentos sejam inferiores aos seus benefícios, efetuando para o efeito análises custo-benefício. Em paralelo, importa promover a harmonização das regras regulatórias, sempre que estas regras não reflitam especificidades nacionais. Para esse efeito, é necessário manter a cooperação entre entidades reguladoras, bem como entre estas e a ACER, a agência de regulação europeia.

Deve ser preservado o papel desempenhado pelos reguladores na criação, no acompanhamento e na agregação dos mercados regionais. Finalmente, devem ser promovidas interações eficazes entre a ACER e os reguladores nacionais para garantir uma repartição equilibrada dos benefícios do MIE pelos consumidores dos diferentes Estados-membros.

3.3. GERAR UM AMBIENTE REGULATÓRIO PROPÍCIO À INOVAÇÃO

A promoção da inovação por parte dos reguladores setoriais é necessária para que o setor beneficie dos avanços tecnológicos anteriormente referidos, tornando “mais inteligente” a gestão das infraestruturas e dos fluxos financeiros e económicos existentes ao longo da cadeia de valor. Estas ações justificam-se porque a necessidade de inovar, inerente às empresas que desenvolvem a sua atividade em mercados competitivos, não surge naturalmente em monopólios, tais como os das infraestruturas de rede do setor elétrico.

Acresce que as alterações em termos de uso das redes, decorrentes de alterações de hábitos de consumo ou de alterações do papel dos agentes na rede, obrigam à utilização de soluções “inteligentes” de gestão do sistema, que privilegiam o software face ao hardware e consequentemente sejam, em última instância, menos onerosas para os consumidores, o que subentende uma revisão das bases de regulação que se têm focado na recuperação de custos. Importará, deste modo, promover as formas de regulação que privilegiam os resultados (*output based*), relativamente às que privilegiam o controlo dos custos (*input based*) (CEER, 2018).

Tendo em conta o impacto temporal da inovação que é tipicamente de longo prazo, designadamente nas infraestruturas de rede (Banco Mundial, 2017), a promoção da inovação deve subentender um acompanhamento faseado do regulador. Em primeiro lugar, é necessário assegurar que a regulação simule a “pressão competitiva” propícia aos processos criadores induzida pela regulação por incentivos (Marques et al., 2014), sem pôr em causa, contudo, a promoção da qualidade de serviço ou a garantia do equilíbrio económico-financeiro das empresas reguladas. Posteriormente, é necessário avaliar os benefícios para o sistema das metodologias aplicadas, designadamente recorrendo à monitorização dos resultados obtidos.

produção de energia elétrica em Espanha e, conjuntamente, com as restantes interligações com Portugal não chega a atingir 6% da sua capacidade instalada.

Mas a eficácia da evolução tecnológica no setor elétrico depende em grande parte da iniciativa de outros agentes, consumidores e empresas não reguladas, pelo que o quadro regulatório não deve criar barreiras ao desenvolvimento de novos serviços que introduzam flexibilidade na rede ou à aplicação de novas tecnologias por esses agentes. Para tal, é necessário promover e acompanhar o desenvolvimento de mercados que facultam esses novos serviços, bem como identificar as atividades que devem ser vedadas às empresas monopolistas de rede. A necessidade de vedar algumas atividades às empresas que operam as redes visa assegurar que estas não influenciam o correto funcionamento dos mercados devido a eventuais conflitos de interesses (CEER, 2015), designadamente quando estão integradas em grupos económicos que atuam em várias atividades dos setores elétricos. Finalmente, a promoção da inovação deve ter igualmente em conta a proteção dos consumidores na recolha, na gestão e na divulgação dos seus dados.

3.4. ATUAR NUMA PERSPETIVA DE LONGO PRAZO, FOCADA NA SUSTENTABILIDADE DO SETOR

A regulação deve assegurar que o contributo dos setores energéticos para a descarbonização da economia se efetue sem comprometer a sustentabilidade económica do setor elétrico.

Em vários países europeus, tal como em Portugal, os apoios à produção de energia elétrica, em especial às energias renováveis, agravaram de forma significativa a fatura de energia dos consumidores de eletricidade. Acresce que estes apoios variam fortemente entre Estados-membros, impedindo que as tomadas de decisões de investimento à produção sejam eficientes à escala europeia, constituindo, deste modo, um entrave à criação do Mercado Interno Europeu de Energia.

No entanto, a maior parte das tecnologias de produção de energia elétrica com base em fontes de energia renováveis atingiu um nível de maturidade suficientemente elevado que permite que o seu desenvolvimento se possa basear mais em soluções de mercado do que em soluções administrativas, tais como as *feed-in tariffs* (ERSE, 2018), que são na prática subsídios à produção. Esta posição já é assumida pela União Europeia, com a publicação da Diretiva (UE) 2018/2001, de 11 de dezembro de 2018, e, de um modo geral, também o é pela legislação nacional publicada nos últimos anos.

Os reguladores europeus deverão promover desenhos dos mercados grossistas que permitam a integração da produção renovável ou a adesão a soluções alternativas, que reflitam as particularidades destas tecnologias. Tipicamente, as tecnologias de produção de energia elétrica com base em fontes de energia renováveis têm uma estrutura de custos que não se coaduna com o funcionamento dos mercados grossistas na Europa. Os mercados grossistas de energia na Europa são mercados do tipo *Uniform Price Auction*, em que os preços da energia elétrica vendida são definidos tendo em conta os custos marginais de produção das tecnologias mais caras. No entanto, os custos das tecnologias de produção de energia elétrica com base em fontes de energia renováveis são principalmente custos fixos, sendo os seus custos marginais de produção quase nulos. A venda de energia elétrica pelos produtores com fontes de energia renováveis nesses mercados pode não permitir recuperar os seus custos de produção, designadamente quando os mercados se caracterizam por uma cada vez maior penetração dessas tecnologias no total da energia transacionada, gerando

um afundamento do preço médio da energia elétrica vendida. Deste modo, deverão ser desenhadas soluções assentes em regras de mercado transparente que permitam recuperar os custos de investimento num período de tempo que reflita o tempo de vida útil do equipamento. A remuneração desta energia através de leilões lançados previamente à realização dos investimentos,¹⁴ que assegurem um preço de venda em linha com os custos de produção alisados durante o período de vida útil do equipamento constitui uma solução alternativa, válida, aos mercados do tipo *Uniform Price Auction*.

Paralelamente, é necessário promover a eficiência energética no consumo, bem como uma reação economicamente racional do consumo de energia aos custos que proporciona, isto é, o *demand side response*, de modo a evitar a necessidade de mais investimentos tanto na produção de energia elétrica, como nas infraestruturas de rede.

4. COMENTÁRIOS FINAIS

Ao longo dos últimos 18 anos, as políticas energéticas nem sempre permitiram garantir a devida proteção dos consumidores, que tiveram, até há pouco tempo, um papel relativamente passivo no desenvolvimento do setor elétrico. Atualmente, as alterações que se verificam e que se perspetivam para o setor elétrico têm nos consumidores os seus principais agentes, justificando, mais ainda, que os reguladores europeus devam reequacionar as suas práticas para com esses agentes.

A mudança do setor elétrico é atualmente impulsionada pelos consumidores, antecipando-se a democratização deste sector (Szulecki, 2018), basilar para o funcionamento das economias e para o bem-estar das populações em geral: “Consumers will be the key drivers acting as the crucial market makers rather than market takers of today” (Sandys et al., 2017).

A democratização do setor elétrico apoia-se na inovação tecnológica que possibilita a diminuição dos custos de produção de energia elétrica com base em fonte de energias renováveis e o incremento da flexibilidade na gestão da procura e da oferta de energia elétrica. Nesse novo contexto, o acesso à energia mais barata por um grande número de consumidores potenciará o maior uso da energia elétrica para novos fins, dos quais se destacam a mobilidade e o aquecimento (designadamente através de bombas de calor) e, conseqüentemente, acelerará a eletrificação das economias, contribuindo para a transição para uma economia ambientalmente mais sustentável.

Num setor que sofrerá muito provavelmente uma atomização de várias das suas atividades, em particular ao nível da produção de energia elétrica, a distribuição de energia elétrica deverá ser, paradoxalmente, uma atividade estratégica para os grandes grupos energéticos, por permitir manter uma ligação privilegiada com consumidores tendencialmente mais autónomos e ativos. A garantia de que as atividades de rede, em particular a distribuição de energia elétrica, sejam devidamente separadas das restantes atividades em todas as suas dimensões económicas (impedindo a alocação indevida de recursos ou garantindo a separação da imagem corporativa) deverá ser uma preocupação reforçada para o regulador, de modo

¹⁴ Tais como os leilões ocorridos em 2019 e em 2020 em Portugal para reserva de capacidade de injeção na rede para produção de energia solar fotovoltaica.

a evitar que as dinâmicas de mercado, que têm permitido a transformação tecnológica e a democratização do setor, sejam destorcidas pela aplicação de estratégias de grupo empresariais.

Ademais, a transição energética não se fará de uma forma instantânea e à mesma velocidade nas diferentes esferas de organização dos setores elétricos, pelo que nem todos os consumidores poderão desde já beneficiar desta transição, sendo expetável que os reguladores contribuam para acautelar os impactes dessa transição. Na verdade, muitos consumidores ficarão ainda integralmente dependentes do fornecimento de energia elétrica através das atuais infraestruturas de rede. A utilização mais pontual das redes de transporte e de distribuição, bem como das centrais de produção convencionais, que se manterão úteis para garantir a resiliência e segurança do sistema, impactará na valorização dessas infraestruturas, principalmente para os consumidores totalmente dependentes do fornecimento de energia através das redes de distribuição.

Finalmente, a promoção da eficiência energética neste contexto constitui outro desafio regulatório. A amplitude das competências técnicas dos reguladores energéticos permite-lhes uma visão holística do setor elétrico, a qual poderá assegurar que as medidas de eficiência energética sejam coerentes com outras que visam a sustentabilidade económica e ambiental deste setor e, conseqüentemente, podem garantir a eficácia dessas medidas sem gerar subsidiasões cruzadas entre consumidores.¹⁵

Em suma, neste novo contexto organizativo, em que os consumidores passarão a ter um papel central, a relevância dos reguladores europeus, nos quais se inclui a ERSE, não deverá diminuir. Pelo contrário, esta transição organizativa não poderá ocorrer sem o devido acompanhamento dessas entidades, de forma a garantir que o quadro regulamentar e legislativo facilite e não obstaculize este processo (Jacobs, 2017).

¹⁵ No entanto, nem sempre a promoção da eficiência energética é incluída pelos reguladores europeus no conjunto dos seus objetivos estratégicos deixando que outros organismos públicos tomem iniciativas nesta matéria. Uma consulta aos documentos produzidos pelo Conselho Europeu dos Reguladores de Energia (CEER) permite observar que o tema da eficiência energética não tem merecido nenhuma reflexão particular. O caso português, materializado na definição do Plano de Promoção da Eficiência Energética por parte da ERSE, constitui uma notável exceção.

REFERÊNCIAS BIBLIOGRÁFICAS

- European Union Agency for the Cooperation of Energy Regulators and the Council of European Energy Regulators – ACER/CEER (2019) *Annual Report on the Results of Monitoring the Internal Electricity and Natural Gas Markets in 2018 – Electricity and Gas Retail Markets Volume*.
- AT Kearney (2017) *Mergers and Acquisitions in Utilities*. <https://www.de.kearney.com/documents/1117166/1117568/Mergers+and+Acquisitions+in+Utilities+2017.pdf/6d21b6d9-761f-1b8d-feba-f7c79382c139?t=151785755271>, acessado a 24/06/2021.
- Banco Mundial (2017) *Practical Guidance for Defining a Smart Grid Modernization Strategy: The Case of Distribution*.
- Baumol, W.J.; Panzar, J.C; Willig, R. D. (1982) *Contestable Markets and the Theory of Industry Structure*. New York, Harcourt Brace Jovanovich.
- Chiradeja, P. (2005) Benefit of distributed generation: A line loss reduction analysis, transmission and distribution. *Conference and Exhibition: Asia and Pacific Conference, IEEE.*, 1-5.
- Chiradeja, P; Ngaopitakkul, A. (2013) The impacts of electrical power losses due to distributed generation integration to distribution system. *Electrical Machines and Systems Conference, IEEE*.
- Comissão Europeia (2011) *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Smart Grids: from innovation to deployment*, (COM202 final).
- Comissão Europeia (2017) *Comission Staff Working Document, accompanying the third report on the Energy Union, Energy Union Factsheet United Kingdom*.
- Comissão Europeia (2019) *Benchmarking smart metering deployment in the EU-28*.
- Council of European Energy Regulators (2015) *The Future Role of DSOs*.
- Council of European Energy Regulators (2017) *Guidelines of Good Practice for Flexibility Use at Distribution Level*, (Consultation Paper Ref: C16-DS-29-03).
- Council of European Energy Regulators (2018) *Incentives Schemes for Regulating Distribution System Operators, including for innovation*.
- Crampes, C.; Prices, W. (2017) *Empowering electricity consumers in retail and wholesale markets*, CERRE (Centre on Regulation in Europe).
- Dai, Z.; Liu, J. T. (2013) Impacts of large dams on downstream fluvial sedimentation: An example of the Three Gorges Dam (TGD) on the Changjiang (Yangtze River). *Journal of Hydrology*, 480, 10-18.
- Esteves, J.; Pousinho, H.; Oliveira, P.; Roldão, P.; Faias, S.; Marques, V.; Santos, A.; Santos, V. (2016?) *Smart grid: Uma Visão da Regulação*. In Castro, N. J. Dantas, G. de A (Eds). *Políticas Públicas para Redes Inteligentes*, Rio de Janeiro, Publit Soluções Editoriais, pp.105-128.
- ECOFYS (2014) *Subsidies and costs of EU Energy Final Report*.
- Entidade Reguladora dos Serviços Energéticos (2017). *Parâmetros de Regulação para o período 2018-2020*.
- Entidade Reguladora dos Serviços Energéticos (2018) *Instrumentos de Oferta e de Procura na Gestão do Sistema Elétrico Nacional*.
- Entidade Reguladora dos Serviços Energéticos (2019) *Tarifas e Preços para a Energia Elétrica e Outros Serviços*.
- European Regulators' Group for Electricity and Gas (2009) *Position Paper on Smart Grids, An ERGEG Public Consultation Paper*.
- Fuchs, G. (2014) The Governance of innovations in the energy sector: Between adaptation and exploration. *Science & Technology Studies*, 27(1), 34-53.

- Gates, B. (2015) *Energy Innovation: Why We Need It and How to Get It*. <https://www.de.kearney.com/documents/1117166/1117568/Mergers+and+Acquisitions+in+Utilities+2017.pdf/6d21b6d9-761f-1b8d-feba-f7c79382c139?t=151785755271>, accedido a 24/06/2021.
- GfK Belgium (2017) *Study on Residential Prosumers in the European Energy Union*. Report by GfK Belgium Consortium, May.
- Goulden, M.; Bedwell, B.; Rennick-Egglestone, S.; Rodden, T.; Spence, A. (2014) Smart grid, smart users? The role of users in demand side management. *Energy Research and Social Science*, 2, 21-29.
- Hatzigiargyriou, N.; Asano, H.; Iravani, R.; Marnay, C. (2007) Microgrids: An overview of ongoing research, development, and demonstration projects. *IEEE power & energy magazines*, 5(4), 78-94.
- International Renewable Energy Agency (2017) *Electricity Storage and Renewables: Costs and Markets to 2030*, October 2017 International Renewable Energy Agency (2018) *Renewable power generation in 2017*.
- Jacobs, S. (2017) The energy Prosumer. *Ecology Law Quarterly*, 43(3), 519-559.
- Lavrijssen, S.; Carrillo Parra, A. (2017) Radical Prosumer innovations in the electricity sector and the impact on Prosumer regulation. *Sustainability, MDPI, Open Access Journal*, 9(7), 1-21.
- Leïße, I. (2013) *Efficient Integration of Distributed Generation in Electricity Distribution Networks* (Doctoral Dissertation). Retrieved from <https://www.ica.lth.se/publications/Theses/LTH-IEA-1071.pdf>
- Lima, I.; Ramos, F.; Bambace, L.; Reinaldo, R. (2008) Methane emissions from large dams as renewable energy resources: A developing nation perspective. *Mitigation and Adaptation Strategies for Global Change*, 13, 193-206.
- Mc Kinsey (2018) Sticker shock: Why utilities now must rethink customer affordability. <https://www.de.kearney.com/documents/1117166/1117568/Mergers+and+Acquisitions+in+Utilities+2017.pdf/6d21b6d9-761f-1b8d-feba-f7c79382c139?t=151785755271>, accedido a 24/06/2021.
- Marques, V. (2003) Poder de mercado e regulação nas indústrias de rede. GEPE, Ministério da Economia.
- Marques, V.; Bento, N.; Moisés, P. (2014) The “Smart Paradox”: Stimulate the deployment of smart grids with effective regulatory instruments. *Energy*, 69(1), 96-103
- Poudineh, R.; Peng, D.; Mirnezami, S. R. (2017) *Electricity Networks: Technology, Future Role and Economic Incentives for Innovation*, Oxford Institute for Energy Services paper EL 27.
- Price, W.; Deller, D. (2015) *Report on Affordability of Utilities’ Services: Extent, Practice*, (CERRE Policy).
- Sandys, L.; Hardy, J.; Green, R. (2017) *Reshaping Regulation Powering from the future*, Grantham Institute, Imperial College, London.
- Szulecki, K. (2018) Conceptualizing energy democracy. *Environmental Politics*, 27, 21-41.
- Wang, C; Nehrir, H. (2004) Analytical approaches for optimal placement of distributed generation sources in power systems. *IEEE Transactions on Power Systems*, 19(4), 2068-2076.

DÉFICE TARIFÁRIO, RENDAS EXCESSIVAS E PRIVATIZAÇÕES

Tariff Deficit, Excessive Rents, and Privatisation

João Confraria

Received for publication: June 15, 2020

Revision accepted for publication: April 9, 2021

RESUMO

O défice tarifário e a possibilidade de haver rendas excessivas têm sido a principal fonte de polémica sobre a intervenção do Estado no setor elétrico nos últimos 25 anos e foram objeto de uma Comissão Parlamentar de Inquérito, cujo Relatório, assim como as declarações de voto dos partidos, são uma fonte de informação bem sistematizada sobre o que se passou. Neste trabalho, discutem-se os fundamentos e as consequências da política de défices tarifários e de políticas que conduziram ao que tem sido considerado como rendas excessivas. Sobre a política de défices tarifários não se encontram fundamentos nem de eficiência, nem de equidade e é duvidoso que se possa justificar com base num argumento de estabilidade tarifária, que foi o argumento mais utilizado formalmente. A criação de rendas excessivas tem no mercado de eletricidade os efeitos negativos associados ao poder de mercado. Numa perspetiva de equilíbrio geral, é possível fundamentar uma política de rendas excessivas com base em critérios de eficiência, mas não foi esta a abordagem seguida pelos governos portugueses que poderá ter estado mais centrada na obtenção de vantagens eleitorais com aquela política. Em sentido diferente, pode-se considerar que em 1995 foi adotado um modelo de regulação contratual que pode gerar a criação de rendas, sem que estas devam ser consideradas como excessivas. Concluindo, qualquer que seja a interpretação dada ao trabalho da Comissão, a sua principal utilidade parece ser a de evidenciar formas de melhorar o processo de decisão pública.

Palavras-chave: Privatizações; preços de eletricidade; rendas, excedente do consumidor.

ABSTRACT

A tariff deficit and excessive rents have been the source of public discussion on energy policy in the last 25 years, and the subject of a recent inquiry by a Parliamentary Commission. The Commission concluded there were excessive rents created by policy options. The rationale for these rents, from the point of view of the policy makers, was that they would increase privatisation revenues and support industrial and environmental policies. The rationale for the tariff deficit was that it would increase price stability. In this paper the Commission's view that there were excessive rents is initially accepted, and these

explanations are discussed from the perspective of public interest and of public choice and economic theories of regulation. It is argued that the best explanation for these policies is not based on the public interest, but on vote seeking behaviour by the governments. It is also suggested that excessive rents in the energy sector might have been sustained on the grounds of public interest, based on efficiency, if complemented by credible commitments to reduce income taxes. However, some doubts remain about the relevant concept of excessive rents. The Commission's report is ambiguous on these. Moreover, the contractual regulation arrangements adopted by Portuguese public decision-makers in the energy sector, since the mid-1990s, may well be the source of informational rents that should not be considered as excessive. In any case, it is argued that the Commission's report suggests, even if implicitly, to improve public decision-making, in the energy sector and in state intervention in general. Keywords: Privatisation; electricity prices; rents; consumer surplus.

JEL Classification: L32; L38; L51; L94.

Agradecimento: Agradeço ao Eng. Rui Sérgio e a um revisor anónimo comentários a uma versão preliminar deste trabalho. No entanto, as opiniões defendidas, assim como qualquer erro ou imprecisão que possa existir, são da exclusiva responsabilidade do autor.

1. INTRODUÇÃO

O déficit tarifário e a possibilidade de haver rendas excessivas têm sido uma fonte de polémica sobre a intervenção do Estado no setor elétrico nos últimos 15 anos. Houve um inquérito parlamentar e a Comissão Parlamentar de Inquérito (CPI) concluiu pela existência de rendas excessivas, mas com divergências significativas entre os seus membros, evidenciadas pelas declarações de voto apresentadas por todos os grupos parlamentares.

Neste trabalho discutem-se os fundamentos e as consequências da política de défices tarifários e das políticas que conduziram ao que a CPI considerou terem sido rendas excessivas aceitando-se o enquadramento legal seguido pela CPI.

Sobre a política de défices tarifários não se encontram fundamentos nem de eficiência, nem de equidade e é duvidoso que se possa justificar com base no argumento habitual de estabilidade tarifária. Essa política parece ser mais facilmente explicada como um aproveitamento pelo governo, com fins eleitorais, de alguns problemas de ação coletiva dos consumidores (eleitores) que valorizariam os seus ganhos imediatos, sem consideração de que os custos futuros seriam superiores.

De acordo com depoimentos apresentados na CPI, a criação de rendas excessivas começou por estar associada à privatização da EDP e, mais tarde, da REN. Neste trabalho discute-se sobretudo o que se admitiu terem sido os efeitos da criação dos CAE e dos Custos de Manutenção do Equilíbrio Contratual (CMEC). No que diz respeito somente ao mercado de eletricidade, aceitando a conclusão de que existiram rendas excessivas, tiveram os efeitos negativos associados ao poder de mercado, subindo os preços de forma ineficiente. No entanto, tendo em conta toda a economia, é possível defender uma política de rendas excessivas com base em critérios de eficiência, sobretudo desde que se assumisse, através de compromissos credíveis, que os custos impostos aos consumidores de eletricidade através de preços mais altos seriam compensados com reduções de impostos, sobre os rendimentos do trabalho, por exemplo. No entanto, não foi esta a abordagem seguida por sucessivos governos que, pelo menos até 2011, não admitiram haver rendas excessivas.

A política na origem das assumidas rendas excessivas parece ter sido iniciada através de um processo de decisão pública relativamente simples. Numa perspetiva de interesse público, o processo pode ser explicado considerando que, decididas as privatizações, por razões que aqui não são discutidas, a maximização da receita a obter era um objetivo de interesse público indiscutível. Estava em causa a receita do Estado, matéria que regra geral não se discute entre nós.

Numa perspetiva diferente, na linha das teorias das escolhas públicas, uma receita maior permitiria ao governo em funções obter recursos para flexibilizar a sua restrição orçamental utilizando-os em objetivos imediatos. Os custos seriam distribuídos através de preços mais altos ao longo dos anos seguintes. Com isto haveria vantagens eleitorais se os consumidores (e eleitores) subestimassem, ou ignorassem, os custos que iriam suportar no futuro. Nesta perspetiva, ao longo dos anos a política de déficit tarifário terá sido complementar, no sentido político, da política de rendas excessivas. Dilatou no tempo os seus efeitos, tornando-os menos visíveis para os consumidores.

Depois desta discussão, que tem como pressuposto a aceitação das conclusões da CPI de que houve rendas excessivas, é analisado o pressuposto. O conceito de rendas excessivas é

complexo e não é evidente que na CPI tenham sido avaliadas todas as suas dimensões, nem todas as consequências dos modelos de regulação do setor elétrico que foram adotados a partir de 1995. E assim, não é de excluir que não tenha havido rendas excessivas – pode ter havido rendas, mas não seria evidente que tivessem de ser classificadas como excessivas, se se entender que esta designação implica que se trate de rendas que devem ser eliminadas.

Não significam estas dúvidas que do relatório da CPI se não retirem conclusões. De facto, quer se aceitem as suas conclusões, quer se entenda que não ficou demonstrada a existência de rendas (ou mesmo que, pelo menos nalguns casos, não havia rendas) o relatório da CPI permite sugerir linhas de atuação futura, não tanto em termos de compensações a pagar aos consumidores de eletricidade, que é uma das preocupações do relatório, mas sobretudo em termos do processo de decisão pública, no mercado elétrico ou em qualquer outro, discutidas como comentários finais deste trabalho.

2. DÉFICE TARIFÁRIO

2.1. ORIGENS DO DÉFICE TARIFÁRIO

Em 1995 estabeleceu-se que a variação das tarifas e preços a clientes finais em baixa tensão não deveria ser superior à taxa de inflação e que o valor dos custos não refletidos nessas tarifas e preços seria repercutido nas tarifas e preços dos cinco anos seguintes.¹ Foi esta a origem do que depois se veio a chamar o défice tarifário no setor da eletricidade. Existe um défice tarifário se as receitas autorizadas para cobrir os custos aceites num determinado ano não são alcançadas com os preços aprovados para esse ano, pagando-se a diferença ao longo dos anos seguintes.

Em 2006 o teto estabelecido em 1995 para os preços da eletricidade, a taxa de inflação, foi pela primeira vez uma restrição ativa na definição dos preços regulados. Nesse ano foi criado um défice tarifário de 335 milhões de euros, para manter o crescimento dos preços ao nível da inflação esperada, de 2,9%, em vez do crescimento de preços de 14,4% que seria necessário para cobrir a variação esperada de custos (CPI, 2019, p. 92).

Em 2006 esta restrição foi eliminada, pelo Decreto-Lei n.º 29/2006, de 15 de fevereiro, que, no contexto da liberalização dos mercados de eletricidade, estabeleceu as bases gerais da organização e funcionamento do Sistema Elétrico Nacional, bem como as bases gerais aplicáveis ao exercício das atividades de produção, transporte, distribuição e comercialização de eletricidade e à organização dos mercados de eletricidade. As regras de formação de preços deveriam ser estabelecidas através do Regulamento Tarifário da ERSE, a autoridade reguladora.²

No entanto, perante a inevitabilidade de novo défice em 2007, o governo voltou a intervir em 2006, definindo no fim do ano novas regras aplicáveis à recuperação do défice tarifário, assim como o regime aplicável aos ajustamentos tarifários apurados em cada ano.

¹ Art. 4.º do Decreto-Lei n.º 187/95, de 27 de julho, com a redação do Decreto-Lei n.º 44/97, de 20 de fevereiro.

² Art. 62.º do Decreto-Lei n.º 29/2006, de 15 de fevereiro, e art. 66.º do Decreto-Lei n.º 172/2006, de 23 de agosto. Este último diploma previa que o défice tarifário acumulado até 31 de dezembro de 2006 fosse recuperado nos cinco anos seguintes.

Basicamente, estabeleceu-se que a variação de preços em 2007 não poderia ser superior em mais de 6% aos preços de 2006, que o défice tarifário acumulado seria recuperado no prazo de 10 anos e que isso se verificaria através da sua integração na tarifa de uso global do sistema relativa à baixa tensão.³ Esta última regra significava que o grupo de consumidores que tinha beneficiado com o défice tarifário em 2006, os consumidores de eletricidade em baixa tensão, era também o grupo que iria suportar os seus encargos.

Em 2008, flexibilizou-se a formação de défices tarifários anuais. Estes poderiam verificar-se, por decisão do ministro responsável pela energia, na sequência de proposta fundamentada da ERSE, a ser apresentada ao governo. Essa proposta deveria dar-se na sequência de uma avaliação fundamentada dos efeitos das flutuações em custos estruturais, como os custos de aprovisionamento de energia.

Depois de 2008 o défice tarifário aumentou sobretudo em consequência de dois fatores: os sobrecustos resultantes do quadro legal que tinha sido criado para a produção de energia com fontes renováveis, que veio a ser designado como produção em regime especial, assim como com a aceitação de diferimentos de pagamentos anuais relativos a um modelo de regulação contratual que estabelecia uma regra de remuneração baseada no que foi designado como Custos de Manutenção do Equilíbrio Contratual (CMEC).⁴ O défice tarifário passou a ser prática habitual em vez de excecional. Do valor inicial de 335 milhões relativo ao diferimento de 2006 atingiu-se um máximo de 5 080 milhões em 2015 (CPI, 2019, p. 110).

2.2. CONSEQUÊNCIAS DO DÉFICE TARIFÁRIO NO MERCADO DE ELETRICIDADE

Para sintetizar os principais aspetos desta política vamos considerar um mercado de eletricidade em que existem somente dois períodos, o período 0 e o período 1. O período 0 está representado na Figura 1. A procura é descrita pela curva D_0 e, na ausência de défice tarifário e de rendas excessivas, os custos unitários aceites pelo regulador são descritos por C_0 . O objetivo do regulador é igualar os preços aos custos unitários aceites em cada um dos períodos.

Sem rendas excessivas nem défice tarifário o mercado está em E_0 . O preço regulado é P_0 e tem-se a produção Q_0 . Se o custo unitário, C_0 , corresponder ao custo marginal social

³ Art. 5.º e art. 6.º do Decreto-Lei n.º 237 B/2006, de 18 de dezembro.

⁴ Sobre a definição de produção em regime especial, vejam-se os artigos 17.º, 18.º e 73.º-A do Decreto-Lei 29/2006 com a redação do Decreto-Lei n.º 78/2011, de 20 de junho. Em 2008, pagamentos relativos a qualquer tipo de encargo, custo ou proveito decorrente do regime estabelecido no Decreto-Lei n.º 240/2004, de 27 de dezembro, que procedeu à definição das condições de cessação dos contratos de aquisição de energia (CAE) e à criação de medidas compensatórias relativamente à posição de cada naqueles contratos. (artigo 2.º do Decreto-Lei 165/2008). No entanto, em 2011, tendo em conta também o aumento do IVA sobre a eletricidade e as medidas tomadas no quadro da crise económica, admitiu-se que os ajustamentos anuais determinados nos termos do Decreto-Lei n.º 240/2004, de 27 de Dezembro, relativos ao ano de 2010 seriam repercutidos nos proveitos permitidos de 2013 do operador de rede de distribuição em média tensão (MT) e alta tensão (AT) (art. 2.º do Decreto-Lei n.º 109/2011 de 18 de novembro). No final de 2012 viria a ser autorizado outro diferimento nos pagamentos anuais relativos aos CMEC de 2011 e de 2012 (art. 2.º do Decreto-Lei n.º 256/2012, de 29 de novembro). Já em 2014, determinou-se que o montante não repercutido dos ajustamentos anuais referente ao ano de 2012 seria repercutido, em partes iguais, nos proveitos permitidos de 2017 e 2018 do operador da Rede Nacional de Distribuição de Eletricidade em média tensão (MT) e alta tensão (AT) (art. 2.º do Decreto-Lei n.º 32/2104, de 28 de fevereiro).

esta afetação de recursos é eficiente. Se corresponder ao custo médio tem-se tipicamente um resultado *second best*, mas que é muitas vezes aceite na regulação, quando existem custos fixos e afundados.

Um défice tarifário é introduzido se o regulador admitir que, no período 0, o custo C_0 é muito elevado porque corresponde a um aumento muito grande relativamente ao preço que se verificava antes. Neste caso, apesar de se reconhecer como relevante para regulação o custo unitário C_0 , é estabelecido um preço C_{do} , inferior a C_0 , e considerado socialmente mais aceitável. Isto implica um prejuízo para a empresa, a ser recuperado no período 1.

Do lado dos consumidores importa ver como entendem esta política. Se só considerarem o que se passa no período 0, ignorando que vão ter de pagar o défice depois, o equilíbrio de mercado passa para E_{do} . O excedente dos consumidores aumenta $A + B$. A área A corresponde a uma transferência de rendimento da empresa regulada para os consumidores e a área B é um ganho que acontece porque há um aumento de consumo, pois a disponibilidade para o pagar é superior ao preço. A empresa regulada perde a área $A + B + C$.

No período seguinte a empresa regulada tem direito a recuperar o lucro perdido. O valor que tem a receber é superior ao ganho dos consumidores no período 0, porque pela quantidade ($Q_{do} - Q_0$) vão pagar mais do que o valor que lhe atribuem, sendo o pagamento em excesso dado pela área C.⁵ A questão é saber como se vai dar essa recuperação – quem paga e como paga. Poderia ser através de um preço mais alto por kwh, ou poderia ser através da introdução de uma componente fixa, independente do consumo no preço. A primeira opção legislativa foi que os défices tarifários de 2006 e de 2007 deveriam ser recuperados através da sua integração na tarifa de uso global do sistema relativa à baixa tensão, nos termos do Regulamento Tarifário.⁶

Daqui resultaram duas consequências. Em primeiro lugar, criou-se uma diferença na tarifa de uso global do sistema, entre a alta e média tensão por um lado, e a baixa tensão, por outro. Procurava-se com isto, naturalmente, garantir que seriam os consumidores de baixa tensão, os beneficiados com a descida de preços em 2006 e em 2007, que pagariam os défices tarifários correspondentes. Em segundo lugar abandonou-se a possibilidade de o défice tarifário ser pago através de um preço fixo, independente do consumo. Pela construção do sistema tarifário, nos termos do Regulamento Tarifário, a tarifa de uso global do Sistema não se traduz por uma componente fixa do preço. Repercute-se nas tarifas através do preço por Kwh e através do preço da potência contratada. Este último é, para cada classe de consumo, independente do consumo – mas varia com a classe de consumo (c.f. Garcia et al., 2016). Neste caso, incorpora-se esta opção regulativa admitindo que há um aumento do custo unitário do período 1, relativamente à situação sem défice tarifário, de C_0 para C_{d1} . O mercado no período 1 está representado na Figura 2. Para facilitar a comparação admite-se que a procura, D_0 , e o custo unitário, C_0 , têm valores iguais aos do período 0 e

⁵ A hipótese de comportamento míope pode ser abandonada facilmente. Admitindo que a procura de cada consumidor é, no período 1, igual à do período 0, se os consumidores interiorizarem que irão pagar no período seguinte para compensar esta descida de preços no período 0, o preço “integral” que enfrentam no período 0 é C_0 – e o seu consumo não se alteraria. Neste caso, não haveria variações de excedentes de mercado, haveria somente uma transferência de rendimento da empresa para os consumidores no primeiro período, dada pela área A.

⁶ Art. 2.º do Decreto-Lei n.º 237-B/2006, de 18 de novembro.

que a taxa de juro é 0. Garantindo-se que a empresa regulada não tem prejuízo no conjunto dos dois períodos, $G = A + B + C$.

A quantidade é Q_{d1} , havendo, relativamente à quantidade sem défice Q_0 , uma contração que é fonte de perda de eficiência. O preço é $C_{d1} > C_0$, ou seja, vai ser superior ao preço que se verificaria sem défice e o aumento de preços no período 1 é superior à redução de preços no período 0, pois o défice tarifário distribui-se por uma quantidade consumida mais pequena. A posição da empresa regulada não varia neste caso. No entanto, ao longo dos dois períodos há uma perda líquida dos consumidores, dada pela soma das áreas C e F. Os consumidores perdem mais do que aquilo que recebem no período inicial, quando se formou o défice. Assim, a política de défice tarifário é uma política de redistribuição do rendimento entre dois períodos, mas, associada a esta transferência de rendimento, é criada ineficiência.

Esta ineficiência depende da elasticidade procura preço direta, da quantidade procurada, das receitas da empresa regulada e da variação de preços em cada um dos períodos. Sobre a elasticidade procura preço direta no mercado de eletricidade português a evidência disponível é escassa. Silva et al. (2017) estimam que a elasticidade procura preço no mercado residencial pode ser relativamente elevada no médio e no longo prazo, com valores entre -0,6 e -0,8 sugerindo que aumentos de preço podem implicar quedas significativas na procura, sendo maiores os efeitos nos níveis de rendimento mais baixos, em que a procura é elástica. Outros estudos, menos detalhados, sugerem uma procura de eletricidade mais rígida com valores de -0,3 e -0,09 (Rodrigues, 2012 e Esperança 2015, respetivamente).⁷

O quadro 1 apresenta os resultados de uma simulação para valores da elasticidade procura preço nestes intervalos, tendo como referência um consumo residencial de 14 mil milhões Kwh, próximo do nível de consumo na segunda metade da década passada e um custo aceite pelo regulador de € 0,105/kwh. Este seria o preço na ausência de défice tarifário (e, no quadro do mercado português, sem os impostos). O valor do défice por Kwh foi definido para que, com uma elasticidade procura-preço direta de -0,3, se tenha um valor do défice tarifário igual ao que se verificou em 2006, no primeiro ano da sua aplicação.

Este exemplo quantifica a ideia anterior, segundo a qual a política de défice tarifário tem um efeito redistributivo, com uma perda para os consumidores. No período 0 há uma aparente redistribuição do rendimento da empresa para os consumidores e no período 1 dos consumidores para a empresa. Na prática, é como se o governo obrigasse os consumidores do período 1 a fazer um empréstimo aos do período 0.⁸ Para esta transferência de recursos no tempo, a sociedade tem uma perda de eficiência €0,07 por cada euro que é emprestado, para uma elasticidade procura-preço de -0,3. Com uma procura mais elástica (rígida), a perda de eficiência associada ao défice tarifário é maior (menor).

⁷ Noutros trabalhos não se encontrou relação significativa entre a procura de eletricidade e o preço (por exemplo, Carmona 2006). Na literatura encontra-se uma grande variação de resultados na estimação da elasticidade procura-preço de eletricidade. Por exemplo, Espey e Espey (2004) encontram um valor mediano para a elasticidade procura preço de curto prazo de -0,28, numa amostra de 36 estudos), mas com valores que vão de -0,004 a -2,01. Para as elasticidades procura-preço de longo prazo a dispersão é semelhante, mas com uma procura mais elástica, com o valor mediano de -0,81 (citado em Bergh, 2008).

⁸ Um governante da época referiu que os consumidores deviam dinheiro à EDP, com o défice tarifário. Mas o quadro legal do défice tarifário pode também ser visto como um empréstimo dos consumidores futuros aos consumidores atuais, intermediado pela empresa regulada, que recebe para o efeito uma remuneração – para este efeito o produtor de eletricidade age como um banco, por imposição do governo.

Figura 1: Efeito do défice tarifário no período 0

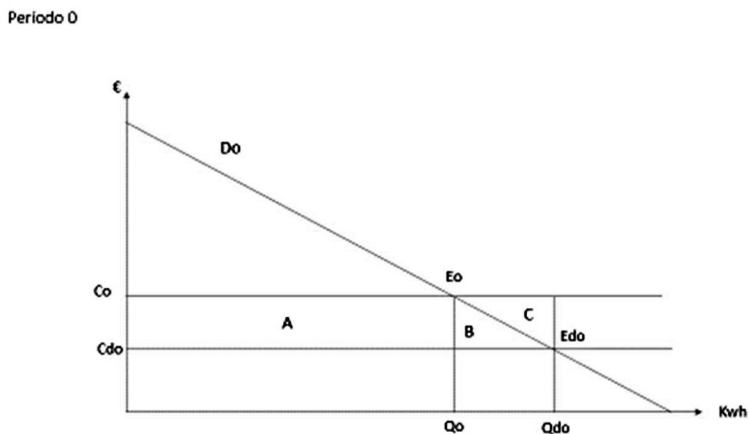
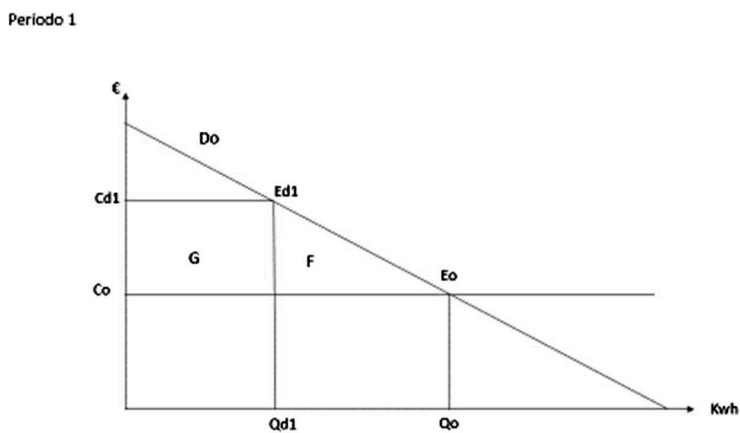


Figura 2: Efeito do défice tarifário no período 1



Quadro 1: Perda de eficiência com o défice tarifário

| Período 0 | | | | | | | | |
|---------------|--------------------|---------------|----------------|----------------|------------|-----------|------|-----------|
| ε | Δc (euros) | c_0 (euros) | q_0 (kwh) | qd_0 (kwh) | cd_0 (€) | A+B+C (€) | C(€) | |
| -0,7 | -2.3% | 0,105 | 14 000 000 000 | 16 100 000 000 | 0,083 | 362,3 | 23,6 | |
| -0,3 | -2.3% | 0,105 | 14 000 000 000 | 14 900 000 000 | 0,083 | 335,3 | 10,1 | |
| -0,09 | -2.3% | 0,105 | 14 000 000 000 | 14 270 000 000 | 0,083 | 321,1 | 3,0 | |
| Período 1 | | | | | | | | |
| ε | Δc (euros) | c_0 (euros) | q_0 (kwh) | qd_1 (kwh) | cd_1 (€) | G (€) | F(€) | C+F/G (€) |
| -0,7 | 0,03325 | 0,105 | 14 000 000 000 | 10 896 666 667 | 0,138 | 362,3 | 51,6 | 0,21 |
| -0,3 | 0,02586 | 0,105 | 14 000 000 000 | 12 965 600 000 | 0,131 | 335,3 | 13,4 | 0,07 |
| -0,09 | 0,02340 | 0,105 | 14 000 000 000 | 13 719 200 000 | 0,128 | 321,0 | 3,3 | 0,02 |

Nota: As áreas C e F são dadas, aproximadamente, por, $\frac{1}{2} \left(\frac{\Delta c}{c_0} \right)^2 c_0 q_0 \varepsilon$ em que ε é a elasticidade procura-preço direta.

A ineficiência causada pelo défice tarifário poderia ser eliminada caso o pagamento do défice no período 1 se verificasse com base num preço fixo independente do consumo. É duvidoso que uma solução deste tipo fosse politicamente viável. Se se tratasse de um montante por consumidor igual para todos, seria uma política regressiva do ponto de vista redistributivo, com os consumidores com níveis de consumo mais baixos a pagar parte do défice dos consumidores com níveis de consumo mais elevados. O problema poderia, no entanto, ser resolvido se cada consumidor no período 1 pagasse o défice que tinha originado no período zero. Neste caso, o pagamento feito no período 1 seria independente do consumo realizado no período 1, dependeria do consumo no período 0 – e neste sentido contribuiria também para que os consumidores não tivessem um comportamento míope, compreendendo no período 0 que sempre teriam de pagar o preço integral da eletricidade – e assim seria também eliminada a ineficiência nesse período. O resultado seria mais eficiente do que o da política que foi adotada.

A discussão pode ser um pouco mais densificada, introduzindo a possibilidade de haver rendas excessivas (Figuras 3 e 4).

Neste caso a regulação faz-se em função dos custos reconhecidos Cr , sendo $Cr-Co$ o valor unitário da renda excessiva. Tal como no exemplo anterior admite-se que a procura e os custos unitários aceites são iguais nos 2 períodos e que a taxa de juro é zero. Sem défice o mercado está em Er. Quando o défice é criado, no período 0, há uma descida de preço de Cr para Cd_0 , sendo a diferença correspondente ao défice tarifário por unidade.⁹ O ganho dos consumidores é $A+B$. Note-se que neste caso existe uma diferença entre a perda real da empresa, $A - D$, e a perda da empresa reconhecida pela autoridade pública, $A + B + C$, pois para a autoridade pública o custo relevante é Cr , mas a expansão do consumo

⁹ Também se poderia admitir que com o défice tarifário o preço ficasse abaixo de Co .

ocasionada pela descida do preço é lucrativa para a empresa regulada devido à hipótese de haver rendas excessivas.

No período 1, o preço está acima de C_r , para que se pague o défice. A perda dos consumidores relativamente ao ponto E_r é $H + F$ e o ganho real dos produtores $H - G$, sendo $H = A + B + C$. No conjunto dos dois períodos os consumidores perdem $C + F$ e as empresas ganham $B + C + D - G$. Neste caso, a descida de preços do período 0 corresponde a um ganho de eficiência nesse período pois as rendas excessivas são diminuídas, mas a perda da empresa é recuperada no período seguinte, com preços mais altos.

Os governos que, ao longo dos anos, adotaram a política de défices tarifários no setor elétrico basearam-se na ideia de que a estabilidade tarifária era um valor em si mesmo, e que os défices tarifários eram uma política adequada para promover a estabilidade tarifária.

Em 2006, por exemplo, perante uma proposta de preços regulados apresentada pela ERSE que implicaria um aumento de preços de cerca de 14% considerou-se que os aumentos propostos, a verificarem-se, teriam efeitos negativos, tanto ao nível da inflação como do poder de compra dos consumidores e assim importaria aumentar para 10 anos o período de recuperação do défice tarifário acumulado por forma a diluir os seus efeitos ao longo do tempo (Preâmbulo do Decreto-Lei n.º 237-B/2006 de 18 dezembro).

Em 2008 foi de novo considerado de interesse público que, em cada ano, pudesse haver desvios entre os preços e os custos e que o défice daqui resultante fosse eliminado de forma gradual ao longo do tempo, mitigando a volatilidade tarifária resultante de fatores como a variação de preços dos combustíveis sólidos, a variação na produção hídrica e, de uma forma mais ambígua das condições relacionadas com a liberalização (Preâmbulo do Decreto-Lei n.º 185/2008 de 21 de agosto.)

Figura 3: Efeito do défice tarifário no período 0, com rendas excessivas

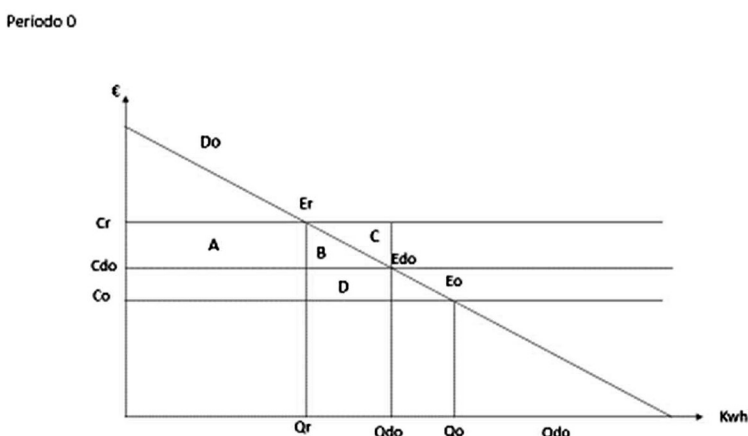
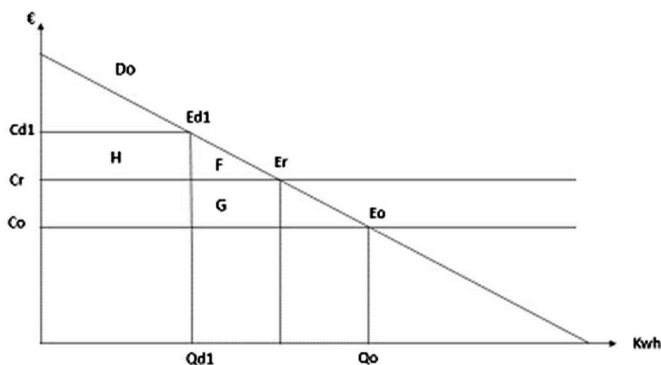


Figura 4: Efeito do déficit tarifário no período 1, com rendas excessivas

Período 1



2.3. ARGUMENTOS DE INTERESSE PÚBLICO PARA A POLÍTICA DE DÉFICE TARIFÁRIO

A análise do déficit tarifário envolve dois aspetos: saber se o déficit tarifário contribui para a estabilidade de preços e saber se faz sentido que a estabilidade de preços seja, neste caso, um objetivo de interesse público.

Quanto ao primeiro aspeto, a resposta não é evidente. Admita-se, no caso representado nas Figuras 1 e 2, que antes do período 0 o custo unitário e o preço eram Cd_0 . No período 0 houve um aumento do custo unitário para Co e a política de déficit tarifário teve como objetivo manter o preço em Cd_0 . No entanto, para pagar o déficit assim constituído o preço no período 1 vai subir para Cd_1 que é maior do que Co . Ou seja, com déficit tarifário tem-se a sequência de preços Cd_0, Co, Cd_1 . Sem déficit tarifário ter-se-ia a sequência de preços Cd_0, Co, Co . Com a política de déficit tarifário troca-se um aumento de preços de Cd_0 para Co no período 0, por um aumento de preços, maior, de Co para Cd_1 , no período 1.

Não é evidente haja estabilidade do nível de preços. Assim, a política de déficit tarifário, entendida como política de estabilidade de preços, é sobretudo uma aposta numa redução de preços no futuro. Seria necessário que no período 1 houvesse uma descida de Co para um valor mais baixo, descida suficientemente grande para que se pudesse estabelecer um preço inferior ao preço Cd_0 mas que permitisse pagar o déficit criado no período 0. Se essa redução de custos não acontecer no período 1, e, por exemplo, se Co se mantiver, os preços têm de subir no período 1 e não há “estabilidade” tarifária. O problema é agravado se Co aumentar no período 1 porque aumentam os custos de produção de energia elétrica.

É claro que o exemplo pode ser mais realista e considerar que o pagamento do déficit não tem de estar concentrado num único período. Neste caso, o “choque” seria certamente menor em cada período, mas porque era distribuído ao longo de vários períodos.

Isto leva-nos ao segundo problema a considerar, o de saber se a estabilidade tarifária, garantida através de preços regulados para o efeito, é um valor de interesse público. Não é evidente que o seja. A orientação dos preços para os custos marginais dos bens e dos serviços é um dos benefícios fundamentais de um mercado concorrencial e em princípio deveria ser um dos benefícios de um mercado regulado. As variações de preços dos bens e serviços devem estar relacionadas com a escassez relativa dos *inputs* utilizados na sua produção para se promover eficiência na produção, no consumo e no investimento. A introdução de regras de estabilidade tarifária como as que temos vindo a discutir quebra essa relação.¹⁰

Discutido o argumento da estabilidade tarifária, importa discutir outros argumentos suscetíveis de ser invocados como argumentos de interesse público.

Nos casos anteriores, a política de défice tarifário é uma política em que basicamente o Estado redistribui rendimento entre cada consumidor e a empresa regulada: no primeiro período o rendimento vai da empresa para os consumidores e no segundo dos consumidores para a empresa ou, dito, de outro modo, cada consumidor subsidia o seu consumo no período 0 com um gasto adicional no período 1. No conjunto dos dois períodos cada consumidor fica pior porque há um custo com a ineficiência provocada pelo défice tarifário. Parece um exercício arbitrário e custoso de gestão, pelo Estado, do rendimento de cada um.

No entanto importa ver se eliminando as hipóteses mais simplificadoras que foram adotadas se encontram outros argumentos para o défice tarifário.

Desde logo, a procura de cada consumidor não deve ser igual em cada um dos períodos: alguns consomem mais no primeiro período, outros menos, outros deixam de consumir e aparecem também novos consumidores. Assim, a política de défice tarifário implica uma transferência adicional de rendimento dos que consomem mais no período final, para os que consomem mais no período inicial.

No entanto, em qualquer um destes casos não se vê um motivo para subsidiar os consumidores de eletricidade num dado momento pelos consumidores futuros de eletricidade, sendo ainda provável que a maior parte dos consumidores futuros coincide com os atuais.¹¹

Em segundo lugar, importa considerar que alguns dos consumidores têm um rendimento baixo e o seu poder de compra seria muito afetado com uma grande subida de preços. Poder-se-ia dizer que procurava proteger estes consumidores de menor rendimento. Mas é provável que muitos dos consumidores de menor rendimento num ano sejam também consumidores de menor rendimento nos anos seguintes, em que vão ter de suportar o aumento de preços necessário para pagar o défice tarifário. Ou seja, não se protegem de facto estes consumidores. E se é possível que nalguns casos o rendimento aumente ao longo do tempo, tornando mais fácil

¹⁰ Se as variações de C_0 fossem resultado não da escassez relativa dos *inputs* utilizados na produção de eletricidade, mas de opções de política económica ineficientes, o objetivo deveria ser corrigir a política económica que estivesse na fonte de distorções nos custos aceites pelo regulador e nos preços.

¹¹ Se os défices tarifários se acumularem ao longo de vários anos e forem reduzidos também ao longo de vários anos, temos uma política que pode ter implicações nos preços ao longo de 15 ou de 20 anos, como aconteceu. No entanto, mesmo ao longo de um período deste tipo a maior parte dos consumidores residenciais são os mesmos. É uma parte pequena a que sai do mercado, por emigração ou por morte, e uma parte pequena a de novos consumidores, por nascimento, constituição de novos agregados familiares, turismo ou imigração. A maior parte dos consumidores está no mercado ao longo do período em que está em vigor a política de défice tarifário. Se o consumo de cada um variou da mesma forma ao longo desses anos, a política de défice tarifário implicou sobretudo uma transferência de rendimento no tempo para cada um deles.

pagar a conta de eletricidade, não é de excluir que alguns consumidores percam rendimento ao longo do tempo, ficando com a vida mais complicada com o défice tarifário.

Uma terceira simplificação tem a ver com a taxa de juro. A empresa regulada é compensada pelo deferimento do recebimento da receita no período zero através de uma taxa de juro estabelecida pelo governo.¹²

Esta medida, inteiramente aceitável, admitindo que desta taxa de juro não resultem nem rendas nem perdas para a empresa regulada, implica um efeito distributivo adicional. Há consumidores que no período zero têm dívidas de consumo e outros que, pelo contrário têm aplicações financeiras. Para os que têm dívidas de consumo, se pagam uma taxa superior à que a empresa recebe pelo défice tarifário, o que é uma hipótese razoável, há um ganho. Para os que apliquem o dinheiro que, no período 0, tenha sido poupado com o défice tarifário haverá uma perda (ganho) se receberem por essa aplicação uma taxa de juro inferior (superior) à taxa de juro que a empresa recebe pelo défice tarifário.

Neste caso, para se proteger os consumidores, a opção mais simples para evitar perdas de ineficiência seria dar a cada consumidor a possibilidade de optar pelo plano de preços relativo ao défice tarifário, explicitando-se devidamente os custos futuros implicados pelo défice. Uma outra opção, era utilizar outros instrumentos para proteger os consumidores mais vulneráveis – poderia ser mais complicado, mas talvez fosse mais ajustado, tanto mais que, como se disse, muitos dos consumidores que têm baixo rendimento no período 0 terão também baixo rendimento no período 1, em que são penalizados.

Sintetizando, a política de défice tarifário não resolve nenhuma falha de mercado, pelo contrário é fonte de ineficiências, tem efeitos redistributivos que em boa parte dos casos são arbitrários e, também por este motivo, não é evidente que possa ser utilizada para proteger as famílias de menor rendimento - um objetivo que seria meritório na presença de aumentos de preços significativos, mas para o qual o défice tarifário não é o instrumento mais adequado. Por fim, não se pode dizer que garanta a “estabilidade tarifária”, que seria um dos motivos para a sua criação.

2.4. OBJETIVOS POLÍTICOS NAS ORIGENS DA REGULAÇÃO

Os argumentos anteriores sugerem que o objetivo de estabilidade de preços formalmente consagrado na lei não teria grande fundamento do ponto de vista do mercado. Ou seja, não parece haver um argumento de interesse público com fundamento empírico.

Uma explicação alternativa aos argumentos de interesse público pode ser sugerida a partir da abordagem sugerida pela teoria das escolhas públicas e pela teoria económica da regulação. Nesta linha, admitimos que o resultado do processo político resulta da interação de todos os interessados, cada um defendendo os seus interesses próprios.¹³

¹² Veja-se, por exemplo, o art. 2.º do Decreto-Lei 237-B/2006, de 18 de dezembro. A fixação da taxa de juro a aplica foi naturalmente objeto de algum debate na CPI: a taxa de juro deveria corresponder ao custo do capital da empresa? Ou deveria corresponder à remuneração de um empréstimo sem risco?

¹³ A presença de associações de consumidores influentes e ativas no mercado, como a DECO, sugere que o problema da ação coletiva pode estar resolvido para um grande número de consumidores. As associações empresariais em princípio deveriam resolver o problema de ação coletiva das empresas suas associadas.

Consideramos basicamente 3 grupos de interesses – o Governo, a empresa regulada e os consumidores. Para a empresa regulada a política de défice tarifário seria em princípio neutra, do ponto de vista dos lucros.¹⁴ Para os consumidores, o interesse no défice dependeria da sua informação sobre os seus benefícios e custos. Os consumidores que antecipassem corretamente os seus custos futuros, perceberiam que iriam ter de pagar mais no futuro. Para estes consumidores não haveria razões para o défice tarifário. No entanto para os consumidores que não antecipassem um aumento futuro na sua fatura de eletricidade a política de défice tarifário poderia parecer vantajosa.

Num caso destes, como em tantos outros, cada consumidor decidiria com grandes limitações cognitivas. Teria um conhecimento da vantagem imediata que resultaria do défice tarifário, mas não é óbvio que processasse a informação sobre os custos futuros. Poderia até esperar que estes custos viessem a ser suportados por outros. Em qualquer caso o défice tarifário seria uma política com benefícios certos e imediatos e custos incertos e futuros.

Neste sentido uma explicação simples para a política do défice tarifário pode basear-se somente num aproveitamento pelo governo das limitações cognitivas dos consumidores (e eleitores) para obter apoio eleitoral. O fim seria manter o poder, na linha da teoria das escolhas públicas. O défice garantia um benefício imediato e a subida de preços verificava-se no futuro, depois de eleições.

Este argumento é compatível com as intervenções do governo sobre os défices tarifários, em 2006, pouco depois das eleições que tinha ganho, e em 2008, pouco antes de novas eleições, para evitar grandes aumento de preços da eletricidade na legislatura.

3. RENDAS EXCESSIVAS E EFICIÊNCIA ECONÓMICA

Em 1995 foram definidas as bases de organização do Sistema Elétrico Nacional, constituído pelo Sistema Elétrico de Serviço Público (SEP) e pelo Sistema Elétrico Independente. No SEP integravam-se a Rede Nacional de Transporte e as instalações de produção e as redes de distribuição exploradas mediante um regime de licença vinculada ao SEP. As instalações de produção da Companhia Portuguesa de Produção de Electricidade (EDP) e as redes de distribuição da EDP– Electricidade de Portugal, (na altura exploradas por empresas separadas dentro do grupo) foram incluídas no SEP.¹⁵

Em 1996 o governo decidiu aplicar às centrais da EDP o modelo dos CAE que, em 1993 e em 1994, tinham sido a base contratual utilizada para a exploração das centrais do Pego e da Tapada do Outeiro, numa altura em que, com sucesso, se pretendia captar investimento

¹⁴ Na prática pode não ter sido, por vários motivos. Em primeiro lugar, a empresa regulada poderia ter ganhos ou perdas com a taxa de juro fixada pelo governo, para capitalizar o valor do défice ao longo do tempo. Além disto, o governo previu desde 2006 a possibilidade de titularização da dívida tarifária pela empresa regulada (veja-se por exemplo o art. 3.º do Decreto-Lei n.º 237-B/2006, de 18 de dezembro). Um assunto que veio a ser debatido na Comissão Parlamentar de Inquérito foi a eventual existência de lucros na titularização desta dívida e a sua repartição com os consumidores. Finalmente há a questão da variação do lucro quando o défice tarifário coincide com rendas excessivas elevadas, descrito na Figura 3.

¹⁵ Artigos 17.º e 28.º do Decreto-Lei n.º 182/95, de 27 de julho.

estrangeiro para a produção de energia. Nestes contratos, a rentabilidade do investimento foi definida na sequência de concurso público internacional.

Por esta via, às centrais da EDP ficou assegurado por um determinado prazo um quadro de remuneração que as protegia de riscos associados a variações de preços dos combustíveis fósseis ou da eletricidade, a maus anos hidrológicos ou à liberalização do setor¹⁶. A rentabilidade prevista era na altura um pouco inferior à que tinha sido estabelecida antes para os CAE das centrais do Pego e da Tapada do Outeiro. Foi ainda prevista a possibilidade de a EDP poder prolongar a exploração dos ativos para além do seu termo contratual.

A extensão dos CAE à EDP tem sido discutida como uma das fontes de rendas excessivas no setor elétrico. Foi considerada como “a maior renda alguma vez criada em Portugal”. Isto, sendo a EDP uma empresa pública. O objetivo seria aumentar o valor da empresa, tendo em vista a sua privatização.¹⁷

Posteriormente, com a liberalização e com a criação do MIBEL, considerou-se necessário extinguir os CAE. O problema era definir o modelo que os iria substituir. Optou-se por um mecanismo que procurava assegurar, num quadro de liberalização, as remunerações que os CAE asseguravam, e que eram os CMEC.¹⁸

A fundamentar a opção tomada terá pesado a ideia de valorização da EDP, por basicamente três motivos: a) o Estado ainda detinha 25% da empresa, que pretendia privatizar; b) a privatização de 75% da empresa tinha sido feita incorporando o valor dos CAE, pelo que adotar uma opção que não garantisse a equivalência contratual com os CAE poderia ter riscos legais elevados e eventualmente prejudicar a credibilidade do Estado em geral; c) importava defender a posição da EDP no quadro do mercado ibérico, para garantir controlo nacional sobre decisões fundamentais.

Este novo mecanismo suscitou dois tipos de problemas. Em primeiro lugar importaria saber se era necessário manter o equilíbrio contratual relativamente aos CAE ou se, perante a liberalização decidida num quadro europeu, o governo poderia adotar um outro modelo eventualmente mais favorável aos consumidores. Em segundo lugar, tendo sido abandonada esta opção, importaria saber se com os CMEC tinham sido introduzidos ganhos adicionais aos que estavam previstos nos CAE.

Uma avaliação das consequências finais dos CMEC não é pacífica. No relatório da Comissão Parlamentar de Inquérito, aprovado por maioria, recomenda-se que haja uma correção dos elementos que “pervertem o objetivo legal de manutenção do equilíbrio contratual”. Além disto recomenda-se a revisão da “sobre-remuneração constituída pela atribuição dos CAE à EDP e mantida pelos CMEC”.¹⁹

¹⁶ Este era de poucos anos no caso das centrais de Setúbal e do Carregado (a fuelóleo), ia até 2017 no caso de Sines (carvão) e situava-se entre 0 e 30 anos no caso das centrais hídricas (Sérgio, 2020)

¹⁷ A expressão é atribuída ao Eng. João Talone. Os Eng. Mira Amaral e Sampaio Nunes, sem mencionar explicitamente rendas excessivas, falaram dos interesses financeiros do Estado (CPI, 2019) pp. 26-27.

¹⁸ Decretos-Lei n.º 185/2003, de 20 de agosto, e n.º 240/2004, de 27 de dezembro.

¹⁹ CPI (2019), pp. 59-60. O Grupo Parlamentar do PSD apresentou uma avaliação dos mesmos factos em que reconhece também a importância dos CAE e dos CMEC para o valor da EDP, a obter pelo Estado nas privatizações, enfatiza os riscos legais de uma opção diferente e valoriza a ideia de se procurar evitar uma desvalorização por via legal da posição da EDP no mercado ibérico. No entanto, não conclui que os CMEC tenham trazido vantagens adicionais à EDP, nem pela necessidade de rever a remuneração atribuída pelos CAE e pelos CMEC (PSD, 2019, p. 53).

A discussão em torno das rendas excessivas que podem ter sido permitidas pelo quadro regulamentar foi mais vasta e envolveu outros aspetos, como a extensão do uso do domínio público hídrico a favor da EDP assim como o quadro regulativo aplicável à REN. Mas pareceu que havia uma grande aceitação da ideia de que houve um objetivo de criar um quadro regulamentar que valorizasse a EDP e a REN tendo em vista as receitas a obter nas várias fases da sua privatização.

Este objetivo não é estranho. Era obrigatório, estava na lei. Entre os objetivos das reprivatizações contavam-se, desde 1990, preservar os interesses patrimoniais do estado e de outros interesses nacionais e reduzir o peso da dívida pública na economia.²⁰ Na revisão de 2011, manteve-se o objetivo de reduzir o peso da dívida, que exige a maximização do encaixe financeiro e, evidentemente, a sua aplicação na redução da dívida.

Para alcançar estes objetivos, o governo podia definir um quadro regulamentar que permitisse às empresas reguladas a privatizar algum poder de mercado e a obtenção futura de lucros económicos, desde que os compradores dessas empresas pagassem antecipadamente ao Estado uma parte desses lucros, através do pagamento de um valor na privatização mais alto do que aquele que se verificaria com um quadro regulativo mais eficiente e mais favorável aos consumidores.

Portanto o argumento de interesse público é claro. A receita do Estado corresponde ao interesse público. Isto significa que, na prática, o estabelecimento de rendas excessivas no mercado, tendo em vista as receitas da privatização, é equivalente a um imposto – pago pelos consumidores ao longo dos anos sendo parte da receita adiantada ao Estado pelos novos acionistas.

Numa perspetiva de análise económica o problema é saber quais são as consequências desta opção na eficiência da economia.

Em princípio dir-se-á que são negativas. Um aumento dos lucros económicos, exige um desvio dos preços em relação aos custos que leva a uma contração do consumo residencial e empresarial.²¹ No entanto, a ineficiência depende da elasticidade procura preço no mercado onde se verificam esses lucros económicos. Como se viu acima, há algum consenso de que a procura de eletricidade é rígida, o que reduz a perda de ineficiência provocada pelo poder de mercado. Se assim for, a renda que resulta do objetivo de maximização da receita do Estado provoca uma ineficiência relativamente reduzida na economia.²²

Um segundo aspeto a considerar é saber o que faz o governo à receita da privatização. Faria sentido utilizá-la para reduzir impostos para reduzir as ineficiências no sistema económico causadas pela fiscalidade se estas fossem elevadas.

²⁰ Alíneas f) e g) do art. 3.º da Lei n.º 11/90, de 5 de abril. O objetivo de reduzir o peso da dívida pública na economia foi mantido com a revisão de 2011 (Lei n.º 50/2011, de 13 de setembro).

²¹ Em limite a competitividade de algumas empresas, com tecnologias relativamente intensivas em energia, pode ser comprometida. Na economia portuguesa a intensidade energética varia muito segundo os setores de atividade, sendo mais elevada nas indústrias químicas, de pasta de papel e papel, cimento, vidro e cerâmica, madeira e mobiliário e transportes (Eurogroup Consulting, 2014).

²² Admitimos aqui que a o outro tipo de ineficiência associado aos monopólios, a falta de controlo de custos, se não verifica porque os mecanismos de governo corporativo são adequados. Mas a explicação pode ser outra, como se verá mais à frente.

A ineficiência provocada pelos impostos pode ser medida pelo custo marginal dos fundos públicos (CMF), definidos como o valor que a economia perde quando se coleta um euro de receita fiscal:

$CMF = (1 + \alpha)$, em que α corresponde à perda de eficiência provocada pelo imposto.

De um ponto de vista de eficiência económica faz sentido aumentar (reduzir) os impostos que provocam menos (mais) ineficiência.

No caso em análise haveria que comparar a ineficiência criada com as rendas excessivas no setor elétrico para aumentar as receitas da privatização, com o aumento de eficiência que se poderia ter reduzindo-se, no mesmo montante, os impostos nos mercados de fatores (por exemplo, o IRS ou o IRC), ou nos mercados de bens e serviços (por exemplo, o IVA). Para o efeito é necessário ter em conta os impactos que as rendas excessivas e os impostos em análise têm nos mercados a montante e a jusante.

Trata-se de um problema de equilíbrio geral e a evidência disponível é escassa. Alguns resultados relacionados com o problema aqui em análise são apresentados por Barrios *et al.* (2013), que procuraram comparar, para cada um dos países da União Europeia, o valor de CMF que resultaria de um pequeno aumento do imposto sobre o trabalho com o que resultaria de um pequeno aumento do imposto sobre o consumo energético das famílias. Neste exercício, não se considerou o impacto que na economia teria a despesa pública que poderia ser financiada com este imposto e, no caso do imposto sobre o consumo energético, não se considerou o benefício ambiental que daí poderia resultar. Os resultados para Portugal, para a média da União Europeia, ponderada pelo PIB de cada país, e para alguns países da União apresentam-se no quadro 2.

Na União Europeia, em média, a coleção de €1 através de um imposto sobre o trabalho implica uma perda de eficiência de €0,90, mas, em caso de imposto sobre o consumo energético a perda de eficiência será de €0,08. Em Portugal, tal como em Espanha, o custo marginal dos fundos públicos é menor do que 1 no caso do imposto sobre o consumo energético, o que sugere que um imposto deste tipo gera um reajustamento dos preços relativos que permite ganhos de eficiência (Barrios *et al.*, 2013).

Estes resultados apoiam dois argumentos:

Sendo necessário um aumento da carga fiscal será preferível que ela se concretize através um imposto sobre o consumo energético e não através de um imposto sobre o trabalho.

Faz sentido substituir impostos sobre o trabalho por impostos sobre o consumo da energia.

Quadro 2: Custo marginal dos fundos públicos para impostos sobre o trabalho e sobre o consumo energético das famílias

| | Impostos sobre o trabalho | Impostos sobre energia |
|----------|---------------------------|------------------------|
| Média EU | 1,9 | 1,08 |
| Portugal | 1,82 | 0,93 |
| Espanha | 1,79 | 0,89 |
| Itália | 1,68 | 1,1 |
| França | 2,41 | 1,42 |

Fonte: Barrios *et al.* (2013).

O segundo argumento será mais sustentável e relevante do ponto de vista de política económica do que o primeiro, pois este dependerá da existência de desajustamentos nos preços relativos que podem ser pontuais e corrigidos com alguma facilidade, mesmo sem intervenções adicionais do governo. Além disto, no segundo argumento está clara também uma compensação dos contribuintes pela carga adicional que sofreriam enquanto consumidores, tornando a política socialmente mais aceitável.

A existência de rendas excessivas no setor elétrico seria compatível com o primeiro argumento e, desde que o Estado beneficiasse dessas rendas e as utilizasse para reduzir impostos, seria também compatível com o segundo.²³ Assim, se o objetivo do Estado na privatização era a maximização da receita, então a privatização deveria ser definida em articulação com a política fiscal e orçamental.²⁴

4. RENDAS EXCESSIVAS E OUTROS ARGUMENTOS DE INTERESSE PÚBLICO

Pode assim defender-se uma política de criação de rendas, equivalentes a um “imposto” adicional sobre o consumo de eletricidade, com base num argumento de promoção de eficiência de toda a economia através de uma redução correspondente de impostos sobre o trabalho.

No entanto, não é óbvio que tenha sido este o objetivo dos sucessivos Governos. Se o tivesse sido, teria sido necessário explicitar as rendas excessivas implícitas no modelo de regulação adotado e um compromisso credível de redução de impostos sobre o rendimento ou sobre bens e serviços. Não se encontram referências à discussão destes problemas. O que há de mais próximo é a ideia de utilizar as receitas das privatizações para reduzir a dívida pública.²⁵

Sem um compromisso credível do governo em ligar receitas das privatizações a reduções de impostos, as privatizações podem ter tido, pelo contrário, um efeito perverso na despesa. Nos anos em que se verificaram contribuíram para flexibilizar a restrição orçamental do Estado pois, mesmo se aplicadas em dívida permitiam reduzir os juros da dívida, e induzir despesa adicional. Mas, mais do que isso, como não havia nenhum limite ao endividamento público, o Estado podia amortizar dívida com receitas das privatizações, cumprindo um

²³ O Estado beneficiou dessas rendas com a receita das privatizações. No entanto, a compatibilidade de rendas excessivas com a segunda ideia diminui ao longo das várias fases de privatização. Foi maior quando se decidiram os CAE, e menor quando se decidiram os CMEC. Neste caso as rendas eventualmente criadas iriam beneficiar sobretudo os acionistas privados que, em 2004, já detinham 75% do capital da empresa. O objetivo de maximizar a receita da privatização incentivava, ao longo das várias fases da privatização, um alinhamento entre os interesses financeiros do Estado e os interesses dos acionistas privados. Se o Estado, já com menor participação, pretendesse mudar de política, essa mudança poderia ter riscos legais. Foi argumentado que a ponderação destes riscos legais pode ter sido um dos motivos para a adoção dos CMEC, embora também tenha sido evidenciado que com os CMEC se pretendia proteger a receita da privatização a fazer (CPI, 2019, p. 28 e seguintes).

²⁴ E também com a política social, tendo em conta o impacto que os preços da energia têm nas famílias de menor rendimento e os eventuais efeitos regressivos de rendas excessivas no setor elétrico.

²⁵ Esta ideia, que parece excelente, não é, de facto, muito interessante. Por exemplo, na sua quarta fase, em 2000, a privatização da EDP gerou uma receita de €1,7 mil milhões. Este valor corresponde a menos de 5 meses de pagamento dos juros de dívida pública desse ano; o impacto no valor da dívida é residual.

objetivo político ou legal, e ao mesmo tempo contrair nova dívida, ficando rapidamente com um endividamento não inferior ao que havia antes da privatização.²⁶

Neste sentido os principais interessados na maximização da receita das privatizações são, em termos de agenda política, os governos em funções porque ganham maior margem para aumentar a despesa pública e utilizá-la para obter apoios eleitorais – isto numa linha de argumentação baseada na teoria das escolhas públicas.

Poderia surpreender a ausência de uma oposição de consumidores e empresas. As suas perdas foram certamente superiores aos ganhos do Estado. Do lado das associações empresariais, é possível que, no quadro geral da defesa dos interesses dos seus associados, o consumo de electricidade tenha um peso relativamente reduzido. Será importante para algumas indústrias, mas para os principais setores de atividade os custos com a electricidade não serão os principais custos que pretendam reduzir e as suas negociações e atividades de pressão sobre o poder político podem estar centradas noutras atividades.

O caso dos consumidores residenciais será diferente. Aqui a falta de ação coletiva poderia estar relacionada com as limitações cognitivas que afetam a decisão de um consumidor típico. No momento em que as medidas são tomadas os benefícios gerais parecem imediatos, as receitas da privatização, e não existe uma perceção adequada dos custos futuros, por falta de informação. Uma outra explicação poderia estar baseada num problema de ação coletiva que os consumidores não conseguissem resolver. Esta é pouco provável, dada a existência de pelo menos uma associação de consumidores bastante ativa nos mercados regulados. Será razoável admitir que as associações de consumidores encontrassem dificuldade em justificar aos seus associados uma eventual oposição a um objetivo de aumentar a receita do Estado e que daria vantagens imediatas aos consumidores.

Por estes motivos terá sido relativamente fácil aos governos fazer opções nas privatizações que significaram que a maximização da receita do Estado prevaleceria sobre interesses dos consumidores empresariais e residenciais sem que destes houvesse grande oposição.

Entretanto, ao longo das sucessivas fases da privatização, os preços aumentariam na sequência das medidas antes tomadas para valorizar a empresa regulada. Para evitar a queda de popularidade associada a estes aumentos de preços um governo em funções poderia ter a tentação de adiar o efeito no tempo desses aumentos, constituindo défices tarifários. Por esta via, a política de défices tarifários pode relacionar-se com uma política de criação de rendas excessivas com o objetivo de reduzir a transparência do sistema de preços, ficando mais difícil para os consumidores perceber as origens políticas dos aumentos de preços.

²⁶ Vale a pena notar, neste contexto, que um período em que se fizeram grandes privatizações, entre 1995 e 2005 a dívida bruta per capita das administrações suplicou, de €5 463 para €10 897. Entre 2005 e 2015 duplicaria de novo, para €22 798 (dados da Pordata). Estes dados não substituem uma análise sistemática do problema, mas não sugerem que as receitas das privatizações levaram a uma redução da dívida.

5. RENDAS E A PRODUÇÃO EM REGIME ESPECIAL

Os argumentos formais de interesse público utilizados para justificar o apoio à produção em regime especial parecem razoáveis: a produção de eletricidade a partir de fontes de energia renováveis e a criação de clusters de empresas na área de energias renováveis permitiriam resolver problemas ambientais e promover a competitividade da indústria nacional numa área em crescimento.

No entanto, há problemas a resolver antes de se aceitar a ideia. O primeiro é estabelecer uma relação entre os custos das opções tomadas e o valor das externalidades ambientais que foram eliminadas e o valor social dos clusters que foram criados. Ou entre os benefícios em termos de competitividade, se se verificaram, e os custos suportados.

Na prática houve aqui também uma “tributação” do consumo de eletricidade para promover objetos de política ambiental e industrial. É até possível que fosse uma forma de financiamento mais eficiente do que a utilização de impostos sobre o rendimento do trabalho, pelos motivos já vistos. No entanto, deveria ser acompanhada de uma avaliação dos custos e dos benefícios de política ambiental e industrial. Isto ajudaria a evitar a confusão que se faz por vezes entre o custo social da produção em regime especial e os seus “sobrecustos”, tal como são calculados pela ERSE, designadamente por não se ter em conta o valor do carbono.

O que não suscitará grandes dúvidas é que a integração do que se designou como “sobrecustos” da produção em regime especial no défice tarifário tornou também mais difícil uma avaliação pelos interessados, empresas em toda a economia e consumidores, dos custos das opções, no início e à medida que foram tomadas.

6. E SE NÃO HOUVE RENDAS EXCESSIVAS?

Definir rendas excessivas, e identificar a sua existência, obriga a considerar problemas com alguma complexidade e não é evidente que todos tenham sido considerados nos trabalhos da CPI.

Para regular os preços, a autoridade pública pode estabelecer um modelo de regulação com preços orientados para os custos, garantindo-se, *ex post* que não há rendas excessivas: em cada ano as receitas igualam os custos, incluindo nestes o custo de oportunidade do capital. No entanto, modelos deste tipo não criam incentivos para as empresas controlarem os custos, ou seja, criam um problema de risco moral. Para resolver este problema podem adotar-se contratos de preço fixo, em que a autoridade pública se compromete com um preço inicial e com uma regra de atualização de preços ao longo de um determinado período. Este período pode ser relativamente longo, por exemplo o período de uma concessão ou de uma licença, caso normalmente designado como de regulação contratual. Ou pode ser mais curto, 3 ou 5 anos tipicamente, correspondendo a uma forma de regulação periódica. O problema de risco moral fica resolvido, sobretudo com períodos de regulação mais longos, pois as empresas têm incentivos a controlar custos e a revelar os seus custos *ex post*. No entanto, há um problema de seleção adversa. O contrato oferecido tem que interessar às empresas, evitando que elas simulem custos superiores aos que podem ter e incentivando-as a controlar custos. Por isso, um contrato eficiente permite muitas vezes a existência de

rendas, que não devem ser consideradas como excessivas pois existem para ultrapassar um problema informacional da autoridade pública.²⁷

Assim, é possível que uma empresa sujeita a regulação contratual tenha, ao longo do período contratual, taxas de remuneração do capital diferentes do que poderia ser considerado como o seu custo de oportunidade do capital. Se forem maiores, podem parecer rendas excessivas – mas não o são necessariamente, podem ser rendas simplesmente rendas informacionais ou de outro tipo, inerentes ao modelo de regulação contratual. Não são defeito, são feição do modelo de regulação. A sua eliminação *ex post* poderia ser origem de uma perda de credibilidade do Estado relativamente aos investidores privados, prejudicando o investimento futuro.

No quadro legal do setor elétrico definido em 1995 o Estado adotou em muitos dos mercados de electricidade uma abordagem baseada na regulação contratual. A regulação periódica, concretizada através de uma autoridade reguladora (e ainda assim, na prática, numa base plurianual), estava limitada à regulação das taxas de uso das redes de transporte e de distribuição.

Identificar rendas excessivas no setor elétrico passaria por uma análise cuidada dos contratos celebrados, tendo em conta a informação existente no momento em que foram celebrados e no quadro do modelo de regulação contratual adotado, incluindo a eventual existência de regras e incentivos destinados a ultrapassar limitações informacionais das autoridades públicas. Esta abordagem não terá sido seguida pela Comissão. Houve um esforço para identificar e quantificar o risco de diferentes alternativas de decisão nalguns momentos e o custo do capital que lhes poderia estar associado, mas não para identificar rendas excessivas entre as rendas eventualmente existentes.

Assim, não se vê bem qual o conceito de renda excessiva utilizado pela CPI, nos vários casos em análise – nem que os vários depoimentos tivessem implícitos os mesmos conceitos de rendas. De forma geral, os vários intervenientes parecem referir-se a situações em que as remunerações obtidas em determinados anos terão sido superiores ao custo de oportunidade do capital nesses anos. Ainda assim, não foi assunto pacífico e houve muitas divergências, por exemplo quanto à equivalência entre os CAE e os CMEC ou ao tratamento regulativo dos direitos de utilização do domínio público hídrico.

Neste contexto, um caso interessante é o do modelo de regulação com base nos CAE, referido como o da “maior renda alguma vez criada”. Será talvez daqueles em que maiores possam ser as dúvidas sobre a existência de rendas excessivas. Os CAE da EDP foram baseados nos CAE que resultaram do concurso público internacional para as centrais do Pego e da Tapada do Outeiro – e reduzindo a taxa de remuneração do capital aceite para estas centrais, como foi referido no âmbito dos trabalhos da Comissão. Foi também estimado na altura (1996), numa apreciação dos relatórios de consultores independentes, que os CAE da EDP só permitiriam recuperar cerca de 90% do valor contratado, considerando a estrutura financeira e as taxas de remuneração dos capitais próprios e alheios admitidos

²⁷ Em teoria, a melhor opção é adotar uma combinação das duas abordagens, embora seja mais frequente a utilização de cada uma delas em separado. Entre outros, para uma síntese, veja-se Joskow (2014).

pelos avaliadores. Só com maior alavancagem financeira, mais típica de parcerias público-privadas, seria possível recuperar a totalidade dos capitais investidos.²⁸

Neste caso, nos depoimentos recolhidos pela Comissão a possibilidade de rendas excessivas estava claramente associada à existência de objetivos políticos de maximização das receitas nas privatizações. Uma situação análoga observou-se também na análise da produção em regime especial, em que se admitiu a existência de um “sobrecusto” associado a objetivos de política industrial e de política ambiental, embora este sobre custo não tenha sido quantificado. Isto é, dizendo com algum exagero, ter-se-á intuído por vezes que houve rendas porque havia objetivos de política económica que seriam mais facilmente alcançados se essas rendas existissem.

7. COMENTÁRIOS FINAIS

Aceitando-se a hipótese de que houve rendas excessivas no setor elétrico, num ou mais dos casos analisados, é razoável tomar medidas para compensar aqueles que as pagaram. É neste sentido que vão boa parte das recomendações da Comissão, ou ainda antes da Comissão, a ideia de haver impostos extraordinários sobre o setor elétrico. Compensar os que perderam, com as compensações pagas pelos que ganharam, não afetaria as decisões de investimento futuras, porque se tratava de remunerações que excediam a remuneração normalmente exigida pelos investidores e faria sentido do ponto de vista de equidade.

No entanto, uma conclusão não menos importante a retirar seria outra: definir medidas que evitassem no futuro o mesmo erro, no setor elétrico ou em qualquer outro setor de atividade.

Para o efeito seria necessário evitar o que tenha estado na origem das rendas. Um caso, em princípio anormal, seria admitir que na origem das rendas estava a falta de capacidade técnica das autoridades públicas para decidir. Aqui as medidas seriam simples, reforçar a capacidade técnica do Estado. Não é de crer que se tenha verificado no setor elétrico (nunca foi mencionado nos trabalhos da Comissão, o que não quer dizer que não possa existir noutros setores de atividade), pelo que importa arranjar outras explicações para as rendas que se achou existirem.

O relatório da CPI sugere objetivos de política económica como a maximização da receita do Estado ou a formação de um cluster de energias renováveis. Estes objetivos são por regra bem acolhidos pelos eleitores, mesmo que na sua origem se encontrem objetivos eleitorais de curto prazo ou a aceitação pelo Estado das posições de alguns grupos de interesses. No entanto, não têm que ser adotados, só porque parecem razoáveis. Se neste caso se entendesse que iriam impor um custo à sociedade, esse custo deveria ter sido devidamente ponderado, face aos benefícios que se esperavam – o que poderia levar á rejeição das políticas em causa. Se, pelo contrário, se aceitasse o argumento de que fazia sentido aumentar impostos sobre

²⁸ Estimativas de 1996 do Eng. Rui Sérgio, vice-presidente da Secção Especializada de Apoio às Reprivatizações, segundo informação não publicada, enviada ao autor. Acresce que a EDP tinha antes alienado, por indicação do Governo, as centrais do Pego e da Tapada do Outeiro. Neste contexto seria também de ter em conta a eventual existência temporária de rendas negativas, na medida em que a capitalização da EDP teve alguns anos valor inferior ao valor contabilístico da empresa (Sérgio, 2020).

o setor elétrico para reduzir impostos sobre o trabalho, então deveriam ter sido assumidos compromissos credíveis para o efeito. Ora, em nenhum caso isto se verificou.

Para resolver este tipo de problemas seria desejável introduzir no processo de decisão pública regras que os eliminassem ou pelo menos atenuassem. Algumas recomendações da Comissão vão, de forma genérica, neste sentido, por exemplo, no que diz respeito à produção em regime especial. Essas regras podem não envolver nenhuma restrição à decisão que as autoridades queiram tomar, mas obrigar somente a que sejam feitos, publicados e discutidos de forma credível, antes de uma decisão pública, estudos que identifiquem, quantifiquem e valorizem todas as consequências dessa decisão. Ou seja, mesmo que se conclua depois desse processo que uma decisão teria consequências negativas que excediam as positivas, ela poderia ser tomada – ficando a autoridade pública com a responsabilidade de o ter feito. Nalguns casos, medidas deste tipo não têm que corresponder a uma grande inovação formal. Trata-se sobretudo de densificar (bastante) alguns princípios já presentes no processo de tomada de decisões, pelo governo ou por autoridades administrativas e que na maior parte dos casos são seguidos de forma ligeira. Numa versão mais restritiva poderia haver um princípio legal de, em determinadas circunstâncias, não poderem ser aceites decisões que originassem para a sociedade mais custos do que benefícios.

Uma conclusão adicional sugerida pelos trabalhos da Comissão é que as privatizações por fases têm riscos quando o objetivo predominante do Estado é a receita. Das medidas tomadas ao longo do tempo, ou da sua ausência, para valorizar a empresa, vão beneficiando em parcela cada vez maior os acionistas privados e em parcela cada vez menor o Estado. O objetivo da receita conduz o Estado a um alinhamento com os interesses privados com benefícios adicionais cada vez mais pequenos, descurando os interesses dos consumidores, desvalorizados em relação ao objetivo que predomina. Não há aqui nenhuma “captura” do Estado, há só a consequência lógica de haver um objetivo que predomina sobre os outros.

Dito isto, o relatório da CPI não perde utilidade caso se não aceitem as suas conclusões e se entenda que não existem rendas excessivas. Desde logo, a ideia de densificar a regulamentação do processo de decisão pública continua a fazer todo o sentido. A eventual ausência de rendas excessivas não significa que se tenham tomado as melhores decisões. É clara a utilidade de densificar o processo de decisão que leva às opções por modelos de regulação contratual e, aceite este modelo, que seja concretizado da melhor forma possível. Isto pode passar por mecanismos que assegurem que à medida que nova informação é conhecida ela possa ser aproveitada pelas autoridades públicas, em novos contratos, ou mesmo nos contratos em vigor. No entanto, nesta hipótese são claros os riscos de haver um problema de credibilidade do Estado e de desincentivo futuro ao investimento, se houver intervenção públicas para eliminar rendas não excessivas, que resultem do próprio processo de regulação contratual.

REFERÊNCIAS

- Barrios, S.; Pycroft, J.; Saveyn, B. (2013) *The marginal cost of public funds in the EU: The case of labour versus green taxes*, Taxation Papers 35, European Commission.
- Bergh, J. (2008) Environmental regulation of households: An empirical review of economic and psychological factors. *Ecological Economics*, 66, 559-574.
- CPI (2019) *Comissão Parlamentar de Inquérito ao Pagamento de Rendas Excessivas aos Produtores de Eletricidade Relatório Final*.
- Carmona, N. (2006) *Modelação econométrica da procura de eletricidade em Portugal Continental: Uma aplicação empírica*. ISEG, Lisboa.
- Esperança, J. (2015) *Estimação da procura residencial de eletricidade em Portugal Continental*. ISCTE Business School, Lisboa.
- Espey, J.; Espey, M. (2004) Turning on the lights: A meta-analysis of residential electricity demand elasticities. *Journal of Agricultural and Applied Economics* 36(1), 65-81.
- Garcia, A. L.; Santana, J.; Resende, M.J.; Verdelho, P. (2016) *Sistema Elétrico: Análise Técnico-Económica*. Lisboa, Lidel.
- Joskow, P. (2014) Incentive Regulation in Theory and Practice. In N. Rose, *Economic Regulation and its Reform*. Chicago, University of Chicago Press, pp. 291-344.
- PSD, 2019, *Declaração de Voto sobre o Relatório Final da Comissão Parlamentar de Inquérito ao Pagamentos de Rendas Excessivas aos Produtores de Electricidade*.
- Rodrigues, M. (2012) *A procura de eletricidade de clientes residenciais*. (Relatório de Estágio), Universidade de Coimbra.
- Sérgio, R. (2020) Notas de leitura, não publicadas.
- Silva, S.; Soares, I.; Pinho, C. (2017) Electricity demand response to price changes: The Portuguese case taking into account income differences. *Energy Economics*, 65(C), 335-342.

Blandina C. R. Oliveira
Polytechnic of Leiria, School of Technology and Management
CEFAGE-UE
blandina.oliveira@ipleiria.pt
orcid.org/0000-0002-4414-3631
Adelino Fortunato
University of Coimbra, CeBER, Faculty of Economics
adelino@fe.uc.pt
orcid.org/0000-0002-4790-6470
DOI: https://doi.org/10.14195/2183-203X_53_5

The Effect of Intermittent Renewable Energy Generation on Electricity Prices: A Literature Survey

Efeito do Incremento da Produção de Energias Renováveis Intermitentes nos Preços da Eletricidade: Revisão da Literatura

Blandina C. R. Oliveira
Adelino Fortunato

Received for publication: June 15, 2020
Revision accepted for publication: April 9, 2021

ABSTRACT

Despite increasing deployment of intermittent renewable energies at lower generation costs, wholesale electricity price has been falling while retail electricity prices go up. This has triggered the debate on the cost-effectiveness of this source of energy. Therefore, the aim of this paper is to present a literature survey on the effect of intermittent renewable energy generation on electricity prices. Researches have used different methodological approaches, different periods and countries to examine the impacts of intermittent renewable energy on electricity prices. Most of the studies found evidence of the merit-order effect, which means that an increase in intermittent source generation would reduce the spot electricity market price. Finally, the few studies that address the retail market found that retail electricity could either increase or decrease.

Keywords: Electricity market; intermittent renewable energy; electricity prices.

JEL Classification: O38; Q41; L11.

Acknowledgement: The authors would like to thank to an anonymous referee for their comments and suggestions, although the usual disclaimer applies.

1. INTRODUCTION

The liberalization of many of the European electricity markets in the early 1990s promoted the development and deployment of intermittent Renewable Energy Sources (RES), both of which have increased significantly. This alternative has become one of the main targets of energy policy. Intermittent Renewable Energy is considered an important way to reduce energy dependence, deal with environmental issues, and to achieve sustainable economic growth.

Countries that have succeeded in increasing their renewable capacities significantly have implemented intensive support policies such as feed-in-tariffs (FIT) or renewable allowance systems which have significant costs. The RES penetration has raised the concern of costs of promotion, which have increased significantly for wind energy and solar PV. Therefore, the impacts of renewable electricity production in energy markets must be understood by assessing its effect on electricity prices and efficiency of renewable support policies.

The electricity production costs are expected to fall for renewables which could be explained by the decreasing costs of investment as technologies evolve. Consequently, this may reduce the need for subsidizing renewable energy technologies.

According to a recent report on energy prices and costs published by the European Commission the wholesale energy prices have fallen due to increasing competition on wholesale market namely from greater amounts of renewable energy.

Despite the fall in generation costs, renewable energy has driven up the retail electricity price and has been an expensive way to achieve greenhouse gas reductions which arises the debate on the cost-effectiveness of this source of energy. The costs of RES support are passed to electricity consumers in their bills and they are often regarded as a main driver of the increase of retail electricity prices.

Given the relevance that this matter has taken on, researchers have been carried out studies mostly applied to wholesale market. However, the studies applied to the retail market are still scarce. Therefore, the main aim of this study is to review the literature on the effect of intermittent renewable energy's generation on electricity prices at wholesale and retail markets.

This paper is structured in the following way. Section 2 presents the theoretical framework and describes the merit order effect. Section 3 shows some relevant studies that analyzed the effect of renewable energy at wholesale electricity price. Section 4 presents some works carried out to measure the effects of renewable energies on retail electricity price. Finally, the last section presents the conclusions and suggestions for future research.

2. THEORETICAL FRAMEWORKS

To understand the impact of renewable technologies on price formation, we start to describe the functioning of liberalized electricity markets. In Europe the liberalized spot electricity markets using a marginalist approach has prevailed.

Nowadays, the power market designs operate on the assumption that electricity generation has a range of positive marginal costs that increase through some rank ordering, as

is the case for thermal generators, based on technologies and fuel sources. This design is based on the construction of an efficient merit order through an implicit auction in the day-ahead market.

Economic theory predicts that high renewable electricity production reduces the wholesale price of electricity, which is commonly known as the merit-order effect. The merit order effect is a detailed analysis of the impact of renewable electricity generation on spot market prices. This effect is only one of several consequences of renewable production on the electricity system. Consequently, it is crucial to determine its size for the economic evaluation of renewable energies.

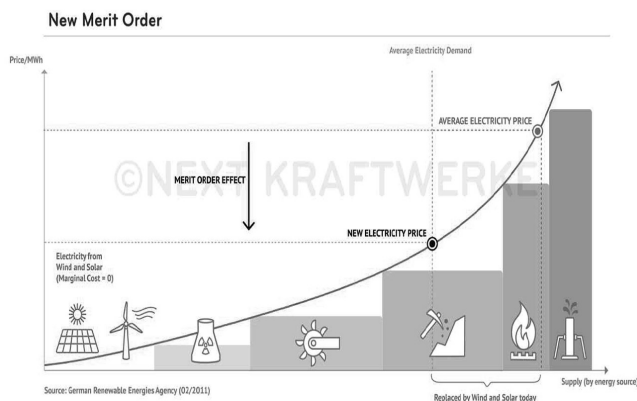
Theoretical considerations introduced by Jensen and Skytte (2002) suggested that renewable electricity production results in lower electricity prices. The merit-order effect means that the price decreases because (additional) renewable electricity bids into the market at lower marginal costs.

2.1. MERIT ORDER EFFECT

In order to supply electricity, different power generation technologies compete with each other according to their availability of supply and their marginal cost of production (fossil fuels (coal or natural gas), nuclear power, renewable energy sources (hydroelectric generators, wind or solar energy)). Conventional approach consists in ranking the power plants of the system in ascending order of their marginal cost of generation. This approach is called the merit order effect.

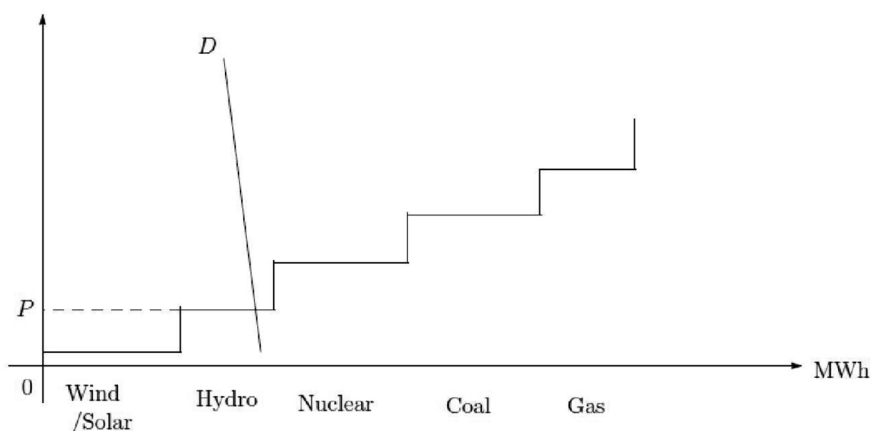
Merit order effect is used to describe the mechanism by which the market price is determined. The electric power price is set by the “merit order”, of which the sources with the cheapest marginal costs will be sold more quickly. Conventional electric power sources (coal and gas) are crowded out along the merit order. Figure 1 shows the merit order effect.

Figure 1: Merit order effect



The merit order effect corresponds to the electricity supply curve. The demand side of the electricity market is assumed to be inelastic. This means that almost none of the buyers cannot respond to the spot price. The intersection between supply and demand curves is obtained the electricity market equilibrium (see Figure 2).

Figure 2: Electricity Market Equilibrium



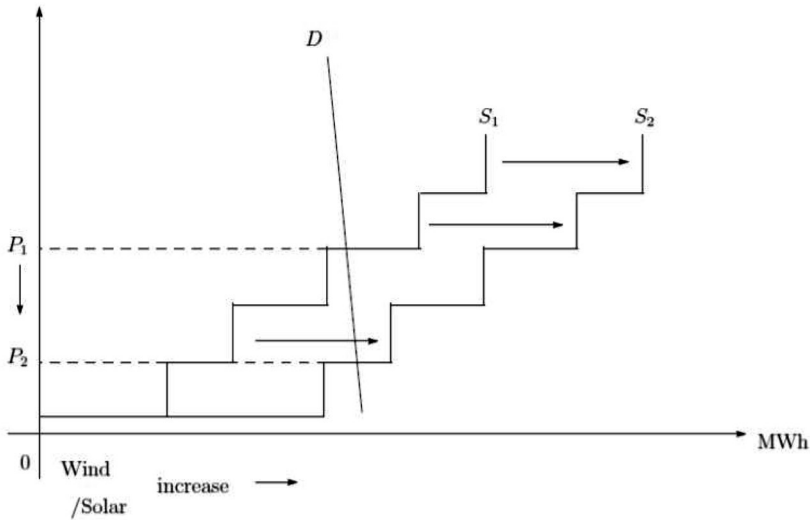
Source: Maekawa et al. (2018), figure 2.

Electricity supply has renewable energies and conventional energies. Renewable energies such as wind and solar are intermittent with large variability and unpredictability. In the case of intermittent renewable energies, the electricity supply heavily depends on environmental conditions such as rainfall, wind power and hours of solar radiation. Because these natural resources can be obtained freely the marginal cost of renewable energies is almost zero.

The merit order effect has gained increasing attention in the literature both on a theoretical and empirical context. Jensen and Skytte (2002) point out that RES generation enters at the base of the merit order function, thus shifting the supply curve to the right and crowding the most expensive marginal plants out from the market, with a reduction of the wholesale clearing electricity price.

An increase in renewable energies pushes the merit order curve (supply curve) S1 to S2 and the spot price falls from P1 to P2 (see Figure 3).

Figure 3: The price effect of increasing renewable energies



Source: Maekawa et al. (2018), figure 4.

3. WHOLESALE MARKET

Theoretical consideration introduced by Jensen and Skytte (2002) suggest that renewable electricity production results in lower electricity prices. This is explained by the fact that renewables bid into wholesale electricity market at almost zero price, and shift the electricity supply curve to the right.

Several researchers carried out empirical analyses on the impact of RES in electricity markets, finding evidence of the merit-order effect. This means that an increase in intermittent sources generation would reduce the spot electricity market price by displacing high fuel-cost marginal generation. RES installations, although they are very capital-intensive, have almost zero marginal generation cost and thus are certainly dispatched to meet demand. More expensive conventional power plants are crowded out and the electricity price declines.

Considering the wholesale market, Würburg et al. (2013) presented an overview of previous studies about the effect of intermittent renewables energies on electricity prices, grouping it in three categories: simulation-based; empirical; and studies with limited information on the price effects of renewables. Despite the different approaches, given the growing availability of ex-post data on electricity prices and renewable capacity, we focus the literature review on empirical approach.

Various papers found empirical evidence of the merit-order effect in various countries, namely Würzburg et al. (2013) for Germany, Gelabert et al. (2011) for Spain, Woo et al. (2011) for Texas and O'Mahoney and Denny (2011) for Ireland. These studies differ with respect to methodology used, types of renewable sources and country analysed, as well as frequency of the data used. However, there is a wide consensus that RES penetration has lowered wholesale electricity prices.

Würzburg et al., 2013 carried out an empirical analysis of the price effect of renewable production for the Austrian-German region. These authors proved that the merit-order effect varies depending on the region and the method chosen. They concluded that the merit-order effect is much larger for smaller markets opposed to larger markets. They found that for each GWh of average hourly predicted renewable energy generation the day-ahead electricity price was reduced by 2% for the German-Austrian market.

Using daily production quantities of different electricity generation types for Spain, Gelabert et al. (2011) investigated its effect on electricity prices during the 2005-2009 period. They concluded that each GW of additional renewable electricity production reduced Spanish electricity prices by roughly 2 €/MWh. A similar result was obtained by Sáenz de Miera et al. (2008) after picking three arbitrary days of February 2006 to perform an exhaustive comparison of electricity prices and wind energy production.

Woo et al. (2011) found that a 1 GWh increase in wind generation (during 15 minutes) decreased Texas balancing electricity prices between 13 and 44 US\$/MWh.

Employing hourly data for Ireland, O'Mahoney and Denny (2011) explained the movements in the shadow price of electricity. They stated that a key challenge internationally is the design of future electricity systems which will bring about emissions savings and fuel security at least cost. Renewable energy is considered a viable alternative to conventional generation with zero or lower emissions. They found that the price of electricity fell by 9.9 €/MWh per GW of wind.

Clò et al. (2015) analyzed to what extent the penetration of solar and wind electricity sources has lowered day-ahead wholesale electricity prices and whether such a reduction has been sufficient to offset the cost of the RES support schemes sustained by final consumers. Based on Woo et al. (2011), Clò et al. (2015) developed a quantitative analysis to assess the extent to which variations in consumption patterns and in the energy mix have had an impact on the national wholesale electricity price. While previous studies focused on wind generation, or treated wind and solar generation jointly, they analyzed separately the differential impact of solar and wind generation on Italian day-ahead wholesale electricity prices over the period 2005-2013. Moreover, as RES production has increased greatly over the period analyzed, they were interested in studying whether the impact of intermittent RES on electricity prices level has varied over time.

Recently, Benhamd et al. (2018) carried out an econometric analysis for Germany, as a country with high penetration of renewable energy sources (RES), in order to investigate impact of wind energy and Photovoltaic feed-in on electricity spot price level, the so-called merit-order effect. They used an ARMA-X- GARCH-X modeling where wind generation and photovoltaic are considered as exogenous variables included in the mean and the variance equation, in order to assess the joint impact of RES on the electricity spot price level as well as on spot price volatility in Germany. They found that wind power and Photovoltaic feed-

in decreases electricity spot price. However, their impact on electricity spot prices volatility are quite different. The solar Photovoltaic power has a lowering on impact electricity price volatility whereas the wind feed-in get worse it.

In short, the previous works showed that higher intermittent renewable energy production was generally and consistently associated with reduce electricity prices, at least in the short term. According to Gelabert et al. (2011), the long-term effects may be different since the new generation mix and short-term lower prices influence investment decisions and thus future prices.

4. RETAIL MARKET

There are several empirical studies on the analysis of the impact of higher levels of intermittent RES generation (triggered by RES support schemes) on electricity prices. These studies on the merit-order effect focused on the impact of such generation on wholesale electricity prices, leading to the conclusion that it results in a reduction on those prices (Sensfuß et al., 2008, Sáenz de Miera et al., 2008, Gelabert et al., 2011, and Würzburg et al., 2013).

Since RES generation is generally supported with an add-on on the wholesale price paid by electricity consumers in their bills, the issue is whether the lower wholesale price is partially or totally offset by the RES support, leading to an increase or reduction in the retail prices. On this matter, there are still few empirical studies and there is no consensus on the conclusions obtained.

The analysis of the impact of RES support on retail electricity prices can be placed in the context of the regulatory analysis of the electricity sector. Electricity is characterized for being one of the most highly regulated sectors. Several regulatory variables can be considered as key drivers of retail electricity prices, including the promotion costs related to RES support, network costs, taxes and levies.

In the European Union (EU), retail electricity prices have significantly increased in the last decade as a result of different regulations raising the concern of policy makers. The growth in the support costs for electricity from renewable energy sources (RES-E) has often been singled out as a main driver of these prices.

Consumer decisions are taken on the basis of retail prices, which have not seen the same decreases as wholesale electricity prices. Inefficiencies within the retail price, such as “tariff structure, taxation and the lack of time-varying options” give incorrect price signals to the consumer.

Gelabert et al. (2011) and Würzburg et al. (2013) found that the reduction in wholesale electricity prices induced by higher RES production offsets the increase in final electricity retail prices induced by RES support schemes (subsidies directly paid by consumers in the final energy bill). They respectively conclude that in Spain and Germany the increase in electricity production from RES has generated a net benefit to consumers.

London Economics (2012) assesses the trends and position of the UK electricity retail market relative to comparable jurisdictions in terms of prices, competition and profitability in the period 1984-2010. The econometric analysis suggests that commodity input prices, fuel mix (electricity) and wages explain a large amount of the variation in retail prices for

both residential and industrial customers across jurisdictions and over time. However, neither of these studies analyze the relative importance of RES-E support with respect to other cost components of retail prices.

Moreno et al. (2012) developed an econometric panel data model to estimate the relationship between the household retail electricity prices and RES in the EU from 1998 to 2009. They concluded that the retail electricity prices increase with the deployment of RES (wind). They found that a 1% increase in renewable energy results in a 0.018% increase in household electricity prices. This small effect is mostly influenced by RES-E support schemes financed by the electricity market. Public RES-E support schemes may effectively mitigate the retail price increase.

Considering the Spanish market and “peak and off-peak prices” separately, Ballester and Furió (2015) found that an increasing share of renewable energy results in lower retail electricity prices in Spain for the period from 2010 to 2013. The opposite is true for the period 2002 to 2009, due to less intense employment of renewable technologies and higher price associated with these technologies during this period. The weight of RES increased from 29% to 59% from 2008 to 2013, consisting of up to 80% of the daily supply on occasion since 2011. They employed a model adapted by Cartea et al. (2005), a stochastic process with mean reversion that includes a discrete jump process, which allows for price volatility to be captured. They found that the relationship between the RES-E share and electricity prices is only significant for peak prices; significant positive relationships between thermal (coal and oil-gas), nuclear, hydroelectric, pumping hydroelectric, combined cycle, renewables and price volatility have been found; renewables are negatively correlated with upward jumps in peak prices and no significant relationship have been found considering off-peak prices.

European Commission (2016) stated that wholesale electricity prices have decreased significantly although retail electricity prices tend to increase due to the network price component, taxes, and levies.

Based on Moreno et al. (2012) and Imura and Cross (2018) analyzed the effects of renewable energy on household electricity prices in liberalized electricity markets in 7 OECD countries. These authors found a strong path dependency for household electricity prices, while market reforms resulted in more significant price decreases than policy anticipated. The results obtained confirm that there is no significant relationship between higher prices and increased renewable deployment. They suggested that renewables are more likely to be traded with neighboring countries than deployed by the host country due to the merit-order effect.

According to Trujillo-Baute, et al. (2018), previous studies are inconclusive about the effect of the intermittent renewable energy share on retail electricity price as country-specific regulatory policy has a significant impact on retail electricity prices. These authors attributed the significant increase in retail electricity prices to the increase in renewable energy sources.

A study carried out by Greenstone and Nath (2019) examined the effects of renewable portfolio standards programs adopted by 29 states and Washington DC. They concluded that these policies raised retail prices significantly and reduced CO₂ emissions only to a small extent. In the European Union as a result of different regulations, retail electricity prices substantially increased in the last decade raising the concern of policy makers. The main reason for such high prices is the growth in the support costs for electricity from renewable energy sources (RES).

Finally, Oosthuizen et al. (2019) analyzed the effect of the increasing intermittent renewable electricity share on retail electricity prices for 34-OECD countries, considering the change in market structure for 23 EU countries. The results showed that the influence of the renewable energy share in the energy mix to retail electricity prices is positive and statistically significant.

5. CONCLUSIONS

In the EU, a recent study on energy prices and costs highlights that the increasing in intermittent renewable energy generation has reduced the wholesale electricity market. Although the final consumers' price has increased as a result of different regulations raising the concern of policy makers and researchers. The growth in the support costs for electricity from renewable energy sources (RES-E) has often been singled out as a main driver of these prices.

Taking into account the aim of this paper, which is to present a literature survey on the effect of intermittent renewable energy's generation on electricity prices, numerous empirical studies have been undertaken on the merit order effect and results obtained are in line with expectations.

That being said, the question is whether the lower wholesale price is partially or totally offset by the RES support, leading to an increase or reduction in the retail prices. The empirical studies carried out to assess this question are still very few in comparison with the merit order effect. In addition, there is no consensus with regard to the conclusions obtained.

Given the interest and relevance of this issue for policy makers, researchers and consumers (households and firms), as future work it is fundamental to carry out further empirical studies using panel data set and applying panel data techniques to evaluate the impact of intermittent renewable energy's generation on wholesale and retail electricity prices.

REFERENCES

- Ballester, C.; Furió, D. (2015) Effects of renewables on the stylized facts of electricity prices. *Renewable & Sustainable Energy Reviews*, 52, 1596-1609.
- Benhmad, F.; Percebois, J. (2018) An econometric analysis of the merit order effect in electricity spot price: The German case, Centre de Recherche en Economie et Droit de l'Energie, Cahier de Recherche n° 18.01.120.
- Cartea, A.; Figueroa, M. (2005) Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4), 313-335.
- Clò, S.; Cataldi, A.; Zoppoli, P. (2015) The merit-order effect in the Italian power market: The impact of solar and wind generation on national wholesale electricity prices. *Energy Policy*, 77, 79-88.
- Cludius, J.; Forrest, S.; MacGill, I. (2014) Distributional effects of the Australian renewable energy target (ret) through wholesale and retail electricity price impacts. *Energy Policy*, 71, 40-51.
- Costa-Campi, M.; Trujillo-Baute, E. (2015) Retail price effects of feed-in tariff regulation. *Energy Economics*, 51, 157-165.
- Del Río, P.; Mir-Artigues, P. (2014) Combinations of support instruments for renewable electricity in Europe: a review. *Renewable and Sustainable Energy Reviews*, 40, 287-295.
- European Commission (2016) *Energy prices and costs in Europe*. Brussels.
- Gelabert, L.; Labandeira, X.; Linares, P. (2011) An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. *Energy Economics*, 33(1), 559-565.
- Greenstone, M.; Nath, I. (2019) Do renewable portfolio standards deliver? Energy Policy Institute at University of Chicago WP 2019-62.
- Hogan, W.W. (2019) Market design practices: Which one are best? *IEEE Power and Energy Magazine*, 17(1), 100-104.
- Imura, A.; Cross, J. (2018) The impact of renewable energy on household electricity prices in liberalized electricity markets: A cross-national panel data analysis. *Utilities Policy*, 54, 96-106.
- Jensen S.; Skytte K. (2002) 'Interactions between the power and green certificate markets'. *Energy Policy*, 30(5), 425-435.
- London Economics (2012) *Energy Retail Markets Comparability Study*. (A report for DECC).
- Maekawa, J.; Hai, B.; Shinkuma, S.; Shimada, K. (2018) The effect of renewable energy generation on the electric power spot price of the Japan electric power exchange. *Energies*, 11, 2215.
- Moreno, B.; López, A. J.; García-Álvarez, M. T. (2012) The electricity prices in the European Union. The role of renewable energies and regulatory electric market reforms. *Energy*, 48(1), 307-313.
- O'Mahoney, A.; Denny, E. (2011) The merit order effect of wind generation in the Irish electricity market, MPRA Paper, University Library of Munich, Germany.
- Oosthuizen, A.; Inglesi-Lotz, R.; Thopil, G. (2019) The relationship between renewable energy and retail prices: Panel evidence from OECD countries, *Economics Research Southern Africa* WP797.
- Ortega-Izquierdo, M.; del Río, P. (2016) Benefits and costs of renewable electricity in Europe. *Renewable and Sustainable Energy Reviews*, 61, 372-383.
- Sáenz de Miera, G.; del Río González, P.; Vizcaino, I. (2008) Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. *Energy Policy*, 36 (9), 3345-3359.
- Sensfuß, F.; Ragwitz, M.; Genoese, M. (2008) The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. *Energy Policy*, 36(8), 3086-3094.

**THE EFFECT OF INTERMITTENT
RENEWABLE ENERGY GENERATION ON
ELECTRICITY PRICES: A LITERATURE
SURVEY**

- Trujillo-Baute, E.; del Río, P.; Mir-Artigues, P. (2018) Analysing the impact of renewable energy regulation on retail electricity prices. *Energy Policy*, 114, 153-164.
- Woo, C.; Horowitz, I.; Moore, J.; Pacheco, A. (2011) The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy*, 39, 3939-3944.
- Würzburg, K.; Labandeira, X.; Linares, P. (2013) Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria. *Energy Economics*, 40, 159-171.

Rita Martins

Univ Coimbra, CeBER, Faculty of Economics
rvmartin@fe.uc.pt
orcid.org/0000-0001-5300-9322

Micaela Antunes

Univ Coimbra, CeBER, Faculty of Economics
micaela@fe.uc.pt
orcid.org/0000-0002-2113-2139

Patrícia Pereira da Silva

Univ Coimbra, CeBER, Faculty of Economics
patsilva@fe.uc.pt
orcid.org/0000-0002-5862-1278

Adelino Fortunato

Univ Coimbra, CeBER, Faculty of Economics
adelino@fe.uc.pt
orcid.org/0000-0002-4790-6470

DOI: https://doi.org/10.14195/2183-203X_53_6

Tarifa Social de Energia: Génese, Incidência e Lições

Social Energy Tariff: Origin, Impact and Lessons

Rita Martins

Micaela Antunes

Patrícia Pereira da Silva

Adelino Fortunato

Received for publication: June 15, 2020

Revision accepted for publication: April 9, 2021

RESUMO

Desde a sua criação em 2010 na eletricidade e em 2011 no gás natural, para combater os efeitos da Crise Financeira Global de 2008-2009, a tarifa social de energia em Portugal sofreu adaptações no sentido do alargamento do leque de beneficiários. Ainda que subsistam detalhes que podem vir a ser melhorados, a introdução da automaticidade no acesso ao benefício em 2016 revelou-se uma alavanca poderosa para alargar o âmbito de famílias contempladas e para explicar o sucesso dos seus impactos. Tendo em conta o balanço desta experiência, procuramos neste artigo usar a pertinência dos seus princípios para recomendar a extensão da tarifa social aos setores das águas, telecomunicações e transportes, no contexto da crise social e económica gerada pela pandemia Covid-19.

Palavras-chave: Tarifa social de energia; acessibilidade económica; acesso universal.

ABSTRACT

Since its implementation in 2010 (for electricity) and in 2011 (for natural gas) as a measure to combat the effects of the Global Financial Crisis of 2008-2009, the Portuguese social energy tariff has been progressively adapted to cover a growing number of beneficiaries. Although there is room for improvement in certain aspects, the introduction of an automatic procedure for applying to the tariff in 2016 has proven to be a powerful lever to expand the number of beneficiaries and explains the success of its impact. Bearing in mind the efficacy of the social energy tariff, in this paper we seek to recommend the extension of social tariff to the water, telecommunications and transports sectors on the grounds of the relevance of its principles, in the context of the social and economic crisis triggered by the Covid-19 pandemic.

Keywords: Energy social tariff; affordability; universal access.

JEL Classification: D18, L51, L94; L95.

Agradecimentos: Os autores agradecem a um revisor anónimo, pelos seus comentários e sugestões. Não obstante, os erros e omissões são da inteira responsabilidade dos autores.

1. INTRODUÇÃO

Na União Europeia (UE), o reconhecimento da importância dos serviços energéticos encontra tradução na classificação do fornecimento de eletricidade e de gás natural como serviços de interesse económico geral (SIEG). Deve, por isso, ser respeitado um conjunto de obrigações de serviço público, entre as quais a universalidade e a acessibilidade económica, que salvaguardem a satisfação das necessidades fundamentais de consumo de energia dos cidadãos.

As tarifas sociais foram criadas na sequência da última crise económico-financeira com o propósito de apoiar o consumo de energia pelas famílias que estariam a passar por um período de maior dificuldade. Enquadram-se no conjunto de medidas de proteção a consumidores vulneráveis, concebidas e aplicadas em linha com o parecer do Comité Económico e Social Europeu (Comissão Europeia – Parecer 2013/C 341/05, 2013, designado “Ação coordenada a nível europeu para prevenir e combater a pobreza energética”). Este instrumento foi concebido para vigorar num horizonte de curto prazo, e tem critérios (rendimentos baixos e/ou determinadas pensões sociais, no modelo atual) de natureza financeira. Para além de serem uma medida de proteção social, as tarifas sociais fazem parte da política de regulação do mercado energético (Dobbins et al., 2016). As tarifas sociais são medidas aplicadas na Bélgica, Bulgária, Chipre, França, Grécia, Espanha, Itália, Roménia e também em Portugal, (Pye et al., 2015).

Em Portugal, a tarifa social de eletricidade foi criada em 2010 e a tarifa social de gás natural em 2011, tendo sofrido, desde então, adaptações no sentido do alargamento do leque de beneficiários. Ao permitir a redução dos encargos energéticos dos consumidores domésticos, contribui para que o preço não seja fator de exclusão no acesso a tais serviços, independentemente da situação económica, social ou geográfica dos consumidores. Serve assim o propósito de assegurar o acesso universal a serviços de qualidade a preços acessíveis.

A tarifa social de energia consiste num desconto sobre as tarifas de acesso às redes, que são reguladas pela Entidade Reguladora dos Serviços Energéticos (ERSE), e que visam refletir os custos com o sistema. Estas tarifas de acesso às redes são pagas por todos os consumidores de energia, independentemente de estarem no mercado regulado ou no mercado liberalizado. Na eletricidade, o desconto a aplicar nas tarifas de acesso às redes em baixa tensão normal incide sobre as tarifas transitórias de venda a clientes finais de eletricidade, excluído o IVA, demais impostos, contribuições, taxas e juros de mora que sejam aplicáveis. No gás natural, a tarifa social corresponde a um desconto na tarifa de acesso às redes em baixa pressão sobre as tarifas transitórias de venda a clientes finais, incidindo sobre as tarifas transitórias de venda a clientes finais de gás natural, excluído o IVA, demais impostos, contribuições, taxas e juros de mora que sejam aplicáveis.

Mais do que nunca, a necessidade de proteção dos consumidores vulneráveis vai acentuar-se devido às consequências decorrentes da retração económica esperada e ao aumento previsível de população economicamente vulnerável, devido aos choques, tanto do lado da oferta como da procura, provocados pelos constrangimentos associados à pandemia por Covid-19. Os serviços energéticos são essenciais a todas as famílias e particularmente relevantes, tanto para garantir o cumprimento do isolamento social como, após o levantamento das restrições mais exigentes impostas pela pandemia.

Neste contexto, o objetivo deste artigo é o de caracterizar a incidência e o perfil dos beneficiários da tarifa social de energia (eletricidade e gás natural) em Portugal para apreciar a sua relevância também para outros SIEG no quadro do agravamento da crise económica e social.

Este artigo está organizado da seguinte forma. Após a introdução, a Secção 2 oferece uma análise sobre a incidência da tarifa social de energia em Portugal. Na Secção 3 é feita uma reflexão acerca do modelo da tarifa social de energia, com vista a compreender os seus méritos e limitações e assim perceber o que pode ser transposto para outros setores essenciais, e como pode ser útil no contexto económico desencadeado pela atual pandemia. Na Secção 4 apresentam-se as conclusões e algumas recomendações.

2. INCIDÊNCIA DA TARIFA SOCIAL DE ENERGIA EM PORTUGAL

A elegibilidade para aceder ao benefício na energia elétrica faz-se pela condição rendimento, se o total anual do agregado familiar não ultrapassar o limiar de 5 808€, acrescido de 50% por cada elemento que não aufera qualquer rendimento, até ao máximo de 10 elementos; e/ou pela condição prestações sociais, se o titular do contrato for beneficiário de determinadas transferências sociais. Os critérios de elegibilidade são mais restritivos no caso do gás natural, uma vez que só a condição prestações sociais dá acesso ao benefício e o leque de prestações tido em conta é também menor que no caso da energia elétrica.

Desde a sua criação, a tarifa social de eletricidade tem sido atribuída a um número crescente de famílias, especialmente a partir do final de 2014, a que não são alheias as alterações introduzidas pelo Decreto-lei n.º 172/2014, de 14 de novembro, em concreto a introdução de um critério rendimento e o alargamento das prestações sociais elegíveis.

Em 2016, o acesso ao benefício da tarifa social de energia passou a ser realizado em Portugal através de um mecanismo de reconhecimento automático, com efeitos a partir de 1 de julho de 2016, centralizado na Direção-Geral de Energia e Geologia (DGEG), no caso de Portugal Continental. Nas Regiões Autónomas, o processo de reconhecimento automático foi atribuído às entidades das respetivas administrações regionais com competências nas matérias em causa.

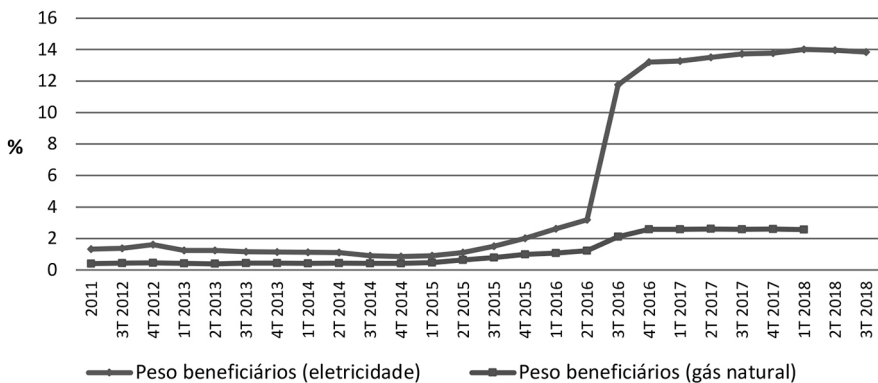
Imediatamente após a entrada em vigor do procedimento automático (do 2.º para o 3.º trimestre de 2016), o crescimento do número de beneficiários foi bastante expressivo, 283%. A partir daí, as evoluções foram muito mais modestas, verificando-se mesmo uma alteração da tendência a partir de 2018 o que poderá significar que a melhoria da situação económica veio atenuar o efeito do procedimento automático. No 2.º e 3.º trimestres de 2018 verificou-se uma ligeira redução do número de beneficiários de tarifa social de eletricidade. Atualmente menos de 800 000 famílias beneficiam deste apoio.

A evolução temporal do número de beneficiários de tarifa social de gás natural apresentou também um crescimento continuado do número de beneficiários (aumentando em cerca de sete vezes em relação aos números iniciais). Tal como na eletricidade, o crescimento mais pronunciado aconteceu na sequência da aplicação do mecanismo de atribuição automática da tarifa social (verificando-se um crescimento de 73% no trimestre subsequente). Ainda assim, os valores são bastante mais estáveis, rondando atualmente as 35 000 famílias beneficiárias.

O crescimento muito acentuado do número de beneficiários subsequente ao automatismo na atribuição da tarifa social parece estar em linha com a relutância em pedir apoio, um dos fatores de risco de pobreza energética (conforme referido em Assist2gether, s.d.) ou com o desconhecimento do benefício e/ou da forma de o solicitar.

Para avaliar a expressão da tarifa social, encontra-se representada na Figura 1 a evolução do peso de beneficiários no universo de consumidores domésticos de energia.

Figura 1: Evolução do peso de beneficiários de tarifa social



Fonte: ERSE e DGEG.

É evidente, através da sua análise, que o peso de beneficiários de tarifa social de eletricidade é superior ao do gás natural. Focando-nos primeiro na energia elétrica, o rácio de beneficiários passou de 1.32% em 2011, aquando da implementação da tarifa social, para 13,84% no 3.º trimestre de 2018. Vila Real registou, em 2018, 20% de beneficiários no total de consumidores domésticos seguido dos distritos do Porto, Braga e Viseu. O aumento mais pronunciado na evolução deste indicador deu-se precisamente no 3.º trimestre de 2016, altura em que entrou em vigor o procedimento automático. O peso de beneficiários no universo de consumidores domésticos foi de 13,92% em 2018. Em relação ao gás natural, o peso de beneficiários de tarifa social foi sempre inferior a 1% até ao final de 2015. A partir daí a evolução foi de forma geral positiva, sendo de 2.57% no último período retratado na figura (1.º trimestre de 2018).

Esta diferença em termos do rácio de beneficiários de tarifa social nos dois tipos de serviços energéticos é explicável, em parte, pelo facto de não haver uniformidade nos critérios de elegibilidade, sendo os da tarifa social de eletricidade mais abrangentes (além das prestações sociais, também incluem o critério rendimento). Acresce a esta explicação o facto de a extensão da rede de distribuição de gás natural ser menor do que a de eletricidade.

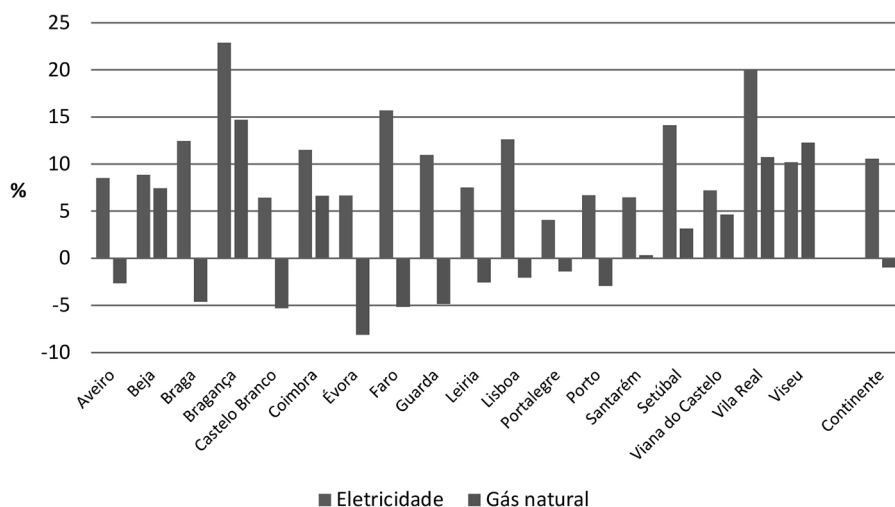
O apuramento mais recente da DGEG revela a existência de 778 1935 famílias beneficiárias de tarifa social de energia elétrica no final do 2.º trimestre de 2020, representando um aumento de 0,25% face ao mesmo período do ano anterior, ou seja, uma estabilização

nos últimos processamentos. Em relação ao gás natural, o apuramento relativo ao final do 1.º trimestre de 2020, aponta para 34 709 famílias beneficiárias, traduzindo uma redução de cerca de 2% face a igual período do ano anterior.

A distribuição do peso de beneficiários de tarifa social de eletricidade em 2018 parece estar em linha com os dados mais recentes do Eurostat relativos à percentagem de população portuguesa em risco de pobreza (considerando como limiar de pobreza 60% da mediana do rendimento total por adulto equivalente), 17% em 2018. De acordo com INE (2019), a taxa de risco de pobreza correspondia, em 2018, à percentagem de habitantes com rendimentos monetários líquidos (por adulto equivalente) inferiores a 6 014€ anuais (501€ por mês), ligeiramente superior ao limiar de 5 808€ de rendimento máximo elegível para acesso à tarifa social.

Para compreender se existem disparidades no país em termos da evolução de beneficiários de tarifa social de eletricidade e de gás natural, representa-se na Figura 2 as respetivas taxas de crescimento, após a entrada em vigor do procedimento automático, nos diferentes distritos de Portugal Continental.

Figura 2: Taxas de crescimento do número de beneficiários de tarifa social de eletricidade e de gás natural, por distrito (3.º trimestre de 2016-4.º trimestre de 2018), Portugal Continental



Fonte: DGEG.

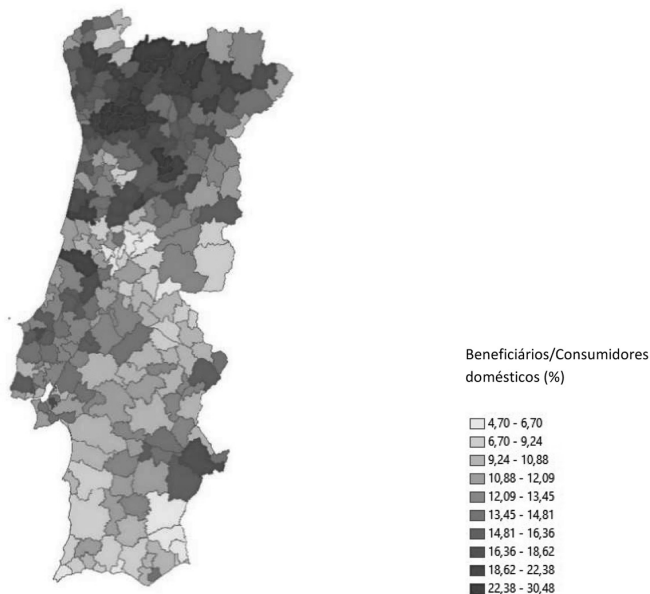
Já com o mecanismo automático de atribuição da tarifa social em vigor, é bem visível o crescimento dos beneficiários de tarifa social de eletricidade em todos os distritos de Portugal Continental, para o período assinalado. No entanto, existem diferenças entre os distritos; desde logo, é notório o crescimento mais elevado do total de beneficiários em Bragança e em Vila Real.

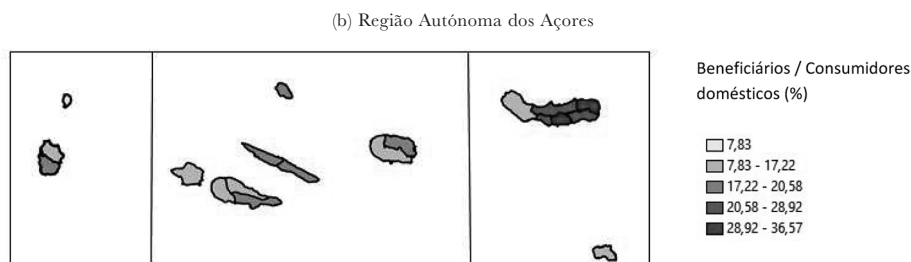
Em relação à taxa de crescimento de beneficiários de tarifa social no gás natural, o comportamento entre distritos é bem mais diferenciado, sendo mais expressivo nos distritos mais rurais de Portugal Continental (como Bragança, Viseu e Vila Real). Verifica-se um decréscimo do número de beneficiários em 10 dos 18 distritos de Portugal Continental, que pode ser explicado pela maior exigência dos critérios de elegibilidade para a atribuição da tarifa social de gás natural, comparativamente ao que sucede na tarifa social de eletricidade, conjugada com a melhoria de condições económicas, de forma geral, no país, entre 2016 e 2018.

Na Figura 3 analisam-se os rácios de beneficiários no universo de consumidores domésticos para uma desagregação territorial ao nível do concelho, para 2018. A Figura encontra-se dividida em três painéis: A – Portugal Continental; B – Região Autónoma dos Açores (RAA); C – Região Autónoma da Madeira (RAM).

Figura 3: Peso dos beneficiários de tarifa social de energia elétrica, por concelho, 2018

(a) Portugal Continental





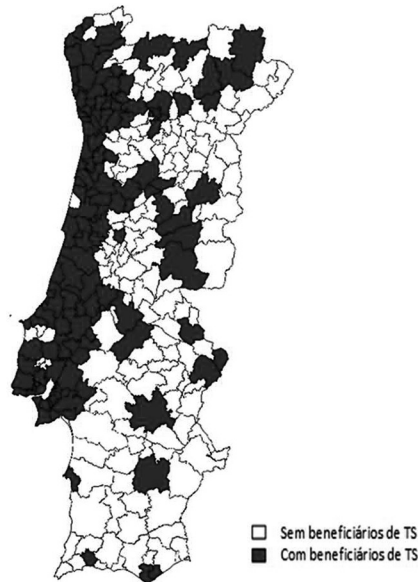
Fonte: Adaptado de Martins et al. (2019b).

É visível um maior peso de beneficiários no universo de consumidores domésticos em concelhos do norte do país em 2018 (3a), indo ao encontro do que se observou ao nível de distritos, na Figura 2. Na RAA (3b) é na ilha de S. Miguel que se concentram os concelhos onde a proporção de beneficiários no universo de consumidores domésticos é maior. Finalmente, na Região Autónoma da Madeira (3c), os concelhos com maior peso de beneficiários localizam-se na zona sul e oeste da ilha da Madeira. Estes resultados corroboram, em certa medida, a constatação de que a taxa de risco de pobreza é mais elevada nas Regiões Autónomas, seguindo-se, no Continente, a região Norte (INE, 2019).

Em relação ao gás natural, no final de 2018 havia beneficiários de tarifa social em 134 dos 278 concelhos de Portugal Continental. Dados relativos ao final do 1.º trimestre de 2020 apurados DGEG¹ dão conta de que existem atualmente 148 concelhos com beneficiários de tarifa social no gás natural.

¹ <https://www.tarifasocial.dgeg.gov.pt/estatistica.aspx>, acedido a 13-05-2020.

Figura 4: Concelhos com e sem beneficiários de tarifa social de gás natural, 4º trimestre de 2018, Portugal Continental



Fonte: Martins et al. (2019b: 25)

3. VANTAGENS E LIMITAÇÕES DA TARIFA SOCIAL DE ENERGIA

A situação atual de pandemia provocada pela Covid-19 está a inverter a tendência de retoma económica dos últimos anos. As projeções mais recentes da Direção-Geral dos Assuntos Económicos e Financeiros da Comissão Europeia, no seu Boletim da Primavera de 2020 (Comissão Europeia, 2020) apontam para uma contração do PIB de 6,8% em 2020 e um crescimento de 5,8% em 2021, indiciando, portanto, que a retoma será incompleta e o PIB permanecerá abaixo do nível de 2019 ainda em 2021. A taxa de desemprego rondará, de acordo com estas projeções, os 9,7% em 2020 e 7,4% em 2021, contrariando a tendência recente de decréscimo neste indicador em Portugal. O Conselho de Finanças Públicas (CFP, 2020) apresentou um cenário ligeiramente mais pessimista. Estes indicadores poderão ainda sofrer agravamentos, fruto da incerteza associada à evolução e à magnitude dos efeitos provocados pela Covid-19, que está a afetar as economias de forma diferenciada, constituindo, pois, um choque assimétrico entre os Estados-Membros da UE.

Certo, no entanto, parece ser o facto de a economia global enfrentar nos tempos próximos uma recessão mais severa do que a imposta pela Crise Financeira Global de 2008-2009. Perante este contexto, é expectável que as famílias mais vulneráveis voltem a enfrentar dificuldades para suportar as suas despesas com os serviços essenciais e, em concreto, os de

energia. Neste sentido, uma reflexão sobre o impacto da tarifa social de eletricidade e de gás natural em Portugal é relevante para perceber as suas potencialidades na mitigação dos efeitos da perda de rendimentos das famílias. Além disso, a compreensão das suas limitações também possibilita a discussão em torno da eventual necessidade da adoção de medidas complementares. A partir daqui retiram-se recomendações acerca das vantagens da eventual transposição do modelo da tarifa social de energia para outros serviços essenciais.

3.1. IMPACTO DA TARIFA SOCIAL DE ENERGIA NA ACESSIBILIDADE ECONÓMICA

Para avaliar o impacto da tarifa social sobre a acessibilidade económica, procede-se a um enquadramento sucinto da literatura sobre esta temática e segue-se uma abordagem empírica, primeiramente a partir de dados microeconómicos efetivos relativos a despesas das famílias e posteriormente recorrendo a simulações.

A avaliação de problemas de acessibilidade económica a serviços essenciais é habitualmente realizada na literatura económica através do cálculo de rácios de acessibilidade económica (RAE) que procuram medir o peso dos encargos com os serviços nos orçamentos das famílias. No caso dos serviços energéticos teríamos:

$$RAE = \frac{\text{Encargos com energia}}{\text{Rendimento}} \quad (1)$$

Ainda que não seja um aspeto isento de críticas, a literatura também fornece referências quanto a limiares para os RAE que não deverão ser ultrapassados, sob pena de representarem um esforço excessivo exigido às famílias. Martins et al. (2019a) e Fankhauser et al. (2008) consideram que RAE para serviços de energia, águas e comunicações superiores ou iguais a 20% são a referência para pesos preocupantes com os respetivos encargos conjuntos nos orçamentos familiares. No que toca ao mesmo tipo de limiar, mas apenas em relação aos serviços energéticos, é comum considerar-se como referência 10% (Sumner et al., 2015; Miniaci et al., 2014; ou Winkler et al., 2011). Recentemente, Brown et al. (2020) e Cook e Shah (2018) sugerem para os Estado Unidos da América um limiar menos conservador, entre 6 e 7%.

Para além disto, um resultado comum aos estudos é a maior vulnerabilidade das famílias com menores rendimentos a problemas de acessibilidade económica, mesmo beneficiando de tarifa social (Winkler et al, 2011). Alguns estudos apontam também para a verificação de rácios com variabilidade por regiões (Sumner et al., 2015) e em função da dimensão dos agregados familiares (Fankhauser e Tepic, 2007).

No que concerne a resultados para Portugal, a literatura é particularmente escassa. Martins et al. (2016), num estudo empírico para os serviços de águas alertam para problemas de avaliação da acessibilidade económica quando baseada em dados agregados. Martins et al. (2019a), num estudo de avaliação da acessibilidade a um conjunto de serviços essenciais, entre os quais a eletricidade e o gás, alertam ainda para a subestimação do problema da acessibilidade económica, quando são utilizados dados da despesa efetiva. Situações que podem ser erradamente interpretadas como não indiciadoras de problemas de acessibilidade são as que

se encontram associadas a subconsumo, por exemplo ligadas a contextos de pobreza. Nestas circunstâncias, baixos valores para o RAE decorrem de situações de desconforto e privação, ao invés de corresponderem a desafogo para pagar as contas com os serviços energéticos.

Os dados microeconómicos utilizados são resultantes da mais recente recolha do Instituto Nacional de Estatística (INE) no âmbito do Inquérito às Despesas das Famílias (IDEF 2015/2016). São apurados rácios de acessibilidade económica conjunta a serviços energéticos (eletricidade e gás), utilizando a despesa total como *proxy* para o rendimento, bem como a percentagem de famílias com RAE superiores a 10%. Os resultados por regiões são apresentados no Quadro 1.

Importa assinalar que com a base de dados disponível não é possível distinguir entre as despesas suportadas com e sem tarifa social. Ainda assim, face ao período a que respeita a recolha de dados (anterior à aplicação do procedimento automático), em que o peso de beneficiários de tarifa social era inferior a 3%, na eletricidade, e cerca de 1% no gás natural, é razoável admitir que a esmagadora maioria das despesas com serviços de eletricidade e gás não beneficiavam do desconto inerente à tarifa social.

Quadro 1: RAE a serviços energéticos em Portugal

| | N.º de observações | RAE Médio | Desvio Padrão | Mínimo | Máximo | % Famílias RAE > 0,10 |
|--------------|--------------------|-----------|---------------|--------|--------|-----------------------|
| Norte | 2 197 | 0,0728 | 0,0437 | 0 | 0,4652 | 20,7 |
| Centro | 1 556 | 0,0760 | 0,0446 | 0 | 0,3808 | 21,6 |
| AMLisboa | 2 536 | 0,0551 | 0,0342 | 0 | 0,3450 | 9,4 |
| Alentejo | 1 244 | 0,0757 | 0,0362 | 0 | 0,2758 | 21,1 |
| Algarve | 1 129 | 0,0633 | 0,0366 | 0 | 0,3487 | 14,3 |
| R.A.Açores | 1 462 | 0,0877 | 0,0469 | 0 | 0,4249 | 32,1 |
| R.A. Madeira | 1 255 | 0,0748 | 0,0432 | 0 | 0,4511 | 21,1 |
| Portugal | 11 379 | 0,0708 | 0,0420 | 0 | 0,4652 | 19,2 |

Fonte: IDEF 2015/2016.

Ainda que os valores médios para os RAE se situem aquém dos limiares problemáticos de 10%, referidos na literatura, é expressiva a percentagem de famílias com rácios acima desses patamares, superior a 19% em Portugal, o que revela que no contexto português há uma proporção elevada de famílias a suportar encargos excessivos com serviços energéticos. Destaca-se também, para além do desvio padrão, a amplitude de valores para o RAE. Pese embora se tomar como referência o limiar de 10%, importa mencionar que se se seguisse uma abordagem menos conservadora, considerando um patamar na ordem dos 6 ou 7%, como em Brown et al. (2020) e Cook e Shah (2018), os RAE médios já seriam todos interpretados como problemáticos, à exceção do da AML.

A RAA apresenta valores médios para o RAE mais preocupantes, aproximando-se de 9%. E mais alarmante é a percentagem superior a 32% de famílias desta região com RAE superiores a 10% (última coluna do Quadro 1). No extremo oposto, encontra-se a região Área Metropolitana de Lisboa (AML) com rácio médio de 5,5% e 9,5% de famílias com problemas de acessibilidade económica. Excetuando as regiões da AML e do Algarve, há mais de 20% de famílias cujos encargos com os serviços energéticos têm um peso superior a 10% dos orçamentos familiares, em todas as outras regiões NUTS II.

A percentagem de famílias que tem de gastar mais de 10% do seu rendimento para fazer face ao pagamento dos serviços energéticos não é, portanto, negligenciável e permite evidenciar a real dimensão do problema da acessibilidade económica aos serviços energéticos por parte das famílias portuguesas, que as médias, por si só, não revelam.

Estes resultados estão em linha com os apresentados em INE (2019), segundo os quais os residentes na Área Metropolitana de Lisboa foram os menos afetados pelo risco de pobreza em 2018 (13,3%), ao passo que na RAA e na RAM registaram-se as percentagens mais elevadas de população em risco de pobreza (31,8 e 27,8%, respetivamente).

A análise agregada apresenta limitações, evidente quando se olha para as percentagens de famílias com rácios excessivos e a amplitude encontrada para os valores do RAE, que deixa antever uma grande disparidade de situações. Quando se calculam os RAE por quintil de rendimento, constata-se a existência de valores próximos do limiar considerado problemático para as famílias do primeiro quintil de rendimento, ainda que, em média, ligeiramente abaixo do patamar dos 10%, sendo o rácio menor nos quintis de rendimento mais elevados, como seria de esperar.

Dado que a base IDEF 2015/2016 não permite distinguir as situações em que as famílias beneficiam de tarifa social das situações em que tal não acontece, e atendendo a que não se dispõe de dados microeconómicos para o período mais recente, apresenta-se no Quadro 2 um exercício de simulação do impacto potencial da tarifa social na acessibilidade aos serviços energéticos.

Para o exercício efetuado, consideram-se, para cada um dos serviços energéticos, dois cenários correspondentes a dois tipos de famílias. Para calcular os encargos, recorre-se ao comparador de tarifários do portal Poupa Energia.

Para o caso da eletricidade, no cenário 1, as hipóteses consideradas são as seguintes: potência contratada de 3,45kVA, tarifa simples, fatura em papel e pagamento Multibanco (MB)/CTT/Payshop. No cenário 2, a potência contratada é de 6,9kVA, com tarifa bi-horária, fatura em papel e pagamento MB/CTT/Payshop.

Quanto ao gás natural, a localização escolhida foi Lisboa, com fatura em papel e pagamento por MB, tendo-se recorrido ao simulador da ERSE.

De entre as alternativas desenvolvidas, a escolha recai sobre a opção mais favorável do ponto de vista do consumidor.

Quadro 2: Simulação de RAE com e sem tarifa social (TS)

| | RAE | |
|--|--------|--------|
| | Sem TS | Com TS |
| Eletricidade | | |
| Cenário E1: Casal sem filhos, consumo anual de 1900 kWh e potência contratada de 3,45kVA ⁽¹⁾ | | |
| 1.1: Ambos os elementos auferem rendimento (5 808€) | 8,3% | 5,5% |
| 1.2: Apenas um elemento auferir rendimento (8 712€) | 5,6% | 3,7% |
| Cenário E2: Casal com 2 filhos, consumo anual de 5000 kWh (2083 kWh em vazio) e potência contratada de 6,9kVA ⁽²⁾ | | |
| 2.1: Ambos os elementos do casal auferem rendimento (11 616€) | 10,0% | 7,0% |
| 2.2: Apenas um elemento auferir rendimento (14 520€) | 8,0% | 5,6% |
| Gás natural | | |
| Cenário G1: Casal sem filhos, com consumo anual 1610kWh, 138 m ³ | | |
| 1.1: Ambos os elementos auferem rendimento (5 808€) | 2,6% | 1,6% |
| 1.2: 1 Apenas um elemento auferir rendimento (8 712€) | 1,7% | 1,1% |
| Cenário G2: Casal com filhos, com consumo anual 3407kWh, 292 m ³ | | |
| 2.1: Ambos os elementos auferem rendimento (11 616€) | 2,4% | 1,6% |
| 2.2: Apenas um elemento auferir rendimento (14 520€) | 2,0% | 1,3% |

Notas: ⁽¹⁾ Dos clientes com potência contratada até 6,9kVA em 2018, mais de metade tinha uma potência contratada de 3,45kVA, de acordo com dados da ERSE (cenário 1). ⁽²⁾ Dos clientes com potência contratada até 6,9kVA em 2018, 30% tinha uma potência contratada de 6,9 kVA, de acordo com dados da ERSE (cenário 2).

Fontes: Comparador de tarifários do portal Poupa Energia (<https://poupaenergia.pt>) e simulador da ERSE (<http://www.erse.pt/pt/simuladores/Paginas/Simuladores.aspx>)

De uma forma geral, é evidente uma redução do RAE quando é aplicada a tarifa social, mais evidente no caso da eletricidade – descendo entre 2 e 3 pontos percentuais (p.p.) consoante o sub-cenário.

Concretizando, o RAE relativo a eletricidade no cenário E1 baixa de 8,3% para 5,5% (sub-cenário E1.1) e de 5,6% para 3,7% (sub-cenário E1.2), representando diminuições de 2,8 e 1,9 p.p., respetivamente.

Quanto ao cenário E2, a aplicação da tarifa social permite baixar o rácio, do valor problemático de 10% para 7% (sub-cenário E2.1.) e de 8% para 5,6% (sub-cenário E2.1.), correspondendo, portanto a descidas na ordem dos 3,1 e 2,4 p.p., respetivamente.

Quanto ao gás natural, o impacto estimado da tarifa social é mais modesto. Concretizando, a tarifa social permite baixar o peso dos encargos, variando entre 0,6 p.p. (sub-cenário G1.2) e 0,9 p.p. (sub-cenário G2.1.).

Se as famílias consideradas nos cenários anteriores consumirem simultaneamente eletricidade e gás natural, então o efeito potencial da tarifa social na descida do RAE, poderá rondar os 3-4 p.p., para níveis abaixo do patamar mais usualmente considerado como problemático, 10%.

3.2. LIMITAÇÕES DO MODELO DA TARIFA SOCIAL DE ENERGIA

No que toca a falhas ou limitações da tarifa social de energia, podem ser apontadas questões relacionadas com diferenças nos critérios de elegibilidade entre eletricidade e gás natural e com a ausência de tarifa social para o Gás de Petróleo Liquefeito (GPL), mais conhecido como gás de botija.

Em relação às condições para beneficiar de tarifa social de gás natural, os critérios são mais restritivos do que na energia elétrica. A via de acesso à tarifa social baseada no limiar de rendimento só vigora no caso da energia elétrica, o que faz com que a incidência da tarifa social no gás natural se resume a uma espécie de extensão dos apoios fornecidos pelas prestações sociais. Além disso, diferentemente do que acontece na eletricidade, só os beneficiários do 1º escalão de Abono de Família são tidos em conta e ficam também de fora os detentores de Pensão Social de Velhice.

A influência desta circunstância pode ser identificada quando se verificou uma ligeira diminuição do número de beneficiários em 2018, como consequência do crescimento e da melhoria da situação económica e social em Portugal. A queda fez-se a um ritmo maior no gás natural, pois alguns consumidores vulneráveis, beneficiários das prestações da Segurança Social, terão diminuído a sua dependência face a elas e esta oscilação teve menor impacto na diminuição do número de beneficiários na energia elétrica. A prevalência da via de acesso ao benefício da tarifa social pelo limiar de rendimento esbateu esta quebra.

Mas, o que levanta maiores situações de iniquidade é sobretudo o facto de não existir tarifa social no GPL, por não existir rede de gás natural em todo o território nacional (apenas 30% dos agregados tem acesso a este tipo de gás). Ou seja, os consumidores domésticos de GPL em condições de vulnerabilidade económica não beneficiam de apoio para o consumo de gás, a que teriam direito se tivessem acesso ao gás natural. Além disso, é sobretudo nos contextos mais rurais que não existe rede de gás natural, a que acresce a essencialidade deste bem, comprovada pela muito reduzida elasticidade preço da procura de GPL (Autoridade da Concorrência, 2017), o que significa que os consumidores de GPL estão particularmente suscetíveis a subidas de preços. O valor negativo da elasticidade rendimento da procura, por seu lado, revela que o gás de botija é um bem inferior, o que reforça a necessidade de o mesmo estar acessível a preços competitivos para as camadas mais desfavorecidas da população.

Pese embora reconhecidas dificuldades decorrentes do GPL ser facilmente transacionável e para a verificação dos critérios de elegibilidade dos consumidores, inerentes à extensão da tarifa social ao GPL, considera-se que esta medida é importante no domínio da equidade, já que beneficiará sobretudo as franjas mais pobres da população, que residem essencialmente fora dos grandes aglomerados urbanos e, nesse sentido, terá um impacto social positivo.

Uma outra limitação da tarifa social tem que ver com o seu impacto reduzido sobre a redução de problemas de pobreza energética. Há uma certa coincidência temporal entre a

introdução do debate sobre o tema da pobreza energética pela Comissão Europeia (Diretivas 2009/72/CE – eletricidade e 2009/73/CE – gás natural) e a adoção de medidas de apoio ao consumo de serviços energéticos por parte de consumidores vulneráveis, como as tarifas sociais. A necessidade de combater a pobreza energética reveste-se de especial importância, visto que esta é apontada como uma das mais predominantes formas de pobreza na UE, essencialmente motivada pelos baixos rendimentos dos agregados familiares, a reduzida eficiência energética e os elevados preços da energia (Thomson et al., 2017). A relevância da pobreza energética e a necessidade de apoiar consumidores economicamente vulneráveis tornou-se particularmente evidente em países como Portugal, que tem uma das piores posições relativas no contexto da UE nesta matéria.

Ainda assim, a evolução registada nas variáveis habitualmente usadas para avaliar, de forma desagregada e indireta, a pobreza energética (percentagem de população incapaz de manter o alojamento aquecido, percentagem de população com atraso no pagamento de *utilities* e peso da população que vive em habitação com problemas de infiltrações, humidade e apodrecimento de pisos, da base de dados European Union Statistics on Income and Living Conditions–EU-SILC) revelou uma melhoria entre 2015 e 2017. Esta melhoria registada acompanhou a evolução observada em termos da abrangência da tarifa social de energia. O desconto inerente à tarifa social liberta recursos que poderão explicar de alguma forma o ligeiro aumento do consumo de energia elétrica por beneficiário (apurado a partir de ERSE, 2018; 2019). Em particular, o benefício via tarifa social poderá contribuir para amenizar a dificuldade em manter a habitação adequadamente aquecida e para reduzir atrasos no pagamento dos encargos com *utilities*. Importa ressaltar, contudo, que uma análise do nexo de causalidade entre o desconto inerente à tarifa social e o nível de pobreza energética implicaria a recolha de dados primários de modo a poder analisar a evolução do consumo energético dos beneficiários de tarifa social, bem como as suas dificuldades relativas ao pagamento de encargos energéticos e à resolução de problemas de infiltrações e humidade nas suas habitações.

Os recursos financeiros libertados pelo desconto concedido pela tarifa social dificilmente poderão ajudar a resolver problemas relacionados com a qualidade do edificado, em concreto problemas de infiltrações no telhado, humidade em paredes e pisos, ou caixilhos de janelas ou chão apodrecidos. Neste sentido, considera-se necessária a adoção de medidas complementares à tarifa social, vocacionadas para intervenções conducentes à mitigação de problemas de pobreza energética e também para a adoção de equipamentos/práticas energeticamente eficientes. Ou seja, a tarifa social parece não estar configurada para a resolução das causas estruturais da pobreza energética, como também conclui o estudo de Farinha Rodrigues et al. (2019).

Na UE, existem medidas destinadas a aumentar a eficiência energética nas habitações (Deller e Waddams, 2015). Em Portugal, existem diversos programas concebidos com esta finalidade, como o Programa Operacional de Sustentabilidade e Eficiência no Uso de Recursos (PO SEUR), direcionado para apoiar renovações de habitações que promovam a eficiência energética; ou o Programa Casa Eficiente 2020, que visa conceder empréstimos em condições favoráveis a operações que promovam a melhoria eficiência energética e hídrica dos edifícios de habitação particular; e o Programa de Reabilitação Urbana (IFRRU), de apoio à renovação urbana de casas com mais de 30 anos, também aplicável a habitação social.

No entanto, são necessárias medidas para vencer a inércia no aproveitamento destes programas, que deverão passar pela simplificação nos processos de candidatura e por um esforço em tornar o benefício económico do acesso a estes programas mais evidente. Para as camadas mais vulneráveis da população e com maior iliteracia são necessárias medidas adicionais, que se articulem com as autarquias no sentido da identificação das habitações a precisar de intervenção e a realização das respetivas requalificações. Carecem de particular atenção as zonas suburbanas, sobretudo, onde a qualidade da construção é fraca e onde habitam famílias com menores recursos.

3.3. POTENCIALIDADES PARA EXTENSÃO DO MODELO A OUTROS SETORES

Em face do atual contexto pandémico, e reconhecida a insuficiência da tarifa social para eliminar ou mitigar a pobreza energética, esta pode vir a agravar-se já no próximo inverno. Se esse momento vier a coincidir com um segundo surto da Covid-19 então as preocupações com a pobreza energética serão acrescidas.

Com a vigência do estado de emergência e posteriormente de calamidade, e a imposição de restrições de convivência social e profissional tornou-se ainda mais clara a importância do acesso universal a serviços essenciais como a energia, as águas, as comunicações e os transportes. Os efeitos da pandemia não afetam todos da mesma maneira, sendo de esperar que as franjas mais vulneráveis da população estejam mais expostas à doença como também às consequências da crise económica entretanto desencadeada. São sobretudo estas camadas da população que devem ser apoiadas no acesso aos serviços essenciais e as tarifas sociais são uma medida já testada para o fazer.

No caso dos serviços de abastecimento de água e de saneamento de águas residuais, por exemplo, o regulador recomenda a aplicação de tarifas sociais aos consumidores domésticos que reúnam praticamente as mesmas condições de elegibilidade exigidas para o acesso à tarifa social de eletricidade. No entanto, não só a tarifa social de água não está generalizada a todos os municípios/entidades gestoras, algo que resulta da estrutura organizacional do setor, com praticamente um operador monopolista por concelho, como a automaticidade não é ainda uma prática comum, Martins et al. (2020). Se há uma lição que se pode retirar do modelo da tarifa social da energia para as águas é que a aplicação deste tipo de tarifa deve ser generalizada a todo o território. Deve ser a condição socioeconómica do consumidor doméstico a conferir o direito a beneficiar da tarifa social e não uma decisão política, dependente do operador ou da gestão autárquica do respetivo município, sob pena de não se corrigirem situações de iniquidade.

No caso particular das telecomunicações, as necessidades criadas pelo isolamento social, o uso intensivo da internet devido às exigências do teletrabalho ou das atividades do ensino à distância expõem a importância de garantir a não exclusão do acesso a serviços de internet.

Também nos transportes, ainda que a existência de passes sociais já acomode preocupações de apoio ao pagamento destes serviços pelas camadas mais vulneráveis, é necessário garantir que o apoio também se verifique fora das grandes áreas metropolitanas.

Face o seu impacto sobre a acessibilidade económica, a tarifa social, tal como aplicada aos serviços de energia, é um instrumento cuja transposição se releva adequada a outros

setores essenciais. É importante não esquecer, neste âmbito, que algumas medidas de apoio extraordinário às famílias estavam associadas ao Estado de Emergência que deixou, entretanto, de vigorar. Além disso, mesmo com a suspensão dos cortes por não pagamento dos serviços decretada pelo governo, durante o estado de emergência e alargada para os meses seguintes, o certo é que depois as famílias têm de pagar, ou seja, a dívida não desaparece.

4. CONCLUSÃO

As tarifas sociais foram criadas na sequência da Crise Financeira Global de 2008-2009 com o propósito de apoiar o consumo de energia das famílias que estariam a suportar maior dificuldade de acesso a este bem essencial. Enquadram-se no conjunto das medidas de proteção a consumidores vulneráveis, concebidas e aplicadas no quadro da União Europeia. Este instrumento foi concebido para vigorar num horizonte de curto prazo, tem critérios definidos (rendimentos baixos e/ou determinadas pensões sociais, no modelo atual) e assume natureza financeira.

A abrangência expressiva da tarifa social de eletricidade e de gás natural em Portugal, em boa parte decorrente da introdução de um procedimento automático, conduziu a um aumento muito acentuado de beneficiários da tarifa social. Em comparação com o que acontecia quando a atribuição do benefício dependia da iniciativa do consumidor, a tarifa social passou a ser a atribuída a muitas mais famílias, tanto na energia elétrica como no gás natural. A automaticidade ajudou a superar falhas de informação e de iliteracia que introduziam inércia na generalização da aplicação da tarifa social a quem reunia condições para dela beneficiar.

Por outro lado, os nossos cálculos permitem concluir que o impacto da tarifa social de energia foi relevante na tentativa de garantir a acessibilidade económica aos serviços energéticos por parte das camadas mais vulneráveis da população. Porém, há ainda limitações a superar, tais como a ausência de apoio ao consumo de gás de botija e a diferença de critérios de elegibilidade para acesso à tarifa social de eletricidade e gás natural.

Neste contexto, os efeitos potenciais da tarifa social permitem formular a recomendação de medidas idênticas para outros setores essenciais, sobretudo nas comunicações, mas também nas águas e transportes. Uma tarifa social nas comunicações é importante para assegurar o acesso universal a serviços mínimos que permitam o ensino à distância ou o teletrabalho e reduzam a infoexclusão. Nos serviços de águas, é necessário garantir a generalização da tarifa social a todo o território e a sua automaticidade.

Esta conclusão é ainda mais premente no momento em que os efeitos de crise económica e social, decorrentes da pandemia Covid-19, se irão fazer sentir com grande intensidade junto de uma franja larga da população atingida pelo desemprego, pela diminuição de rendimentos e pela pobreza.

REFERÊNCIAS BIBLIOGRÁFICAS

- Assist2gether (s.d.) *Report on National and European Measures Addressing Vulnerable Consumers and Energy Poverty*, European Union.
- Brown, M.; Soni, A.; Lapsa, M.; Southworth, K. (2020) Low-income energy affordability: Conclusions from a literature review, Oak Ridge National Laboratory, US Department of Energy, March 2020, ORNL/TM-2019/1150.
- Comissão Europeia (2020) European Economic Forecast, Spring 2020, Institutional Paper 125, May, Luxembourg, Publications Office of the European Union.
- Conselho de Finanças Públicas–CFP (2020) *Perspetivas Económicas e Orçamentais 2020-2022*, Relatório n.º 7/2020, junho, Lisboa, Conselho de Finanças Públicas.
- Cook, J.; Shah, M. (2018) Reducing energy burden with solar: Colorado’s strategy and roadmap for States. National Renewable Energy Laboratory, NREL/TP-6A20-70965.
- Deller, D.; Waddams, C. (2015) Affordability of utilities’s services: extent, practice, policy, Research Paper 2: Affordability in the EU, Based on Eurostat and ITU Data, 22 October 2015, Center on Regulation in Europe (CERRE).
- Dobbins, A.; Nerini, F. F.; Pye, S. (2016) *Measures to Protect Vulnerable Consumers in the Energy Sector: An Assessment of Disconnection Safeguards, Social Tariffs and Financial Transfers*, INSIGHT-E Programme, Policy Report, December, European Commission.
- ERSE (2018) *Tarifas e Preços de Gás Natural para o Ano Gás 2018-2019*. ERSE, Lisboa.
- ERSE (2019) *Caracterização da Procura de Energia Elétrica em 2019*. ERSE, Lisboa.
- Farinha Rodrigues, C.; Nunes, F.; Vicente, J.; Escária, V. (2019) *Relatório Final do Estudo sobre “A Pobreza Energética em Portugal”* realizado pelo ISEG para a EDP. Lisboa, EDP.
- Fankhauser, S.; Tepic, S. (2007) Can poor consumers pay for energy and water? An affordability analysis for transition countries. *Energy Policy*, 35(2), 1038-1049.
- Fankhauser, S.; Rodionova, Y.; Falcetti, E. (2008). Utility payments in Ukraine: Affordability, subsidies and arrears. *Energy Policy*, 36, 4168-4177.
- Instituto Nacional de Estatística – INE (2019) *Rendimento e Condições de Vida 2019 (Dados Provisórios): Informação à Comunicação Social*, Lisboa, 26/11/2019.
- Martins R.; Quintal C.; Cruz, L; Barata E. (2016), Water affordability issues in developed countries: The relevance of micro approaches. *Utilities Policy*, 43, 117-123.
- Martins, R.; Quintal, C.; Antunes, M. (2019 a) Making ends meet: Actual versus potential joint affordability of utility services. *Utilities Policy*, 56, 120-126.
- Martins, R.; Pereira da Silva, P.; Antunes, M.; Fortunato, A. (2019 b) Estudo sobre a Aplicação da Tarifa Social de Energia em Portugal. Lisboa, ADENE.
- Martins, R.; Antunes, M.; Fortunato, A. (2020) Regulatory changes to the Portuguese social tariff regime: Carrying water in a sieve. *Utilities Policy*, 64(C).
- Miniati, R.; Scarpa, C.; Valbonesi, P. (2014) Fuel poverty and energy benefits: The Italian case. In Jon Strand (Ed.), *The Economics and Political Economy of Energy Subsidies*, Munich, CESifo Seminar Series.
- Pye, S.; Dobbins, A.; Baffert, C.; Brajković, J.; Grgurev, I.; Miglio, R.; Deane, P. (2015) *Energy Poverty and Vulnerable Consumers in the Energy Sector Across the EU: Analysis of Policies and Measures*, INSIGHT-E Programme, Policy Report. European Commission.
- Sumner, P.; Pridmore, A.; Duff, A. (2015) *Understanding affordability pressures in essential services*. Report, UK Regulators Network-UKRN.
- Winkler, H.; Simões, A. F.; Rovere, E. L.; Alam, M.; Rahman, A.; Mwakasonda, S. (2011) Access and Affordability of Electricity in Developing Countries. *World Development*, 39(6), 1037-1050.

AUTONOMOUS VEHICLES AND PUBLIC TRANSPORTATION

Veículos Autónomos e Transporte Público

Daniel Murta

Received for publication: June 15, 2020

Revision accepted for publication: April 9, 2021

ABSTRACT

Full automation Autonomous Vehicle technology (AV), is announced as possible, and safe, in the future. Widespread deployment would deliver sharp reductions in accidents and parking needs, two boons for society, particularly in cities. AV's role in total travel, hence congestion, energy and emissions is ambiguous, possibly negative. Relieving drivers from driving, better serving those unable to drive, will strain the system. Unless shared AVs reduce vehicle ownership (difficult, uncertain), increase sharing and do not tempt people away from Public Transportation – calling for planning, taxing and subsidizing. Otherwise, less cost, more convenience might double traffic.

Keywords: Autonomous Vehicles; carpooling; ecosystem; public transportation.

JEL Classification: L91; L92; L98; R41

RESUMO

A tecnologia de Veículos Autónomos (AV) é anunciada como possível e segura no futuro. A implantação generalizada proporcionaria reduções drásticas nos acidentes e nas necessidades de estacionamento, duas vantagens para a sociedade, principalmente nas cidades. O papel do AV no total que é viajado e, portanto, no congestionamento, energia e emissões é ambíguo, podendo ser negativo. Livrar os motoristas da guiar, servir (melhor) os que não guiam, vai sobrecarregar o sistema. A menos que os AVs partilhados reduzam a taxa de motorização (difícil, incerto), aumentem a partilha e não afastem as pessoas do transporte público – o que requer planeamento, tributação e subsídios. Caso contrário, com menos custo e mais conveniência, o tráfego pode duplicar.

Palavras-chave: Veículos autónomos; Mobilidade partilhada; Ecosistema de transporte; Transporte Público.

1. INTRODUCTION

Land based passenger transportation under 500 kms (a bit over 300 miles) is dominated today by the personal car. Vehicle registration was still growing in 2014 (Davis et al., 2016), and Vehicle Miles Travelled (VMT), in U.S., were also rising as recently as 2016.¹ That includes cities, where only in the biggest does Public Transportation (PT) enjoy any relevance, and even there as a clear non-leader (Button, 2014).²

However, the biggest players in technology and the cyberspace, Alphabet, Facebook, Amazon, and in the transport industry, Uber and most car manufacturers, envision a future with driverless vehicles. Amazon and other competitors in the commerce and distribution business hope to automate all road based freight transportation, which currently comprises all city freight traffic, and sizably leads from trains in intercity traffic land based transport.

AV technology, if widely adopted, promises to be disruptive in four ways: (a) dramatically reducing car crashes, avoiding most of the 1,25 Million traffic deaths each year,³ associated morbidity (injuries and disabilities), and economic toll; (b) vastly improving car use efficiency, promoting less car ownership in favor of ‘transport as a service’, possibly doing away with 90 per cent of cars; (c) consequently, reducing parking needs; and (d) displacing a lot of jobs, from professional drivers to the ‘crash economy’,⁴ along with those related to producing and servicing a much smaller number of vehicles.

This paper aims to analyze the broad opportunities and challenges for welfare from this seismic change, focusing on internal contradictions that may undermine progress.

The perspective will be one borrowing from Industrial Organization (consumers, markets, supply and regulation), Public Economics (welfare, externalities, fiscal instruments, health issues accidents and emissions), Urban Economics (land use and commuting patterns), and Transport Economics.

A number of broad assumptions will be put forward to frame the analysis:

(a) full AVs will only be allowed when they are safe; with safety being defined as better than the average human driver under all road, traffic and weather conditions;

(b) the ubiquity of the internet, web-enabled smartphones and cloud-based services will expand; the smartphone will continue to rival the personal car as a position good;

(c) online shopping will continue to rise, at the expense of the traditional, face-to-face sort; according to the ‘Economist’ magazine,⁵ “Over the past decade global e-commerce has been expanding at an average rate of 20% ... its share of total retail trade, at 8.5% worldwide, was still modest...”;

(d) people who walk, cycle or use PT, instead of the car, will not rise from secondary relevance or less, to any dominant role or share;

¹ Source: Federal Highway Administration; US Department of Transportation, retrieved at ‘<https://www.statista.com/statistics/185579/us-vehicle-miles-in-transit-since-1960/>’.

² Chapter 2, section 2.3, pages 30-32.

³ Source: World Health Organization, in ‘http://www.who.int/gho/road_safety/mortality/en/’.

⁴ Set of firms, organizations that profit/exercise activity in dealing with the human/material consequences of accidents, broadly in the Health, Auto and Insurance sectors.

⁵ Oct 26th, 2017, ‘E-commerce takes off’, special report, print edition.

(e) new forms of transportation, if they are to partially replace the current leader – the car – will have to provide not only clear added value in either time, convenience or cost, but also to do so in most, if not all circumstances – carrying kids, shopping, luggage, leisure, non-urban, longer range travel, bad weather, among other things.

(f) demographically, the best that can be predicted is for current trends to continue, with very low, low or lowering fertility rates in all continents bar Africa, and an ageing society, through better health care and nutrition, and declining smoking habits; and

(g) economically, stagnant real median incomes, across rich economies, should continue, due to the forces of job automation/replacement dampening workers' negotiating power.

The structure will work out as follows: section two is a review of literature; section three, a perspective of where transportation and society are now and where they would be in the next quarter century, if AV technology does not 'take over' – the baseline scenario; in the fourth, the nature of the possible, alternative AV paradigm is described, the players, management of and regulatory framework. A fifth section concludes.

This paper will seek to review the literature and contribute to the ongoing economic debate on the future of transportation, by (a) reflecting on an AV scenario for all of road based transportation, not only urban, and including freight, at least for the scale of a country; and (b) addressing the central tension between the need for a dynamic ride sharing (DRS) – essentially a taxi with multiple sharing clients – system to succeed in delivering better value to lure people out of car ownership, and not to steal too many people from mass transportation (PT), otherwise congestion persists.

2. LITERATURE REVIEW

Given that AV technology, is certain to be costly and, since it is not yet deployed, demand for it will have to start from zero, and be large enough for huge economies of scale to bring costs down, it is thought to have to rely mainly shared vehicles. Otherwise, the cost will be bearable only to the very rich, whose valuation of time and safety would demand such an upgrade (the personal AV). Since the current promise of AVs is quite recent, the first articles reviewed focus not on AV, but on car sharing and pooling. Literature uses the word 'sharing' in the usual sense, but the expression 'Car Sharing' usually means a client, individual or a family, hiring a car for a period, during which it is its sole user; the car is 'shared' in the sense that it serves several customers, but one at a time. 'Pooling', refers to a practice where participants take turns, each bringing his vehicle to serve him and the group (pool). Both Pooling and modern DRS (explained above) expect a car to serve several users at the same time.

Correia and Viegas (2011) ran a Web-based survey to assess personal disposition to Carpool, around the Lisbon Metropolitan Area (4 Million people live in and around Portugal's capital city, Lisbon). Carpooling has been associated with employer-based initiatives, especially during the 2nd World war, and after the Oil shocks in the 70s', but has tailed off to insignificance, with the return to normal, affordable fuel prices. The results found respondents were, in general, not keen to join a car pooling club; those with children not at all; money savings, including in parking, were the main reason in favor, so poorer people declined less; sharing with both colleagues or strangers was found negative.

Chan and Shaheen (2012)'s work goes through past, present and future of ridesharing in the U.S., stressing very low market share: schemes lacked of critical mass (enough users) or profitability (private ones) Only the more stable patterned commuters afford the loss of flexibility. Safety and privacy concerns are important, so closed systems inside companies or campuses have met less resistance. Large employers, public or private, can succeed.

Vij et al. (2013)'s paper uses a survey to estimate latent modal preferences. Cluster analysis reveals three portraits of users: the 'auto-male', mostly uses his car, is employed and is available only for short walking distances; the 'time-sensitive multi-modals', frequently female, not always employed, probably take lifts from partner/husband, but use other modes if sufficiently convenient; 'time-insensitive multi-modals', are employed and buy a [PT] monthly pass. The article dwells on the relation between everyday behavior and long-term choices like owning a car, regularly buying a pass, stressing the potential for better targeting of public incentives.

Jorge and Correia (2013)'s work reviews literature on car sharing systems, also reporting very low market share. Up-takers are mostly young, urban people, cost and/or environmentally conscious. The logistical and economic problems from managing stocks from the different locations, with the minimum fleet to provide good service (available car), and the minimum human resources to move (empty) cars around, are taxing. Availability is further linked in a two-way relation with the number of rides (demand), leaving a vicious cycle all too probable, if critical mass isn't achieved. Authors raise an important issue, carpooling has severe sociological barriers, to do with strangers, safety and security, and a crucial question: regarding its role for the environment [emissions, efficiency, energy use], does car sharing steal from car owners or PT?

Green et al. (2017)'s article peeks at younger generations' minds about cars, by interviewing 16- to 21-year individuals from small towns in Britain. Their PT system offers poor quality, in comfort, frequency and depth of destinations. But the car is expensive to acquire and keep; accidents do happen and hurt, enduringly; car licenses are expensive, hard to obtain and can be lost, through the point-based system. Sharing is sometimes obligatory, or a way to cut costs. It is a generation less enthralled with the car and driving per se.

The rest of the review deals with AV proper.

Thierer and Hagemann (2014)'s paper focuses on Intelligent vehicles and Driverless Cars, to argue for innovation to be let to flourish unhindered, for the potential huge benefits to come to fruition. They rank less accidents and congestion, more fuel economy, fewer parking needs, mobility to non-drivers and greater convenience and productivity from time freed up from driving as the main 'promises'. But warn about the risk that widespread adoption might not be achieved, if regulation, too restrictive, too fragmented or contradictory stands in the way. They present NHTSA five level categorization of vehicle control automation:⁶ 0, is for no automation, 1 is for function-specific automation, one or several non-combined tasks, 2 is combined-function, when at least two primary control functions are jointly automated, e.g. adaptive cruise control and lane centering; 3 is for limited self-driving, where

⁶ U.S. Department of Transportation Releases Policy on Automated Vehicle Development, NATIONAL HIGHWAY TRAFFIC SAFETY ADMIN. (May 30th, 2013), URL: www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+Department+of+Transportation+Releases+Policy+on+Automated+Vehicle+Development.

full control may be surrendered, but return to driving is both optional and required in less than normal situations, 4 full self-driving automation dispenses with all driving controls.

Anderson et al. (2016)'s groundbreaking book, first published in 2014, covers AVs' promises and perils. Crashes are 90% human-related. Mobility for those unable/unwilling to drive, poorly served by mass transit PT, taxing relatives/friends, or paratransit agencies for incapacitated, inefficient in their inherent lack of scale. Land use, both by longer commutes becoming acceptable without driving, and vast savings on parking, which they cite as taking up to 31% of Central Business District's space. Effects on energy and emissions, from smoother riding, reducing distance, platooning,⁷ less crashing, search for parking. Lighter vehicles if safety features are allowed to become less stringent (again, because of rarer accidents). On the other hand, VMT can increase, from the discussed gains in in-vehicle time, empty travel repositioning of shared/pooled vehicles, and extra travel from those unserved segments, also including too young or old or ill to drive. VMT can further increase for those finding AVs (shared or pooled) cheaper or more convenient than current forms of car use, and/or affordable enough relative to PT. More VMT aggravates congestion. While finding, on balance, that benefits can outweigh costs, the authors call for more research on both, and on how they're distributed between individual user/operator and society.

The need to avert more VMT and the existential threats to PT are central this article.

Burns et al. (2012)'s work sets out to generate synthetic trips for three distinct scenarios: a mid-sized American city, a low-density suburb, and Manhattan, NY. The vehicles are shared, centrally dispatched, but with no carpooling (DRS); personal cars remain. They obtain less costs for mile in all cities, lowest in the first; also, significant reduction in the number of vehicles, higher in the first two. Smaller, purpose built, shared AVs are even cheaper to run. Manhattan's simulation kept mass PT.

Similarly, Fagnant and Kockelman (2014)'s article models how a fleet of Shared AVs (SAVs) would perform in Austin, Texas. Again, no DRS is planned. Following a service (including pick-up), AVs would reposition in places that balanced low parking costs and faster availability. They model just a 3,5% replacement of current trips by conventional cars, resulting in each SAV replacing 12 cars, serving 31-41 persons a day, with an average wait of 20 seconds (more cars could be replaced, but with higher waiting times). On the other hand, 11% more travelled distance takes place, due to repositioning, albeit still with sizable reductions in emissions. They tie some of the gains to the transport system from lower costs per mile, on them not generating excessive extra travel – again this central topic.

Fagnant and Kockelman (2015) compute numbers on a national (American) scale. It focuses on crashes, and their costs, both narrowly and broadly defined. Citing (Hayes 2011), saying fatality rates (per person-mile) could drop to those of aviation and railways, or 1% of current ones, and FHWA's⁸ estimate of 25% of congestion being caused by traffic incidents, half of which from crashes, in Cambridge Systematics 2005's report. Economic costs from accidents in the U.S. are 277 billion USD, double those attributable to congestion. The authors agree with previous research that AVs would drastically reduce parking needs, but

⁷ Platooning is an experimental technique that uses technology-enabled high coordination within a convoy of trucks, to have them travel as close as possible to each other, saving fuel with the reduction of drag.

⁸ U.S. Federal Highway Administration.

would increase mileage, mainly due to previously excluded users. They calculate the effects from a market penetration 90% of AV enabled vehicles, of which 10% would be shared (not DRS), each displacing ten standard vehicles. VMT would go up by 10%. Problems facing (mass) deployment of AVs are discussed: security, real and perceived; the threat of hacking; privacy, since AVs generate comprehensive information. This also brings new opportunities to nudge road use and behavior, charging more in congested hours and routes. The savings for society, from less cars, congestion, parking and accidents are significantly higher than those for the individual. Particularly so if ownership isn't curtailed, and early small-scale adoption doesn't drive costs down by much. Some of the congestion-saving improvements will stem from vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication, the first being borne by AV owners.

Spieser et al. (2014)'s paper, take a congested big city-state, Singapore, and design a model to predict the performance of a system of AVs that would substitute all modes of travel currently in use, whilst assuming no DRS. The basic scenario yields a need for just 1/3 of the current number of vehicles. A comparison of total cost of ownership between personal car and AV is included, and estimates for total cost of time. Along with the financial costs (where fuel is just 6%), these form the Total Cost of Mobility (TMC). They conclude the fleet of shared AVs (SAVs) would come much cheaper than personal car travel: costs in time would drop by a third, costs from using/owning the vehicle by two thirds, and global TMC by almost half (-46,4%). VMT would increase, mainly because of empty travel.

Luis Martínez's report (ITF, 2015) carries out a similar exercise for a mid-sized, mildly congested capital – Lisbon. Drawing on data from surveys and PT operators, they too construct a synthetic model of the city's transport needs, which they proceed to experiment with total discarding of cars, buses and (human operated) taxis. Crucially, they keep the metro PT system in the main scenario. They put forth what they call Autobots, AV taxis called by app., not shared with different users; and Taxibots, with DRS but severely limited on the number of people, waiting time, total additional time (waiting plus travel time) and extra distance. Results are presented according to 24-hour weekday averages or morning peak. They do not attempt to model additional demand that could arise from more convenience or smaller cost. The model, with Taxibots (DRS) and metro, predicts that 90% of the vehicles could be out, while fulfilling the same needs. Those percentages naturally drop with less pooling, less cars' replacement, or if the metro is removed. Travel volume, with these services and the need for empty car repositioning, rises in all scenarios – in the “greener” (no cars, pooling and metro) only 6,4% over a 24h weekday, and 8,8% at the morning peak. These increases point to congestion remaining. Authors conclude this simulation yields vast improvements in vehicle numbers and parking needs, but not in VMT or congestion; they find it crucial for “impacts of additional demand to be mitigated”, through fiscal nudging. Borrowing from predictions above on the impact of AVs in accident reduction, their impact on congestion and on the potential for dwindling parking needs freeing space for extra lane capacity, which the authors do not do, it is possible to accept congestion improving or at least not worsening.

A follow up report (ITF, 2017), retains most of the premises – no cars, taxis or buses, no tourists/travelers from afar – and extends the scope of the analysis to the Greater Lisbon (suburban) area. Mass Transit PT here comprises the metro system, coupled now with

suburban rail. The AV actors are slightly tweaked – there are smaller Shared Taxis and larger Taxi-Bus – both accepting DRS. Their role of feeding to and from Mass Transit transport systems is hugely expanded, reducing congestion, time and cost, improving convenience and efficiency. The running of the two AV fleets would be done centrally. The main scenario would result, again, in significant reductions to fleet size and parking needs. Here though, it would also increase the use of Mass Transit PT and end congestion, reducing VMT in the roads/streets, with superior performance, in terms of time and cost. Plausible paths towards gradual adoption were modeled, starting with car users who either pay for parking or for the monthly pass, along with all bus users and some of the taxi users. Earlier market penetration stages deliver much smaller benefits to the system, but successful demonstration can be key to future willingness to switch. Constraining parking, both supply and maximum allowed time, is a way to nudge users towards leaving the car. Another path would be progressively banning suburban residents from bringing cars to the city (thus discriminating them from city residents); it was shown it could work. SAVs were modelled to accept requests for exclusive rides, which hurts efficiency. Multiple dispatchers were simulated, according to vertical market segmentation, also found to impair on performance. Recognizing the exclusive rides' (taxi) effect on an AV-only system, as similar to that of a personal car, they conclude their role is negative, but manageable if kept expensive and small.

Wadud et al. (2016)'s work, on the travel, energy and emissions' impacts from AVs, assumes they will substitute cars and taxis, increasing their use and, without the driver's role, possibly evolve to new sizes, more in line with actual occupation, in turn becoming more efficient. Migration from the Public Transport towards AVs is not covered, nor is DRS (pooling). Under various scenarios, significant savings in energy and emissions are found to be achievable, with cuts to nearly half of the base case. Partial automation can deliver sizable gains. Full automation, though, is more vulnerable to potential downside risks, namely two that have to do with the valuation of time. If significant traits of the population come to value it highly, they are likely to demand more performance (including average speed) which steeply rises energy consumption and emissions (one would add, greater risks for accidents, and less tolerance for waiting, in Pooling). The other risk, of opposite nature, is how invehicle time is valued in AVs – if, through comfort, entertainment or sleep, it is viewed as less costly, greater distances will become tolerable, leading to longer commuting, again increasing total VMT. Time savings (congestion, parking, accidents) reduce one of the largest components of generalized cost. Pricing can play a part – currently, car owners are faced with a high fixed/low marginal cost that stimulates driving. Future pricing models presenting consumers with a higher marginal cost could prevent growth in VMT. This holds for exclusive rides in SAVs and for personal cars, where demand management must be considered. Authors conclude warning overall effects may be positive or negative, meriting further study, planning and policy.

In another article, Wadud (2017) focuses on the costs in personal transportation and commercial freight transportation – operators, to whom driving is just a constraint (legal limits for professional drivers) and a cost (overall labor costs and those attributable to human error related accidents). From these, factors influencing early adoption of AVs are discussed: the sector with the strongest case for automation is freight; next, they predict the most affluent and/or intensive car users will be the first to switch, in line with the previous

focus on the valuation of time – both total and in-vehicle. Regarding partial, early adoption, differences are pointed out, between the ‘everything somewhere’ model whereby, in a limited spatial scale, probably urban, AVs are allowed, but not outside it, which limits freight use to urban small freight, and leaves people having to switch vehicles, probably keeping their non AVs; and the ‘something everywhere’ model, typically level 3 automation, where driving is surrendered in motorways or in good weather, but not always, everywhere, with no driver redundancy for freight operators, nor for personal transportation, either shared or private. Both developments are likely to precede full AVs, which is relevant for planning.

Krueger et al. (2016)’s paper, explores future preferences for AVs, by conducting a survey on willingness to take part in DRS (pooling) and SAVs. Main factors were found to be travel cost, time and waiting time. Cluster analysis yielded four classes: a quarter was composed of exclusive car users; another also had car users, but available for complementary walks; next were cars users, pliant to walking and to (mass) PT, but not to cycling; the final quarter was evenly split, between car users that used PT, but neither walked nor cycled, and people willing to use the full range of modal choices. In line with research above, the probability of switching to either SAVs or DRS, was found to be higher in current ‘multi-modals’ and, generally, among younger people.

Litman (2018)’s report could be mistakenly thought of as contrarian to AV technology adoption, but it does deliver a dispassionate, balanced analysis, as well as sobering facts and insightful views on the interests behind the AV ‘revolution’. Previous vehicle technologies have taken a long time to adopt, not only because reliability had to be achieved and cost had to drop, but because cars are replaced slowly, they last long. Typically, they are the most important household purchase, after housing. The parallels made by AV enthusiasts, with the rapid world adherence to smartphones and tablets, ignores they cannot hurt or kill, neither owners nor people around them – cars can and do. The author identifies three main actors with financial interests in the (AV) industry, as being behind most of the optimism: investors in AV technology; candidates to become dispatchers of AVs and suppliers of AV enabling equipment. Litman also points to the likelihood of additional travel being induced, increasing congestion. Relying on public infrastructure, AVs are bound to require costly upgrades, especially to do with communication. Opposing preferences between individuals and society may arise, regarding speed and safety, the latter in terms of own safety vs fellow citizens’, urging public choices. While the AV experience can reduce stress, improve mobility and, thus, productivity, there are several caveats: to shared vehicles, cleaning and vandalism costs will exist; minimizing them will involve sacrificing comfort in favor of “hardened” interiors, or privacy by using cameras, or both; sharing also carries security fears and risks; the success of sharing may reduce support and funding to PT, which may result both in less options and social exclusion, since PT is typically the cheapest mode. So long as human-driven vehicles coexist with SAVs, benefits such as HOV lanes will present a loss to the non-AVs. The ease with which people relax and work in trains cannot be taken for granted in urban transit, where stop-go-turn is more frequent than not – when the novelty expires, the experience may be seen as “more like an elevator than a spaceship”. There are potential harms to people not using the technology: if faster and travelling closer together, AVs can leave less room for pedestrians and cyclists; if slower, for more smoothness or to deal with intensive or trickier situations, they can slow

traffic as a whole. Costs for purchase and maintenance of these sophisticated platforms, are bound to be high, dampening deployment, but for the most affluent. Ownership has several motives, including prestige, convenience, particularly in less than optimal conditions, comfort (not sharing, leaving your objects, including work-related), that have made sharing a limited option, both in range and appeal. Taxis, buses and freight are heavily weighted down by driver costs, thus being likelier early adopters. Policies will be central, requiring fairness and acceptability, to guide AV deployment. A compromise will have to be struck between being cheap and convenient enough to win a weighty part of car users, to lower costs through scale, and not hurting PT. In turn, PT will require funding in order to stay cheap (or to become cheaper). Delivering superior performance, through higher speeds, raises accident risks, energy consumption, and emissions.

3. BASELINE SCENARIO (NO AV)

Starting with freight, since it is less complex: medium to long range inter-city road-based freight is central in Europe and Asia, while very important and increasingly so in the U.S. It uses more energy and emits more per ton. than rail; it adds to congestion directly, by the amount of space it takes, but also indirectly, because trucks move more slowly and clumsily, and disrupt more whenever they are involved in accidents. They will continue to grow with the economy, despite the European Union's efforts to promote more rail; and the competitive threat in some (longer) distances, weights and volumes (smaller) from air freight Assumptions 7 and 8, on low fertility, ageing society and stagnant real incomes point towards feeble GDP growth, so road freight's growth will likely be moderate. Short-haul, city freight transportation is important now, contributes to congestion, albeit less than personal transportation, because it is more evenly distributed across the day, and some can be rescheduled to off-hours (the same holds for road-carried freight, but there, its impact is larger on the roads). Its growth, present and into the future, is fueled by three converging trends:

Humanity is moving to cities, now host to more than half, more so in the developed world, and since they provide better jobs, income, and quality of life, the trend will continue;

Online shopping creates billions of parcels to deliver, and expands total commerce; Affluence and social networking spread 'fads' and the consumer society will keep consuming ever more, both online and through classic channels.

Looking into a future without AV, one can see freight as weighing (down) more on the transport system's performance in urban traffic. Roads, nonetheless, are strained by trucks, forcing people to commute ever earlier.

In personal transportation, the car leads everywhere, for every function, on the road. For leisure, air travel is important. In regional/rural settings, PT has either withered, under the downward spiral of less demand, less frequency, less comfort, or ceased altogether. Some coaches remain relevant connecting cities using motorways, as do trains, but mostly appealing to the minorities that use PT regularly (also in the cities). Society, some say, has reached

an era of “peak car” use, perhaps by over-estimating the numbers of the environmentally conscious, but certainly helped by the convergence of two trends:

Society’s ageing means people move less; some are unable/unwilling to drive; most tend to have more fears, which makes ‘mixing in’ in PT more daunting; their vulnerabilities outside the home are both real and perceived as such;

Younger people, almost always online, are going out less; stagnant incomes, and persistent unemployment, particularly in Europe, reinforce savings to be had by staying in, thus reducing transport needs.

The lure of the car is much diminished for younger generations (see Green et al., 2017), who see it as expensive, dangerous, ineffective in dense/congested metropolis with good Public Transportation. So, the car may have reached a (very high) plateau in market share, from which it may not grow, or may gently descend. It still enjoys a clear leadership, solidly founded on a host of strengths that make it the best ‘all-rounder’, leaving alternatives “very good for sunny days and... other people”.

Public transportation will continue to be relevant, the bigger the cities and either the wealthier or the more congested. The first two because they can host a better PT system, having the scale to have a dense (convenient) network, and the wealth to build it. In congested cities, because Public Transport becomes the smart option, having always been the cheapest (although congestion is a sign of intense car use). Especially where housing is expensive (relative to incomes), a large number of people find they can’t afford a premium mode.

Taxis, and sharing stand, in vertical differentiation terms, between car ownership and PT, will continue to enjoy some growth, since the internet, and apps. make them easier to book, organize and plan. Uber’s and its rivals’ foray into Taxi’s market power will carry on delivering lower prices, bringing some additional demand. Sharing schemes, although more intelligent, and ‘real time’, will struggle to evolve from their (current) irrelevance, because comfort, privacy, and security will still favor the default option – the car.

Walking and cycling will keep their constructive role on an efficient transport system, alleviating congestion, emissions, energy and land use on the shortest distances, as feeders to PT, and complements to the car. Some people’s view on their fitness, quality of life, health, ‘carbon footprint’/sustainability, or even on saving money, will continue to exist. But because congestion and pollution are externalities (no direct link between individual choice and outcome), and time, safety and convenience, including in bad (weather and other) conditions, favor personal transportation, their number will continue a minority.

4. TOWARDS AN AUTONOMOUS VEHICLE ECOSYSTEM

Much of what has been presented and commented in the literature review, and on the baseline scenario, suggests what an AV enabled transport system should look like. To discuss and help plan for AVs, a leap into the future is necessary, with all the fragilities and uncertainties thereof. The essence of an AV ecosystem is the deep interconnectedness of all modes, and their performance, both to their customers and the system. Like

in the previous section, in trying to work from the simple to the complex, first, freight will be addressed.

4.1. FREIGHT

Freight companies, either dedicated or those having substantial volumes to carry, like Amazon and its Asian competitors, Walmart and its rivals, face huge pressure to deliver unprecedented amounts of freight, and number of parcels, at the lowest cost, and the shortest time. Several ‘offer’ free delivery and can’t charge much for ‘premium’/‘express’ delivery. Therefore, they are placing high hopes on two technologies: AVs and drones. Both dispense with drivers. Freight companies and those related to big commerce, will surely be early adopters of AV technology, several having announced it, and currently spending to that effect. Their path will interact with the rest of the system in four ways:

Spreading more of the work around the day, cutting accidents, they’ll contribute somewhat to de-congestion;

Oiling the growth of online shopping (that includes delivery), they’ll reduce people’s transport needs, and a major reason to own a car (shopping);

By achieving sizable economies of scale, and scope, they’ll help drive down the cost of delivery to a point where personal transport as a service, SAVs or DRS, will offer, cheaply, the extra of luggage, shopping and even work related tools, to be handled separately by freight companies, arriving ‘just in time’; Litman (2018) cited this point (work related tools) as an important reason for people to cling to the car;

They need AV technology to spread, to be deployed massively, to make a mark in the (bigger) market for personal transportation, to drive down the technology’s costs.

Common to both freight and personal transportation is the objection of AVs, being driverless, leaving people bereft of help in and out (very young or old, disabled, etc.), and cargo loading and unloading. But during all journeys, the helping/loading isn’t taking place. It can only be vastly more efficient to have people at one or both ends of journeys, performing those tasks. The problem is real, the solution just doesn’t have to travel inside the vehicle.

4.2. PASSENGER TRANSPORTATION

Most authors agree AV’s highest hurdle is for a strong demand to emerge for services as yet nonexistent, enabling, through scale, the costs of a complex, developing technology to come down, and become competitive. Beyond freight, which also needs this to happen, there are the richest of citizens, who may want a driverless car, whatever the costs. Uber and its competitors will want to switch to AV because it’s cheaper. This article argues these three segments, ‘masters of freight’, ‘plutocrats’, ‘modern-day’ Taxis will not bring enough scale for AV to deliver its full promise – better, safer and cheaper.

Widespread adoption of AVs promises two valuable boons to the transport system:

Sharply reduce accidents, because 90% of them are caused by human-related mishaps, and all reckless, unlawful behavior will be programmed away;
Reduce parking needs, either on shared fleets or private cars – both AVs can find their ways to efficient, cheaper places to park and fleets can be either too busy to need much parking, or can find it worthwhile to cruise while waiting for the next service.

Both these outputs contribute to de-congestion: the first was estimated to cause a quarter of it; the second, would make traffic flow smoother and free urban space for possible extra lanes. But for the technology to take off, it must reduce own-car use – they are the leaders, the majority of vehicles, and run with an average of 1,2 people for their (usual) five seats; while not increasing VMT, since that would aggravate congestion, energy demands, and emissions.

ITF's reports (ITF, 2015) and (ITF, 2017), show what an AV ecosystem must have to be successful: a SAV and DRS supply that wins over most of car users (car ownership, even in an AV era, must be presumed to continue, at least for the wealthiest); a mass PT that continues to de-congest, is fed by SAVs and DRS, and is not unduly encroached by their convenience and cost.

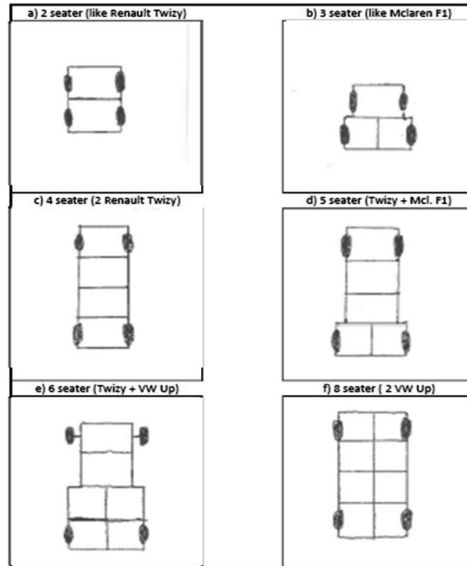
The key to commercial success for any public transport company is summed up in the phrase “keep them packed and busy” – using the fleets as intensively as possible, and with the highest occupancy rate.⁹ In scheduled transport, that increases frequency, lowering waiting and total time. Private car sharing and pooling schemes have floundered because of insufficient demand, as has Public Transportation in regional/rural settings.

Dynamic ride sharing, or pooling, as modelled in the ITF reports, where one or two people call for a vehicle, that then stops at one or two extra locations, to pick up two or three additional passengers, and drives them to their final locations, without the stress of driving or parking (and the latter's cost) has crucial strengths. It discourages car ownership and de-congests traffic, as long as it doesn't empty PT. But car users don't want to share or pool. They feel good and familiar with a lot of features their car provides. So, next is a sketch of the desirable features for a good, promising DRS service, to be followed by how those would meet car owners' demands.

A driverless public vehicle, will have four wheels, probably with independent electric or fuel-cell-powered drivetrains and, thus, all visible space will be for passengers; that will allow between one to four rows of seats. Besides enjoying a discount/rebate for allowing the vehicle itself to be shared, raising somewhat waiting time, journey length and total time (within boundaries), the norm will be to have one seat per row/door; and all the layouts with two seater rows are to attend people who are together – if empty, they can take a solo passenger. This way, strangers don't share the same row, ever (comfort/safety). Also, it's safer to exit on the right, so with a single seat-row, no one is forced to step outside (rain, wind, cold?), for the person on their left to exit. Figure 1, illustrates possible seat configurations, under the above restrictions, and placing single seat rows first, since they spoil less of the view.

⁹ On occupancy, taxis are an exception, since in their classical pricing model, extra passengers don't pay more.

Figure 1: Six possibly more popular layouts



Darwinian experimentation, with real demand, will determine which of the above, or other, configurations will prove most in demand. As far as low waiting time, few detours and extra miles (all with compensation), vertically differentiation, should dictate the menus the market will accommodate. Never sharing a seat row with strangers addresses multiple objections car users usually pose: security, not mixing with annoying or violent passengers; taking kids, seen as incompatible with pooling; bad weather situations. Naturally, the more pooling, miles and minutes one tolerates, the less one is charged. The most stringent of customers (no sharing), do not de-congest. Hence, they must be priced so as not to become the norm. Freight companies will complement all but the lightest and smallest of shopping/luggage with cheap, on time, delivery. So, how can a such a DRS centered and differentiated scheme measure up to personal (car) transportation?

Personal cars, being the incumbent leader, are the default option. It serves all kinds of travel (making assignment of costs to each kind beyond the effort of most users); its high fixed cost is often not questioned (assignment, again), and the manageably low marginal cost spurs intensive use; of the less immediate or visible motives, the pleasure, and power and privacy around driving are bolted into one's life experience; it is a back-up instrument for emergencies, such as hospital/maternity, or evacuation from natural or other disasters. Contrary to ITF's simulations and (Spieser et al., 2014)'s in Singapore, and others,¹⁰

¹⁰ All studies project the demise of the personal car as an exercise on the effects to congestion, and the gains from an AV alternative, not as a prediction.

personal cars needn't be totally removed from the system; non shared/pooled public AVs, running empty miles during repositioning, can be just as congesting, and the wealthiest will continue to own AVs. They will, however, pay their way: in parking, both at home and at work, for ownership, at purchase and regularly, and for the use, more if at congested hours or routes. The bulk of the (less affluent) car users, though, ought to be tempted by personal transportation as a service propositions, for regular, practical journeys, as long as it is clearly cheaper, more convenient.

It is being assumed that AVs can go anywhere, not being confined to cities, even if not from the outset. Otherwise, a changeover, somewhere, would be needed – a losing proposition, in replacing the car. Also, at some point of their successful deployment, human driving will be barred from public roads (most authors agree), since they will come to be seen as posing unwarranted risks to others; when AVs are safe at all times, deaths and injuries from human error become unacceptable. As far as driving pleasure, there are two kinds to consider:

The one from a sightseeing journey, where not driving allows all occupants to fully appreciate the scenery and interact;

The one from sports driving, with all the extra risks, more serious to others,¹¹ is already forbidden, but currently takes place (in hiding from law enforcement); several entities profit – sports car are costlier, consume more fuel, and governments tax both; AVs force society to face there's no room for sports driving in public roads; at the same time, they open up a large opportunity for a market to provide tracks of all kinds.

So, the main actors, freight companies or those handling a lot of freight, AV enabling suppliers, tech giants vying for the dispatching role, fleet owning companies, possibly in alliance or owned by car manufacturers, will push for AV to succeed, and that will have to come at some expense of the personal car. Car companies may worry about simulations where 90% of all vehicles become useless, but there are countervailing factors:

AVs will be more complex to produce and service, so can be more profitable;

Some personal car ownership will remain, with infinite scope for vertical differentiation, including amenities;

The continuum of vertical differentiation inside sharing and DRS will, in the upper segments, dent the savings on vehicle numbers, because bigger fleets are necessary to cater for demanding customers (less tolerant of waiting, sharing or mid-stops);

Shared, DRS vehicles will be used so intensively, they will be serviced and replaced more frequently, helping various revenue streams;

All the extra convenience, and lower cost, that will be put into ensuring the success of the SAV, DRS model will generate extra demand, adding to that of previously non-served or poorly served publics – young, old, disabled, etc.

¹¹ Both because they are 'innocent' to the cause, and don't share the pleasure.

4.3. PRICING AND EXTERNALITIES

The challenge to attract as much demand as possible for a new service, by sophisticated, profit-maximizing companies has best and most recently been seen in telecoms and all the new services, generations of devices and protocols (3G, 4G, etc.). Those markets present companies with direct competition and very powerful lock-in, network effects, that grant market power if a loyal customer base to your company grows to become large. It is clear management of the AV traffic and dispatch of SAVs will have to be centrally run by a sole company or agency, at least in the scale of a relevant market – it may be smaller than a whole country, but not as small as any small city. The private company or companies offering transport as a service are to be expected, in trying to win over as many costumers as possible, and having them use the system intensively, to offer ‘all you can eat menus’ – subscription type pay-structures – or, in PT terms, a kind of monthly pass. It has happened in telecoms, with voice calls, SMS, and internet traffic, and it has worked, in expanding the various markets from non-existence.

This practice is particularly well suited to counter personal transportation, because a subscription puts the customer facing zero marginal cost, spurring consumption. Car owners, who have to be won over, also face low marginal costs to travel. But there are two, connected perils surrounding this possible development that merit continuing evaluation, and possible counter-acting intervention:

Overusing transport services aggravates negative externalities in emissions and congestion, unlike in telecoms, where (over)consumption, the number of hours spent on the phone, tablet, etc. does not clog traffic, or leave a higher carbon footprint.¹²

Threatening PT – efforts put into making a foray into personal transportation, time, privacy, convenience and cost, risk tempting people away from public transportation; this is a major problem with AV mass deployment, most authors agree; the problem is not one of nostalgia or pity for a suffering PT – it is that mass transit PT de-congests, whilst cars, AV or not, do not, unless in the most efficient, tolerant of pooling, segments.

Time is one of the top elements on modal choice and evaluation – this positive/negative externality ties AVs to PT inextricably – if SAVs and personally owned AVs steal too much custom from PT, two effects are clear: VMT rises substantially, worsening congestion, hence part of the promised improvement; and PT loses critical mass, hurting frequency (thus convenience) and profitability, possibly survival – it has happened in regional traffic.

What can be done? It depends on the scale of the transport system, and on how many layers it has: in a village, there may not exist any PT, so competition from DRS to personal AVs is welcome, bringing more efficiency to the system. Possibly, configurations of vehicles emerge with room for luggage, work-related tools, even shopping, if it is not economical for freight companies to serve such markets. Overconsumption should only be dealt with (taxing, managing demand), if and when the whole country chooses to. The principle is

¹² Although the server farms, that power the net and the ‘clouds’, are becoming ever more energy demanding.

VMT reduction should cost/hurt everyone the same, everywhere, unless too much congestion makes it urgent.

In a small town, there won't be metro, but there could be buses, and/or rail connecting to a bigger city. For buses to remain attractive, at least in the main routes which offer more frequency, two complementary approaches must be in place:

Keeping PT cheap, maintain pricing structures ranging from single tickets, quantity discount for multiple tickets, various discounts for identifiable publics, and monthly passes, in essence, the flexibility to cater for as many transport needs as possible, while ensuring a significant discount over the price from SAV options; if the marginal price in SAVs is zero, through subscription, zero may have to be the cost in PT (Germany has recently pondered such a move, in dealing with crippling congestion and pollution);¹³ since cars and car use (including fuels and parking) have always been taxed, there is scope for cross-subsidization;

Monitoring, regulating, at times curtailing, the aggressiveness of SAV companies' drive for extra custom; subscription-based models usually include a caveat that consuming beyond a threshold costs more; regulators can intervene on these boundaries, to force people to face marginal social cost pricing, so as to manage demand and over-consumption.

The commuting PT, though typically only alleviating congestion in the bigger city it connects to, should likewise be fiscally protected, with SAVs acting as feeders to commuting PT, and seamless changeover.

In a large city, with all the possible layers, from personally owned AVs, SAVs with segmented DRS, buses, metro, pedestrians and cyclists, SAVs continuing experimentation with how much people accept, in pooling (waits, detours, extra time), vehicle configurations, in terms of capacity, may endanger buses. Or leave just a fraction of successful lines – it has happened when metro systems are introduced.¹⁴ But the central tension and 'battleground' is between the Metro and SAVs, with the implications referred above. Huge, valuable, scarce space will be freed by significantly reduced parking needs from a SAV-dominated, public street using system. Multiple candidates for that space uses exist: extra lanes, if congestion persists, cyclist lanes, wider sidewalks; for outdoor parking lots, public parks; or anything else, including sale, for in-door parks. Some AV private cars will remain, but will have to arrange for parking, both at residence and work. Street parking (as opposed to stopping for drop-offs) may disappear. Private cars will be confined to the very rich few, so upmarket malls may offer parking, to cater for these lucrative customers. For large, cultural or sports events, metro and buses are the most efficient candidates. If the latter are no longer around, SAV companies will use the maximum capacity vehicles they have. Tourists and other infrequent visitors to the city (on business, for example), probably skip mass transit PT systems, with their fixed routes and timetables, but use SAVs. Group organized tourism, will be able to

¹³ The Guardian, February ,14th, 2018, URL: www.theguardian.com/world/2018/feb/14/german-cities-to-trial-free-public-transport-to-cut-pollution.

¹⁴ Oporto, Portugal's second city, has a recent Metro system that led the supply of Buses to contract.

save considerably, by planning with SAV companies, the most efficient set of vehicles, routes and hours that set of predictable, shared journeys need.

4.4. POLICY IMPLICATIONS

Widespread deployment of AVs will come about through piecemeal erosion of personal car use, in favor of DRS alternatives. A set of policies to nudge people in that direction include:

Creating a standard measure of transport inefficiency, such as the number of hours lost to congestion,¹⁵ to breed consciousness on the problem and support for policies to improve; that index should be publicly visible and object of debate; then, goals could be set in performance gains from alleviating congestion, quantifying how change from less cars, more people in each car, DRS, PT, walking and cycling – would achieve; similarly, an index on road mortality and morbidity could help in winning acceptance for limits on human driving;

Giving steep discounts, or temporary gratuities in public transportation to youngsters/students, inviting them to try out the PT experience and away from car culture; giving away some money's worth of PT or SAV travel, to all household members, whenever their fleet is reduced (a one-off); giving away a free trial in PT, for anyone with a car, for example, one month once every two years;

Parking should be made gradually more expensive, for everyone, including residents and workers; rebates should be worked out based on two cumulative criteria: wealth/ability to pay; modal choice/availability of alternatives;

Expand HOV lanes and explore with high occupancy tolls (discounts), thus increasing the money gains from sharing;

Police wrong driving behavior, including substance abuse; enforce points-based licenses; together, these policies remove relevant part of offenders that cause more than their share of accidents, possibly making them captive to PT or SAVs.

4.5. ECONOMICS AND MARKET STRUCTURE IN AN AV ECOSYSTEM

Being public transport systems, fleets of SAVs should be private, local or regional monopolies, granted by competitive tender contracts, unless a way is found for competition not to interfere with fleet efficiency. They would deal directly with customers, with relationships ranging from the one-off tourist, through the more or less casual user, to the intensive, committed client. In offering transport as a service, they will handle fueling, servicing, parking, and overall management of the fleets, in their various sizes (seat range). SAVs will be able to leave their city or region, possibly paying their way as they do, because they must be able to accept, and perform, any service, as long as it starts in their 'domain'. Separate from SAV companies, there will be public or private dispatchers, in local or regional monopoly, to act

¹⁵ Much as the federal debt clock in the U.S. or the P.M.I. indexes.

as flight controllers do presently for air flight: gathering, and processing V2I information, managing the dispatch of SAVs to achieve the maximum of efficiency. They will validate, and manage privately owned AVs' journeys. These companies/agencies, much as the fleet companies, will require oversight, and regulation. Their algorithms must be transparent, as well as their pricing behavior. Taxation will have very fine-grained information, from which to base its incidence: miles travelled, times of day, routes taken (more or less congested), and vehicle occupancy.

Vertical differentiation will naturally take place both in privately owned AVs (size and luxury) and SAVs (waiting, sharing, detours, extra time), and the taxman can intervene, in order for demand to be managed, keeping PT's role in de-congesting streets and roads, and promoting DRS, as long as all pay their way (in proportion to what each takes in parking and public road space, energy, and emissions).

5. CONCLUSION

Widespread deployment of AV technology, for full automation, will deliver sharp reductions in accidents and parking needs, two boons for society at large and cities in particular. AV paradigm's role in Vehicle Miles Travelled, hence congestion, energy and emissions in comparison with today is ambiguous, which means it can be negative. The cornerstone of AVs improving on the status quo rests on sharing taking center stage, offering privacy (not sharing rows with strangers) and heavy discounts on flexibility (how much bother you tolerate for sharing). To lower the costs for AV, through scale, the personal car's leading position must be conquered. Also, shared AV's (SAVs) convenience must not endanger Public Transportation, at least in cities, where they are needed to decongest. To ensure that survival, PT will need continued support, and SAVs companies may need taxation or regulation to moderate/manage their quest for custom, namely curtailing the aggressiveness of 'all you can eat' menus.

Further research in this developing and prospective field include consumers' tastes and choices and technological developments on safe deployment of full automation and policy implications for SAVs.

REFERENCES

- Anderson, J. M.; Karla, N.; Stanley, K. D.; Sorensen, P.; Samaras, C.; Oluwatola, O. (2016) *Autonomous Vehicle Technology: A Guide for Policymakers* (Santa Monica, CA., RAND Corporation, 214).
- Burns, L. D.; Jordan, W.; Scarborough, B. (2012) *Transforming personal mobility*, The Earth Institute, New York Columbia University, 42.
- Button, K. (2014) *Transport Economics*. Cheltenham, Edward Elgar.
- Cambridge Systematics, I. (2005). *Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation*, Federal Highway Administration.
- Chan, N.; Shaheen, S. (2012) Ridesharing in North America: Past, present, and future. *Transport Reviews*, 32(1), 93-112.
- Correia, G.; J. Viegas (2011) Carpooling and carpool clubs: Clarifying concepts and assessing value enhancement possibilities through a stated preference web survey in Lisbon, Portugal. *Transportation Research, Part A*, 45, 81-90.
- Davis, S.; Diegel, S.; Boundy, R. (Eds.), (2016). *Transportation Energy Data Book. Oak Ridge*. Tennessee, Center for Transportation Analysis.
- Fagnant, D.; Kockelman, K. (2014) The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. *Transportation Research, Part C*, 40, 13.
- Fagnant, D.; Kockelman, K. (2015) Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations. *Transportation Research, Part A*, 77, 167-181.
- Green, J.; Steinbach, R.; Garnett, E.; Christie, N.; Prior, L. (2017) Automobility Reconfigured? Ironic seductions and mundane freedoms in 16-21 year old's accounts of car driving and ownership. *Mobilities*, 15.
- Hayes, B. (2011) Leave the driving to it. *American Scientist*, 99(5), 362-366.
- ITF, C. (2015). *Urban Mobility System Upgrade: How shared self-driving cars could change city traffic*. OECD, 36.
- ITF, C. (2017). *Transition to Shared Mobility - How large cities can deliver inclusive transport services*. OECD, 56.
- Jorge, D.; Correia, G. (2013) Carsharing systems demand estimation and defined operations: A literature review. *European Journal of Transport and Infrastructure Research*, 13(3), 201-220.
- Krueger, R.; Rashidi, T.; Rose, J. (2016) Preferences for shared autonomous vehicles. *Transportation Research, Part C*, 69, 343-355.
- Litman, T. (2018) *Autonomous Vehicle Implementation Predictions: Implications for Transport Planning*. Victoria, BC, Victoria Transport Policy Institute.
- Spieser, K.; Treleaven, K.; Zhang, R.; Frazzoli, E.; Morton, D.; Pavone, M. (2014). Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore. In Gereon Meyer, Sven Beiker (Eds.). *Lecture Notes on Mobility*, Massachusetts, MIT.
- Thierer, A.; Hagemann, R. (2014) *Removing Roadblocks to Intelligent Vehicles and Driverless Cars*. Mercatus Center, George Mason University, Mercatus Working Paper.
- Vij, A.; Carrel, A.; Walker, J. (2013) Incorporating the influence of latent modal preferences on travel mode choice behavior. *Transportation Research, Part A*, 54, 164-178.
- Wadud, Z. (2017) Fully automated vehicles: A cost of ownership analysis to inform early adoption. *Transportation Research, Part A*, 101, 163-176.
- Wadud, Z.; Mackenzie, D.; Leiby, P. (2016) Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles. *Transportation Research, Part A*, 86, 1-18.

Aggregation in the Portuguese Water Industry: The Case of Retail Service Operators

As Agregações das Operadoras de Água e o Caso das “Baixas” em Portugal

Sérgio Hora Lopes

Received for publication: April 16, 2020

Revision accepted for publication: March 24, 2021

ABSTRACT

The regional aggregation of water service operators has been considered a possible solution for improving these services, for enhanced service quality and efficiency in the performance of the entities responsible for providing the service. This paper seeks to study the aggregations that have occurred in Portugal in the last decade, in addition to review some of the most recent and relevant literature published on the subject. Such recent aggregation differs significantly from first and second-generation aggregation that occurred between the late 1980s and mid-2000s, being essentially aimed at (i) integrating only water distribution operators (i.e. direct providers, also known as retail operators), and (ii) integrating smaller and less populated territories.

Keywords: Water services; aggregation of operators; direct provider; and bulk operators; Portugal

JEL Classification: L52; Q25.

RESUMO

A agregação regional de operadores de serviços de águas tem vindo a ser considerada como uma possível solução organizativa para a melhoria destes serviços, quer para uma melhor qualidade de serviço quer para uma maior eficiência no desempenho das entidades responsáveis pela prestação do serviço. O objetivo deste trabalho é o de, para além de referir alguma da literatura mais recente publicada sobre o assunto, estudar as agregações que têm sido feitas em Portugal na última década. Estas são bem diferentes das da primeira e segunda geração, verificadas no final dos anos 80 até meados da 2000/2010, pois dirigem-se essencialmente para a integração de operadores com distribuição e não produção (baixa e não alta) além de integrarem territórios muito mais reduzidos e muito menos populosos, desenvolvidas num contexto diferente e condições de sucesso igualmente diversas.

Palavras-chave: Serviços de água; agregação de operadores; alta e baixa; Portugal.

Acknowledgment: This work was conducted within the scope of the activities developed by APDA, the Portuguese association of water operators. Although it is my sole responsibility, without the precious collaboration of the colleagues of the APDA this work would not have been elaborated. To all my sincere thanks. I am also grateful to João Cabrita for his help with English.

1. INTRODUCTION

The responsibility for the provision of water services has historically been a municipal competence and has been understood essentially as a local service. However, the world is constantly changing. Demographics, land occupation, the environment, the economy, technologies, the way we communicate, the values that determine our behaviors, the needs (of nature and humans) change in an accelerated way and water services are an integral part of this change. These transformations, or the accelerated evolution as it is often referred to, has had a huge impact on the way international institutions, scientific and professional organizations of the water sector have been adjusting their thinking on how to analyze the relationship between territory and operators, and on how to articulate themselves aiming at greater efficiency and better quality of service. Policymakers from several countries have also been sensitive to these matters, carrying on important changes in several countries, such as England and Wales, Italy, the Republic of Ireland or Hungary, just to name some European countries. One example of such changes is a clearer process of spatial aggregation in the way services are provided. Aggregations are a current theme and will most likely be on the agenda of water utilities' management for a long time, particularly in Portugal. This study aims at fundamentally analyzing the aggregation currently in place amongst direct providers of water and wastewater drainage, and at addressing, in a preliminary way, the advantages and limitations of such an ongoing process. In this context, we will present an analysis of some recent related literature, a brief historical reference on aggregations taking place in the context of the Portuguese democracy, present a specific study on aggregations, and a synthesis of the main conclusions.

The analysis of the aggregations in this work is focused on the retail service providers' activity, which had its most dynamic period during the course of the XXI Constitutional Government (2015-2019), even though it began in previous governments. Before detailing on the analysis of what happened in regards to these aggregations, we will expand on a few theoretical considerations about aggregations, including, among other aspects, the different types, their justifications, the conditions of success and the factors of failure. To do so, we will consider the international literature and the Portuguese experience. The international literature is already vast and there is even a very comprehensive review by Ferro (2017) that we strongly recommend for those who want to deepen into the theme. The Portuguese experience and the accumulated reflection on the subject are already notorious, with the country been often presented as an example at the international level.

2. AGGREGATIONS IN SOME RECENT STUDIES

The recent World Bank study of 2017 *“Aggregation of Water Supply and Sanitation Utilities - Joining Forces for Better Services?”* where several international experiences are analyzed (including two Portuguese), defines aggregation “as the process by which two or more service providers consolidate some or all of their activities under a shared organizational structure, regardless of whether or not it implies physical connection of infrastructures and whether the original service providers continue to exist or not.”

This definition of aggregation can cover a huge variety of situations. We may be discussing a merger of services as a whole, or the share of a Water Treatment Plant (WTP) or Waste Water Treatment Plant (WWTP), or even the simple creation of a common billing or accounting system. This work considers the most comprehensive concept of aggregation, that corresponds, as generically happens in Portugal, to the integration of providers, thus leading to the creation of a new entity to which previous service providers transfer the activities they were responsible for.

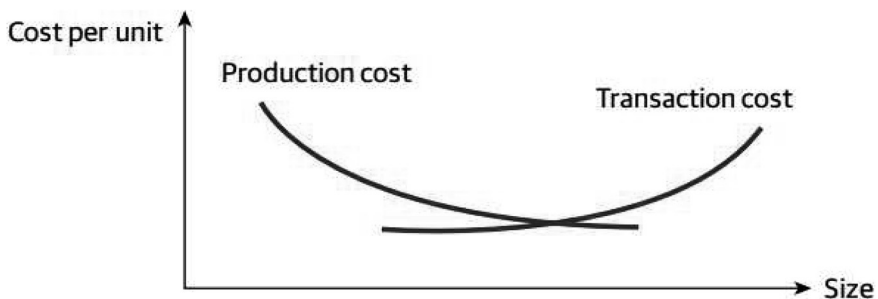
The reasons that justify the aggregation of the service providers are several but, in one way or another, are associated with the possibility of creating synergies brought about by cooperation. Depending on the type of aggregation, literature mentions numerous advantages for the creation of a new managing body, merger or absorption of several services into a single one. The current document discusses in a synthetic way, the three advantages that seem the most relevant:

- the increase in economic efficiency, that translates into lower production costs of both O&M and investment (CAPEX), due to positive internalities, economies of scope, process and scale, through the use of networks that serve several municipalities, enhancing the number of connected consumers. The major beneficiaries will be, in particular, consumers/users who will be able to see the sustainability of the operators with the most appropriate prices guaranteed, and the economy in general that will benefit from greater effective availability of resources for the country as a whole;
- the improvement of the operators' performance – because of the opportunities that the increase in size potentiates. This may occur due to the possibility of the new operators to professionalize (the much-claimed “*entrepreneurialization*”, referred to practically in all documents on the sector in recent decades, and which has taken place very slowly in Portugal), the availability of more qualified human resources and better capacity of organization, the quality of management, etc. The service provider can extend and improve the range of services, and reduce costs potentially benefiting consumers with tariffs better suited to household incomes, in addition to increase environmental best practices, at a time when pollution, resource scarcity and the environmental crisis are so widespread. This positive externality must now be seen as one of the main strategic objectives in the reform of the water sector; and
- the increase in territorial and social cohesion, both at national and regional level. The exercise of aggregation is always based on bringing together in a single entity what was previously separated into different players. By joining territories, among other advantages, it is possible to minimize the consequences of demographic and wealth differences across the various spaces; and take advantage of the management capabilities existent in the most advanced operators. This is an essential exercise of solidarity to guarantee that, in the end, there is equity. Overall, this positive externality of aggregations and the need to efficiently implement the large increase in EU cohesion funds made available from the early 1990s, resulted in one of the main key reasons for the creation of multi-municipal systems (bulk systems).

However, aggregations may bring negative consequences and incur into the so-called transaction costs (see Figure 1). These can be punctual, single or permanent, and long-term ones. The costs of technical and economic studies prior to the kick-off decision are examples of punctual costs. In the preparation of the aggregation there are negotiation costs necessary to reach an agreement between the parts. As it is well known, this type of costs can be relevant and derives from several factors such as the power or information asymmetries between parts, or compensations for key municipalities.

From the various long-term transaction costs, there are some that derive from the solution found not configuring an optimal dimension, either by falling short, or by going beyond the adequate dimension. The issue of optimal size is of the utmost importance when it comes to drawing an aggregation from the point of view of efficient resource use. However, the factors that involve aggregation go far beyond economic ones and, using a concept most appreciated by “conventional” economists, are not based on agents’ “rational behavior”. It should be highlighted, in any case, that regardless the rational the agents are, determining an optimal dimension is an extremely difficult task. Each space to be added is different, either for demographic (quantity and density), natural (orography, precipitation, etc.), economic, social, cultural, or political reasons. As a result of these complexities, the aggregated entity may fall short or beyond what is desirable in terms of efficiency, due to either external environment or internal reasons (bureaucracy, lay-offs, demotivation, etc.). Ferro (2017, p.17), in the above-mentioned work, refers that concerning the English and Welsh case “there is evidence that the average water and sanitation company is characterized by economies of scale.” These costs imply final solutions, designated as subroutines and characterized by the reduction of production costs, not compensating for the increase in transaction costs resulting from the allocation of resources.

Figure 1: Inverse relationship between production and transaction costs



As stated above, the success of aggregations should not be taken for granted, quite the contrary. Hull (2013) cites a study by Abott and Cohen (2009), in which, by analyzing several studies on economies of scale and scope, they concluded that, in 26 examples, 13 show evidence for the existence of economies of scale, five (all in England and Wales) of diseconomy, five in which economies of scale were exhausted and another five are

inconclusive. As for economies of scope, this study reports that in 10 cases, five show positive evidence, two of diseconomies and three are inconclusive. It is interesting to note that in the two types of economies analyzed, 50% of the cases are positive and in the remaining, there are either diseconomies or the results are inconclusive. In conclusion, it can be said that aggregations are useful in certain cases but not in all, and should therefore not be seen as a universal panacea.

Thus, since there are economies and diseconomies of scale, reaching the equilibrium point is a highly complex task. Even though the most realistic/pessimistic person may say that it is impossible at all to reach the optimal point, there is always a “second best”.

An important issue to discuss about the advantages and weaknesses of aggregations is the need to understand the importance of economies of scale, the size and density of aggregation and how, in many cases, greater or lesser efficiency is more determined by the size and density of aggregation than by the economies of scale. This subject has a particular acuity in the case of the retail operators and, as far as it seems, has not gained much attention in Portugal so far. Given its relevance, it deserves a thorough analysis.

In his article with a particularly suggestive title (“Urban infrastructure: Density matters, not just size”, Wenban-Smith, 2006), and in more detail in his doctoral thesis (Wenban-Smith, 2009) the author shows that, in certain cases, dimension is not the main determinant of utilities’ efficiency. He mentions in his thesis that “Water distribution costs are more significant than water production costs. For example, in the case of BWC (alias of the company analyzed), although the distribution operating costs are approximately the same as the operating costs of production, the distribution capital costs are about twice as high. The effects of scale on the distribution therefore deserve careful attention.”

In water services we have fixed facilities, WTP, for example, and a distribution network that crosses the whole territory. According to Wenban-Smith, the existence of economies of scale at the level of fixed installations is evident, but using an indicator that combines volume and distance, he concludes that “unit distribution costs are characterized by economies of scale in relation to volume, but diseconomies in relation to the average distance of properties” (p. 3). He concludes that higher distribution densities reduce unit distribution costs, while lower densities increase them.

Wenban-Smith concludes that “Interaction with production costs then means that (a) greater urban density (“Densification”) is characterized by economies of scale in both production and distribution; (b) greater territorial dispersion (“Dispersion”) leads to dissaving in distribution; (c) “Suburbanization” (expansion to peripheral areas of lower density) is in the middle, with approximately constant scale returns, if production and distribution are together; and (d) the expansion of “constant density” leads to small economies of scale. Keeping water supply costs (per capita) low seems to depend on both density and size.”

For those who know the sector well, the author’s conclusions are not exactly a novelty. It is already well established that the “per capita” costs of supplying a dense area are much lower than in sparse area. The question that arises when we study a direct provider aggregation, for which the relevant costs are distributional, is that the consequences of aggregation can be diverse. When we group areas with similar densities, the effects of the dimension will benefit all integrated parts. When we aggregate areas with different densities the question arises in a different way, and it is not clear what the result will be. Even assuming that the

size effects will result in a reduction in the average cost, benefiting society as a whole due to a better allocation of resources, the denser areas, with lower per capita costs will see, due to the scattering effect, their costs increase. In contrast, in the less dense areas, with higher per capita costs, the overall costs will decrease. This means that there will always be losers and winners. We face a complex problem which is to assess the consequences at macro (national) and micro levels and, for regions belonging to the aggregation, to be able to make a cost-benefit analysis for aggregating or not.

Another aspect frequently mentioned in the literature on aggregations is the need for anchors, or at least the advantageous of their existence. This concept is now widely disseminated with slightly different meanings but always with the common idea that there are certain operators that can give the necessary viability to the aggregation. Without them, it would be, if not impossible, at least more difficult to achieve that viability. With regard to the water sector, the anchor can be defined as a territorial space with a larger dimension (and density) than the other partners and, therefore, from a quantitative point of view, significantly expands the demand and scale of aggregation. But there is also a qualitative aspect, because scale is normally associated with more know-how, technological and management capacity for that utility. These aspects can work as factors of dissemination of knowledge and modernization for the aggregation.

It can be concluded, therefore, that one does not face an exclusively quantitative factor when it comes to anchoring. For example, there may be aggregations in which a more populous municipality is the one with less infrastructures, and therefore requires a disproportionately large investment, both in absolute and relative terms. There may also be a large operator with very large idle capacity, highly ineffective, that will introduce entropy into the aggregation. In these cases, it is clear that the potential turnover that a large municipality/operator brings will not generate, at least in the short and medium term, any kind of positive effect. Designating as anchor municipalities those with these characteristics may not be correct. Moreover, it should be noted that according to the already mentioned “Aggregation of Water Supply and Sanitation Utilities - Joining Forces for Better Services?”, the most important is the existence of a “champion” because: “Having a stable champion throughout the aggregation often improves the likelihood of success.” As we will see, among other aspects, the absence of anchor municipalities that can be considered as “champions” is one of the problems in Portugal.

3. BRIEF HISTORICAL NOTE ON AGGREGATIONS IN PORTUGUESE DEMOCRACY¹

The political debate about the assignment of management competences for water services in Portugal immediately after the-April 25th Revolution, evidences the existence of two trends, underlying the different organizational rationale that continues until today, although in a more mitigated way after the 1990s. One is a technical-economic trend, in favor of a water management based on the creation of companies covering large regional territorial spaces,

¹ This section is largely tributary of the study of Pato (2011). About the same subject, another reference for Portugal is APDA (2018).

including several municipalities. The other trend supports the strengthening of municipal (local) power, with the attribution of water supply and sanitation management competencies to municipalities, thus continuing a traditional practice, both in Portugal and in most Western countries, but with greater autonomy, particularly financial, than that existing during the *Estado Novo* regime. The relevance of this issue was brought to light due to the underdevelopment of the water sector in Portugal, and because between 1971 and 1975 there were cholera outbreaks, with 40 deaths for 2,371 cases of the disease in 1975. Also, back then it was discussed how power would be distributed in territorial terms (deconcentration, decentralization and regionalization) in a country that was under a highly centralized authoritarian regime for so many years.²

The I Constitutional Government followed a more technical-economic orientation, having proposed ³a “new configuration that involved a profound change in relation to the political formula that made responsibility for water and sewage services dependent on local authorities: it is now clear the intention to develop an integrated vision of the problem at the national level, through the creation of regional basic sanitation companies capable of integrating its three strands (water, sanitation and solid waste) and obtaining economies of scale, scope and process. Its implementation should be coordinated by the General Directorate for Basic Sanitation (DGSB), created under this Government and in articulation with the municipalities”. This proposal was rejected by the opposition parties that advocated the full autonomy of local power. Let us focus on two speeches in the Portuguese National Assembly that summarize this anti-aggregation position of municipalities. The PSD (center right party) asked for the ratification of the proposal and her deputy Mrs. Helena Roseta said that the creation “*of a regional entity that has the sophisticated name of installation commissions of basic sanitation providers, and that in essence is no more or less than the embryos of future public sanitation companies, directly opposes to the law’s attributions of municipalities in relation to basic sanitation. Indeed, the law of local authorities gives them the power to act in this matter, being and all the actions of this sector an assignment of the local authorities. Therefore, it is not reasonable that an organic law should create commissions to install public companies that would withdraw this attribution from local authorities.* Also, the deputy of the Communist Party, Veiga de Oliveira, criticized the “*centralist and technocratic*” route presented by the Government, stating that the respective proposal, oriented from “*Terreiro do Paço*”, *with a few employees scattered throughout the province, a few bodies of sophisticated technical apparatus and maybe a few public companies with powers and skills, obviously, for planning and solving everything, would be the solution for basic sanitation. Nothing more wrong and dangerous. Quite the opposite to solve the enormous problems that affect us all, related both to the consolidation of democracy and to the economic, social and political progress of our country.*”

Thus, the proposal to create basic sanitation regions was abandoned and the basic sanitation responsibility was practically and exclusively assigned to the local government. Although there was some evolution, namely in the water supply, it was particularly slow and the improvements in the performance evaluation indicators have fell short of what would

² See Pato (2011).

³ Colonel Baltazar Barroco was the Secretary of State for Water Resources and Basic Sanitation, hence this reform was designated at the time by the name of his political officer, Military Engineer.

⁴ In a reference to the Portuguese Government.

be desirable. The sector was stagnant and a good part of the more than three hundred operators were unable to fulfill their mission. With few exceptions, the water sector was under-endowed with financial and human resources and that was in fact one of the many concerns of a recent democratic local power that, although having introduced a different dynamic in the territorial management, was unable to have a strategic vision for the sector.

Baptista and Matos (1994) concluded that, from an institutional point of view, the fundamental problem was the difficulty in having an integrated regional vision and to articulate the three hundred and sixty five municipalities responsible for the management of networks in their geographical areas. This had as a consequence the *“excessive disaggregation of systems, with a large number of small or very small systems, which necessarily implies a poor and non-optimized application of investments due to the absence of economies of scale”* (p. 14). Moreover, given that the responsibility is essentially municipal and the municipal budgets were insufficient to bear the costs, the financial situation of the sector would face two problems. On the one hand, ‘there are systematic financial difficulties of the municipalities in the implementation of the competences conferred on them’ and, despite the diversity of funding forms, it is still frequent *“the inability of municipalities to guarantee their self-financing share”* (p. 16) to benefit from European Community programs. On the other hand, there was a lack of a realistic tariff policy, *“with the usual practice of water and wastewater tariffs below the actual cost of service..., or even the absence of tariffs”* (p. 16), which has prevented the municipalities’ revenue increase.

However, in the early 1990s, the first studies about the establishment of regional sanitation companies in the Estoril Coast and of water supply companies in the Cávado region had already been launched, through the collaboration of the Ministry of the Environment and the EPAL (the Portuguese company responsible for the supply and distribution of water in the Lisbon council). This required the articulation of the central government with the municipalities involved, resulting in the preparation of the work that culminated in 1993 in the amendment of the legislative sector’s framework, in particular with the opening up of the sector to private initiative and, with regard to our theme, the creation of multi-municipal systems. Using this new legal framework, several public companies were created, mostly owned by Águas de Portugal (AdP), a state-owned holding company created during this period, where municipalities had a shareholding, aimed at managing integrated systems with several municipalities. It is important to note that these new companies were only responsible for the “bulk” part of the industry. The water distribution and sewage collect (the retail part) remained under the municipal sphere without any territorial aggregation.

In addition to the aforementioned problems faced by the sector, in 1986 had become a member of the European Union Community (EEC), which had brought new demands since the EEC legislation was much stricter than the national one. In addition, more financial resources were made available, which increased the possibility of investing in the sector. It was clear to Portugal that EEC funds for economic and social cohesion had to be used, particularly when in the early 1990s the Cohesion Fund (CF) was created, anticipated by the Cohesion Financial Instrument.

During the second half of the 1990s, multi-municipal companies contributed to a significant improvement in the regions covered by these companies, even though there was a slow evolution of municipal utilities (direct management and concessions). Not all problems were solved. Quite the contrary. The multi-municipal systems did not cover the entire country. And

in retail systems, the tariff deficit remained an almost widespread reality, thus preventing the ability to adequately finance many of the operators. The improvement of the service was not a universal reality, the water scarcity tended to worsen, and the utilities' response were not always fast and effective (the water losses were too high e.g.). From the analysis carried out at the time, it was concluded that the problems reached the greatest severity mainly in the direct management of the systems at the distribution level. There were persisting difficulties, namely regarding limited financial and human resources, investment below what was necessary to extend service coverage and quality, inadequate tariffs, low numbers of users and low population density, and insufficient turnover for utilities to gain critical mass. To overcome these issues, it was necessary to complete the policy measures started in the 1990s.

At the beginning of the XXI century a medium-term strategic instrument was developed, defining the policy for the sector, the PEAASAR 2000-2006, which essentially focused on extending the model of multi-municipal systems from the previous decade. But the success of the policy would be accomplished in the subsequent plan – PEAASAR II (2007-2013)⁵, with new aspects in the strategy for the sector. *“Depending on the diagnosis developed, it was then considered essential to reorient the overall strategy for the sector in order to develop a model for extending the partnerships between the State and the Local Authorities which should now be oriented, essentially, for the resolution of the problems of the distribution level, which also intended the revision of the legal regime of the existing institutional forms for the management of water and wastewater services, as well as the creation of a new system of partnership contracts to be established between the State and local authorities.”*

“It is therefore recommended that the sector’ organizational model should be based on the integration of the distribution corresponding multi-municipal systems (without prejudice for free initiative of municipalities in inter-municipal systems), also assuming the possibility of creating new municipal systems integrated with the participation of the State, either through the AdP or through the concessions of the multi-municipal systems.”

At the end of the 2001/2010 decade, an inter-municipal company (*Águas do Ribatejo*, EM) and the first Public-Public Partnership (PPP) between AdP and municipalities (ADRA, SA), in the Aveiro region were created. At the beginning of the following decade, in 2013, another PPP (*Águas da Região do Noroeste*) was created within the multi-municipal company *Águas do Norte*.

In 2014 a new plan, PENZAAR 2020 — A new strategy for the Water Supply and Wastewater Sanitation sector, was approved, under the XIX Constitutional Government covering the period from 2014 to the end of the decade. It showed some continuity in relation to the previous planning instrument, the PEAASAR II, by stating that it “did not identify the need to define new management models in relation to existing ones in the current legal framework”, regardless of the need to make several improvements but not related to what that matters here, aggregations. At the end of the decade, under the XX Constitutional Government, a clear political will to increase the number of aggregations, whether in the form of PPPs or inter-municipal companies, was revealed. Between 2019 and early 2020, seven new entities were created, only at the “distribution” level or vertically integrated”, tripling their number.

⁵ Approved in 2007 by the XVII Constitutional Government.

The current phase is characterized by a strong political will from the XX Government and a greater acceptance on the part of the municipalities for new and different forms of organizations in the sector. This results from a better understanding of the fact that there are objective aspects that make the scale important, not only to gain critical dimension, but also to increase operators' efficiency, provide services of increased quality and better respond to future challenges such as the ones related to climate change and resource scarcity. This phase intends to be, in some way, the answer to the problems posed after the 25th April Revolution, when there was an awareness that isolated, many municipalities would not be able to create sustainable operators in a context in which regionalization was a forever postponed project.

In summary and analyzing only what happened after the entry of Portugal in 1986 in the EEC, it can be said that, respecting the municipal attributions in the water sector, the legislation and the practice have been adapting to the economic, financial and political requirements. In the 1990s, with the creation of multi-municipal systems, there was a separation of activities between "bulk" and "retail", a form of "unbundling" so popular in those years. Therefore, in many territories of the country, the bulk activities, with a territorial aggregation, was handed over to the central government, although in partnership with the municipalities, the former being responsible for all activities. In a first phase, mainly in urban areas and in the Algarve, and in a second phase – already in this century –, it has extended to practically the whole country. The systems covered water supply and wastewater. It was also at the end of the first decade of this century that the first PPPs appeared, creating the legal framework that allowed governments to encourage the constitution of management entities in a business way, aiming essentially at the retail activity, often at the municipal level and occasionally with the participation of Águas de Portugal. It is on the latter that this empirical investigation is focused on.⁶

In a first phase, the process took place essentially in urban areas and in the Algarve. In a second phase, already in this century, it has extended to practically the whole country. The systems covered water supply and wastewater services. It was also at the end of the first decade of this century that the first public-public partnerships emerged, creating the legal framework that allowed governments to encourage the creation of corporate retail operators, often at the municipal level and occasionally with the participation of Águas de Portugal. It is on the latter that the present empirical investigation focuses.

⁶ The "unbundling" that can be translated by separation, disaggregation, is used in utilities as a way to autonomize activities, separating activities to introduce competition in monopolistic sectors. The World Bank has often proposed it, for example, for the energy sector. The creation of multi-municipal systems in water and waste management that took place in the 1990s in Portugal, is part of this conception.

4. THE AGGREGATIONS UNDER ANALYSIS⁷

We will analyze spatial and demographic factors (size and population density of aggregations) that affect economic efficiency, as well as qualitative factors (institutional, organizational) and others that influence the quality of service and management (costs, etc.). It is evident that, being a first approach and not having carried out any case study based on empirical analysis on the internal structures of each aggregation, the study is essentially based on demographic data.

We consider five cases of aggregations in activity at the beginning of 2020. The first is from 1997, the case of Águas do Planalto, SA, created by the Association of Municipalities of the Planalto Beirão, being an inter-municipal concession that covers five municipalities. The second case is ADRA – Águas da Região de Aveiro, SA, a public-public partnership started in 2009, between Águas de Portugal, SA and 10 municipalities in the Ria de Aveiro area. Still with the same management model we have ADAM- Águas do Alto Minho, SA that started the activity in 2019. With a different model, although within the type of partnership mentioned above, there is the water system of the Northwest region, an autonomous operation within AdNorte – Águas do Norte, SA, covering eight municipalities. Another model, an inter-municipal company, is the case of Águas do Ribatejo, EM, from Lezíria do Tejo region, which was created at the end of 2007 with six municipalities and that extended to seven municipalities in 2011. These operations cover 30 municipalities with a total of about one million people (10% of Portuguese population). It is interesting to note that these five aggregations are managed according to four different models with an appreciable diversity of management arrangements, because only ADRA and Águas do Alto Minho can be considered similar. The “distribution” operation of Águas do Norte, SA, although separated from the “bulk” system, is managed by the same company, and there is, therefore, a *verticalization*, that is, a common coverage of the bulk and retail activities. Águas do Ribatejo is fully vertical. In other words, we are facing five operations managed by four different models, guaranteeing a diversity of situations.

In addition to the cases already mentioned there are five companies created in 2019 that started their activity in 2020: AIN-Águas Interior Norte, EIM, AMBG – Águas Baixo Mondego e Gândara, Tejo Ambiente EIM, APIN-Empresa Intermunicipal de Ambiente do Pinhal Interior and, more recently, Águas do Alto Alentejo, EIM. The number of municipalities covered is 41 and the population of about 450,000 inhabitants. The most prominent note in relation to these five cases is that they are all of the same model of organization, inter-municipal company. Is the diversity referred in the preceding paragraph an issue?

Let us study the ten cases, and start by analyzing the aspects related to the dimension. Not going into great detail it should be said that in the water market the dimension is essentially created by population. The advantages of aggregation, therefore, begin with an increase in the number of inhabitants covered by the operator. Table 1 makes a synthetic analysis of what is happening in Portugal.

⁷ The data used in this analysis were collected from the websites of INE - Instituto Nacional de Estatística (https://www.ine.pt/xportal/xmain?xpgid=ine_main&xpid=INE), Pordata (<https://www.pordata.pt/>), ERSAR (<http://www.ersar.pt/pt>) and from interviews and other personal contacts with colleagues from the Secretariat of State for the Environment, Águas de Portugal and the entities mentioned in this study.

Table 1: Population

| Agreggations | No. Mun. | Area km ² | Population Evolution (Inhabitants and per cent) | | | | | Mun. Pop Maximal | Mun. Pop Minimal |
|--|----------|----------------------|---|---------|---------|----------------|----------------|------------------|------------------|
| | | | 1981 | 2011 | 2017 | 2017- -1981 | 2017- -2011 | | |
| ADAM Águas do Alto Minho, SA | 7 | 1,588 | 206,778 | 203,808 | 194,599 | -5.9% | -4.5% | 85,017 | 8,617 |
| AdNorte (SMM) | 8 | 1,654 | 292,526 | 300,102 | 286,901 | -1.9% | -4.4% | 68,524 | 18,653 |
| ADRA Águas da Região de Aveiro, SA | 10 | 1,473 | 289,115 | 340,329 | 335,521 | 16.1% | -1.4% | 77,631 | 10,307 |
| Águas do Planalto, SA | 5 | 1,051 | 86,119 | 71,534 | 67,426 | -21.7% | -5.7% | 26,942 | 8,948 |
| Águas do Ribatejo, SM | 7 | 3,215 | 140,781 | 149,153 | 144,003 | 2.3% | -3.5% | 35,208 | 7,155 |
| AMBG Águas Baixo Mondego e Gândara | 3 | 618 | 63,184 | 57,581 | 54,851 | -13.2% | -4.7% | 25,399 | 11,932 |
| APIN Emp. Intermunicipal de Amb. Pinhal Interior | 11 | 1,938 | 104,607 | 87,764 | 82,462 | -21.2% | -6.0% | 13,999 | 3,890 |
| Tejo Ambiente EIM | 6 | 1,500 | 123,657 | 113,176 | 107,424 | -13.1% | -5.1% | 44,406 | 3,778 |
| Águas Interior Norte, EIM | 8 | 1,559 | 126,619 | 106,999 | 105,857 | -20.1% | -5.4% | 51,417 | 3,677 |
| Águas do Alto Alentejo EIM | 10 | 3,732 | 71,820 | 53,561 | 48,097 | -33.0% | -10.2% | 15,285 | 2,910 |

Source: See footnote 7.

The first finding is that the population is decreasing in the current decade in all cases studied. As it would be expected, the situation is more serious in the interior Norte, Beiras and Alentejo, but even the industrial and urban area of Aveiro is confronted with the same problem.

The second, is that the aggregations carried out correspond to very different operations: some allow to anticipate strong economies of scale, others not so much. Of course, this cannot be seen in abstract: for example, the Águas do Ribatejo EM, with less than half the population than ADRA, can partly compensate for the lack of dimension by extending the service to the entire urban water cycle. AdNorte has a relatively higher than the average population of the aggregations studied, but for more than one third (corresponding to Santo Tirso and Trofa municipalities), the service provided is only wastewater treatment and drainage, and the water supply is of the responsibility of a private concession. That is, only a case-by-case analysis will give us well-founded answers. In any case, many aggregations do not seem, in a quick approach, to be in the best position to ensure operational efficiency

and management gains derived from the scale. In addition, density does not satisfy Wenban-Smith's (2009) concerns and there are not exactly "champion" municipalities.

The third outcome is that even if we may be far from achieving any "optimal dimension", the truth is that we are much better than we were before. In the Area of APIN each utility had a population served between 4,000 and 14,000 inhabitants, but the new entity will have almost 80,000 potential users. Also, other aggregations will end up with dozens of different operators of less than five to ten thousand users. It may not be a giant step but it's a step forward.

A last and brief reference goes to the areas of the aggregations, which vary widely, similarly to the population, ranging from 618 km² to 3,732 km², although in most cases it is between one thousand and two thousand km². Although the economies of scale in water services result from the volume of population served, the extension of the territory is not a non-negligible factor for efficiency, since the greater or lesser concentration of the population and population density have an influence on economic efficiency. These issues are analyzed in Table 2, which provides synthetic information on population density in the cases under analysis, in a dual (inter and intra aggregation) perspective, as well as the comparison with some municipalities with autonomous water services management located in metropolitan areas or district capitals.

Table 2: Population density

| Agregations | Medium Density Hab/Km ² | Mun. Dens. Maximum | Mun. Dens. Minimal | Relation Max/Min. | Relation Max/ Medium |
|---|---------------------------------------|-----------------------|-----------------------|----------------------|----------------------------|
| Águas do Planalto, SA | 64 | 95 | 36 | 2.7 | 1.48 |
| ADRA Águas da Região de Aveiro, SA | 228 | 527 | 89 | 5.9 | 2.31 |
| Águas do Ribatejo, SM | 45 | 131 | 13 | 10.4 | 2.92 |
| AdNorte (SARN) | 173 | 524 | 64 | 8.2 | 3.02 |
| ADAM Águas do Alto Minho, SA | 123 | 267 | 47 | 5.6 | 2.17 |
| ADIN Águas Interior Norte, EIM | 65 | 166 | 14 | 11.9 | 2.55 |
| AMBG Águas Baixo Mondego e Gândara | 89 | 111 | 66 | 1.7 | 1.25 |
| Tejo Ambiente, EIM | 72 | 148 | 16 | 9.2 | 2.06 |
| APIN Emp. Intermun. de Amb. Pinhal Interior | 43 | 125 | 10 | 12.1 | 2.93 |
| AAA Águas do Alto Alentejo | 13 | 20 | 8 | 2.5 | 1.54 |
| | | | | | |

| Agregations | Medium Density Hab/Km ² | Mun. Dens. Maximum | Mun. Dens. Minimal | Relation Max/Min. | Relation Max/ Medium |
|-------------------|--|-----------------------|-----------------------|----------------------|----------------------------|
| Braga | 988 | n.a. | n.a. | n.a. | n.a. |
| Coimbra | 420 | n.a. | n.a. | n.a. | n.a. |
| Gondomar | 1,255 | n.a. | n.a. | n.a. | n.a. |
| Seixal | 1,734 | n.a. | n.a. | n.a. | n.a. |
| Sintra | 1,206 | n.a. | n.a. | n.a. | n.a. |
| Vila Nova de Gaia | 1,780 | n.a. | n.a. | n.a. | n.a. |
| Viseu | 196 | n.a. | n.a. | n.a. | n.a. |

Source: See footnote 7.

The first note from Table 2 is the huge difference between the densities of aggregation operations and those from suburban municipalities in metropolitan areas. It is not something we did not know, but it is always worth pointing out. When compared with district capital municipalities (last seven lines of Table 2), in which urban concentration is evident, these differences in density are no longer so evident; also worth of note are the differences between the various aggregations in terms of population density: from 13 to 228 inhabitants/km². Returning to the Wenban-Smith line of reasoning, surely there will be important differences in inter-aggregation economic efficiency.

Looking at what is happening intra each of the aggregations, we see once again the presence, in some cases more than in others, of a multiplicity of situations. In several aggregations, municipalities with high population density are joined with low density ones, as in ADRA and AdNorte (this, with the already pointed out caveat that in the two municipalities with higher density, only the activity of wastewater is covered, thus being more difficult to recover costs), and, although less marked, in Águas do Alto Minho. In theory, we can see a loss of efficiency in some of the previously existing utilities, which is offset by greater efficiency resulting from the aggregation.

As already mentioned, the existence of “anchors” are considered advantageous for the success of this type of operations. “Anchors” are the municipalities that due to various factors (e.g. dimension, productivity, good organization, technological evolution) when associated with others, give scale and generate innovation in the aggregate set. It is not easy to say at first sight whether an operator/municipality will be an “anchor” of a future aggregation, which can only be concluded “*a posteriori*”, but the analysis of the size of the municipality and its relative weight in the aggregation can help to understand whether or not it can become an “anchor”. As we saw before, “champions” may play an important role, but one question arises when we look at the cases of the previous aggregations: where are the “champions”?

Of course, we have no case of municipalities with more than a hundred thousand consumers, with financial, management and technical capacity.

Thus, we did a quantitative exercise to analyze the existence or not of “anchors” in the Portuguese aggregations, analyzing the information available for the 75 municipalities of the ten aggregations of the sample, very representative of our reality.

The first finding is that there are no “champions” and, in their absence, we classify the municipalities into three types, according to population and density criteria (the only objectives available) as follows:

Gold: population above 50,000 inhabitants and density greater than twice the aggregate average;

Silver: population above 50,000 inhabitants and density higher than the aggregate average;

Bronze: population above 30,000 inhabitants or density greater than 1.5 times that of aggregation.

The choice of intervals was defined based on the generally small size of the municipalities and low density differences, an essential factor in reducing costs, as we saw before. No municipality alone would essentially change the configuration since it is the necessary a greater scale. Let us focus on the results of the analysis.

The aggregations of Águas do Planalto, AMBG – Águas Baixo Mondego and Gândara, APIN – Empresa Intermunicipal de Ambiente do Pinhal Interior, do not have any municipality that meets the criteria for being “anchor”, of any kind. In the case of ADAM – Águas do Alto Minho, there is a municipality that meets the criteria of anchor *Gold*, Viana do Castelo; the same happened with Vila Real in AIN – Águas Interior Norte, EIM,

Santo Tirso, in AdNorte, also meets the criteria for the *Gold* group, although as mentioned above, only wastewater is included in the whole; another similar municipality is the Trofa corporate group, which is part of the *Silver* group. Amarante, with more than 50,000 inhabitants, but with a low density, practically equivalent to the global average, is part of the *Silver* anchor group.

ADRA has three municipalities that are part of the *Silver* group: Aveiro and Ovar, mainly by size, with respectively more than 75,000 inhabitants and 55,000 inhabitants, and Ílhavo, which, from the 75 municipalities analyzed, is the one with the highest density, 527 inhabitants/km². This is the aggregation that presents itself, at least quantitatively, as the most balanced one.

In Águas do Ribatejo, the only “anchor” is Torres Novas, small in absolute terms, but with some relative weight in terms of population in the aggregation, because it has almost three times the average density (131 Hab/Km², which compared with the average 45 Hab/Km²), integrating into the *Silver* group. In the case of Tejo Ambiente, Ourém and Tomar have more than 30,000 inhabitants each, but although they are close, they do not have 1.5 times the density of the aggregation, so in all 75 municipalities they are the only that are considered *Bronze* anchors.

The scenario presented about anchors is not the most favorable, but we must keep in mind that as the poet and Cardinal José Tolentino de Mendonça (2020, p. 117) says, “Exactly because we are not just beings of the present, we don’t just have anchors; we also have wings.”

5. CONCLUSION

From what has been stated in the previous sections it is possible to draw some conclusions and assess the issues that these aggregations raise for the development of water services in Portugal, especially in low-density regions.

1. Let us start with economic efficiency via reducing production costs. From this point of view, what can be concluded from the literature is that, both regionally and nationally, comparing the *ex-ante* and *ex-post* situations, everything points to the final situation being better than the initial one. In terms of intra-aggregation analysis, the evaluation is more complex and would require a deeper analysis. If the final solution corresponds to the optimal economic dimension, every intervenient will win. But, as already mentioned, this is not easy, given the multiplicity of factors, such as the size of the population and territory, the available water resources (quantity, location and type), orography, the sophistication of existing utilities, revenues and their distribution, among others, that influence efficiency. Thus, the existence of cross-subsidization will certainly contribute to the impossibility of verifying a Pareto Optimal. But the cases studied show that the possible negative financial effects for the municipalities that are negatively affected, those called “anchor gold”, are always mitigated through the application of compensatory mechanisms from external elements to the aggregation. In fact, with the contribution of external support, it is intended that the intra-aggregation is “balanced” and thus equity in the treatment of all territories and their representative bodies is guaranteed.⁸ In national terms, the post-aggregation situation will be better, because the size of these aggregations will not be sufficient to lead to diseconomies of scale.⁹

2. But is the effective purpose of aggregations the economic efficiency understood only as a short-term cost reduction for all players. Undoubtedly, reducing costs by size is an important factor in justifying the need to create aggregations, but we must be able to see that aggregations aim at broader, holistic objectives even, going far beyond the immediate view of problems from an exclusive cost perspective. The essential objectives to be pursued for water services to ensure their mission are generally: to ensure the preservation of the environment and to provide the population with adequate and affordable water services. For this,

⁸ This issue should be properly examined, but, as far as we know, it has not caught so far the attention of researchers. Probably the small number of operations still existing limits the empirical information base, so it is likely and useful that it will deserve more attention in the near future.

⁹ According to Ferro (2017, p. 9), there is “an ample empirical literature [showing] that agglomeration leads to economies of scale in the WSS industry, at least up to a certain level. Furthermore, it seems that large utilities tend to operate at a lower unit cost and perform better than smaller ones do. Yet there are also counter examples that show that sometimes scale is beyond the optimal efficiency size. An exhaustive examination of empirical literature on economies of scale is presented next. Empirical studies indicate that economies of scale are prevalent in small-size providers and diseconomies of scale appear in very large ones. In the middle, there seems to be a continuum of possible minimum efficient scale.”

the sector and its operators need to have at least a business and professional management, be endorsed with financial autonomy and their own decision-making capacity/ability, be economically and financially sustainable, have sufficient human, financial and organizational resources not only to manage day-to-day activities, but also to keep up with the sector's permanent innovation and the complex challenges it has to overcome¹⁰. A water utility in the 21st century must have a holistic vision and a long-term horizon. However, without an appropriate dimension that gives it what is normally called a critical mass, certainly the operators will not have the necessary tools to fulfil their mission.¹¹

3. The arguments put forward, *often off-record*, by many municipal policy-makers who, having entered into negotiation processes for the aggregations, then abandoned it, relate to increases in tariffs which aggregations will “inevitably” bring. In general, this argument is fallacious because the existence of low tariffs does not result from cost efficiency but from the application of tariffs that do not recover costs and thus do not guarantee the sustainability of the operation, nor avoid dependence on other budgetary benefits. In these cases, the quality of the service provided is also far below that legally required, as can be seen in the annual service evaluation reports drawn up by the regulator. The tariff increase is not the result of aggregation, but of pricing policy options. Another argument which is also put forward, usually by larger municipalities, is the refusal of possible cross-subsidization. The issue is complex, as we have already mentioned in point 1, and is highly linked to political motives, often electoral, short-term oriented. For economic reasons, the long-term interest of all, even the citizens of these municipalities, is set aside.^{12, 13}

¹⁰ This issue gains particular acuity when we are looking at cases of geographical spaces where services are less evolved. In the developing world, aggregation is primarily a mean to deliver better services rather than to lower costs. To some extent, several Portuguese water service providers are still far from European average performance, perhaps because the cases studied correspond to aggregations of operators with limited dimensions and densities.

¹¹ Two brief notes on business management. Firstly, business management should not be confused with private management, in any way used. This generates discussions about the corresponding option, while the defense of the former is generalized by those who study the sector: academics, international organizations or experts. Secondly, a note on national legislation and the importance that is given to the sector's business-oriented management. There is virtually no legal diploma on the subject in which this reference does not exist. Moreover, the Administrative Code of 1936, in its article 146, referred to the possibility of municipalities creating Municipal services to organize “in industrial form” the activities of “capture, conduct and distribution of drinking water” at their responsibility. Seen in: <https://www.fd.unl.pt/Anexos/Investigacao/2195.pdf>.

¹² This is noteworthy, as there is a permanent tendency to impute to the utility resulting from the aggregation the responsibility for tariff increases. However, that reasoning omits whether the average tariff previously applied by the municipalities alone was much lower than the required one, resulting in a tariff deficit which increases year after year.

¹³ The analysis of the relationship between increased tariffs and election results deserves a detailed analysis, because it is not clear the existence of a direct relationship between them. Analyzing only what happened in Portugal, there is a discrepancy between, for example, the extreme cases of Barcelos and Paços de Ferreira, in which it was evident the impact of tariff increases on the loss of elections of the mayors responsible for the increases. There are other cases, where the increases were politically well prepared and explained: in the Aveiro region, the tariff increases following the creation of ADRA did not have any evident electoral effect; or Santa Maria da Feira, in which, despite the significant increases in tariffs during the first years of the concession, there was also no obvious effect over the election's result; or even Águas do Ribatejo, in which there was a very significant increase only in some municipalities, which had initially very low rates, with subsequent re-election of all executives who joined the aggregation. In any

4. As mentioned before, the size, although necessary, is not sufficient for the success of the aggregation. It is necessary to minimize transaction costs and, therefore, to not go beyond the point at which the reduction of production costs is exceeded by the increase in transaction costs. This means that the advantages of the dimension are limited. Then, the central issue is to ensure better management, not to grow without need or with limited need for a larger number of customers, greater sales volume or increase in the number of workers. What is intended is not a mere exercise of addition of municipalities, but that of creating a new organization: quantity has to be accompanied by quality. In other words, it is essential that the transition to a new management paradigm – with the characteristics already mentioned of autonomy, professionalization, sustainability, innovation, among others – follow the extension of the geographical scope of action of the utility.

5. It is well known that the problems that the sector has to overcome are increasingly complex. Also well-known are the technical, financial and human weaknesses of most of the operators involved – as we have seen, even after aggregation, are generally very small and have limited human and financial resources – and therefore the choice of the most appropriate management model is decisive for the success of the aggregation. In Portugal, there are different types of aggregations, larger and smaller, with and without anchors, endorsed with greater or lesser tradition of professionalization, etc. However, it is especially in cases where the shortcomings, mostly technical and human, are more evident, that the way in which aggregation is designed is more important, so that the operator resulting from the aggregation is endorsed with the necessary means to meet the short, medium and long term challenges. It is crucial, decisive and imperative to ensure that aggregations have the necessary competences. However, in many of the cases we have examined, it does not seem credible that this can be achieved only through the use of municipal capacities. Without municipalities finding partners that add value to the new utility, be it of any type, public or private, the success of the aggregations is not guaranteed.

case, nothing can be concluded and this question deserves to be studied, namely to understand the reasons that explain these differences in behaviour.

REFERENCES

- Abott M.; Cohen B. (2009) Productivity and efficiency in water industry. *Utilities Policy*, 19, 233-244.
- APDA (2018) Water and Sanitation in Portugal: The Market and Prices. Comissão Especializada de Legislação e Economia (CELE), APDA.
- Baptista, J. M.; Matos, R. (1994). O Saneamento básico em Portugal. In LNEC (1994) *Instrumentos de Apoio a Uma Política de Desenvolvimento Sustentável em Saneamento Básico*, Vol. I, Lisboa, LNEC.
- CELE/APDA (2018) Água e Saneamento em Portugal: O Mercado e os Preços. APDA, Lisboa.
- Ferro, G. (2017) Literature Review- Global Study on the Aggregation of Water Supply and Sanitation Utilities. Washington, DC, World Bank.
- Hull, M. (2013) *Infrastructure & Utility Economics*. Professional Economics Limited, London.
- Mendonça, J. T. (2020) *O Que é Amar um País: O Poder da Esperança*. Quetzal, Lisboa.
- Pato, J. H. (2011) *História das políticas públicas de abastecimento e saneamento de águas em Portugal*. ERSAR.
- Wenban-Smith, H. (2006) *Urban infrastructure: Density matters, not just size*. Retrieved from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.522.7455&rep=rep1&type=pdf>
- Wenban-Smith, H. (2009) *Economies of scale, distribution costs and density effects in urban water supply: A spatial analysis of the role of infrastructure in urban agglomeration*. PhD thesis. London School of Economics and Political Science (LSE).
- World Bank Group (2017) *Joining Forces for Better Services: When, Why, and How Water and Sanitation Utilities Can Benefit from Working Together*. Washington, DC, World Bank.
- XIV Governo Constitucional (2000) PEAASAR-Plano Estratégico de Abastecimento de Água e Saneamento de Águas Residuais 2000-2006.
- XV Governo Constitucional (2014) PEAASAR-Plano Estratégico de Abastecimento de Água e Saneamento de Águas Residuais 2007-2014, 2007 XIX Governo Constitucional.

I
IMPRESA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS
U

• U



C •