IMPRENSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS

ASSESSMENT METHODOLOGIES

ENERGY, MOBILITY AND OTHER REAL WORLD APPLICATION

PEDRO GODINHO
JOANA DIAS
EDITORS

The use of assessment methodologies is becoming increasingly important in the decision making process under very different environments. There are increasing concerns about making the best possible use of the limited available resources. There is also an increasing will of justifying the decisions in a sound way. This is even more important when considering public spending, public scrutiny and governance accountability. The use of recognized assessment techniques is actually mandatory in many situations, considering environmental, social, economic and sustainability concerns.

This book gathers contributions from several authors, advancing the state-of-the-art and presenting examples of applications of assessment methodologies in energy, mobility and other real world applications.

This book appears in the context of a workshop organized under project EMSuRE - Energy and Mobility for Sustainable Regions. EMSuRe is a project from the University of Coimbra, funded by QREN, Mais Centro and the European Union, which started in 2013 and will end in 2015.

I N V E S T I G A Ç Ã O

IMPRENSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS

Coordenação editorial
Imprensa da Universidade de Coimbra
Email: imprensa@uc.pt
URL: http://www.uc.pt/imprensa_uc
Vendas online: http://livrariadaimprensa.uc.pt

Concepção gráfica
António Barros

Infografia da Capa
Carlos Costa

Infografia
PMP

Execução gráfica
Norprint - a casa do livro

ISBN
978-989-26-1038-2

ISBN DIGITAL
978-989-26-1039-9

DOI
http://dx.doi.org/10.14195/978-989-26-1039-9

Depósito legal
399997/15

Obra publicada com o apoio de:

mais CENTRO
Programa Operacional Regional do Centro

QREN QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL PORTUGAL 2007.2013

UNIÃO EUROPEIA
Fundo Europeu
de Desenvolvimento Regional

IMPRENSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS

ASSESSMENT METHODOLOGIES

ENERGY, MOBILITY AND OTHER REAL WORLD APPLICATION

PEDRO GODINHO
JOANA DIAS
EDITORES

Página deixada propositadamente em branco.

CONTENTS

This book appears in the context of the "Workshop on Assessment Methodologies – energy, mobility and other real world applications" (WAM 2015), organized under project EMSuRe - Energy and Mobility for Sustainable Regions. The editors are indebted to several persons whose assistance made this book possible.

We wish to thank the leader of project EMSuRe, Professor António Gomes Martins, and the coordinator of the work package WP5 "Sustainability Assessment and Planning for Sustainability", Professor Pedro Nogueira Ramos, who trusted the editors to conduct the process that led to this book, giving them all the assistance they could have wished for throughout all stages of this process.

The members of the Scientific Committee of the WAM 2015 provided invaluable assistance in several important scientific decisions concerning the workshop and the book. All the referees who accepted to review the submitted papers were also of fundamental importance to make this book possible.

WAM 2015 benefited from the conditions provided by the ITeCons bulding. We wish to thank the ITeCons Director, Professor António Tadeu, who graciously allowed us to use the building, and also the ITeCons staff, who provided invaluable assistance in the local organization of the workshop.

The success of WAM 2015 was of huge importance to allow the publication of the book. We want to thank all those persons who helped organize the workshop and the participants in WAM 2015, in particular those who organized sessions in the workshop.

We also want to thank Imprensa da Universidade de Coimbra, in particular Maria João Padez, for supporting us in the publication process.

Finally, the authors are the most important actors in the book. We thank all the authors who submitted papers to this book, not only those with accepted papers but also those whose papers we were not able to include in this book.

This book was supported by project EMSuRe - Energy and Mobility for Sustainable Regions (CENTRO-07-0224-FEDER-002004).

PREAMBLE

Pedro Godinho and Joana Dias[1]

The use of assessment methodologies is becoming increasingly important in the decision making process under very different environments. There are increasing concerns about making the best possible use of the limited available resources. There is also an increasing will of justifying the decisions in a sound way. This is even more important when considering public spending, public scrutiny and governance accountability. The use of recognized assessment techniques is actually mandatory in many situations, considering environmental, social, economic and sustainability concerns.

This book gathers contributions from several authors, advancing the state-of-the-art and presenting examples of applications of assessment methodologies in energy, mobility and other real world applications.

This book appears in the context of a workshop organized under project EMSuRE - Energy and Mobility for Sustainable Regions. EMSuRe is a project from the University of Coimbra, funded by QREN, Mais Centro and the European Union, which started in 2013 and will end in 2015.

Energy has been an important global concern, given its crucial role in all kinds of activities. The depletion of traditional non-renewable

[1] University of Coimbra, Faculty of Economics
E-mails: pgodinho@fe.uc.pt, joana@fe.uc.pt

energy sources has led to the search for alternatives and for more efficient energy usage, and to the development of renewable energy sources. An integrated management of the supply-side and the demand--side is fundamental for achieving sustainable development strategies that can be kept in the long term without the risk of running out of energy sources.

Energy is strongly intertwined with mobility, with transportation systems accounting for a significant share of energy consumption across the world. Therefore, it is important to address energy and mobility strategies in an integrated fashion, taking into account their mutually influencing aspects. Energy and mobility strategies should contribute to sustainable development, avoiding environmental damage and resource depletion without impeding progress.

Energy and mobility concerns are relevant not only at a global level but also at a regional level. Efficient energy and mobility strategies may lead to important advantages to specific regions, and be drivers to a fast and sustainable development.

The holistic perspective that energy and transportation systems should be addressed in combination, giving a strong focus to the regional level, is at the inception of the EMSuRe project. The assessment of solutions, strategies and models, from different perspectives, is an important component of EMSuRe and, in fact, one of the six work packages of the project concerns “Sustainability assessment and planning for sustainability”.

The workshop had the main objective of bringing together researchers working in the broad area of project appraisal and impact assessment methodologies. The workshop received contributions from researchers and practitioners both from within and from outside of the EMSuRe project. The assessment methodologies considered in the workshop included not only applications to energy and mobility, but also other real world applications. Nearly 50 contributions were submitted to this workshop, and the present book is the result of a peer-review process that led to the selection of 15 full articles.

The chapters of this book reflect the diversity of areas in which assessment methodologies play an important role, starting with the core topics of the EMSuRe project – energy and mobility – and extending to very different areas, like health and happiness.

In Chapter 1, Clara Bento Vaz and Ângela Ferreira assess the technical efficiency of 26 European countries between 2009 and 2013, evaluated through the mix of energy consumption, labour force and gross fixed capital formation on the Gross Domestic Product of each economy. The authors resort to a Data Envelopment Analysis framework, which is complemented by bootstrapping to obtain statistical inferences. The authors also explore the effect of the European Union (EU) targets regarding energy efficiency, renewable energy share and the greenhouse gas emissions. They conclude that the efficiency of the economies increased approximately 13%, on average, since 2009, and that the efforts regarding the EU energy policy targets have not threatened the improvement of the efficiency.

In Chapter 2, Patrícia Pereira da Silva, Blanca Moreno and Ana Rosa Fonseca use econometric techniques to analyse the impact of the European Union emission trading system on stock market returns. The emission trading system has been implemented in three phases, and the authors focus on the first year and a half of the third phase, from January 2013 until the end of July 2014. In this third phase, emission allowances were auctioned. The authors analyse the interactions between the prices of emission rights and the stock returns of different sectors of the Spanish market. They conclude that there are positive long-run impacts on the power, cement and petroleum sectors, and negative impacts on the iron and steel sectors.

In Chapter 3, Nuno Carvalho Figueiredo and Patrícia Pereira da Silva assess some determinants of the integration level of the South West Europe regional electricity spot markets. The authors consider the French, Spanish and Portuguese markets, and analyse the importance of interconnection capacity and two weather variables: temperature and wind. The authors apply a vector autoregression model, concluding that the variables related to interconnection capacities do not improve the model, but average temperature slightly contributes to improve the model and wind speed shows a significant negative influence in spot electricity prices. Strong integration was found between the Portuguese and Spanish markets, leading the authors to conclude that the existing market setup is efficient and contributes to the integration of spot electricity markets. A weak integration level was found between the Iberian market and the French market.

In Chapter 4, Guillermo Ivan Pereira and Patrícia Pereira da Silva present a governance analysis of energy efficiency in European Union member states. The authors consider several country-level indicators, and perform both a European Union-wide analysis and a global ranking of the member states. The authors conclude that more efforts are required to ensure that member states follow and adopt existing legislation, alongside with developments on the existing financial support mechanisms, human capacities and institutional structures. Additionally, governance performance is shown to vary widely from member state to member state.

Chapter 5 combines energy and transportation issues, addressing the identification of optimal strategies for developing batteries for electric vehicles. Joana Fialho, Pedro Godinho and João Paulo Costa consider the case in which multiple levels of resources may be used to undertake the development process, and the level of resources may change according to the way the task is developing. The authors then use a procedure that aims at identifying the strategies that maximize the net present value of the cash flows associated to the task. The approach used by the authors allows them to assess and, at the same time, determine the best strategy, using a procedure combining dynamic programming with Monte Carlo simulation. The procedure is used in several scenarios, allowing the definition of general rules that may help managers in their planning activities.

In Chapter 6, Joana Dourado and Ana Bastos Silva address transportation issues, namely the assessment of the impact of traffic calming schemes focused on the reduction of traffic fatalities. The authors propose the use of microsimulation, which allows the assessment of the impact of such schemes without disturbing the real system. Microsimulation-based models enable the representation of the reality with high detail, simulating the real network conditions and developing a perform analysis with computer representations. The authors apply these concepts to a case study located in a Coimbra neighbourhood. A microsimulation model is used to evaluate the effects of traffic calming schemes and a methodology for implementing such measures in residential areas is presented. The evaluation of the results shows that the application of traffic calming measures would not generate a significant increase of the inter-zonal travel times, achieving however significant reductions in through traffic volumes and speeds within the case study area.

The growing importance of intelligent transportation systems based on automated data collection frameworks is the focus of Chapter 7. Jorge Freire de Sousa, João Mendes-Moreira, Luís Moreira-Matias and João Gama analyse the performance indicators used to evaluate schedule plan reliability. The authors start by presenting a brief review on improving the network definition based on historical location-based data and then review on automated data collection-based evaluation techniques of the schedule plan reliability, discussing the existing metrics. The authors undertake a critical analysis of the performance indicators used in the evaluation of the schedule plan reliability, providing a contribution to shape the approaches developed by the research community for improving the quality of public road transportation operations based on data collected by automated data collection systems.

In Chapter 8, Sandrina Filipe, Pedro Godinho and Joana Dias analyse the impact of macro-economic uncertainty in the economic evaluation of road infrastructures. The authors consider a road infrastructure awarded by the Portuguese government to a private consortium in 2008, the construction of which was based on economic analysis made on 2008. Then the authors analyse what would be changed in the 2008 analysis if it would have been possible to foresee the macro-economic changes that occurred afterwards, updating the analysis according to the information available at the beginning of 2015. The authors conclude that the changes in the macro-economic situation that were not foreseen in 2008 would be enough to change the recommendation produced by the economic evaluation, and they discuss how the methodologies could be improved in order to reach more robust decisions.

Chapter 9 applies assessment methodologies to health decision making. Patrícia Antunes, Pedro Lopes Ferreira and Lara Noronha Ferreira consider preference based measures applied in the health context. The authors argue that the preference based measures currently in use require further research in order to improve cost-benefit analysis and explore alternative preference elicitation methods for decision making. The authors describe several techniques: the Discrete Choice Experiment, the Visual Analogue Scale, the Time Trade-Off and the Standard Gamble. They argue that comparative studies show that Discrete Choice Experiment is particularly promising, but this technique still presents challenges and requires improvements.

Chapter 10 considers a very different type of assessment: the assessment of happiness. Maria da Conceição Pereira and Daniel Martins analyse the effect of education on subjective well-being, using data from the European Social Survey for Portugal. A large set of mediating variables is considered, in order to test the channels through which education affects subjective well-being: higher lifetime earnings, higher professional status, lower risk of unemployment, higher social capital and better health. The authors conclude that education does not exert a direct effect on well-being, but it influences well-being through most of the mediating variables. Additionally, they find that secondary education provides a wider range of benefits than higher education.

In Chapter 11, Stephan Printz, René Vossen and Sabina Jeschke consider the profitability estimation focused on benefits (PEFB)-method. This method is based on a holistic, participation-oriented approach for the evaluation of costs and benefits of an investment, integrating long term non-monetary aspects and also mapping the chronological sequence of an investment within the organisation's target system. The authors consider PEFB with regard to the conceptual design and its application to strategic investment decisions. Within this framework, a survey on 110 participants is conducted, identifying a structural deficit in the conception of the methodology regarding the estimation of probabilities due to personal affection. Based on this survey, some recommendations for application of the PEFB method are made.

In Chapter 12, Pedro Ramos, Luís Cruz, Eduardo Barata, André Parreiral and João Pedro Ferreira apply input-output analysis in the assessment of the impacts of different kind of shocks on Portuguese (and regional) industry's outputs and other correlated variables. The authors build a bi-regional model that divides the country into two regions: the NUT II Centro of Portugal and the rest of the country. The model is based on a general flexible procedure, termed MULTI2C. The authors use rectangular matrices with 431 products and 134 industries, considering also different types of households according to their main source of income. The major conclusion is that the shock effects do not confine to the place where they hit, so it cannot be assumed that all the benefits or costs of a regional occurrence take place in that same region. In the case analysed by the authors, a redistribution of the income generated in the Centro region, transferring it from the households that live mainly from labour income to the firms, has an expansionist effect in the Portuguese economy, but a discrepant effect on regions: the Centro region incurs a burden while the Rest of the Country reaps the benefits.

Input-output analysis is also the focus of Chapter 13. Carmen Ramos analyses the consequences of sectorial aggregation, considering two perspectives: the effects of aggregation on the amount of information contained in an input-output matrix, and the changes caused in the output multipliers. The results thus obtained are then applied to the input-output table for Asturias, corresponding to the year 2010. The author concludes that the relative loss is negative for some aggregations and that the pure linkage value does not vary when sectors are aggregated.

Chapter 14 addresses the use of social accounting matrixes (SAMs). Susana Santos uses such matrices and the underlying empirical and theoretical description of the socio-economic activity of a country to provide the methodology for the study of the distribution and use of income. A SAM for Portugal in 2011 is constructed, using the National Accounts and adopting its nomenclatures. Complemented with data for 2010, the structural features of the distribution and use of income in Portugal are identified with indicators for the functional and institutional distribution of generated income and for the distribution and use of disposable income.

An assessment of the macroeconomic effects of changes in income that can be captured with SAMs is made through an experiment, illustrated by an application for Portugal, with a change in the taxes on income and wealth paid by the households to the government.

Chapter 15 addresses the application of multi-criteria decision aid to the assessment of dairy farms sustainability. Sandra Silva, Luís Alçada-Almeida and Luís Dias consider dairy farms from the Entre Douro e Minho region in Portugal and define a multidimensional sustainability concept encompassing seven environmental and seven socio-economic criteria. The criteria are evaluated with the aid of experts from different fields of knowledge. The multi-criteria classification method ELECTRE TRI is applied to sort dairy farms into three categories, with socio-economic and environment viability being separately considered. The results are then combined to summarize the farms sustainability according to four categories, and a visualization of the spatial behaviour of the classification is performed with the aid of a Geographic Information System.

This book does not have the pretension of bringing to the reader an exhaustive list of assessment methodologies and applications. Considering the broad range of applications, and the diversity of existing methodologies, that would in fact be an unachievable goal. This book will, however, be an important repository of information both for acquainted readers, who will find contributions going beyond the state-of-the-art, and also for readers that are taking the first steps in this interesting field of applied research, bringing to their knowledge different applications, methodologies and points of view.

Measuring Technical Efficiency of European Countries using Data Envelopment Analysis

Clara Bento Vaz[1] and Ângela Paula Ferreira[2]

Abstract

This study proposes a Data Envelopment Analysis (DEA) framework to assess the technical efficiency of 26 European countries in the last five years, under the ongoing 2020 energy policy. DEA is used to estimate efficiency which is complemented by bootstrapping to obtain statistical inferences. Further, we explore the relationship between the targets regarding energy efficiency, renewable energy share and the greenhouse gas emissions and, in addition, the electricity prices derived from the energy system on the efficiency levels of European countries through a panel data truncated regression with bootstrapping. It is observed that the bias-corrected efficiency of the economies increased approximately 13%, on average, since 2009. The results achieved bring into view that the efforts regarding the energy policies developed in each country to follow 20-20-20 targets, have not threatened the improvement of their efficiency.

1. Introduction

In the last decades, the European Union (EU) has been promoting an integrated approach to climate and energy policy, aiming combat climate change, increase the EU's energy security and strengthen its competi-

[1] Polytechnic Institute of Bragança, Campus de Santa Apolónia, Apartado 1134, 5301-857; CGEI/INESC TEC – INESC Technology and Science, Portugal, e-mail: clvaz@ipb.pt

[2] Polytechnic Institute of Bragança, Campus de Santa Apolónia, Apartado 1134, 5301-857, Bragança; CISE - Electromechatronic Systems Research Centre, University of Beira Interior, Covilhã, Portugal, e-mail: apf@ipb.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_1

tiveness. There is also a geostrategic objective, which aims to reduce excessive dependence on imported fossil fuels. In fact, the long-lasting instability in many fossil fuel-producing countries increases the price of energy and reinforces the need to find alternatives. As a consequence, a paradigm shift has arisen in the global energy sector, foreseeing a sustainable and environmentally friendlier development. The climate and energy package in EU has been set through binding legislation, which establishes the following three key objectives for 2020, known as the 20-20-20 targets (EU, 2009):

- 20% reduction in greenhouse gas emissions from 1990 levels;
- Raising the share of energy consumption produced from renewable resources to 20%;
- A 20% improvement in the energy efficiency.

Within this context, the main objective of this work is to determine the technical efficiency of EU countries for the last five years with available data, *i.e.*, from 2009 till 2013. The level of the technical efficiency is evaluated through the mix of energy consumption, labor force and gross fixed capital formation (measuring the value of fixed assets) on the Gross Domestic Product (GDP) maximization of each economy.

From an economics point of view, each economy is efficient if it increases its GDP by decreasing the used resources, through technological, behavioural and economical changes. In the scope of energy efficiency, it encompasses the energy reduction for a given service or level of activity (WEC, 2008), *i.e.*, each economy should promote the energy efficiency by eliminating the redundant energy consumption. In terms of environmental perspective, it is important that the growth of the economy is based on the sustainable use of resources to produce outputs.

Additionally, this work also explores factors concerning the penetration of renewable energy in total energy consumption, the greenhouse gas emissions and the electricity prices on the efficiency estimates.

The technical efficiency of the economies under analysis is evaluated through Data Envelopment Analysis (DEA). The potential of DEA in energy studies has been gradually investigated by researchers due to its ability in combining multiple factors, being an important method for benchmarking in the energy sectors, particularly in the electricity area (Zhou, *et al.*, 2008). DEA allows the identification of the countries with best practices, from which linear combination defines the best frontier technology. By reference to this frontier, the technical efficiency for each country is determined, through the use of the fundamental resources as energy, labor and capital to maximize the production of the GDP.

Since DEA is a deterministic approach, the construction of the best frontier can be affected by random noise or measurement errors in the data. To assure the robustness of the efficiency estimates, the DEA assessment is complemented with the appropriate bootstrapping to derive statistical inference. To explore the relationship between the efficiency estimates and the impact of the renewable energy penetration, the greenhouse gas emissions and the electricity prices, a panel data truncated regression is used.

The remainder of this paper is organized as follows. Next section describes the literature review and the following one presents the performance assessment methodology used. Fourth and fifth sections present the data used and the results achieved. Finally, last section summarizes the paper findings.

2. Literature review

In the energy literature, there is no consensus on the appropriate method for defining and measuring energy efficiency. It is common to proxy energy efficiency using a simple average indicator such as the ratio of energy to GDP or the reciprocal (Khademvatani, Gordon, 2013). This traditional energy efficiency indicator takes energy consumption into account as a single input that produces an economic output; therefore, some other key inputs are ignored, such as capital and labor (Zhang, *et al.*, 2011). Energy consumption must be combined with other inputs to produce an economic output, such as labor or capital. Boyd and Pang (2000) argue that energy efficiency improvement relies on total factor productivity improvement. To overcome the disadvantage of the partial energy efficiency indicator, an increasing number of researchers have adopted the DEA for the assessment of total factor energy efficiency in different countries from the viewpoint of technical efficiency. Some examples of such studies are (Hu, Kao, 2007), (Hu, Wang, 2006), (Chien, Hu, 2007), (Zhang, *et al.*, 2011), (Zhou, *et al.*, 2012) and (Menegaki, 2013). To consider statistical noises, Zhou, *et al.* (2012) present a parametric stochastic frontier analysis (SFA) approach to measure total factor energy efficiency performance of OECD countries. The SFA method assumes that deviations from the efficient frontier are composed by inefficiency and statistical noise terms, however it is a parametric technique that requires a-prior specification of the functional form of the best practice frontier as opposed to non-parametric nature of DEA. Although there are many studies investigating the total factor energy efficiency, its relationship with the use of renewable energy sources has been studied by fewer authors, for instance (Zhou, *et al.*, 2012) and (Menegaki, 2013). In order to fulfil some gaps existing in the literature, we assess the total factor energy efficiency of EU economies, investigating the relationship between their efficiency and the penetration of renewable energy sources. This analysis is performed through the suitable bootstrap framework (Simar, Wilson, 1998; 2007) to obtain statistical inferences on efficiency estimates derived from DEA models.

3. Performance assessment methodology

The total factor energy efficiency for each economy is evaluated through the technical efficiency derived from the DEA model, introduced by Charnes, *et al.* (1978). DEA is a non-parametric approach to assess the relative efficiency of a homogeneous set of Decision Making Units (DMUs) in producing multiple outputs from multiple inputs. DEA is used to assess the technical efficiency of the European countries in producing GDP by taking into account the production factors. DEA allows the identification of the best practices DMUs and their linear combination defines the frontier technology which envelops all DMUs observed in the production possibility set (PPS). For the inefficient DMUs located inside the PPS, the magnitude of the inefficiency is derived by the distance to the frontier and a single summary measure of efficiency is calculated.

Consider a set of n DMUs j (j=1,...,n), each consuming m resources (inputs) x_{ij} (x_{1j},...,x_{mj}) to produce s results (outputs) y_{rj} (y_{1j},...,y_{sj}). For an output maximizing perspective and assuming the most productive frontier observed, defined by constant returns to scale, the relative efficiency of the assessed DMU_o can be evaluated using the linear programming model (1):

$$Max\left\{h_o | x_{io} \geq \sum_{j=1}^{n} \lambda_j x_{ij}, i = 1, \ldots, m,\ h_o y_{ro} \leq \sum_{j=1}^{n} \lambda_j y_{rj}, r = 1, \ldots, s, \lambda_j \geq 0\right\} \quad (1)$$

The optimum solution of model (1), h_o^*, corresponds to the maximum factor by which the outputs levels can be expanded with the current level of resources. The relative efficiency (θ) of the assessed DMU_o is estimated by the reverse of h_o^*. DEA enables to identify the efficient DMUs which have the best practices and the inefficient units which activity can be improved. The efficiency measure is equal to 100% when the unit under assessment is efficient, whereas lower scores indicate the existence of inefficiencies. For inefficient units, it is also possible to obtain, as by-products of the DEA efficiency assessment, a set of targets for becoming efficient, which are feasible points observed on the frontier.

To correct the DEA efficiency estimates for bias, the bootstrapping method is used according to Simar and Wilson (1998), which is suitable for use with DEA efficiency estimates, ranging from 0 to 1. Efron (1979) proposed bootstrapping based on the idea of resampling from an original sample of data to derive replicate datasets from which statistical inference can be performed. Resampling directly from the original data (naïve bootstrap) provides a poor estimate of the data generating process (DGP) as the efficiency estimates computed by model (1) are truncated, with upper value equal to one and there may exist several estimates equal to unity. Simar and Wilson (1998) proposed the smoothed bootstrap method suitable to DEA, estimating the original densities of the non-parametric efficiency scores using kernel smoothing methods combined with a reflection method (Silverman, 1986) by mimicking the DGP. This procedure was implemented using the statistical software R including the FEAR library, developed by Wilson (2008). Thus, for each DMU is derived a confidence interval for θ, the bias and the bias-corrected efficiency, $\hat{\theta}$. These scores are used to assess the countries performance.

To explore the factors that can be associated with good efficiency levels of the economies we use the bootstrap-truncated regression formulated according to the double bootstrap method (algorithm #2) proposed by Simar and Wilson (2007), in which efficiency scores are bootstrapped in the first stage, as explained before, and then the second step is performed based on the bootstrap truncated regression. This approach is used to investigate the penetration of renewable energy into the total energy consumption, the greenhouse gas emissions and electricity prices on efficiency levels of countries. Additionally, a panel data truncated model controlling for time effect is used. Considering the country j (j=1,...,n) in the time period t (t=1,...,m), the impact of the regressors, defined by variables z_{jt} on efficiency score θ_{jt}, is assessed by the following model:

$$\theta_{jt}=\alpha_o+\delta_t+z_{jt}\beta+\varepsilon_{jt} \quad (2)$$

where δ_t is a vector of dummy variables for each year, α_o denotes the intercept and β corresponds to the vector of regression coefficients to be estimated and ε_{jt} is the error term with a $N(0,\sigma_\varepsilon^2)$ distribution with a truncation at $(1-\alpha_o-\delta_t-z_{jt}\beta)$. The θ_{jt} corresponds to the efficiency of country j, in year t, estimated using model (1) and corrected by bootstrapping.

4. Data description

This study considers the panel dataset of EU-28 countries, for which data ranges from 2009 till 2013. Data on labor, energy consumption, capital and GDP are collected from the Eurostat database. The GDP is defined at market prices, in Purchasing Power Standard (PPS) per inhabitant. Total energy consumption encompasses gross inland energy consumption by renewable and nonrenewable energies, in tons of oil equivalent (toe) per inhabitant. Capital is the real gross fixed capital formation at market prices, in PPS per inhabitant. Labor is assessed by the employment rate, representing employed persons as a percentage of the labor force which is the total number of people employed and unemployed (persons aged 15 to 74). Before proceeding with the analysis, the existence of outliers is checked as the location of the best--practice frontier is sensitive to extreme observations. From the original 28 countries, based on the available data, 2 were considered outliers (Luxembourg and Irland). Additionally, some data about Greece is not available in 2013, and, consequently excluded from the assessment. Thus, the final sample includes 129 observations, from 26 countries, excluding data from Greece in 2013 year.

Descriptive statistics of the variables across countries, in the corresponding year, are summarized in Table 1.

Table 1: Descriptive statistics of the data under analysis.

	Capital (PPS)		Labor (%)		Energy (toe)		GDP (PPS)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
2009	4253.8	1106.1	0.91	0.04	3.35	1.13	21069.2	6260.1
2010	4230.8	1134.1	0.90	0.04	3.47	1.30	21961.5	6510.4
2011	4357.7	1222.2	0.90	0.04	3.37	1.21	22646.2	6513.3
2012	4326.9	1272.0	0.89	0.05	3.31	1.18	23080.8	6460.0
2013	4264.0	1223.0	0.89	0.05	3.29	1.25	23344.0	6485.6

In the time span under study, the mean GDP per capita raised about 2% while energy per capita and labor usage decreased by 1.7% and 1.9%, respectively, and the capital remained almost unchanged. In terms of standard deviations, the higher values are observed with respect to capital, energy used and GDP, which show that the countries are heterogeneous. The highest difference among countries is observed on the amount of total energy used per capita which implies that European countries may have different energy strategies to support their economy.

5. Empirical results: DEA efficiency assessment

The technical efficiency for each country is assessed by model (1), evaluating the capacity of each economy in maximizing the GDP produced taking into account the fundamental inputs (energy, labor and capital). The relative efficiency of a country in a given year is estimated by comparison to the best practices observed during the period analysed, ranging from 2009 to 2013 years. Based on the inputs-output mix of the model, the CRS and VRS frontiers are very close, and the CRS assumption is considered in the DEA model (Dyson, *et al.*, 2001). These efficiency estimates provide insights into potential improvements by taking into account statistical inference derived through bootstrapping framework.

The correction of the DEA efficiency estimates for bias, has been performed by using 2000 bootstrap samples. Table 2 summarizes results for the technical efficiency, bias-corrected efficiency, standard error and bias. Bias-corrected efficiencies reveal that magnitude of the corrected efficiencies are slight lower than the original efficiencies, although this variation is small. As the absolute value of bias estimates are larger than the standard error estimates for each country, the bias-corrected efficiency estimates are preferred to the original efficiencies, since they represent a more precise estimate of the true efficiency.

It is possible to conclude that since 2009 the bias-corrected efficiency of the economies increased approximately 13%, on average, which is a relevant improvement. These results suggest that efforts regarding the energy policies developed in each country to follow 20-20-20 targets have not threatened the improvement of their efficiency. Looking now at the distribution of efficiency scores across countries, given in Table 3, it is observed that each country follows the same improvement trend as described above. Considering as benchmarks the countries that have the bias-corrected efficiency equal or higher than 95%, Denmark was the single benchmark in 2010, Austria, Netherlands and United Kingdom became benchmarks in 2011, Sweden has been established as benchmark in 2012 and Germany became efficient in 2013.

Table 2: Results of original and bootstrapped average efficiency estimates.

Year	2009	2010	2011	2012	2013
Original eff. Score	75.9%	79.3%	81.5%	84.2%	86.0%
Bias-corrected eff.	74.0%	77.3%	79.5%	82.1%	83.6%
Bias	-3.5%	-3.5%	-3.2%	-3.2%	-3.4%
St.dev.	1.9%	1.9%	1.7%	1.7%	1.7%

Considering the efficiency average from 2009 till 2013, it is also observed a stability on the performance of Denmark, Austria, Germany, Netherlands, Sweden and United Kingdom (with average efficiency scores higher than 92%). Belgium, Cyprus, Spain, France, Italy, Finland, Malta and Portugal have an average performance, presenting an efficiency ranging from 82% to 89%. Cyprus has the highest increasing improvement, about 25%, followed by Portugal which efficiency has increased by 21%, over the period analysed. Finally, Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Poland, Romania, Slovakia, Croatia, Lithuania and Slovenia have an efficiency ranging from 51% and 75%. In this group, Lithuania has the highest efficiency (75%).

From this group, Bulgaria, Romania, Lithuania, Poland and Croatia have the highest increasing improvement in efficiency, over 20%. Bulgaria presents the lowest average efficiency, 51%, although it has improved the efficiency by 26% during the study period. Also noteworthy is the fact that Estonia, with an average efficiency equal to 60%, did not improve its efficiency in the considered time period.

Regarding the efficiency scores levels observed in the last year of the time period, 2013, we compare benchmarks with inefficient countries, using similar number of units in each group, *i.e.*, benchmarks are Denmark, Austria, Netherlands, United Kingdom, Sweden and Germany and inefficient countries under analysis are Bulgaria, Czech Republic, Estonia, Latvia, Poland and Romania. Figure 1 shows that inefficient countries have about 55% of the GDP per capita observed in the benchmarks, but the same reduction percentage is not observed on the resources: labor is similar for the two groups, but the inefficient countries require more percentage in terms of energy and capital, per capita, to produce GDP.

Table 3: Bias-corrected efficiency scores across countries and time span.

	Year					Average
	2009	2010	2011	2012	2013	
Austria	87%	90%	95%	97%	98%	93%
Belgium	83%	89%	90%	92%	92%	89%
Bulgaria	45%	49%	52%	53%	56%	51%
Cyprus	75%	78%	84%	91%	94%	84%
Czech Republic	62%	62%	64%	66%	68%	64%
Germany	88%	90%	92%	93%	95%	92%
Denmark	89%	96%	95%	96%	96%	94%
Estonia	59%	64%	61%	58%	59%	60%
Greece	80%	84%	86%	90%	[1]	85%
Spain	83%	84%	86%	90%	93%	87%
Finland	82%	85%	88%	88%	87%	86%
France	80%	82%	85%	86%	87%	84%
Croatia	63%	69%	71%	74%	77%	71%
Hungary	63%	66%	71%	73%	74%	69%
Italy	82%	83%	84%	89%	91%	86%
Lithuania	67%	74%	74%	80%	80%	75%
Latvia	61%	65%	68%	68%	71%	67%
Malta	85%	82%	91%	91%	95%	89%
Netherlands	90%	92%	95%	96%	97%	94%
Poland	59%	62%	66%	70%	72%	66%
Portugal	74%	79%	82%	88%	89%	82%
Romania	59%	59%	60%	65%	73%	63%
Sweden	86%	92%	94%	97%	98%	93%
Slovenia	66%	71%	75%	77%	78%	73%
Slovakia	65%	68%	66%	72%	74%	69%
United Kingdom	94%	94%	95%	95%	98%	95%

[1]Data is not available.

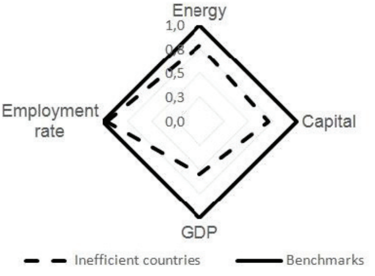


Figure 1: Benchmarks and inefficient countries comparison in 2013.

In order to investigate the implementation of the energy policy, concerning the attainment of 20-20-20 targets, imposed to EU countries, the following analysis explores the energy efficiency associated with economic efficiency and greenhouse gas emissions in terms of their average rate in the study period for the two groups previously defined (benchmarks and inefficient countries). The greenhouse gas emissions are measured by the ratio between energy related greenhouse gas emissions (carbon dioxide, methane and nitrous oxide) and gross inland energy consumption. Considering the results obtained, presented in Figure 2, it is possible to conclude that benchmarks have developed a higher effort in reducing greenhouse gases emissions and have decreased the amount of energy used to produce one unit of economic activity (energy used per unit of GDP). Thus, the DEA model reflects in great extent the energy efficiency promoted by European energy policy.

Regarding the other target, the penetration of renewables energies in the total energy consumption, based on the results for the latest year under analysis (2013 year), it is observed that both groups have developed similar efforts in increasing the renewable energy share of the total energy consumption (Figure 3). Specifically, there are found good and bad performers concerning the exploitation of renewable energy in both groups.

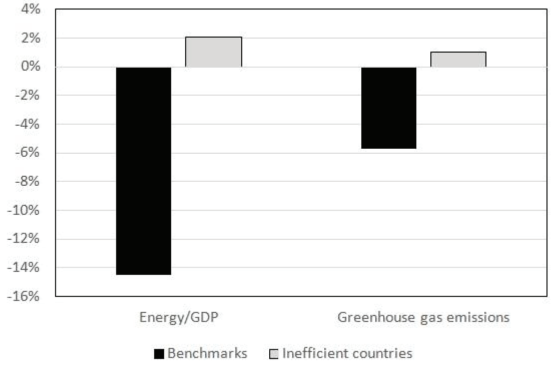


Figure 2: Comparison of the average rate of Energy efficiency and greenhouse gas emissions for benchmarks and inefficient countries from 2009 till 2013.

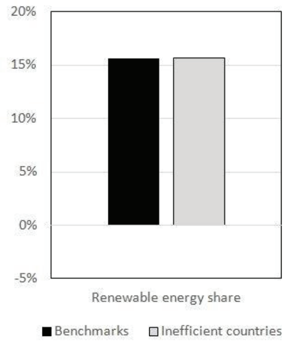


Figure 3: Renewable energy share for benchmarks and inefficient countries, in 2013.

Further, we explore the relationship between the targets regarding renewable energy share and the greenhouse gas emissions on the efficiency levels of European countries. This analysis also includes the electricity prices charged to final consumers measured by €/kWh. In some cases, because of financial constraints imposed by high energy prices, consumers may decrease their energy consumption through a reduction in their energy services. Such reductions do not necessarily result in increased overall energy efficiency of the economy. Nevertheless, energy prices assess an output of the energy system observed in each economy.

This evaluation is performed using truncated regression with bootstrapping (Simar, Wilson, 2007), formulated according to (2). The model uses the bias-corrected efficiency for country *i*, in year *t*, as the dependent variable, and as regressors variables related to amount of renewable energy used, the greenhouse gas emissions and electricity prices observed in each country. Due to the use of a panel data truncated model, the year is also used as dummy variable to control the time effect (the reference level is 2009 year).

Table 4: Truncated regression analysis results.

Variable	Coefficient	Std. Err.	p-value
Year 2010	0.040	0.045	0.379
Year 2011	0.072	0.048	0.130
Year 2012	0.101	0.052	0.054
Year 2013	0.125	0.055	0.022
Greenhouse gas emissions	-0.007	0.003	0.012
Electricity prices	0.093	0.342	0.786
Renewable energy usage	0.094	0.035	0.007
Constant	1.327	0.241	0.000

Table 4 summarizes the results from the panel data truncated model in terms of coefficients, standard errors and p-values. The total number of observations was 129. The truncated regression model is statistically significant (χ^2 Test with p-value equal to 0.0008), with a pseudo R^2 equal to 0.17. From the obtained results, it is possible to confirm that countries increased significantly the efficiency level from 2012 (at 10% level of significance). Apparently, the electricity prices increase with the efficiency level, although this effect is not significant. It is observed a negative and significant coefficient for the greenhouse gas emissions, meaning that the most efficient countries have lower levels of greenhouse gas emissions. The opposite occurs concerning the amount of renewable energy used which has a positive and significant impact on the efficiency of the economies. The most efficient countries use higher levels of renewable energy. Thus, it is possible to increase the share of renewable energy among total energy supply without deteriorating the technical efficiency of the economies, which is in agreement with the findings by Chien and Hu (2007).

6. Conclusions

This study proposes a DEA framework to assess the technical efficiency of 26 European countries from 2009 till 2013. In order to fulfil some gaps existing in the literature, a panel data truncated regression is used to explore the impact of the usage of renewable energy, the greenhouse gas emissions and the electricity prices on the efficiency estimates. This analysis is performed through the suitable bootstrap framework (Simar, Wilson, 1998; 2007) to obtain statistical inferences on efficiency estimates derived from DEA models.

It is possible to conclude that since 2009 the bias-corrected efficiency of the economies increased approximately 13%, on average, which is a relevant improvement. Three groups were identified taking into account the bias-corrected efficiency. The benchmarks are Denmark, Austria, Netherlands and United Kingdom, Sweden and Germany which have the bias-corrected efficiency equal or higher than 95%. The medium performers are Belgium, Cyprus, Spain, France, Italy, Finland, Malta and Portugal, presenting an efficiency ranging from 82% to 89%. The inefficient countries are Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Poland, Romania, Slovakia, Croatia, Lithuania and Slovenia with an efficiency ranging from 51% to 75%.

Comparing the benchmarks with the lowest inefficient countries, by using similar number of units in each group, we investigate the implementation of the energy policy, concerning the attainment of 20-20-20 targets, imposed to EU countries. It emerges clearly that benchmarks have developed a higher effort in reducing greenhouse gases emissions and have decreased the amount of energy used to produce one unit of economic output, since 2009. Therefore, the DEA model reflects in great extent the energy efficiency promoted by European energy policy. Regarding the penetration of renewables energies in the total energy consumption observed in 2013, it is observed that both groups have developed similar efforts in increasing the renewable energy share of the total energy consumption. Specifically, there are found good and bad performers concerning the exploitation of renewable energy in both groups.

Regarding the relationship between the efficiency levels of the EU economies and the targets on renewable energy share and greenhouse gas emissions, it is observed that the most efficient countries have lower levels of greenhouse gas emissions and higher renewable energy shares.

Concerning the effect of electricity prices on the efficiency estimates, results are not significant.

Based on the results achieved, we can conclude that the efforts regarding the energy policies developed in each country to follow 20-20-20 targets, have not threatened the improvement of its technical efficiency.

In future developments, we propose to extend the methodology to incorporate other countries, to assess their performance concerning the sustainable and environmentally friendlier development. Other perspective can be to model other factors than can affect the energy efficiency of the economies.

References

Boyd, Gale A.; Pang, Joseph X.; (2000). Estimating the linkage between energy efficiency and productivity. Energy Policy. Vol. 28(May 2000). pp. 289-296.

Charnes, A.; Cooper, W. W.; Rhodes, E.; (1978). Measuring the efficiency of decision making units. European Journal of Operational Research. Vol. 2, n.º 6. pp. 429-444.

Chien, Taichen; Hu, Jin-Li; (2007). Renewable energy and macroeconomic efficiency of OECD and non-OECD economies. Energy Policy. Vol. 35(July 2007). pp. 3606-3615.

Dyson, R. G., et al.; (2001). Pitfalls and protocols in DEA. European Journal of Operational Research. Vol. 132(July 16, 2001). pp. 245-259.

Efron, B.; (1979). Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics. Vol. 7, n.º 1. pp. 1-26.

EU; (2009). Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources. Official Journal of the European Union. pp. L 140/16: L140/62.

Hu, Jin-Li; Kao, Chih-Hung; (2007). Efficient energy-saving targets for APEC economies. Energy Policy. Vol. 35(January 2007). pp. 373-382.

Hu, Jin-Li; Wang, Shih-Chuan; (2006). Total-factor energy efficiency of regions in China. Energy Policy. Vol. 34(November 2006). pp. 3206-3217.

Khademvatani, Asgar; Gordon, Daniel V.; (2013). A marginal measure of energy efficiency: The shadow value. Energy Economics. Vol. 38(July 2013). pp. 153-159.

Menegaki, Angeliki N.; (2013). Growth and renewable energy in Europe: Benchmarking with data envelopment analysis. Renewable Energy. Vol. 60(December 2013). pp. 363-369.

Silverman, B. W.; (1986). Density Estimation for Statistics and Data Analysis. London: Chapman and Hall.

Simar, L.; Wilson, P. W.; (1998). Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models. Management Science. Vol. 44, n.º 1. pp. 49-61.

Simar, L.; Wilson, P.W.; (2007). Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics. Vol. 136(2007). pp. 31-64.

WEC; (2008). Energy Efficiency Policies around the World: Review and Evaluation. World Energy Council. Available from: http://www.worldenergy.org/

Wilson, Paul W.; (2008). FEAR: A software package for frontier efficiency analysis with R. Socio-Economic Planning Sciences. Vol. 42(December 2008). pp. 247-254.

Zhang, Xing-Ping, et al.; (2011). Total-factor energy efficiency in developing countries. Energy Policy. Vol. 39(February 2011). pp. 644-650.

Zhou, P.; Ang, B. W.; Poh, K. L.; (2008). A survey of data envelopment analysis in energy and environmental studies. European Journal of Operational Research. Vol. 189(August 16, 2008). pp. 1-18.

Zhou, P.; Ang, B. W.; Zhou, D. Q.; (2012). Measuring economy-wide energy efficiency performance: A parametric frontier approach. Applied Energy. Vol. 90 (February 2012). pp. 196-200.

Towards an Auction System in the Allocation of EU Emission Rights: Its Effect on Firms' Stock Market Returns

Patrícia Pereira da Silva[1], **Blanca Moreno**[2] **and Ana Rosa Fonseca**[2]

Abstract

The impact of the European Union Emission Trading System (EU ETS) on firms' stock market returns relies on the system employed to allocate emission allowances. This impact has been analysed in the literature for the EU ETS Phase I and II periods under which allowances were given for free. However, the effect during the current phase, Phase III, where the allocation of emission permits is based on an auction system has not yet been analysed and discussed.

In this framework, this paper discloses the results of a research aimed at investigating the interactions between the stock market returns of Spanish industry sectors under EU ETS and emission rights prices during the first year and half of Phase III. A cointegrated Vector Error Correction analysis is employed for the period covering January 1st 2013 until July 31st 2014.

The analysis presents statistically significant positive long-run impact of EU ETS on power sector, cement and petroleum and negative impact on iron and steel sectors. No short-run interactions were found for the sectors analysed.

Keywords: European Union Emission trading system; Energy costs; Vector error correction analysis

---

[1] Faculty of Economics - University of Coimbra
INESC Coimbra - Institute for Systems Engineering and Computers at Coimbra, e-mail: patsilva@fe.uc.pt

[2] Faculty of Economics and Business - University of Oviedo, e-mail: morenob@uniovi.es, arfon@uniovi.es

http://dx.doi.org/10.14195/978-989-26-1039-9_2

1. Introduction

The European Union Emission Trading System (EU ETS) has been implemented in three phases with different systems to give out emission allowances: Phase I, from January 2005 to December 2007, where emission rights were allocated for free; Phase II, from January 2008 to December 2012, where over 90 percent of the allowances were given out for free (usually referred as 'grandfathering'); and the current phase, Phase III, from January 2013 to December 2019, during which the allocation of emission permits are given out predominantly through auctions. As stated by Ellerman et al. (2014), the EU ETS is possibly the principal market-based application of economic principles in the climate domain and the largest cap-and-trade program yet implemented.

The introduction of an EU ETS instrument can produce impacts on the stock market returns, profitability and international competitiveness of companies covered by the EU ETS, such as: power plants; oil refineries; ferrous metallurgy; cement clinker or lime; glass including glass fibre; ceramic products by firing and pulp, paper and board.

In that sense, European Union carbon emissions allowances (EUA) price fluctuations can disturb companies' stock market value through cash flows and expected returns. Firstly, carbon dioxide prices could influence cash flows of companies as they can incorporate their carbon emission allowance costs in their sale offers. Therefore, a variation in pollution prices would be reflected in output prices as well as in costs. Secondly, Litterman (2013) and Pindyck (2013) indicate that carbon emissions generate "carbon risk" as they could lead to a climate disaster impacting the prosperity of next generations. Since polluting firms are exposed to carbon risk, they will require higher expected returns relative to non-polluting firms.

Therefore, the final effect of carbon emissions allowances prices on firms' profitability is ambiguous, as it depends on the ability of firms to pass to consumers the increase in marginal cost through higher output prices and the uncertainty about carbon risk. For instance, Sijm et al. (2006) indicate that pass through rates of German and Netherland power companies vary between 60 and 100 percent of carbon costs depending on the carbon intensity of the marginal production unit. In the same way, Ostreich and Tsiakas (2014) show that the initial two phases of EU ETS generated a carbon premium in stock returns of German firms explained by the higher cash flows due to the free allocation of carbon emission allowances and the carbon risk factor of firms with high carbon emissions.

The existing empirical studies do not converge to a shared position as many of the studies are country-region specific and results also rely on the modelling method, the period studied, the system used to give out the emission allowances, the market structure, the used econometric tool and scenarios analysis adopted impeding the generalization of their findings.

Some scholars have concluded that the EU ETS has had a positive effect on companies: Smale et al. (2006), Demailly and Quirion (2009), Goulder et al. (2010), among others suggest that, in many sectors, firms make net profits due to the impact on product prices combined with the free allocation of allowances during Phase I and II.

There is another view that the EU ETS has had little or no effect on companies: Anger and Oberndorfer (2008) indicate that the initial allocation of allowances did not affect revenue and the employment of German firms over the period of 2005–2006.

Moreover, the effect of the EUA prices on companies also depends on the economic sector studied but also considering the same sector results can be ambiguous: Bushnell et al. (2013) found that the decline in the price of carbon allowances had the highest negative impact on the stock returns of carbon-intensive industries; Smale et al. (2006) applying the Cournot representation of an oligopoly market to five sectors for UK (steel, aluminium, cement, newsprint, and petroleum) concluded that most participating sectors would be expected to profit; Demailly and Quirion (2009) discovered that profitability of EU steel makers raised under a 20 euro per ton CO_2 price and the amount of allowances allocated for free during Phase I, but Chan et al. (2013) concluded that EU ETS had no impact on the revenue performance of cement and iron and steel industries.

For the case of electricity sector, Oberndorfer (2009), Veith et al. (2009), Keppler and Cruciani (2010), Mo et al. (2013) and Chan et al.

(2013) found that EUA price changes and stock returns or revenue of the European electricity corporations are shown to be positively correlated. However, the particular effect of EUA price changes on electricity corporations´ stock returns varies across countries (Oberndorfer 2009 found a significantly small negative relationship for Spain), EU ETS phase (Mo et al. 2013 found a positive and negative correlation during phase I and II respectively) or power generation technology (Bode, 2006).

It has been showed that several studies have investigated this impact under EU ETS Phase I and II when the allowances were allocated for free. However, the effect during current Phase III, when the allocation of emission permits is given out predominantly in auctions, is still unknown. In 2013, EU ETS began its ninth year of operations after having progressed from a system with 25 national caps and decentralized allocation based on national allocation plans and dealing with CO_2 emissions alone towards a centralized system including several greenhouse gases (GHGs) and presenting an EU-wide cap indefinitely declining at an annual rate of 1.74%. Having entered Phase III a prevalence of free allocation has given way to a mixture of auctioning and free allocation based on benchmarking for sectors believed to be at risk of carbon leakage, with full auctioning for all sectors as the medium-term goal (Ellerman et al. 2014).

In the context of this debate, this study analyses whether and to what extent the EUA prices may be linked with polluting sectors' stock market returns in Spain during current Phase III. We use a cointegrated Vector Error Correction (VECM) model. Multifactor market models are widely used to study the effect of EU emission allowance prices (and others variables such as fuel prices and electricity prices) on corporate value change (Mo et al. 2013, Veith et al. 2009, Oberndorfer 2009). However, dynamic interactions among variables may play a fundamental role (see Paolella and Taschini 2008 and Keppler and Mansanet-Bataller 2010) so a multivariate analysis of simultaneous equations would avoid the endogeneity problems by treating all variables to be endogenous. Specifically, we use a cointegrated Vector Error Correction analysis, which allows the estimation of long-run equilibrium relations and short-run interactions between stock market returns and carbon emission prices.

The daily sample period used in our analysis ranges from the 1st of January 2013 to the 31st of July 2014.

The present study contributes to the literature on this topic in five ways: (*i*) to enrich the body of empirical literature on this matter, (ii) to offer a comprehensive empirical investigation of the effect of the EU ETS on stock returns during the current Phase III, (*iii*) to provide useful information to policy makers on which sectors show additional impacts on their competitiveness due to the European Emission Trading Scheme, (*iv*) to contribute with practical lessons for countries who are contemplating "cap and trade" systems; and (*v*) to provide findings that may be important for market investors.

The remainder of the paper is organized as follows. Section 2 presents a brief description of the methodology used including the multifactor model specification and the cointegrated Vector Error Correction analysis. Section 3 describes the data used. Section 4 presents the empirical results and VECM estimation. Finally, Section 5 concludes and discusses policy implications.

2. Methodology: A cointegrated Vector Error Correction Model

Multifactor market models are widely used to study the effect of any possible factor on corporate value change. In fact, Veith et al. (2009), Oberndorfer (2009) or Mo et al. (2013) have used a multifactor market model to investigate the impact of EUA price changes on firms' stock returns.

The basic model employed takes the following form:

$R_{it}= \beta_0 + \beta_1 R_{mt}+ \beta_2 P_t^{EUA} + u_{it}$ (1)

R_{it} represents the return on the stock index of the i^{th} sector or firm, R_{mt} the return of the market portfolio, P_t^{EUA} the price of EU emission allowances changes and u_{it} a disturbance term with $E(u_{it})=0$, $var(u_{it})=\sigma^2$.

Moreover other authors as Lee et al. (2012) or Moya-Martínez et al. (2014) include the long-term interest rate to incorporate market expectations.

Many empirical analyses indicate that stock returns are closely related to the price of oil (Lee et al. 2012, Moya-Martínez et al. 2014 for Spanish case) and gas (Acaravci et al. 2012) so other influencing factors such as fuel prices are included in the basic model. For example, Veith et al. (2009) include oil and natural gas prices as control variables and Oberndorfer (2009) also includes the electricity price in the regression equation.

We notice that electricity prices are very important for the Spanish industry, as electricity usually represents a significant proportion of the total energy cost for the industry (51.7% of the total energy consumption, according to The Energy Consumption Survey of the Spanish Statistical Institute, 2013).

The Energy Consumption[3] Survey offers data broken down for 96 activity sectors, (excluding power activity) which enables a detailed study of the industrial reality, as it provides information regarding which of the different types of fuel are the most significant in consumption.

The following Table 1 presents the 10 sectors with the highest energy consumption, representing more than 50% of the total consumption in the extractive and manufacturing industry. As it is showed, all the sectors included in the EU ETS belong to the sectors with the highest energy consumption.

Table 1: The 10 sectors with the most energy consumption (2013, thousands of euros).

Activity sectors	Consumption	% of the total
Manufacture of basic chemical products	1.390.747	12.5
Manufacture of basic products in iron, steel and ferroalloys	811.478	7.3
Production of precious metals and other non-ferrous metals	680.329	6.1
Manufacture of pulp, paper and cardboard	574.848	5.2
Petroleum and natural gas industries	533.319	4.8
Manufacture of ceramic products for construction	410.838	3.7
Manufacture of plastic products	403.323	3.6
Meat industry	377.215	3.4
Manufacture of glass and glass products	329.063	3.0
Manufacture of cement, lime and plaster	277.861	2.5
TOTAL	11.086.198	100.0

[3] Energy consumption is measured in monetary terms, at current prices. Therefore, its evolution considers both the evolution of the amounts consumed and the evolution of the prices of the different energy products.

Figure 1 represents the distribution of energy consumption, with electricity representing a very significant percentage of the total consumption for a large number of industrial sectors.

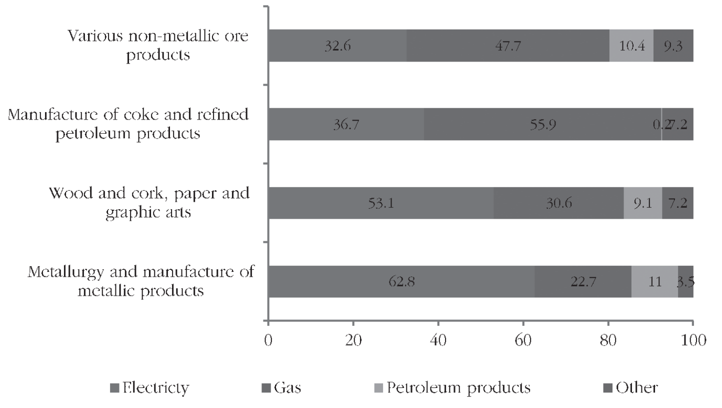


Figure 1: Percentage distribution, by type of energy and activity grouping (2013)

Moreover, gas it is also one of the most used fuels in industrial activity. In fact, belonging to the 10 sectors with the highest percentage of use of gas are those covered by EU ETS as it is shown in Table 2: manufacture of ceramic products for construction (75.0% of the total), manufacture of glass and glass products (58.7%) sectors, manufacture of ceramic products, except those used for construction (58.7%), petroleum and natural gas industries (55.6%).

Table 2: The 10 sectors with the highest percentage of use of gas year (2013)

Activity group	% use of gas
Manufacture of ceramic products for construction	75.0
Textile finishing's	61.7
Sugar, coffee, tea and infusions and confectionery	60.8
Manufacture of artificial and synthetic fibres	60.0
Manufacture of ceramic products, except those used for construction	58.7
Manufacture of glass and glass products	58.7
Manufacture of abrasive products and non-metallic mineral products n.e.c.	56.9
Petroleum and natural gas industries	55.6
Paints, varnishes, printing ink and mastics	52.6
Treatment and coating of metals	43.1

Regarding coal, a very residual use of it as a source of energy was observed, with the exception of the manufacture of cement, lime and plaster sector, where it represented 31.5% of the total energy consumption by companies practising that activity in 2013. Moreover, among the 10 sectors with the highest percentage of use of petroleum products is manufacture of elements made of concrete, cement and plaster (where petroleum represent the 42.5 % of the total energy consumption of this activity).

Regarding the energy used by power sector (which is not included in the Spanish Energy Consumption Survey), the last Spanish Energy Balance information published by International Energy Agency (2014), shows that electricity power station consumed 53 Mtoe of primary energy fuels. The percentage distribution of primary energy used by power stations is shown in Figure 2.

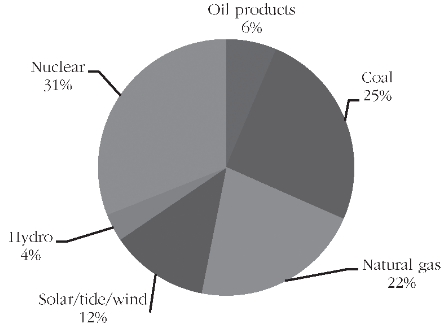


Figure 2: Distribution of primary energy used by power stations (2012).

Thus, the initial model can be specified as follows

$$R_{it} = \beta_0 + \beta_1 R_{mt} + \beta_2 P_t^{EUA} + \beta_3 P_t^{Oil} + \beta_4 P_t^{Gas} + \beta_5 P_t^{Coal} + \beta_6 P_t^{Elect} + \beta_7 r_t + u_{it} \quad (2)$$

P_t^{Oil}, P_t^{Gas}, P_t^{Coal} and P_t^{Elect} being the oil, gas, coal and electricity prices, respectively, and r_t the long-term interest rate.

The above multifactor model presumes the direction of causality, however dynamic interactions among variables may play a fundamental role thereby making estimation erroneous.

For instance, Paolella and Taschini (2008) have shown that oil and gas prices drive emission allowance costs: if, for example, the price of coal rises compared to gas, gas will be preferred by electricity producers, resulting in lower CO_2 emissions. Consequently, it follows a reduction of EUA demand and its price.

Moreover, the practice of marginal cost pricing in power generation implies that fuel and EUA prices play a fundamental role in the electricity price formation process through wholesale markets (see Freitas and Silva 2013, 2015 for the Iberian case) and consequently the final electricity prices for consumers (see Moreno et al. 2014 and Moreno and García-Álvarez 2013 for the Spanish case).

In the same way, Mansanet-Bataller et al. (2007) and Keppler and Mansanet-Bataller (2010) have analysed the interplay between daily carbon, electricity and gas price data with EU ETS concluding that fuel prices changes may be drivers of the EUA price itself.

Moreover, fuel and carbon emissions and fuel price fluctuations may also affect the interest rate. Rising energy prices are often indicative of inflationary pressures that central banks typically control by raising interest rates.

Therefore, when dynamic interactions among variables exist, multivariate analysis of simultaneous equations is the only technique that avoids the endogeneity problems by treating all variables (electricity price, fuel prices, EUA prices,...) as endogenous.

Multivariate analysis has been developed using either vector autoregressive models (VAR) or cointegrated models, which are also named Vector Error Correction models (VECM). These models offer a system of equations that expresses each variable in the system as a fraction of the lagged values of all the variables of the system- including its own lagged values.

As noted by Engle and Granger (1987), there are strong beliefs that economic data are non-stationary, meaning that any particular price measure over time will not be tied to its historical mean. So, modelling that kind of data by a levels VAR model appears to be inadequate, because of spurious regression risk, thus requiring one of two solutions: *i*) modelling a VAR in first differences which may impose the risk of losing relevant information about long-term relationships; *ii*) specifying a VECM, if the variables show a very interesting property, namely the cointegration. The latter alternative, if possible, has the advantage of allowing the simultaneous analysis of the long-run interactions and the short-term adjustments to the equilibrium relationship.

The cointegration concept, introduced by Engle and Granger (1987), means that individual economic variables may be non stationary and wander through time, but they are expected not to be completely independent of each other. That is, similar economic forces influence each variable and it is expected that the different variables will be tied together. In a more formal way, it is possible that two or more variables are non-stationary and wander through time, but a linear combination of them may, over time, converge to a stationary process. Such a process, if present, may reflect the long-run equilibrium relationship, and is referred to as the cointegration equation. According to Engle and Granger (1987), cointegrated variables must have an error correction representation in which an error correction term (ECT) must be incorporated into the model. Accordingly, a VECM is formulated to reintroduce the information lost in the differencing process, thereby allowing for long-run equilibrium as well as short-run dynamics.

Since the influential work of Engle and Granger (1987) several procedures have been proposed for testing the null hypothesis that two or more non-stationary time series are not cointegrated, meaning that there exist no linear combinations of the series that are stationary. One approach is to use likelihood ratio tests based on estimating a VAR. This approach was first proposed by Johansen (1988) and refined further by Johansen and Juselius (1990), Johansen (1991), Johansen (1992) and Johansen (1994). Johansen`s approach provides a unified framework for estimation and testing in the context of a multivariate VECM.

The cointegration test procedure specifies a VAR of order *k*, without imposing any restriction a priori, in the form of error correction model (ECM). Assuming the existence of cointegration, the data generating process P_t can be appropriately modelled as a VECM with *k-1* lags (which is derived from a levels VAR with *k* lags). Consider a VAR of order *k* with a deterministic part given by μ_t. One can write the p-variate process as $P_t = \mu_t + A_1 P_{t-1} + A_2 P_{t-2} + \cdots + A_k P_{t-k} + \varepsilon_t$. Taking the variables in first differences, with Δ as the difference operator ($\Delta P = P_t - P_{t-1}$) than $P_{t-i} \equiv P_{t-1} - (\Delta P_{t-1} + \Delta P_{t-2} + \cdots + \Delta P_{t-i+1})$ and one can re-write the process as:

$$\Delta P_t = \Pi P_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta P_{t-i} + \mu_t + \varepsilon_t \quad (3)$$

Where: $\Pi = \sum_{i=1}^{k} A_i - I$, $\Gamma_i = -\sum_{j=i+1}^{k} A_j$ and $\varepsilon_t \sim Niid(0, \Sigma)$

In Eq. (3) P_t represents a vector of p non-stationary endogenous variables and the matrix Π contains information about the long-run relationship among endogenous variables and can be decomposed as $\Pi = \alpha\beta'$, whereas β represents the cointegration vectors and α the matrix with the estimations on the speed of adjustment to the equilibrium. The matrix α is called an error correction term, which compensates for the long-run information lost through differencing. The rank of matrix $\Pi(r)$ determines the long-run relationship. If the rank of the matrix Π is zero ($r = 0$), there is no long-run relationship and the model above is equal to a VAR in differences. If the matrix Π has the full rank ($r = p$), then it is invertible, meaning that the processes P_t are stationary I(0) and a normal VAR in levels can be used. The cointegration relationship occurs when the order of the matrix is between 0 and p ($0 < r < p$) and there are (p x r) matrixes α and β such that equation $\Pi = \alpha\beta'$ holds. In this case, P_t is I(1) but the linear combination $X_t = \beta' P_t$ is I(0). If, for example, $r = 1$ and the first element of β was $\beta = -1$, then one could write the linear combination as $X_t = -P_{1,t} + \beta_2 P_{2,t} + \cdots + \beta_p P_{p,t}$ which is the equivalent to saying that long-run equilibrium relationship among variables of vector P_t is expressed as $P_{1,t} = \beta_2 P_{2,t} + \cdots + \beta_p P_{p,t} - X_t$. This long-run relationship may not hold all the time, however the deviation X_t is stationary I(0). In this case, Eq. (4) can be written as follows:

$$\Delta P_t = \alpha\beta' P_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta P_{t-1} + \mu_t + \varepsilon_t \quad (4)$$

If β is known, then X_t would be observable and all the remaining parameters could be estimated by OLS. In practice, the procedure estimates β first and then the rest.

An error correction model provides two alternative channels of the interaction among variables: *i*) short-run effects of the variables are captured similar to the VAR of differences, whose parameters are estimated in the matrix Γ_i; *ii*) the long-run effects enter the model with the term $\Pi\ P_{t-1}$ or $\alpha\beta\ ´\ P_{t-1}$. μ_t is a vector of deterministic terms (constant and trend) and ε_t is a vector of innovations reflecting new information emanating from each of the variables.

Estimation typically proceeds in two stages: first, a sequence of tests is run to determine r, the cointegration rank. Then, for a given rank the parameters of Eq. (4) are estimated. The rank of Π (row rank of β) determines the number of cointegration vectors. Usually two tests on the eigenvalues are used to determine r: Trace Test and λmax Statistics.

3. Data and variables

The daily sample period herein used covers the first year and a half of the second year of the third phase of the EU ETS: 1st January 2013 to 31st July 2014.

We are aware of the fact that low frequency data (monthly or weekly) is often preferred to daily data which may induce errors-in-variables problems. However, the sample period of our analysis is relatively short (one year and a half), so by using weekly or monthly data were used our sample size would be too small to perform traditional time series analysis. We use daily data for working days.

Information on daily stock prices during 2013-2014 comes from the Datastream Database. We used the adjusted close price corrected by capital increases and splits. Data about the capitalization of the companies have been obtained in Bolsas y Mercados Españoles (www.bolsasymercados.es). We take the daily stock market price of companies affected by EU ETS for which financial market data is available for the whole sample.

We eliminated those companies that state as their main or core activity an activity related to renewable energy. The final sample consisted of 27 firms which have been clustered by sectors (as it is shown in Table 3): oil refineries; ferrous metallurgy; cement clinker or lime; glass including glass fibre; ceramic products by firing and pulp, paper and board.

The daily return weighted sector has been calculated using as a weighting factor the market capitalization of each company at year-end compared to the market capitalization of all companies in the same sector.

This process allows us to obtain an aggregate daily stock market return for each sector: R_{ccg}, R_{power}, R_{paper}, R_{petrol}, R_{metal}.

As cement, ceramic and glass only account for one of the companies listed on the Spanish Stock Market, we have grouped them into a unique series, which we have named R_{ccg} (weighted aggregate of the daily stock market return for cement, ceramic and glass sectors).

The proxy for the Rm market portfolio used is the *Índice General de la Bolsa de Madrid*, the biggest Spanish market index.

The yield on 10-year Spanish Treasury bonds is used to assess the interest rate *r*.

The electricity series P^{elect}, from OMEL, is the day-ahead price (€/MWh) for the peak load regime. The peak price is the hourly average of spot prices quoted from 8:00 h to 20:00h.

The natural gas price P^{gas} (€/MWh gas) is the spot price from the Zeebrugge Hub (European virtual trading point, Belgium).

The coal price P^{coal} (€/ton.) is the spot index API#2 (CIF ARA Delivered to the Amsterdam/Rotterdam/Antwerp region).

The EUA price series P^{EUA} (€/ton.) is the spot price quoted at EEX – European Energy Exchange (Leipzig, Germany).

The oil Crude price P^{oil} is the Oil Dated Brent (€//BBL).

Table 4 summarizes the main descriptive statistics of the variables.

Table 3: Companies* traded on the Madrid Stock Exchange that belong to sectors affected by EU ETS.

Company	Economic sector
Cementos Portland Valderrivas, S.A.	Cement
Uralita, S.A.	Ceramic
Vidrala S.A.	Glass
Centrais Ele. Brasileira S.A. Electrobas Enersis S.A. Empresa Nacional de Electricidad SA Endesa S.A. Iberdrola S.A.	Power
Grupo Empresarial Ence S.A. Iberpapel Gestion S.A. Miquel y Costas & Miquel S.A.	Paper
Petroleo Brasileiro S.A. Petrobras Repsol YPF S.A.	Petroleum refineries
ArcelorMittal S.A. Vale S.A.	Ferrous Metallurgy

*Companies representing 85% of the total market capitalization of each sector.

Table 4: Summary descriptive statistics

Variable	Units	Phase III				
		Mean	Median	Min.	Max.	Stnd.Dv
r	%	4.12	4.22	2.58	5.43	0.73
p^{EUA}	€/ton	4.84	4.78	2.68	7.11	0.85
p^{elect}	€/MWh	43.29	46.20	0.79	91.89	16.34
p^{gas}	€/MWh gas	3.05	3.03	2.33	5.85	0.51
p^{oil}	€//BBL	81.68	81.10	75.12	90.52	3.24
p^{coal}	€/ton	59.74	59.23	53.03	69.30	4.18
R_m		940.29	945.42	760.72	1143.30	105.47
R_{CCG}	€	22.18	21.39	14.22	29.68	4.52
R_{power}	€	11.38	10.33	8.74	16.48	2.35
R_{paper}	€	11.10	11.00	8.39	14.09	1.73
R_{petro}	€	9.49	9.55	8.09	11.00	0.68
R_{metal}	€	13.38	13.52	9.824	16.49	2.017

We transformed the variables into their natural logarithms, except for the interest rate.

4. Empirical results

The estimation method proceeds as follows: i) unit root tests are conducted to test for the order of integration in individual price series, *ii*) assuming the tests conclude that the series are I(1), the cointegration rank is determined, and *iii*) a VECM is estimated.

4.1. Preliminary tests: Unit-root-testing

Before deciding on applying either a VAR or VEC model, we need to test for the presence of a unit root. The series in the current study are tested for the presence of unit root by the Augmented Dickey-Fuller unit root prior test. The null hypothesis for this test is the presence of a unit root in the time series; the alternative hypothesis is the time series being generated by a stationary process. The results of the testing are presented in Table 5.

Table 5: Unit root testing of variables using the Augmented Dickey-Fuller test (variables - except r - in natural logarithms)

	Level		diff.	
	Statist.	p value	Statist.	p value
r	-0.111	0.947	-4.363	0.000
p^{EUA}	-2.818	0.056	-6.172	0.000
p^{elect}	-2.782	0.061	-7.957	0.000
p^{gas}	-1.913	0.327	-5.570	0.000
p^{oil}	-2.240	0.192	-6.636	0.000
p^{coal}	-1.571	0.497	-20.083	0.000
R_m	-0.287	0.925	-5.032	0.000
R_{CCG}	-3.149	0.024	-6.717	0.000
R_{power}	-0.450	0.898	-5.135	0.000
R_{paper}	-1.504	0.532	-8.576	0.000
R_{petro}	-2.591	0.095	-4.391	0.000
R_{metal}	-1.671	0.445	-4.216	0.000

The test indicates that, at 1% of significance, all the series contain a unit root (integrated of order 1) and therefore must be differentiated for the purpose of the current research.

Then, we obtain the growth rate of the relevant variables by their differenced logarithms.

4.2. Econometric model

Given the order of integration of the variables used, a general VECM specification can be formulated for each sector j as:

$$\Delta P_t = \alpha\beta' P_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta P_{t-1} + \mu_t + \varepsilon_t \quad (5)$$

P_t is a (7x1) vector of prices (endogenous variables) measured at time *t*: $P_t = [r_t, R_{mt}, R_{jt}, P_{telec}, P_{toil}, P_{tgas}, P_{tcoal}]$. R_{mt}, R_{jt}, P_{telec}, P_{toil} and P_{tcoal} are in natural logarithms, α and β are (7xr)[4] matrix, whereas β and α represent the cointegrating vectors and the matrix with the estimations on the speed of adjustments to the equilibrium, respectively. Γ_i is a (7x7) matrix with the estimations of short-run parameters relating price changes lagged *i* periods. μ_t is a (7x1) vector of constant and ε_t is a (7x1) vector of innovations.

4.3. Cointegration testing

The first step of the modelling procedure is to determine the lag relationship among the price series in levels VAR (used to generate Eq. (5)). Both the AIC (Akaike Info Criterion) and HQC (Hannan and Quinn Criterion) loss metrics suggest the appropriate VAR lag length is two[5] K=2 (Table 6).

---

[4] Where r is the number of cointegrating vectors.

[5] As the VAR is specified in first differences, the number of lags lag in the VECM should be one (k-1).

Since the unit root tests reveal that the series are integrated of order one, a need arises to check whether these time series contain a common stochastic trend. Performing the Johansen cointegration test does this. If the I(1) variables do not exhibit cointegration relations, we opt for a VAR model analysis.

The existence of cointegration relations is shown in Table 7. The null hypothesis states that the amount of cointegrating vectors is equal to r; the alternative hypothesis is that the number of cointegrating vectors is greater than r.

The tests of cointegration were implemented with the technique based on the reduced rank regression introduced in Johansen (1991). Since the VAR model contains exogenous variables, the Osterwald-Lenum (1992) and Johansen (1995) asymptotic critical values are no longer valid, and we therefore use the asymptotic critical values provided in Mackinnon et al. (1999).

Table 6: Lag length in endogenous variables

Sector j	Lags	AIC	SC	HQ
	0	-12.541	-12.460	-12.509
	1	-32.429	-31.698*	-32.139*
Cement	2	-32.464	-31.083	-31.917
	3	-32.495*	-30.465	-31.691
	0	-13.305	-13.224	-13.273
	1	-33.162	-32.431*	-32.872*
Power	2	-33.233	-31.852	-32.685
	3	-33.256*	-31.226	-32.452
	0	-12.942	-12.861	-12.910
	1	-32.743	-32.012*	-32.454*
Paper	2	-32.7872	-31.406	-32.240
	3	-32.819*	-30.789	-32.014
	0	-12.915	-12.834	-12.883
	1	-32.874	-32.143*	-32.585*
Petro	2	-32.954	-31.574	-32.407
	3	-32.989*	-30.959	-32.184
	0	-12.378	-12.297	-12.346
	1	-32.160	-31.429*	-31.870*
Metal	2	-32.181	-30.800	-31.633
	3	-32.196*	-30.166	-31.392

Note: Model with constant and a maximum of 20 lags

Table 7: Cointegration tests Phase III

	H_0:	Trace test			λmax -max eigen value test		
Sector	r=	Statistics	Critical Value	p-values	Statistics	Critical Value	p-values
	1	0.161	125.615	0.000	0.161	46.231	0.000
Cement	2	0.137	95.754	0.005	0.137	40.078	0.000
	3	0.052	69.819	0.583	0.052	33.877	0.700
	1	0.137	125.615	0.000	0.137	46.231	0.002
Power	2	0.128	95.754	0.000	0.128	40.078	0.001
	3	0.072	69.819	0.004	0.072	33.877	0.164
	1	0.168	125.615	0.000	0.168	46.231	0.000
Paper	2	0.091	95.754	0.092	0.091	40.078	0.104
	3	0.056	69.819	0.436	0.056	33.877	0.567
	1	0.197	125.615	0.000	0.197	46.231	0.000
Petro	2	0.108	95.754	0.053	0.108	40.078	0.015
	3	0.052	69.819	0.603	0.052	33.877	0.690
	1	0.157	125.615	0.000	0.226	52.363	0.000
Metal	2	0.087	95.754	0.057	0.157	46.231	0.000
	3	0.073	69.819	0.259	0.087	40.078	0.144

4.4. Sectorial VECM estimation

With the cointegrated rank and optimum number of lags determined, the parameters of model (5) for each sector can be estimated.

Following, the VECM estimations for each EU ETS affected sector are presented. We only report the estimation of the coefficients that are significant at 1% (***), 5% (**) or 10% (*) significant levels - with the exception of the coefficients related to EUA prices that are always shown. The interpretation of the results is focused in the effect of EUA change prices on stock prices changes.

4.4.1 VECM estimations for cement, ceramic and glass sector

The results reported in Table 8 for the cointegrated vector β, which is normalized on R_{CGt-1}, r_{t-1} and R_{mt-1} show that the long-run relationships between EUA price change and stock market price change for cement, ceramic and glass sector are important (but small) during phase III. Since the coefficients can be interpreted as price elasticities, therefore, a EUA price rise of 1%, would, in equilibrium, be associated with a stock price for the sector increase of 0.0187%.

Table 8: VECM parameter estimates for cement, ceramic and glass sector

Cointegration relationships

$R_{CG\,t-1}$	r_{t-1}	$R_{m\,t-1}$	P^{EUA}_{t-1}	P^{ELECT}_{t-1}	P^{COAL}_{t-1}	P^{OIL}_{t-1}	P^{GAS}_{t-1}	Const.
1.000	0.000	0.000	-0.923***	-0.369***	-0.887**	2.014***	-0.982***	-
0.000	1.000	0.000	-	2.382***	-	-	7.138***	-
0.000	0.000	1.000	-0.598***	-0.442***	-1.085***	1.222**	-0.853***	-

Short run dynamics

	$\Delta R_{CG\,t}$	Δr_t	$\Delta R_{m\,t}$	ΔP^{EUA}_t	ΔP^{ELECT}_t	ΔP^{COAL}_t	ΔP^{OIL}_t	ΔP^{GAS}_t
$EC1_{t-1}$	-0.052***	-	-	-	-	-	-	0.161***
$EC2_{t-1}$	-	-	-	0.033***	-	-	-	-0.050***
$EC3_{t-1}$	0.049**	-	-	0.128**	1.403**	-	-	-0.387***
$\Delta R_{CG\,t-1}$	-	-0.408**	-	0.351**	-	-	-	-
Δr_{t-1}	-	-0.139***	0.027***	-	-	-	-	-
$\Delta R_{m\,t-1}$	-	-1.526***	0.094*	-0.533**	-	-	-	-
ΔP^{EUA}_{t-1}	-0.026	-	-	0.085*	-	-	-	-
ΔP^{ELECT}_{t-1}	-	-	0.002*	-0.008	0.105**	-0.002**	-	-0.012**
ΔP^{COAL}_{t-1}	-	-	-	-	-	-	-	-0.872***
ΔP^{OIL}_{t-1}	-	-	-	-	-	-	-	-
ΔP^{GAS}_{t-1}	-	-	0.024*	-	-	-	-	0.180***
Const.	0.002***	-0.005**	-	-	-	-	-	-

*stands for estimates significantly different from 0 at a 10% level, ** stands for estimates significantly different from 0 at a 5% level and *** stands for estimates significantly different from 0 at a 1% level.

The short-run parameters in the VECM indicate that the EUA price change does not have an effect on stock market returns of the cement, ceramic and glass companies. The short run parameter corresponding to EUA price for Phase III is -0.026 but it is not significant.

The results are similar to those found by Chan et al. (2013) who examine the impact of EU ETS on firms' unit material costs, employment and revenue during 2005–2009. They concluded that EU ETS had no impact on the performance of cement and iron or steel industries.

4.4.2. VECM estimations for power sector

The results reported in Table 9 for the cointegrated vector β , which is normalized on R_{CGt-1}, r_{t-1} and R_{mt-1} show that the long-run relationships between EUA price change and stock market price change for the power sector are important (but small) during Phase III. Since the coefficients can be interpreted as price elasticities, therefore, a EUA price rise of 1%, would, in equilibrium, be associated with a stock price for the sector increase of 0.03% during phase III.

Table 9: VECM parameter estimates for power sector

Cointegration relationships								
$R_{power\ t-1}$	r_{t-1}	$R_{m\ t-1}$	P^{EUA}_{t-1}	P^{ELECT}_{t-1}	P^{COAL}_{t-1}	P^{OIL}_{t-1}	P^{GAS}_{t-1}	Const.
1.000	0.000	0.000	-0.809***	-0.386***	-0.702**	0.202***	-1.719***	-
0.000	1.000	0.000	-	1.758***	-	-	6.955***	-
0.000	0.000	1.000	-0.677***	-0.419***	-1.109***	1.514**	-0.878***	-

Short run dynamics								
	$\Delta R_{power\ t}$	Δr_t	$\Delta R_{m\ t}$	ΔP^{EUA}_t	ΔP^{ELECT}_t	ΔP^{COAL}_t	ΔP^{OIL}_t	ΔP^{GAS}_t
$EC1_{t-1}$	-0.040***	-	-	0.105***	-	-	-	0.031***
$EC2_{t-1}$	-	-	-	0.053***	-	-	-	-0.029***
$EC3_{t-1}$	-0.020**	-	-	0.122**	0.873**	-	-	-0.131***
$\Delta R_{power\ t-1}$	-	-0.736**	-	-	-	-	-	-
Δr_{t-1}	-	-0.145***	0.030***	-	-	-	-	-
$\Delta R_{m\ t-1}$	-	-1.130***	0.045*	-0.605**	-	-	-	-
ΔP^{EUA}_{t-1}	-0.015	-	-	0.097*	-	-	-	-
ΔP^{ELECT}_{t-1}	-	-	0.002*	-	0.109**	-0.002**	-	-0.013**
ΔP^{COAL}_{t-1}	-	-	-	-	-	-	-	-0.741***
ΔP^{OIL}_{t-1}	-	-	-	-	-	-	-	-
ΔP^{GAS}_{t-1}	-	-	0.022*	-	-	-	-	0.173***
Const.	0.002***	-0.005**	-	-	-	-	-	

*stands for estimates significantly different from 0 at a 10% level, ** stands for estimates significantly different from 0 at a 5% level and *** stands for estimates significantly different from 0 at a 1% level.

The results are similar to those found by empirical literature. Oberndorfer et al. (2006) examined the impacts of the EU ETS on competitiveness in Europe and concluded that for the power sector the impacts were modest. In the same way, Chan et al. (2013) concluded that EU ETS was associated with increased material costs and revenue of the power industry during 2005–2009. Also by using a Cournot representation, Bonenti et al. (2013)

evaluated the impact of EU ETS on the Italian electricity market profits under different allocation scenarios of allowances (free and auctions), concluding that the generators would be expected to profit in an oligopolistic market as they are able to transfer almost all their emission costs to the final price paid by consumers. In addition, Veith et al. (2009) by using a modified multifactor market similar to Eq. 2 and 2005-2007 data of 22 electricity companies estimated a coefficient β_2 equal to 0.006.

However, the existing empirical studies do not converge with our research findings, as some scholars have concluded that the EU ETS has a negative effect on power companies. For instance, Mo et al. (2013) indicate that positive EUA prices generated corporate value depreciation during phase II. By using a modified multifactor market similar to Eq. 2 and 2008 and 2009 data of 48 electricity companies, they estimate a coefficient β_2 equal to -0.0334. Moreover, Oberndorfer (2009) found that although EUA price changes and stock returns of the most important European electricity corporations were positively related, Spanish electricity corporations exhibit a significant (but small as far as the size of the estimated coefficient is concerned) negative relationship.

Jaraitė and Kažukauska (2013) found that the first years of the EU ETS (2002-2010) couldn't be associated with excess profits for electricity producers.

Regarding the short-run parameters in the VECM, these indicate that the EUA price changes do not have an effect on stock market returns of power sector in the EU ETS Phase III.

We would like to point out that, in the long-run, the electricity sector as a whole has modest gains from the introduction of the EU ETS instrument, but these results could change if the electricity sector was grouped by companies according to its main generation technology (Bode 2006).

4.4.3. VECM estimation for paper sector

The results reported in Table 10 for the cointegrated vector β , which is normalized on R_{CGt-1}, r_{t-1} and R_{mt-1} show that there are not significant long-run relationships between EUA price change and stock market price change for paper sector during the current phase. Moreover, the short-

-run parameters in the VAR indicate that the EUA price changes do not have an effect on stock market returns of the sector.

Meleo (2014) discusses the main factors affecting competitiveness coming from the EU-ETS, considering the case of the Italian paper industry. He points out that paper firms cannot easily move prices to cover environmental costs because paper demand is price-elastic and paper products have several substitutes such as plastic goods.

Table 10: VECM parameter estimates for paper sector

Cointegration relationships								
$R_{paper\ t-1}$	r_{t-1}	$R_{m\ t-1}$	P^{EUA}_{t-1}	P^{ELECT}_{t-1}	P^{COAL}_{t-1}	P^{OIL}_{t-1}	P^{GAS}_{t-1}	Const.
1.000	0.000	-1.696***	-	0.384***	0.803***	-	0.662***	-
0.000	1.000	2.095**	-	1.320***	-2.113*	-	5.004***	-

Short run dynamics								
	$\Delta R_{paper\ t}$	Δr_t	$\Delta\ R_{m\ t}$	ΔP^{EUA}_t	ΔP^{ELECT}_t	ΔP^{COAL}_t	ΔP^{OIL}_t	ΔP^{GAS}_t
$EC1_{t-1}$	-0.024**	-	0.009**	-0.073**	-1.216***	0.014*	-	0.180***
$EC2_{t-1}$	0.007***	-	-0.004*	0.026***	-	-	-	-0.057***
$\Delta R_{paper\ t-1}$	-0.101**	-	-	-	3.124*	-0.065*	-	-
Δr_{t-1}	-	-0.146***	0.028***	-	-	-	-	-
$\Delta R_{m\ t-1}$	-	-1.597***	0.118**	-0.523**	-	0.084*	-	-
ΔP^{EUA}_{t-1}	-	-	-	-	-		-	-
ΔP^{ELECT}_{t-1}	-	-	-	-	0.115**	-0.003*	-	-0.014***
ΔP^{COAL}_{t-1}	-	-	-	-	-	-	-	-0.908***
ΔP^{OIL}_{t-1}	-	-	-	-	-	-	-	-
ΔP^{GAS}_{t-1}	-	-	0.025**	-	-	-	-	0.171***
Const.	0.002**	-0.006**	-	-	-	-	-	-

*stands for estimates significantly different from 0 at a 10% level, ** stands for estimates significantly different from 0 at a 5% level and *** stands for estimates significantly different from 0 at a 1% level.

4.4.4. VECM estimations for petroleum refineries sector

The results reported in Table 11 for the cointegrated vector β, which is normalized on R_{CGt-1}, r_{t-1} and R_{mt-1}, show that the long-run relationships between EUA price change and stock market price change for oil refineries sector are important (but small) during phase III. A EUA price rise of 1%, would, in equilibrium, be associated with a stock price for the sector increase of 0.0261%.

The short-run parameters in the VAR indicate that the EUA price changes do not have an effect on stock market returns for this sector.

Table 11: VECM parameter estimates for oil refineries sector

Cointegration relationships								
$R_{petro\ t-1}$	r_{t-1}	$R_{m\ t-1}$	P^{EUA}_{t-1}	P^{ELECT}_{t-1}	P^{COAL}_{t-1}	P^{OIL}_{t-1}	P^{GAS}_{t-1}	Const.
1.000	0.000	0.000	0.231***	-0.186***	-0.363*	-	-	-
0.000	1.000	0.000	-	2.236***	-	-	7.182***	-
0.000	0.000	1.000	-0.493***	-0.397***	-0.896***	-	-1.054***	-

Short run dynamics								
	$\Delta R_{petro\ t}$	Δr_t	$\Delta R_{m\ t}$	ΔP^{EUA}_t	ΔP^{ELECT}_t	ΔP^{COAL}_t	ΔP^{OIL}_t	ΔP^{GAS}_t
$EC1_{t-1}$	-0.032***	-	-	-	0.755**	-0.023***	-0.031***	-0.215***
$EC2_{t-1}$	-0.015***	-	-	0.044***	-	-	-	-0.050***
$EC3_{t-1}$	-0.068***	-	-	0.257***	0.794*	-	-	-0.161***
$\Delta R_{petro\ t-1}$	0.117*	-0.637***	0.173***	-	-	-	-	-
Δr_{t-1}	0.026**	-0.148***	0.029***	-	-	-	-	-
$\Delta R_{m\ t-1}$	-	-1.205***		-0.688**	-	-	-	-
ΔP^{EUA}_{t-1}	-0.011	-	-	0.101**	-	-	0.020*	-
ΔP^{ELECT}_{t-1}	-	-	-	-0.009*	0.111**	-0.002**	-	-0.013**
ΔP^{COAL}_{t-1}	-	-	-	-	-	-	-	-0.927***
ΔP^{OIL}_{t-1}	-	0.535*	-	-	-	-	-	-
ΔP^{GAS}_{t-1}	-	-	-	-	-	-	-	0.175***
Const.	-	-0.006**	0.001*	-	-	-	-	-

*stands for estimates significantly different from 0 at a 10% level, ** stands for estimates significantly different from 0 at a 5% level and *** stands for estimates significantly different from 0 at a 1% level.

Although the petroleum refineries sector´s stock market returns are positive related to EUA price changes as a whole according to Reinaud (2005) the results for each refinery may depend of the way they obtain electricity. The refinery produces its own electricity, it purchases its electricity from the grid or it produces its energy needs from an Integrated Gasification Combined Cycle. Reinaud concludes that refinery margins in the Mediterranean area are lower if power is purchased from the grid.

Moreover, results could also change according to the main product of the refinery: Babusiaux (2003) reveals higher emission contents for diesel than for gasoline for large number of scenarios. In certain cases, negative gasoline marginal contents are even obtained.

4.4.5. VECM parameter estimates for ferrous metallurgy sector

The results reported in Table 12 for the cointegrated vector β, which is normalized on R_{CGt-1}, r_{t-1} and R_{mt-1} show that the long-run relationships between EUA price change and stock market price change for ferrous metallurgy sector are important (but small) during phase III. In fact an EUA price rise of 1%, would, in equilibrium, be associated with a stock price decrease for the sector of -0.0174%.

The short run parameter corresponding to the EUA price is 0.0158 but it is not significant.

Demailly and Quirion (2009) examine the impact on iron and steel industry under a euro 20 per ton CO_2 price. They found that profitability depends on the amount of allowances allocated for free. As they point out, even though more allowances are going to be allocated for free in the future, the decrease in the profitability of the ferrous metallurgy sector would be modest.

Table 12: VECM parameter estimates for ferrous metallurgy sector- Phase III

				Cointegration relationships				
$R_{metal\ t-1}$	r_{t-1}	$R_{m\ t-1}$	P^{EUA}_{t-1}	P^{ELECT}_{t-1}	P^{COAL}_{t-1}	P^{OIL}_{t-1}	P^{GAS}_{t-1}	Const.
1.000	0.000	-4.086***	1.241***	1.272***	1.995*	-3.724*	2.644***	-
0.000	1.000	1.918**	-	1.528***	-	-	5.422***	-
				Short run dynamics				
	$\Delta R_{metal\ t}$	Δr_t	$\Delta R_{m\ t}$	ΔP^{EUA}_t	ΔP^{ELECT}_t	ΔP^{COAL}_t	ΔP^{OIL}_t	ΔP^{GAS}_t
$EC1_{t-1}$	-0.014***	-	-	-0.045***	-0.268***	-	0.004*	0.032***
$EC2_{t-1}$	0.010***	-	-	0.039***	-	-	-	-0.032***
$\Delta R_{metal\ t-1}$	-	0.370**	-0.079**	-	-	-	-	-
Δr_{t-1}	-	-0.144***	0.027***	-	-	-	-	-
$\Delta R_{m\ t-1}$	-	-1.882***	0.154***	-0.494**	-	-	-	-0.551**
ΔP^{EUA}_{t-1}	0.016	-	-	-	-	-	-	-
ΔP^{ELECT}_{t-1}	-	-	-	-	0.106**	-0.002**	-	-0.013**
ΔP^{COAL}_{t-1}	-	-	-	-	-	-	-	-0.693***
ΔP^{OIL}_{t-1}	-	-	-	-	-	-	-	-
ΔP^{GAS}_{t-1}	-	-	0.024**	-	-	-	-	0.143***
Const.	-	-0.006*	-	-	-	-	-	-

*stands for estimates significantly different from 0 at a 10% level, ** stands for estimates significantly different from 0 at a 5% level and *** stands for estimates significantly different from 0 at a 1% level

5. Conclusions and policy implications

Over the course of the EU ETS history, the system has been expanded in scope to include both additional countries and new sectors. Furthermore, links have been established with both the permit trade under the mechanisms of the Kyoto Protocol and non-EU national emission trading systems. The existence of 2 billion unused allowances at the end of Phase II of the EU ETS, approximately 20% of the five-year cap, is frequently quoted as the cause of the current low price of EUAs and an indication of some fundamental shortcoming in the design of the EU ETS.

It was in this set up that this study investigated the interactions between the stock market returns of Spanish industry sectors under EU ETS and emission rights prices during the on-going Phase III. The witnessed divergence amongst allowances distributed and allowances used has been the reason for the debate about "back-loading" that lead the examination concerning the EU ETS in 2013, as well as for the proposal made in January 2014 to establish a Market Stability Reserve. Together, these measures would diminish the amount of allowances available in the immediate period while placing the reserved allowances back into circulation at a later time (European Commission, 2012, 2014).

By using daily data from January 2013 to July 2014 and a cointegrated Vector Error Correction (VECM), the results obtained indicate that the EUA price change does not present short-run effects on stock market returns on the sector during the current phase. However, the long-run relationships between EUA price change and stock market price change depending on the considered sector- as it is showed in the second column of Table 13 (Long-run cointegration relationship Π).

A statistically significant positive long-run impact of EU ETS on power sector stock market return is found. In fact, an EUA price rise of 1%, would, in equilibrium, be associated with a stock price for the power sector increase of 0.03% for Phase III.

Moreover, statistically significant positive long-run impact of EU ETS on cement and petroleum and, on the other hand, negative impact on iron and steel sectors were found during phase III. Thus, an EUA price rise of 1%, would be associated with a stock price change of 0.0187%, 0.0261%, and -0.0174 % for cement, oil refineries and ferrous metallurgy sectors, respectively.

A long-run link between carbon prices and stock market on the paper sector is found no to be supported during the studied period.

The long-run cointegration relationship (Π) can be decomposed as $\Pi = \alpha \beta'$, whereas β represents the cointegration vectors and α the matrix with the estimations defining the speed of adjustment of the daily stock market return of each sector to the long-run equilibrium. The matrix α is called an error correction term (EC), which compensates for the long-run information lost through differencing. The presence of cointegration requires at least one of the coefficients of the error correction terms to be statistically significant. This condition is observed throughout the VECM model. As it is shown in Table 3 all the sectors have at least the estimation of two error correction terms statistically significant. High absolute values of the error correction terms indicate that a sector is largely able to correct the disequilibrium within one day. In that sense, the estimated values are very small for all the sectors, so we can conclude that no sectors are able to correct the disequilibrium within one day. In general, the correction terms are higher for Cement, ceramic and glass sector, i.e. the adjustment is faster. For example, the EC1 indicates that about 5.2% of the disequilibrium is corrected within one day in this sector.

Table 13: Long-run and Error correction terms from VECM estimations- Phase III

Economic sector	Long-run cointegration relationship	Error Correction terms*		
		EC1	EC2	EC3
Cement, ceramic and glass	0.0187	-0.052		0.049
Power	0.0300	-0.040	-0.013	-0.020
Paper	0.0000	-0.024	0.007	
Petroleum refineries	0.0261	-0.032	-0.015	-0.068
Ferrous Metallurgy	-0.0174	-0.014	0.010	

Estimated coefficients that are significant at 1%, 5% or 10% significant levels.

As it is shown, electricity sector is the one where EU allowances price changes has the highest long-run impact on stock market returns, following by petroleum refineries. An essential element of the effect of EUA price increase on each sector stock market returns is the capacity that companies belonging to each sector have to increase product prices. In general, the cost incurred to respond to the EUA constraint should reduce profit margins, all other things equal. However, companies can incorporate their carbon emission allowance costs in their sale offers. The market structure, the demand elasticity to price variations or the number of substitutes of the principal product of the firm, among others, influence the grade of the pass-through of environmental costs on prices. Regarding the electricity sector, although the Spanish electricity market liberalization compels to introduce competition into electricity generation, the old integrated monopolies continue with the control of the electricity production (Endesa e Iberdrola). Although the concentration of the electricity generation market decreased in 2010 and 2011 to a moderate level, below that of other European countries, it is still high. According to the last report on the development of competition in gas and electricity markets published by the Spanish Energy Commission (Comisión Nacional de la Energía, 2012), the Herfindahl and Hirschman index (HHI) associated to the domestic market is of 1400. Moreover, if we use a more conservative definition that uses only the group of technologies setting the price in the Spanish market (combined cycle gas, coal, reservoir hydro), the HHI stands at 1900. However, there continues to be a difference between new entrants and incumbents in terms of the degree of vertical integration (the former sell most of their energy on the spot market, while the latter enter into bilateral contracts –mainly between companies of the same group– that were primarily associated with nuclear and hydro power plants in 2011).

In addition, the demand elasticity to electricity price variations is very low because of the special characteristics of electricity output (such as no storability or the existence of capacity constraints in the short term offer) and the number of substitutes limited to gas.

As we mentioned before, the market structure, among others, influence the grade of the pass-through of environmental costs on prices. In that sense, when comparing results of the impact of EU allowances price on sectors' stock market returns from different countries they might differ as countries could have different market conditions for the same sector. For example, European countries with the highest market shares of the largest generator in the electricity market (as a percentage of the total generation) are Cyprus (100%), Malta (100%), Estonia (87%) or France (84%), and with the lowest market shares are Spain (24.5%) and Poland (17.3%) according to 2013 data (Eurostat).

Moreover, Ireland and United Kingdom have introduced a carbon floor price in April 2013. This means that polluter producers in Great Britain must pay a premium when the EU allowances prices become lower than the price floor. Thus, although sectors covered by EU ETS in Ireland and United Kingdom are the same than in other countries, the carbon price might differ, so we must be cautious when making sectoral comparison with other countries without this system.

It is important to note that these findings should be viewed in alignment with the specific period and EU ETS phase analysed. We would like to point out that during Phase III the allocation of emission allowances are given out predominantly through auctioning, starting from a proportion of the 20% in 2013 and reaching a 70% level in 2020 (European Commission, 2009). Thus, the results we obtain based on data from January 2013 to July 2014 could vary, following an increase in the proportion of allowances auctioning as we approach 2020. For instance, concerning the power sector, switching to a higher proportion auctioning would leave the electricity sector as a whole better off than before the introduction of the EU ETS (Keppler and Cruciani 2010).

In the future, we plan to extend the research herein presented in the following two directions. First, while the present paper focuses on the analyses at a global sectoral level, we have not examined the impact on stock market returns at firm level, thus a more detailed study by firms of each sector would be of added value.

Second, we plan to include updated data from July 2014 in our analysis so that we can acquire a more comprehensive depiction of the impacts of the third phase of the program.

Acknowledgements: The first two authors would like to acknowledge that this work has been partially supported by the R&D Project EMSURE (Energy and Mobility for Sustainable Regions) – CENTRO 07 0224 FEDER 002004. The first author would like to acknowledge Fundação para a Ciência e Tecnologia, under project grant UID/MULTI/00308/2013, and the Energy for Sustainability Initiative of the University of Coimbra.

References

Acaravci, A. Ozturk, I., Kandir, S.Y. (2012): Natural gas prices and stock prices: Evidence from EU-15 countries, Economic Modelling 29(5), 1646–1654.

Anger, N., Oberndorfer, U. (2008): Firm Performance and Employment in the EU Emissions Trading Scheme: An Empirical Assessment for Germany. Energy Policy 36, 12-22

Babusiaux, D. (2003): Allocation of the CO2 and Pollutant Emissions of a Refinery to Petroleum Finished Products, Oil and Gas Science Technology, Vol 58, No. 6, 685-692

Bode S. (2006): Multi-period emissions trading in the electricity sector e winners and losers. Energy Policy 34(6), 680-91.

Bushnell, J.B., Chong H., Mansur E.T. (2013): Profiting from Regulation: Evidence from the European Carbon Market, American Economic Journal: Economic Policy, 5(4): 78-106.

Chan S.R, Li S.Zhang, F. (2013): Firm competitiveness and the European Union emissions trading scheme, Energy Policy 63, 1056–1064.

Comisión Nacional de la Energía (2012): Report on the development of competition in gas and electricity markets. Period 2008-2010 and Preview 2011. Madrid: Spain

Demailly, D., Quirion, P. (2008): European Emission Trading Scheme and Competitiveness: A Case Study on the Iron and Steel Industry, Energy Economics 30, 20-27.

Ellerman, D., Marcantonini, C. Zaklan, A. (2014): The EU ETS: Eight Years and Counting, Robert Schuman Centre for Advanced Studies Working Paper 2014-04, European University Institute.

Engle, R., Granger, C. (1987): Co-Integration and Error Correction: Representation, Estimation, and Testing, Econometrica 55, 251-276.

European Commision (2009): Directive 2009/29/EC of The European Parliament and of the Council of 23 April 2009 amending Directive 2003/87/EC so as to improve and extend the greenhouse gas emission allowance trading scheme of the Community.

European Commission (2012), Report from the Commission to the European Parliament and the Council: The State of the European Carbon Market in 2012, COM(2012) 652 final, Brussels: 14.11.2012.

European Commission (2014), "A Policy Framework for Energy and Climate from 2020 to 2030: Communication from the Commission to the European Parliament the Council, the European Economic and Social Committee, and the Committee of the Regions,"COM(2014) 15 final, Brussels, January 22, 2014.

Freitas C.J., Silva P. P. (2013): Evaluation of dynamic pass through of carbon prices into electricity prices - a cointegrated VECM analysis, International Journal of Public Policy 9, 65-85

Freitas C.J., Silva, P. P. (2015): European Union Emissions Trading Scheme impact on the Spanish electricity price during Phase II and Phase, Utilities Policy, 33: 54 - 62.

Goulder, L., Hafstead, M., Dworsky, M. (2010): Impacts of Alternative Emissions Allowance Allocation Methods Under a Federal Cap-and-Trade Program, Journal of Environmental Economics and Management 60, 161-181.

International Energy Agency (2014): Energy Balances of OECD Countries.

Jaraitė, J., Kažukauskas, A. (2013): The profitability of electricity generating firms and policies promoting renewable energy, Energy Economics 40, 858–865.

Johansen, S. (1988): Statistical analysis of cointegrating vectors, Journal of Economic Dynamics and Control 12, 231–254.

Johansen, S. (1991): Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Econometrica. 59, 1551–1580.

Johansen, S. (1992): Determination of cointegration rank in the presence of a linear trend, Oxford Bulletin of Economics and Statistics 54, 383–397.

Johansen, S. (1994): The role of the constant and linear terms in cointegration analysis of nonstationary variables, Econometric Reviews 13, 205–229.

Johansen, S. (1995): Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Oxford University Press, New York.

Johansen, S. and Juselius, K. (1990): Maximum likelihood estimation and inference on cointegration – with applications to the demand for money, Oxford Bulletin of Economics and Statistics 52.169–210.

Keppler, J.H. and Mansanet-Bataller, M. (2010): Causalities between CO2, electricity, and other energy variables during phase I and phase II of the EU ETS, Energy Policy 38, 626–632.

Keppler, J.H., Cruciani, M. (2010): Rents in the European power sector due to carbon trading, Energy Policy 38, 4280–4290

Lee B-J, Yang C.W., Huang B-N. (2012): Oil price movements and stock markets revisited: A case of sector stock price indexes in the G-7 countries, Energy Economics 34, 1284–1300.

Litterman, B. (2013): What is the Right Price for Carbon Emissions? Regulation, Summer 2013, 38.43.

Mackinnon, J., Haug, A., Michelis, L. (1999): Numerical distribution functions of likelihood ratio tests for cointegration, Journal of Applied Econometrics, Vol. 14, No. 5, pp.563–577.

Mansanet-Bataller, M., Pardo, A., Valor, E. (2007): CO2 prices, energy and weather. The Energy Journal 28, 73–92.

Meleo, L. (2014). On the determinants of industrial competitiveness: The European Union emission trading scheme and the Italian paper industry, Energy Policy, in press

Mo J.L., Zhu, L., Fan Y. (2012): The impact of the EU ETS on the corporate value of European electricity, Energy 45, 3-11.

Moreno, B.; García-Álvarez, M.T. (2013): The role of renewable energy sources on electricity prices in Spain. A maximum entropy econometric model, Strojarstvo 55, 149-159.

Moreno, B.; García-Álvarez, M.T.; Ramos, C.; Fernández-Vázquez, E. (2014): A General Maximum Entropy Econometric approach to model industrial electricity prices in Spain: A challenge for the competitiveness, Applied Energy 135, 815-824.

Moya-Martínez P., Ferrer-Lapeña R., Escribano-Sotos F. (2014): Oil price risk in the Spanish stock market: An industry perspective, Economic Modelling 37, 280-290

Oberndorfer, U. (2009): EU Emission Allowances and the Stock Market: Evidence from the Electricity Industry, Ecological Economics 68, 1116.1129.

Oberndorfer, U., Rennings, K., Sahin, B. (2006): The impacts of the European emissions trading scheme on competitiveness in Europe. Zentrum für Europäische Wirtschaftsforschung/ Center for European Economic Research.

Oestreich, A.M., Tsiakas, I. (2014): Carbon Emissions and Stock Returns: Evidence from the EU Emissions Trading Scheme. Available at SSRN: http://ssrn.com/abstract=2189497

Osterwald-Lenum, M. (1992): A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics, Oxford Bulletin of Economics and Statistics 54, 461–471.

Paolella MS, Taschini L (2008): An econometric analysis of emission allowance prices, Journal of Banking & Finance 32, 2022-2032

Pindyck, R.S. (2013). Pricing Carbon When We Don't Know the Right Price.Regulation, Summer 2013, 43.46.

Reinaud, J. (2005): The European refinery industry under the EU Emissions Trading Scheme, International Energy Agency, IEA Information Paper.

Sijm, J.P.M., Neuhoff, K., Chen, Y. (2006): CO2 Cost Pass-Through and Windfall Profts in the Power Sector, Climate Policy 6, 49-72.

Smale, R., Hartley, M., Hepburn, C., Ward, J., Grubb, M. (2006): The impact of CO2 emissions trading on firm profits and market prices, Climate Policy 6, 31-48,

Spanish Statistical Institute (2013). Spanish Energy Consumption Survey 2011. INE: Madrid, Spain; 2013

Veith, S., Werner, J., Zimmermann, J. (2009): Capital Market Response to Emission Rights Returns: Evidence from the European Power Sector, Energy Economics 31, 605-613.

Página deixada propositadamente em branco.

EXPLANATORY VARIABLES ON SOUTH-WEST SPOT ELECTRICITY MARKETS INTEGRATION

Nuno Carvalho Figueiredo[1,3] and Patrícia Pereira da Silva[1,2,3]

Abstract

This paper aims to assess some determinants on the integration level of the South West Europe regional electricity spot markets created under the initiative launched by the European Regulators Group for Electricity and Gas. The integration of the European South-West regional electricity spot market relies on the physical interconnection between two pairs of Transmission Systems: Portugal-Spain and Spain-France. Interconnection capacity is thought to be critical to ensure electricity market integration and was therefore studied. Two other determinants were included in our study, corresponding to weather conditions: temperature and wind. Whilst temperature is thought to have influence on demand, wind on the other hand should have influence on available generation. Results obtained from a vector autoregression model specification show interconnection capacities related exogenous variables to not improve the model, whilst average temperature contributes slightly to improve the model. It is to note that the introduction of the exogenous variable average wind speed improves all models' specification. Strong integration was found between MIBEL markets, leading to the conclusion that Price Splitting mechanism is efficient and contributes to the integration of spot electricity markets. Also, as demonstrated by Grange-causality and impulse response analysis, there is a weak integration level between MIBEL and Powernext.

Keywords: Electricity market integration; South-West REM; VARX modeling

[1] INESC Coimbra

[2] Faculty of Economics of the University of Coimbra

[3] Energy for Sustainability Initiative - University of Coimbra, Portugal
Sustainable Energy Systems – MIT-Portugal Program
e-mail: nuno.figueiredo@portugen.com, patsilva@fe.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_3

1. Introduction

The Council Directive 90/547/EEC of 29 October 1990 on the transit of electricity through transmission grids (European Union, 1990a) and Council Directive 90/377/EEC of 29 June 1990 concerning a procedure to improve the transparency of gas and electricity prices charged to industrial end-users (European Union, 1990b), provided the first steps for the creation of the internal European electricity market (Bower, 2002).

The European Directive 2003/54/EC and lately the European Directive 2009/72/EC reviewed the European Directive 96/92/EC which for the first time established common rules for the various electricity markets in Europe, based on the liberalisation of the sector without prejudice of the public service required and the access by the generators and consumers to the transmission and distribution grids (Jamasb and Pollitt, 2005). These requirements are guaranteed by regulating authorities established in each country (Silva and Soares, 2008).

The European Directive 2001/77/EC repealed by the European Directive 2009/28/EC called for the promotion of electricity generation by renewable energy sources (RES) in Europe in order to reduce dependency on imported fossil fuels and to allow the reduction in greenhouse gas (GHG) emissions. The RES electricity (RES-E) generation capacity in Europe was 239.2 GW by 2010 with 52.1% hydroelectric, 25.7% wind, 17.86% biomass, 3.3% solar, 0.93% geothermal and 0.08% tidal or wave generation (Jäger-waldau et al., 2011). The RES-E generation technologies are in different stages of development which explain the different shares of deployment achieved in each technology (Brown et al., 2011). The large deployment of RES-E generation in Europe was achieved by strong financial support mechanisms (Meyer, 2003), like feed-in tariffs, fiscal incentives, tax exemptions and other (Jager et al., 2011).

To guarantee the supply of electricity, reduce costs, maintain competition and ensure security of supply, whilst respecting the environment, are the objectives set for European energy policies. However different degrees of market opening and development of interconnectors between electricity transmission grids across European countries are observed. European countries took necessary measures to facilitate transit of electricity between transmission grids in accordance with the conditions laid down in the Directives. The adequate integration of national electricity transmission grids and associated increase of electricity cross-border transfers should ensure the optimization of the production infrastructure.

In 2006 the European Regulators Group for Electricity and Gas (ERGEG - currently the Agency for the Cooperation of Energy Regulators – ACER established by European Commission Regulation 713/2009 of 13 July 2009) launched seven Electricity Regional Initiatives (Karova, 2011; Meeus and Belmans, 2008) for the creation of seven Regional Electricity Markets (REMs): Baltic States (Estonia, Latvia, Lithuania); Central East (Austria, Czech Republic, Germany, Hungary, Poland, Slovakia, Slovenia); Central South (Austria, France, Germany, Greece, Italy, Slovenia); Central West (Belgium, France, Germany, Luxembourg, Netherlands); Northern Europe (Denmark, Finland, Germany, Norway, Poland, Sweden); South West (France, Portugal, Spain); and France-UK-Ireland (France, Republic of Ireland, UK). The objective for the creation of these REMs was to provide an intermediate step for the consolidated European Electricity Market (CEER, 2015; ERGEG, 2006).

Consequently, the aspect of transmission costs determination plays an important role and its allocation methods are usually either Flat Rate based or Flow-based. Flat rate methods are simple to calculate and implement, however, according to (Galiana et al., 2003) unfair to generators that use less capacity and extent of the transmission lines.

On the other hand, flow-based costs are most commonly used due to their dependence on the capacity and extent used by each generator of the transmission lines. Explicit auctioning, where interconnector capacity is sold to the highest bidder or implicit auctioning, which integrates electricity and transmission markets and also called Market Splitting/ Price Coupling, are both used across Europe (Coppens and Vivet, 2006).

In the Spain-France interconnection, the method of explicit auctioning is used however the mechanism of Market Splitting is applied to the Portuguese-Spanish interconnection.

In this framework, an initiative, denominated Price Coupling of Regions (PCR) was launched at the Florence Regulatory Forum in 2009 by three power exchanges: Nordpool, EPEX and MIBEL (Europex, 2009), to be implemented by the end of 2012. In the mean time additional members joined the initiative, APX-Endex, Belpex and GME, reaching the 2860 TWh/year of potential electricity trading (Europex, 2011) and to be fully implemented by the end of 2014.

Building on previous work by the authors (Figueiredo and Silva, 2012), the objective of this paper is to assess some determinants on the integration level of the South West REMs. In Sections 2, 3 and 4 brief summaries about the French, Spanish and Portuguese electricity markets are provided. Section 5 follows with a small description of the cross-border interconnections available between the considered electricity markets. In Section 6 data used in this study is presented and discussed and in Section 7 VAR model specifications are presented. Analysis and results are presented and discussed in Section 8 and, finally, in Section 9 final remarks can be found.

2. The French electricity market

In France, there was no privatisation and no unbundling (Newbery, 2005). With several acts of legislation, France managed to carry out the electricity sector reform (through the law 2000-108 of the 10th of February) without restructuring the main operator. An electricity market was created around a public monopoly (Glachant and Finon, 2004). The law 2000-108 of the 10th of February also created the "Commission de Régulation de L'énergie" (CRE) the French regulator (Journal Officiel de la République Française, 2000).

It was considered that competition would come from abroad through the interconnections with the various European countries. The main reasoning behind this was the vast nuclear power capacity and associated low variable cost electricity (Newbery, 2005). EDF has currently 97.2 GW of installed electric capacity, of which 63.7 GW are of nuclear power and owns the complete transmission grid (EDF Group, 2010).

The French electricity day-ahead market, Powernext, started operation in November 2001 (Bower, 2002) and by January 2006 explicit capacity auctions on interconnections was introduced (Commission de Régulation de L'énergie, 2011). The market coupling between France, Belgium and the Netherlands was launched in November 2006 and in November 2010 this was extended to Luxembourg and Germany (after the merger of Powernext and EEX, the new EPEX Spot and EPD futures in April 2009). "Powernext Intraday" and "Powernext Continuous" markets were introduced in July 2007.

On the 7th of December 2010 the law 2010-1488 was issued, the "Loi NOME", establishing a new model for the electricity market. The main objective of this law was to effectively open the market by resolving the problem of the competitors' access to competitive sources of electricity. It ensures the transitory right of access to the Historical Nuclear Regulated Electricity (ARENH) by alternative suppliers at a regulated price and volume, which are both determined annually by CRE with a maximum volume limitation by law to 100 TWh/year (Journal Officiel de la République Française, 2010).

Additional details on the French electricity market can be found in Lévêque (2010).

3. The Spanish electricity market

An agreement was reached between the authorities and the electricity companies late in December 1996 (Ministerio de Industria y Energía - Spain, 1996), allowing for the electricity sector reform.

The law for the electricity sector issued in November 1997 established the electricity sector regulation with the objectives to guarantee the supply, the quality of supply at the minimum possible cost while respecting the environment. The existing public service was replaced by the guarantee of supply for all consumers; the electrical sector was privatised on the generation and commercialisation sides and regulated on the transmission and distribution sides (Boletín Oficial del Estado - Spain, 1997). The transmission system was assigned to Red Eléctrica de España (REE) and in January 1998 an electricity spot market was introduced in Spain (OMEL).

After successive delays the Iberian electricity market (MIBEL) started operation in July 2007 and by 2008 the corresponding spot electricity market comprised 88% of total demand (Zachmann, 2008).

Additional details on the Spanish electricity market can be found in Crampes (2004), Furió & Lucia (2009) and Garrué-Irurzun & López-García (2009).

4. The Portuguese electricity market

The Decree-law 7/91 of the 8th of January established the conversion of the Portuguese public electricity company Electricidade de Portugal (EDP) into a private company still owned by the state. This would allow the unbundling of the Portuguese electricity sector and later privatisation.

The re-privatisation of EDP was started in 1997 after the issue of the Decree-law 56/97 of the 14th of March which determined on the first phase the sale of 29.99% of its capital and was followed by several other phases, the last one in 2012.

The transmission system operation was assigned to Redes Energéticas Nacionais (REN), created in 1994, under the ownership of EDP. In the end of the year 2000 the Portuguese state acquired 70% of REN from EDP. Only in 2007 the initial phase of REN's privatization (Redes Energéticas Nacionais, 2012a) took place. Currently EDP still owns a 5% share in REN (Redes Energéticas Nacionais, 2012b).

The Portuguese regulator for the energy sector (ERSE) was created in 1995 with the Decree law 187/95 of 27th of July (Diário da República Portuguesa, 1995) and has since then been adjusted through several other laws to the requirements of the energy sector and EU requirements (Silva, 2007).

The Iberian electricity market was only a reality in July 2007 after several years of preparation and negotiation between the Portuguese and the Spanish states. The MIBEL is composed by a spot (OMIE) and a bilateral (OMIP) electricity markets (Conselho de Reguladores do MIBEL, 2009).

Additional details on the Portuguese electricity market can be found in Amorim et al. (2010).

5. Interconnections between Portugal, Spain and France

Interconnections offer numerous advantages under normal operating conditions, such as optimal power station daily production, increasing opportunities for operation with renewable energies, the creation of competition and improvement of supply security.

However interconnectors are limited and have constraints due to physical behavior. Electrical current behaves like a fluid in a pipe; it flows through the easiest path. Therefore we have high voltage grids interconnected through many interconnectors placed in different geographic positions, which originate unidentified flows not necessarily related with cross-border contracts. Also, a consumer that contracted with one generator across the border will probably receive electricity from a different generator. All this physical properties of high voltage grids can create congestion of transmission lines and interconnectors causing the so called Loop Flow Problem (Coppens and Vivet, 2006).

Constraints have then to be managed by the Transmission System Operators (TSO) and specifically cross-border exchanges in electricity have to comply with European Community Regulation 1228/2003/EC of 26 June 2003 and later with European Community Regulation 714/2009 of 13 July 2009. These Regulations established initially a set of rules for cross-border exchanges in electricity, in order to enhance competition, establish a compensation mechanism for cross-border flows of electricity, setting principles on cross-border transmission charges and allocating available capacities of interconnections (European Union, 2003). With the latest Regulation the creation of the European Network of Transmission System Operators (ENTSO) was established, aiming to prepare network codes to guarantee an efficient transmission network management, together with allowing trade and supply of electricity across borders.

Transmission Costs allocation methods can be Flat Rate based or Flow-based. Flat rate methods are simple to calculate and implement, however unfair to generators that use less capacity and extent of the transmission lines (Galiana et al., 2003).

Flow-based costs are most commonly used due to their dependence on the capacity and extent used by each generator of the transmission lines. Explicit auctioning, where interconnector capacity is sold to the highest bidder or implicit auctioning, which integrates electricity and transmission markets and also called Market Splitting/Price Coupling, are both used across Europe (Coppens and Vivet, 2006).

During the period of this study, the method of explicit auctioning is used in the Spain-France interconnection, however the Market Splitting mechanism is applied to the Portuguese-Spanish interconnection.

The Spain-France electrical interconnection currently consists of four HV lines: Arkale-Argia, Hernani-Argia, Biescas-Pragneres, Vic-Baixas. These have a total commercial exchange capacity of 1,400 MW for transits from France to Spain and 1,100 MW for transits from Spain to France.

To fulfill the requirements of the European Commission a new HV line is being built by INELFE, a consortium with equal shares of the Spanish National Grid (Red Eléctrica de España - REE) and French National Grid (Réseaux de Transport d'Électricité - RTE). This new line will double the current interconnection capacity and will be in operation by 2015 (INELFE, 2011). The development of the interconnection capacity will allow a better market integration and provide additional security of electricity supply, being considered a critical factor to ensure integration (Everis and Mercados EMI, 2010).

The Portugal-Spain electrical interconnection currently consists of eleven HV lines, of which the last two have in practice no use, with an average capacity of 1800 MW: Alto Lindoso – Cartelle 1, Alto Lindoso – Cartelle 2, Lindoso – Conchas, Lagoaça – Aldeadávila, Pocinho – Aldeadávila 1, Pocinho – Aldeadávila 2, Pocinho – Saucelle, Falagueira – Cedillo, Alqueva – Brovales, Tavira – P. Guzman.

A new interconnection line between Viana do Castelo and Fontefria is planned to be constructed and forecasted to be in service by 2017, which with several other internal line reinforcements will allow the completion of the interconnection capacity between Portugal and Spain of 3000 MW, essential for the joint Iberian electricity market MIBEL (Redes Energéticas Nacionais, 2015).

6. Data

Day-ahead spot electricity prices in €/MWh (base, peek and off-peak), obtained from *Datastream*, were used in this study from the 1st of January 2012 to the 31st of December 2014. The data for the day-ahead base, peak and off-peak spot electricity prices is plotted in Figure 1.

Price spikes are observed in electricity markets, which confirms the high volatility behaviour of electricity spot prices, as in Goto and Karolyi (2004), Hadsell et al. (2004) and Higgs (2008). The limited possibility of storage, the physical characteristics of simultaneous electricity production and consumption, technical constraints in transmission and generating plants are the main reasons for these spikes (Coppens and Vivet, 2006; Silva and Soares, 2008).

After transforming the prices into their natural logarithms, to obtain directly the elasticity values from the parameter estimates, summary statistics were calculated (Table 1). Skewness and kurtosis values also indicate non-normal distribution, which is confirmed by JB statistic.

Unit root tests were made to all daily-log spot electricity prices. As per Table 2 we observe that all time series are considered to be stationary at 5% agreeing with findings in Park et al. (2006) and Bunn and Gianfreda (2010)

Daily average interconnection capacities were obtained from the corresponding system operator (REN, REE and RTE) and are plotted in Figure 2.

Daily weather data was retrieved from the website www.wunderground.com: maximum and minimum ambient temperatures (in degrees Celsius) and average wind speed (in km/h) for each country of the SWE REMs (Figure 3). Given the large number of installed wind power plants in the SWE electricity markets, it is believed to be a good approximation to use averaged weather variables across the existing weather stations linked to the www.wunderground.com website. In this way a country average is calculated for every hour and then averaged for every day. Maximum and minimum ambient temperatures were then used to calculate Heating Degree-days (HDD) and Cooling Degree-days (CDD) according to the UK Meteorological Office method (Mourshed, 2012; UK Climate Projections 2009, 2013).

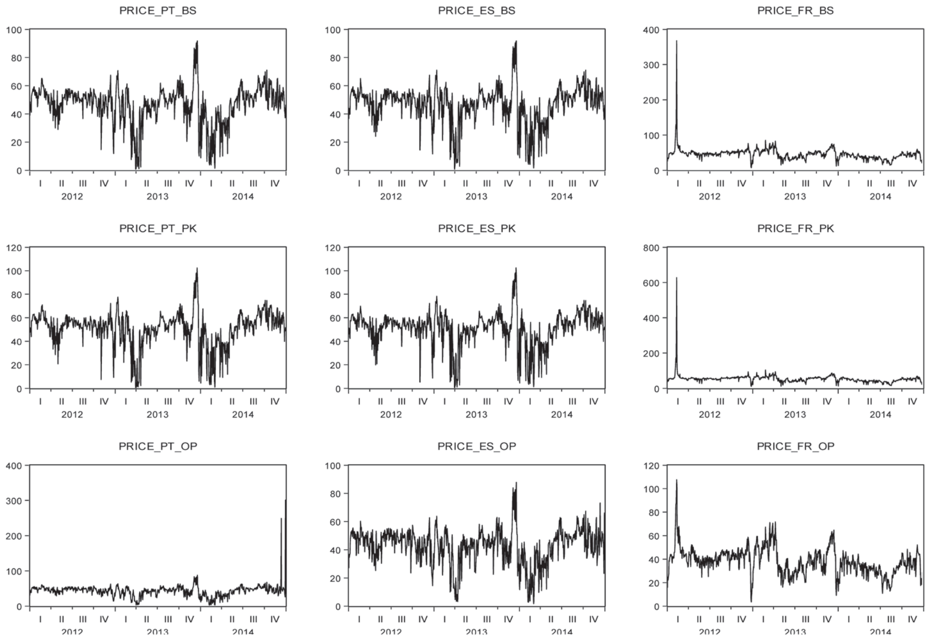


Figure 1: Day-ahead base spot electricity prices – Price_PT (Portugal), Price_ES (Spain) and Price_FR (France)

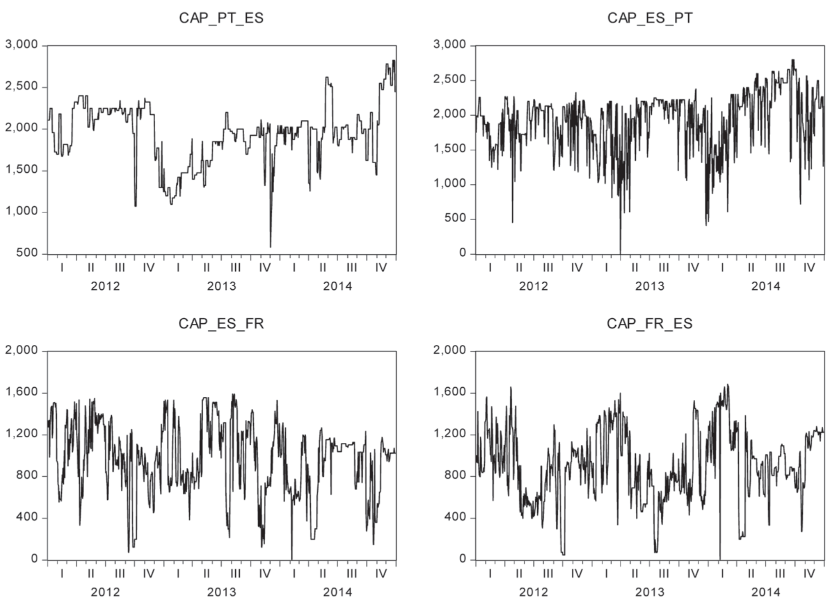


Figure 2: Import and export interconnection capacities between Portugal-Spain and Spain-France [MW]

Table 1: Summary Statistics

	PRICE_PT_BS	PRICE_ES_BS	PRICE_FR_BS	PRICE_PT_PK	PRICE_ES_PK	PRICE_FR_PK
Mean	46.56047	46.75766	45.57176	50.50936	50.88954	52.8457
Median	49.2	49.15	45.71	53.4	53.42	53.31
Maximum	91.89	91.89	367.6	102.42	102.42	627.59
Minimum	0.79	0.79	7.11	0.05	0.05	10.67
Std. Dev.	13.64581	13.32416	17.45264	15.17455	14.85422	25.77347
Skewness	-0.91986	-0.884998	8.5148	-1.038287	-1.049672	14.57907
Kurtosis	5.079319	5.340286	151.869	5.358144	5.676603	318.9676
Jarque-Bera	251.4779	280.8955	732495.9	322.1068	377.5185	3284872
Probability	0	0	0	0	0	0
Observations	783	783	783	783	783	783

	PRICE_PT_OP	PRICE_ES_OP	PRICE_FR_OP	PT_HDD	ES_HDD	FR_HDD
Mean	43.08826	42.55379	38.2928	2.251454	3.330681	4.43684
Median	45.195	45.015	38.16917	1.282996	1.819442	3.077545
Maximum	301.285	87.805	107.6158	12.00704	14.10248	22.43006
Minimum	1.745833	1.72	3.538333	0	0	0
Std. Dev.	17.47134	12.71196	11.61162	2.363278	3.599877	4.265026
Skewness	5.878117	-0.644137	0.70576	1.149727	0.905703	0.953528
Kurtosis	87.1428	4.479365	6.233589	3.893398	2.752871	3.153219
Jarque-Bera	235494.4	125.5465	406.1317	198.5441	109.0414	119.4185
Probability	0	0	0	0	0	0
Observations	783	783	783	783	783	783

	PT_CDD	ES_CDD	FR_CDD	PT_WSAVG	ES_WSAVG	FR_WSAVG
Mean	0.874864	1.071543	0.900092	3.529939	5.966993	2.601971
Median	0	0	0	2.908853	4.976699	2.429253
Maximum	9.009729	8.91479	10.60773	12.22416	19.2016	10.84552
Minimum	0	0	0	0.275082	0.606043	0.042626
Std. Dev.	1.472858	1.751397	1.724336	2.205962	3.488747	1.39227
Skewness	2.252565	1.754339	2.685284	0.945743	1.29511	1.141516
Kurtosis	8.432273	5.339177	10.9702	3.41704	4.545661	5.79865
Jarque-Bera	1624.914	580.1562	3013.477	122.3973	296.8325	425.5825
Probability	0	0	0	0	0	0
Observations	783	783	783	783	783	783

	CAP_PT_ES	CAP_ES_PT	CAP_ES_FR	CAP_FR_ES
Mean	1943.556	1914.928	968.0859	907.9869
Median	2000	1991.667	1025	893.7917
Maximum	2825	2800	1592	1686
Minimum	583.3333	0	0	0
Std. Dev.	354.9334	422.2091	343.1235	338.215
Skewness	-0.354533	-0.830241	-0.412352	-0.138102
Kurtosis	3.246843	3.979948	2.663116	2.711556
Jarque-Bera	18.39094	121.2834	25.89208	5.203325
Probability	0.000101	0	0.000002	0.07415
Observations	783	783	783	783

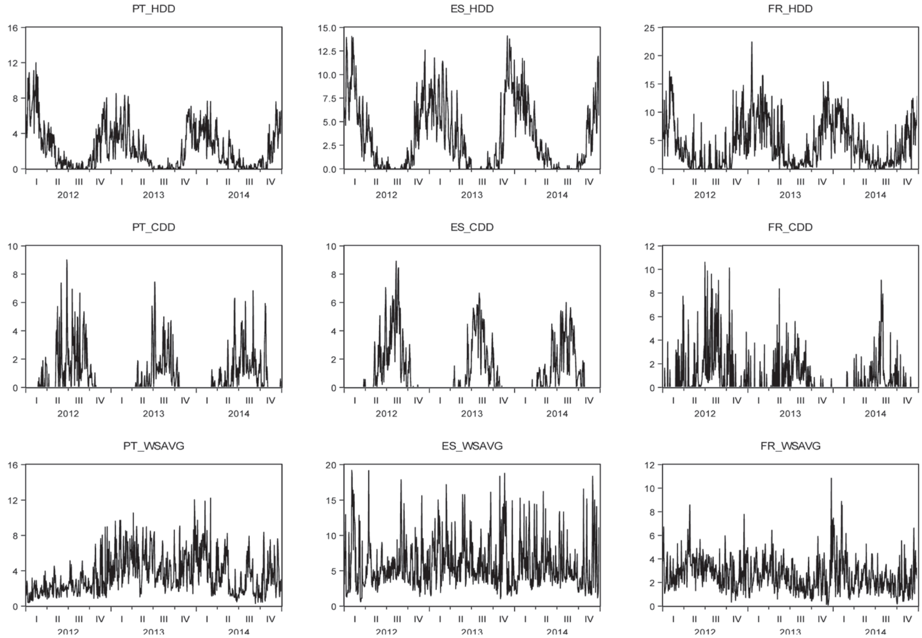


Figure 3: HDD, CDD [ºC] and Average Wind Speed [km/h] in Portugal, Spain and France

It is to note the big variability in the average wind speed. Literature reports some related issues, as such: transport of excess production, electrical system fault endurance, available and flexible standby generating capacity and effective control or curtailment of wind power production (Benatia et al., 2013; Franco and Salza, 2011; Söder et al., 2007).

Table 2: Unit root tests

	L_PRICE_PT			L_PRICE_ES			L_PRICE_FR		
	Base	Peak	Off-peak	Base	Peak	Off-peak	Base	Peak	Off-peak
ADF test	-5.510	-5.081	-5.592	-5.743	-5.230	-5.855	-8.130	-6.507	-8.020
(p-value)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
PP test	-12.141	-11.938	-10.518	-12.837	-13.498	-10.083	-8.370	-10.173	-8.135
(p-value)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)

7. Model Estimation

The VAR model has proven to be especially useful for describing the dynamic behavior of economic and financial time series and for forecasting. It is known to provide superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models. In addition to data description and forecasting, the VAR model is also used for structural inference and policy analysis (Lütkepohl, 2005; Sims, 1980). In structural analysis, certain assumptions about the causal structure of the data under investigation are imposed, and the resulting causal impacts of unexpected shocks or innovations to specified variables on the variables in the model are summarized. These causal impacts are usually summarized with impulse response functions, as is performed in this work.

A VARX model was then considered to proceed with the evaluation of the determinants in the electricity market integration, due to its ability in capture the linear interdependencies among multiple time series.

Considering a VARX model for the three log prices:

$$\mathbf{Y}_t^{(z)} = \mathbf{C}^{(z)} + \sum_{i=1}^{p} \mathbf{A}_i^{(z)}\mathbf{Y}_{t-i}^{(z)} + \mathbf{B}^{(z)}\mathrm{X}_t^{(z)} + \mathbf{u}_t^{(z)} \quad (1)$$

where z is the base, peak or off-peak model, $\boldsymbol{Y}_t^{(z)} = (l\_Price\_PT_t^{(z)},\ l\_Price\_ES_t^{(z)},\ l\_Price\_FR_t^{(z)})'$ the day-ahead electricity price matrix, $\boldsymbol{X}_t^{(z)}$ the exogenous variables matrix, $\mathrm{C}^{(z)}$ are (3x1) constant matrices, $\boldsymbol{A^{(z)}}_i$ and $\mathrm{B}^{(z)}$ are (1x3) coefficient matrices and $\mathrm{u}^{(z)}{}_t$ are (3x1) matrices of unobservable error terms. In order to determine the order of each the models, successive VAR models were estimated by a sequential test procedure, starting with the estimation of the models with *p =15* lags and calculating-down for lower lags the Schwarz Bayesian criterion (BIC) and the Hannan-Quinn criterion (HQC).

In Table 3 the best values for the endogenous variable lags where criteria are minimised are presented. For each model a lag exclusion Wald test was performed in order to detect lags where the respective coefficients do not present significance in the model, which were then removed as indicated. Autocorrelation testing in all models was performed (Davidson and Mackinnon, 2004).

Table 3: Lag selection for estimated models

	Price VAR model				Price VARX model			
	Base							
Lag Length Criteria	Lags	SC	HQ	Lag removed	Lags	SC	HQ	Lag removed
	5	-1.908569	-2.087095*	NA	8	-1.906685	-2.330686*	5
	Off-peak							
Lag Length Criteria	Lags	SC	HQ	Lag removed	Lags	SC	HQ	Lag removed
	4	-1.8528	-1.997853*	3 and 5	4	-2.02962	-2.319726*	2 and 3
	Peak							
Lag Length Criteria	Lags	SC	HQ	Lag removed	Lags	SC	HQ	Lag removed
	9	-0.18936	-0.501781*	NA	9	-0.156834	-0.614308*	NA

8. Analysis and discussion of results

Weather conditions have impacts on both demand and supply of electricity. The estimated VARX model provides insights of the related dynamics between the considered exogenous variables and spot electricity prices.

CDD and HDD are considered proxies for electricity demand, therefore a positive contribution in the models is expected. The results shown in Table 4 demonstrate that the exogenous variables related with CDD do not contribute too much for the model specification, whereas HDD improves the model in some cases. Positive significant contributions are found for the Spanish HDD in both spot electricity prices for Portugal and Spain. However, it is interesting to note that the Portuguese HDD has significant negative contributions to these same prices, which might be related with weather dynamics (which were not modeled here) rather than price dynamics. The French HDD only provides a positive contribution to the spot electricity price in France. Furthermore, during peak periods only the Spanish HDD significant positive contributions on both Iberian spot electricity prices remain.

A relevant improvement in model specification is found by incorporating as exogenous variables the average wind speeds. It is expected that average wind speed contributes negatively to spot electricity prices due to the normally low marginal prices bid into spot markets. This is actually seen for almost all Portuguese, Spanish and French average wind speeds where significant negative contributions to spot electricity prices are found. Some small positive contributions are found, however weaher dynamics might explain these. It is to note that an increase of 1 km/h in the Portuguese average wind speed contributes to a 4.46% decrease in the base Portuguese spot electricity price (Table 4). Furthermore, there is a 3.06% negative contribution of the French average wind speed to the base Portuguese spot electricity price. However given the small existing interconnection between France and Spain, this contribution might be related with weather dynamics rather than arbitrage between markets.

It would also be expected that the growth in ATC would contribute to a higher level of arbitrage, thus with negative effects to spot electricity prices. However, ATC significant contributions to spot electricity prices do not have a major contribution to the model specification.

As per Figure 4 all three models satisfy the stability condition of no roots outside the unit circle.

Table 4: Wind [km/h], HDD [°C] and CDD [°C] significant coeficients in the VARX model

Price	Base			Off-peak			Peak		
	Wind_PT	Wind_ES	Wind_FR	Wind_PT	Wind_ES	Wind_FR	Wind_PT	Wind_ES	Wind_FR
L_PRICE_PT	-4.46%	-0.70%	-3.06%	-3.98%	-0.89%	-2.89%	-4.40%		-2.85%
L_PRICE_ES	-3.92%	-0.82%	0.19%	-3.45%	-1.12%	-3.06%	-4.04%		-3.58%
L_PRICE_FR	-1.18%	0.49%	-2.60%	-1.34%	0.54%	-3.18%	-1.12%	0.55%	-2.36%
Price	Base			Off-peak			Peak		
	HDD PT	HDD ES	HDD FR	HDD PT	HDD ES	HDD FR	HDD PT	HDD ES	HDD FR
L_PRICE_PT	-2.28%	2.12%		-1.89%	1.58%			2.93%	
L_PRICE_ES	-2.40%	2.20%		-2.12%	1.54%			3.20%	
L_PRICE_FR			0.59%			0.65%			
Price	Base			Off-peak			Peak		
	CDD PT	CDD ES	CDD FR	CDD PT	CDD ES	CDD FR	CDD PT	CDD ES	CDD FR
L_PRICE_PT									
L_PRICE_ES									
L_PRICE_FR		0.98%			1.26%				

Table 5: ATC [MW] significant coeficients in the VARX model

Price	Base				Off-peak			
	ATC PT-ES	ATC ES-PT	ATC ES-FR	ATC FR-ES	ATC PT-ES	ATC ES-PT	ATC ES-FR	ATC FR-ES
PRICE_PT		0.02%	0.01%			0.02%	0.01%	
PRICE_ES	-0.01%	0.02%	0.01%	-0.01%	-0.01%	0.02%	0.01%	
PRICE_FR		0.00%	0.00%			0.00%	0.00%	

Price	Peak			
	ATC PT-ES	ATC ES-PT	ATC ES-FR	ATC FR-ES
PRICE_PT		0.02%		
PRICE_ES		0.02%	0.01%	
PRICE_FR			-0.01%	

Granger Causality tests to the time-series variables and impulse response analysis displaying the responses of each daily-log price time-series to a standard error shock in one of the time-series were carried out to the models considered and are presented, respectively, in Table 6 and in Figure 5 to Figure 7.

Outcomes in Table 6 show that both MIBEL market prices fail to Granger-cause the Powernext market prices on a pairwise relation. This can likewise be observed in the impulse responses of the French spot electricity market prices, which are practically inexistent to shocks in any one of the MIBEL spot electricity market prices. In spite Powernext market prices Granger-cause the MIBEL_ES price in all models, the impulse response analysis indicates a very weak effect. Additionally, there is a Granger-causality relation between Powernext and MIBEL_PT base and peak prices, yet fairly weak as confirmed by the impulse response analysis.

Within Iberia both MIBEL prices Granger-cause each other in all base, peak and off-peak models, which confirms the good integration between both Iberian electricity markets. This is also seen in the impulse response plots with strong responses of the Spanish spot electricity price to shocks in the Portuguese spot electricity price in all base, off-peak and peak models and vice-versa.

9. Final Remarks

The level of integration and some determinants on the South West Europe regional electricity market are evaluated in this study in order to assess the degree of accomplishment of the building of the European Internal market as aimed by the consecutive European Directives.

A future common competitive electricity market is aimed by European policy giving guidance to Member-State policy and statutes. The Electricity Regional Initiative was later on launched along this long process to attain the common electricity market. Simultaneously, the promotion of electricity generation by renewable energy sources was similarly an objective in Europe, reducing the dependency on imported fossil fuels and allowing GHG emissions mitigation. Large deployment of RES-E generation in Europe has been achieved through strong financial support mechanisms.

Results obtained from the model specification herein presented show that HDD contributed to a slight improvement to the model significance. Furthermore, it is to note that the introduction of the exogenous variable average wind speed improves all models specification.

The average wind speed both in Portugal and Spain show a significant negative influence for the Portuguese and Spanish base, off-peak and peak log-daily spot electricity prices. Regarding the French average wind speed, there is a significant negative contribution in all Powernext base, off-peak and peak log-daily spot electricity prices.

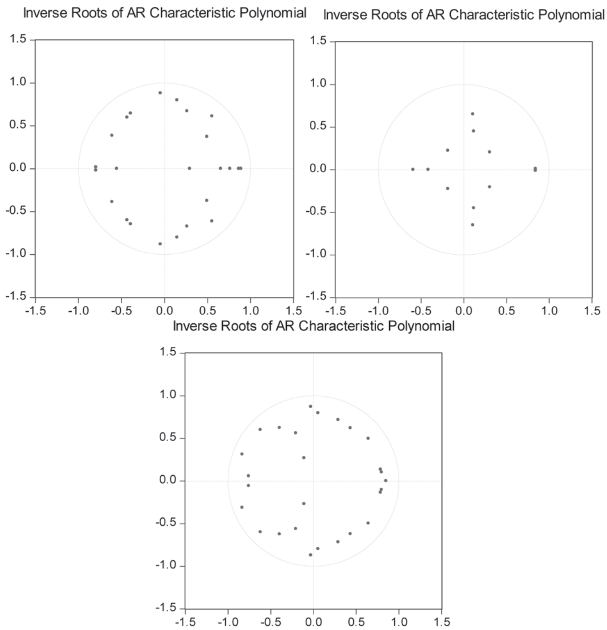


Figure 4: Unit circle plot for base, off-peak and peak models (left to right)

Table 6: Granger Causality test output

Base				Off-peak				Peak			
Dependent variable: L_PRICE_PT_BS				Dependent variable: L_PRICE_PT_OP				Dependent variable: L_PRICE_PT_PK			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
L_PRICE_ES_BS	64.21	7	0.00	L_PRICE_ES_OP	7.48	2	0.02	L_PRICE_ES_PK	82.37	9	0.00
L_PRICE_FR_BS	9.85	7	0.20	L_PRICE_FR_OP	6.28	2	**0.04**	L_PRICE_FR_PK	16.57	9	0.06
All	75.22	14	0.00	All	13.69	4	0.01	All	96.15	18	0.00
Dependent variable: L_PRICE_ES_BS				Dependent variable: L_PRICE_ES_OP				Dependent variable: L_PRICE_ES_PK			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
L_PRICE_PT_BS	59.98	7	0.00	L_PRICE_PT_OP	21.92	2	0.00	L_PRICE_PT_PK	81.87	9	0.00
L_PRICE_FR_BS	12.44	7	0.09	L_PRICE_FR_OP	3.49	2	0.17	L_PRICE_FR_PK	14.41	9	0.11
All	74.34	14	0.00	All	26.27	4	0.00	All	93.57	18	0.00
Dependent variable: L_PRICE_FR_BS				Dependent variable: L_PRICE_FR_OP				Dependent variable: L_PRICE_FR_PK			
Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.	Excluded	Chi-sq	df	Prob.
L_PRICE_PT_BS	3.49	7	**0.84**	L_PRICE_PT_OP	2.65	2	**0.27**	L_PRICE_PT_PK	7.57	9	**0.58**
L_PRICE_ES_BS	1.43	7	**0.98**	L_PRICE_ES_OP	3.71	2	**0.16**	L_PRICE_ES_PK	7.32	9	**0.60**
All	11.49	14	**0.65**	All	5.28	4	**0.26**	All	22.52	18	**0.21**

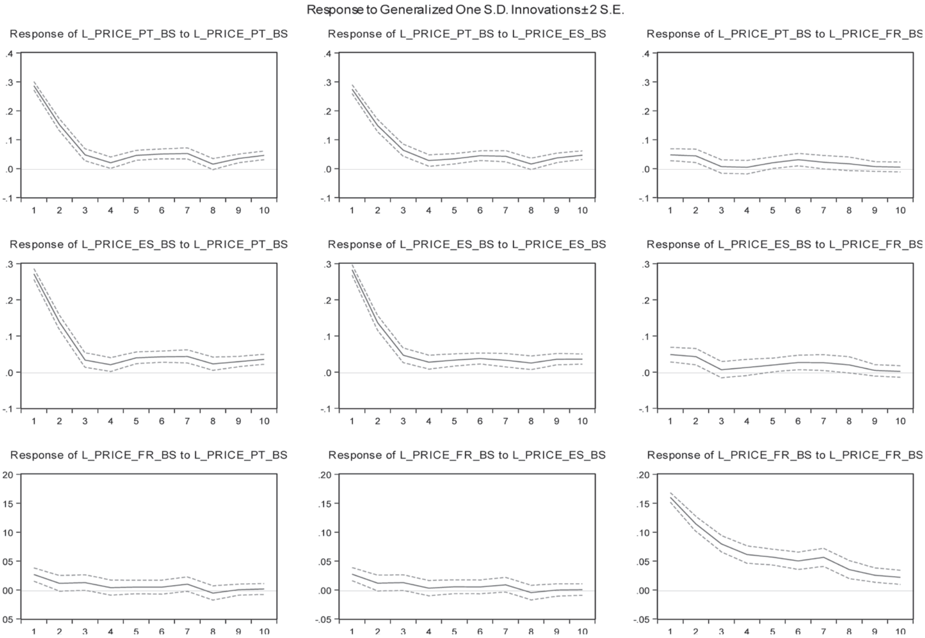


Figure 5: Impulse response plots for daily-log base price models

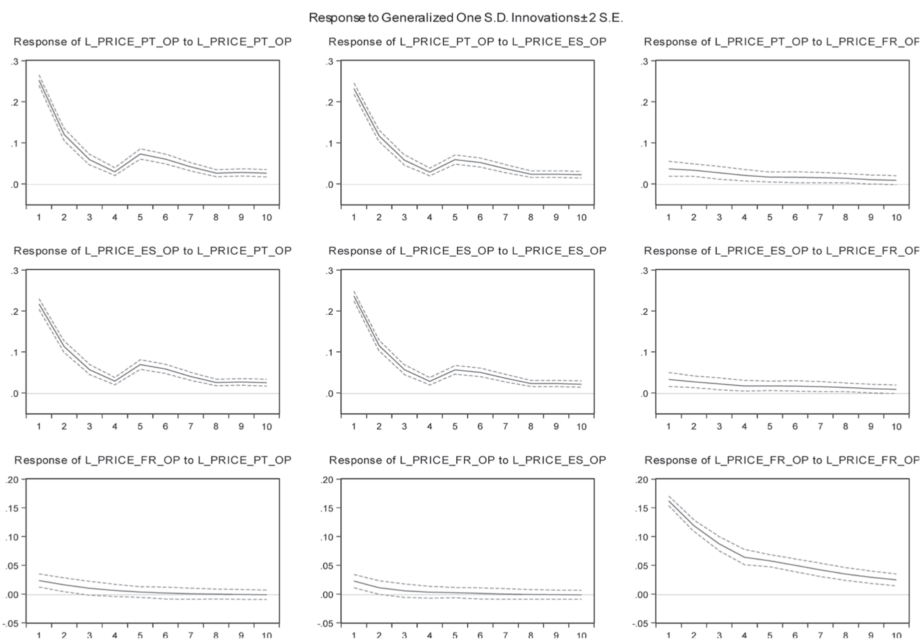


Figure 6: Impulse response plots for daily-log off-peak price models

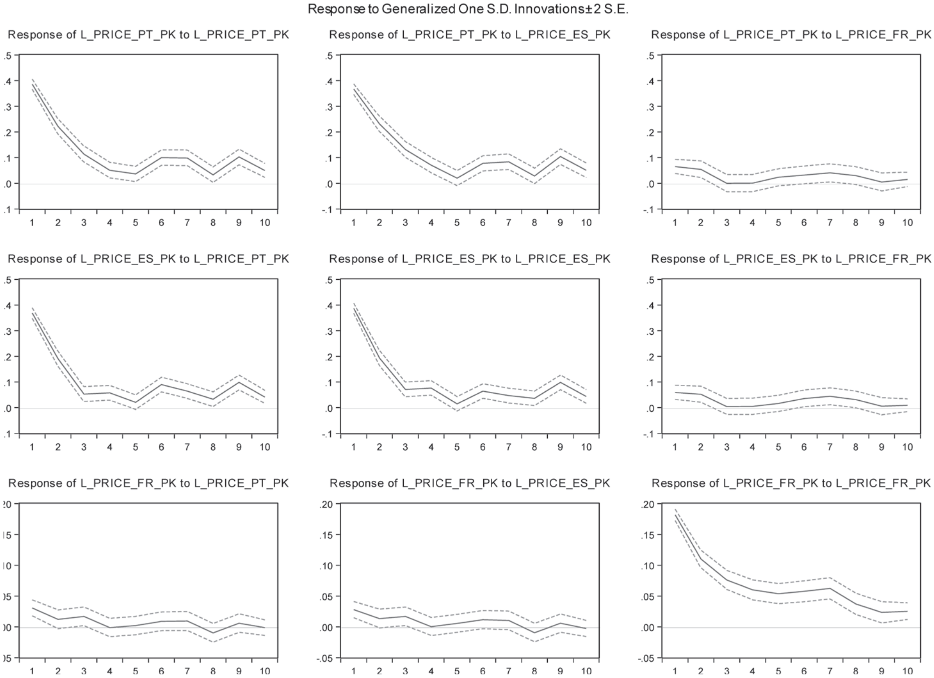


Figure 7: Impulse response plots for daily-log peak price models

Strong integration was found between MIBEL markets, leading to the conclusion that the existing market setup is efficient and contributes to the integration of spot electricity markets. Also, as demonstrated by Grange-causality and impulse response analysis, there is a weak integration level between MIBEL and Powernext.

Findings in this article related with the impact of wind generation on interconnected markets are aligned with conclusions reached by several studies, albeit relying on a novel data and modelling approach. Results herein shown, highlight the importance of efficient electricity market design and effective renewable policies in facilitating RES-E penetration as in Klessmann et al. (2008), Milligan et al. (2009), Cruz et al. (2011) or Cutler et al. (2011). The significant negative contribution of wind speed on electricity spot market price does not mean that the electricity consumer price also decreases. As a matter of fact Silva and Cerqueira (2013) recently established a significant positive impact of 1.8% increase in consumer price for each 1% of RES-E. The incentives for the deployment of RES-E (Meyer, 2003) should be reviewed as it has been established to be in some cases a high burden on consumer electricity prices (Amorim et al., 2010; Sáenz de Miera et al., 2008), distorting the desired effect of the electricity spot market price decrease.

It is relevant to emphasize that in spite of the fact that the interconnection capacities available do not improve model specification *per se*, these are extremely important to transport the electricity generated by renewable sources and more specifically wind generation. The lack of sufficient interconnection capacity between France and Spain is likely to explain the non-significance of Spanish average wind speed on the French electricity market price. Additional work is currently being pursued to tackle this issue with more depth.

Having the current Internal Energy Market Directive aim as guideline, conclusions found in this study support that coupling and interconnection capacity expansion should be continuously sought between the French and Spanish electricity markets in order to achieve a full functioning South West Electricity regional market.

Acknowledgements: We would like to recognise the kind assistance provided by the Portuguese TSO REN for making data available and Carlos Freitas for the support given in the paper review.

This Work has been framed under the Energy for Sustainability Initiative of the University of Coimbra, supported by the R&D Project EMSURE, Energy and Mobility for Sustainable Regions (CENTRO 07 0224 FEDER 002004) and also partially supported by the Portuguese Foundation for Science and Technology under project grant[s] PEst-OE/ EEI/UI308/2014.

References

Amorim, F., Martins, M.V.M., Pereira da Silva, P., 2010. A new perspective to account for renewables impacts in Portugal, in: 2010 7th International Conference on the European Energy Market. IEEE, pp. 1–6.

Benatia, D., Johnstone, N., Haščič, I., 2013. Effectiveness of Policies and Strategies to Increase the Capacity Utilisation of Intermittent Renewable Power Plants. OECD Environ. Work. Pap. OECD Publ. 1–49.

Boletín Oficial del Estado - Spain, 1997. Ley del Sector Electrico. Spain.

Bower, J., 2002. Seeking the Single European Electricity Market - Evidence from an Emprirical Analysis of Wholesale Market Prices. Report, Oxford Inst. Energy Stud. 1–42.

Brown, A., Müller, S., Dobrotková, Z., 2011. Renewable Energy - Markets and Prospects by Technology. Report, Int. Energy Agency 1–62.

Bunn, D.W., Gianfreda, A., 2010. Integration and shock transmissions across European electricity forward markets. Energy Econ. 32, 278–291.

CEER, 2015. Electricity Regional Initiative [WWW Document]. URL http://www.ceer.eu/portal/page/portal/EER_HOME/EER_ACTIVITIES/EER_INITIATIVES/ERI (accessed 5.4.15).

Commission de Régulation de L'énergie, 2011. Observatoire des marchés de l'électricité et du gaz - 4e trimestre 2011.

Conselho de Reguladores do MIBEL, 2009. Descrição do Funcionamento do MIBEL.

Coppens, F., Vivet, D., 2006. The single European electricity market : A long road to convergence. Natl. Bank Belgium, Work. Pap. 1–53.

Crampes, C., 2005. The Spanish Electricity Industry : Plus ça change .... Energy J. 26, 127–154.

Cruz, A., Muñoz, A., Zamora, J.L., Espínola, R., 2011. The effect of wind generation and weekday on Spanish electricity spot price forecasting. Electr. Power Syst. Res. 81, 1924–1935.

Cutler, N.J., Boerema, N.D., MacGill, I.F., Outhred, H.R., 2011. High penetration wind generation impacts on spot prices in the Australian national electricity market. Energy Policy 39, 5939–5949.

Davidson, R., Mackinnon, J.G., 2004. Econometric Theory and Methods, illustrate. ed. Oxford University Press.

Diário da República Portuguesa, 1995. Decreto-lei 187/95 de 27 de Julho. Portugal.

EDF Group, 2010. Activity & Sustainable Development. Report, 1–108.

ERGEG, 2006. The Electricity Regional Initiative: Making Progress Towards a Single European Market. Electr. Reg. Initiat. - Fact Sheet 1–5.

European Union, 1990a. Directive 90/547/EEC of 29 October 1990 on the transit of electricity through transmission grids. Off. J. Eur. Communities L 313, 30–33.

European Union, 1990b. Directive 90/377/EEC of 29 June 1990 concerning a Community procedure to improve the transparency of gas and electricity prices charged to industrial end-users. Off. J. Eur. Communities L 185, 16–24.

European Union, 2003. European Community Regulation 1228/2003/EC of the European Parliament and of the Council of 26 June 2003 on conditions for access to the network for cross-border exchanges in electricity. Off. J. Eur. Union L 176, 1–10.

Europex, 2009. Multi Regional Day-Ahead Price Coupling - Towards Implementation of the PCG Target Model. Conf. Proceedings, Florence Regul. Forum 1–5.

Europex, 2011. Status of Implementation of PCR for Day-Ahead Price Coupling. Conf. Proceedings, Florence Regul. Forum 1–7.

Everis, Mercados EMI, 2010. From Regional Markets to a Single European Market.

Figueiredo, N., Silva, P.P. da, 2012. Integration of South-West Spot Electricity Markets: An Update, in: Conference Proceedings, 12th IAEE - European Energy Conference. pp. 1–14.

Franco, A., Salza, P., 2011. Strategies for optimal penetration of intermittent renewables in complex energy systems based on techno-operational objectives. Renew. Energy 36, 743–753.

Furió, D., Lucia, J.J., 2009. Congestion management rules and trading strategies in the Spanish electricity market. Energy Econ. 31, 48–60.

Galiana, F.D., Conejo, A.J., Gil, H.A., 2003. Transmission network cost allocation based on equivalent bilateral exchanges. IEEE Trans. Power Syst. 18, 1425–1431.

Garrué-Irurzun, J., López-García, S., 2009. Red Eléctrica de España S.A.: Instrument of regulation and liberalization of the Spanish electricity market (1944-2004). Renew. Sustain. Energy Rev. 13, 2061–2069.

Glachant, J.M., Finon, D., 2004. A Competitive Fringe in the Shadow of a State Owned Incumbent : The Case of France. Energy J. 181–204.

Goto, M., Karolyi, G.A., 2004. Understanding Electricity Price Volatility Within and Across Markets. Dice Center, Work. Pap. 2004-12, 1–41.

Hadsell, L., Marathe, A., Shawky, H. a., 2004. Estimating the Volatility of Wholesale Electricity Spot Prices in the US. Energy J. 25, 23–40.

Higgs, H., 2008. Modelling spot prices in deregulated wholesale electricity markets : A selected empirical review, Working Paper,.

INELFE, 2011. INELFE website [WWW Document]. URL http://www.inelfe.eu/?-rubrique25-&lang=en

Jager, D. de, Klessmann, C., Stricker, E., Winkel, T., Visser, E. de, Koper, M., Ragwitz, M., Held, A., 2011. Financing Renewable Energy in the European Energy Market. Report, Ecofys 1–264.

Jäger-waldau, A., Szabó, M., Monforti-ferrario, F., Bloem, H., Huld, T., Arantegui, R.L., 2011. Renewable Energy Snapshots 2011. Report, Eur. Comm. JRC - iet 1–54.

Jamasb, T., Pollitt, M., 2005. Electricity Market Reform in the European Union : Review of Progress toward Liberalization & Integration. Energy J. 26, 11–41.

Journal Officiel de la République Française, 2000. LOI nº 2000-108 10 février relative à la modernisation et au développement du service public de l'électricité.

Journal Officiel de la République Française, 2010. LOI n° 2010-1488 du 7 décembre portant nouvelle organisation du marché de l'électricité.

Karova, R., 2011. Regional electricity markets in Europe: Focus on the Energy Community. Util. Policy 19, 80–86.

Klessmann, C., Nabe, C., Burges, K., 2008. Pros and cons of exposing renewables to electricity market risks—A comparison of the market integration approaches in Germany, Spain, and the UK. Energy Policy 36, 3646–3661.

Lévêque, F., 2010. La nouvelle loi française de l' électricité : un barrage contre le marché et l' Europe. Cern. Work. Pap.

Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer.

Meeus, L., Belmans, R., 2008. Electricity Market Integration in Europe, in: Conference Proceedings, 16th Power Systems Computation Conference, Glasgow, Scotland. pp. 1–5.

Meyer, N.I., 2003. European schemes for promoting renewables in liberalised markets. Energy Policy 31, 665–676.

Milligan, M., Lew, D., Corbus, D., Piwko, R., Miller, N., Clark, K., Jordan, G., Freeman, L., 2009. Large-Scale Wind Integration Studies in the United States : Preliminary Results. Report, 8th Int. Work. Large Scale Integr. Wind Power Transm. Networks Offshore Wind Farms 1–5.

Ministerio de Industria y Energía - Spain, 1996. Protocolo para el Establecimiento de una Nueva Regulación del Sistema Eléctrico Nacional. Spain.

Newbery, D., 2005. Refining Market Design, in: Proc. Implementing the Internal Market of Electricity: Proposals and Time-Tables. SESSA - Sustainable Energy Specific Support Assessment.

Park, H., Mjelde, J.W., Bessler, D.A., 2006. Price dynamics among U.S. electricity spot markets. Energy Econ. 28, 81–101.

Redes Energéticas Nacionais, 2012a. REN-2012-grupo-ren_historia.pdf [WWW Document]. URL http://www.ren.pt/vPT/Investidor/GrupoREN/Historia/Pages/investidor_grupo_ren-historia.aspx

Redes Energéticas Nacionais, 2012b. Estrutura Accionista [WWW Document]. URL http://www.ren.pt/vPT/Investidor/InformacaoaoAccionista/EstruturaAccionista/Pages/InvestidorGrupoRENEstrutura.aspx

Redes Energéticas Nacionais, 2015. Capacidades indicativas de interligação para fins comerciais para o ano de 2015 2015, 1–72.

Sáenz de Miera, G., del Río González, P., Vizcaíno, I., 2008. Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. Energy Policy 36, 3345–3359.

Silva, P.P. da, 2007. O sector da energia eléctrica na União Europeia: Evolução e Perspectivas. Coimbra University Press.

Silva, P.P. Da, Soares, I., 2008. EU spot prices and industry structure: assessing electricity market integration. Int. J. Energy Sect. Manag. 2, 340–350.

Silva, P.P., Cerqueira, P., 2013. Drivers for household electricity prices in the EU : a system-GMM panel data approach. Proc. ICEE - Energy Environ. bringing together Econ. Eng. 1–12.

Sims, C.A., 1980. Macroeconomics and reality. Econometrica 48, 1–48.

Söder, L., Hofmann, L., Orths, A., Holttinen, H., Wan, Y., Tuohy, A., 2007. Experience From Wind Integration in Some High Penetration Areas. IEEE Trans. Energy Convers. 22, 4–12.

Zachmann, G., 2008. Electricity wholesale market prices in Europe: Convergence? Energy Econ. 30, 1659–1671.

Energy Efficiency Governance in the European Union Member States – Analysis on current status

Guillermo Ivan Pereira[1,3] **and Patrícia Pereira da Silva**[1,2,3]

Abstract

Improving energy efficiency in the European Union is a complex task, which requires the commitment of Member States to be accomplished. The existing 20% energy efficiency target for 2020, recently reinforced for 2030 towards a 27% energy consumption reduction goal creates a higher pressure to deliver the potential benefits for the economy, environment and society. This research paper presents a governance analysis, as a proxy on the ability of Member States to contribute to the existing energy efficiency targets. The governance analysis conducted for the EU-28 highlights that more efforts are required to ensure that Member States follow and adopt existing legislation, alongside with developments on the existing financial support mechanisms, human capacities and institutional structures. Furthermore, individual country analysis depicts a misalignment on the governance performance for the EU-28 Member States.

Keywords: Energy Efficiency, European Union, Governance, Energy Policy.

[1] Energy for Sustainability Initiative - MIT Portugal Program - University of Coimbra

[2] Faculty of Economics - University of Coimbra

[3] INESC Coimbra - Institute for Systems Engineering and Computers at Coimbra
e-mail: gpereira@student.dem.uc.pt, patsilva@fe.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_4

1. Introduction

Creating stimuli towards a more energy efficient European Union (EU) has been a key pillar in the existing policy agendas. Member States have been called to collaborate further on this effort through the latest Energy Efficiency Directive (EED) released in 2012 (European Parliament, 2012b) to ensure the achievement of the proposed energy efficiency (EE) targets by 2020, of 20% reduction of primary energy consumption set on the EU 2020 goals (European Commission, 2011b). Implementing actions that contribute to greater EE in the EU is crucial.

Whilst the EU strives to follow the plan for energy and climate for 2020, the European Commission (EC) has already defined the pipeline beyond 2020 and towards 2030. In this new policy framework the EE ambitions are greater, the EE communication released in 2014 (European Commission, 2014d) proposed a target for increasing EE by 27% in 2030. This evolution is backed by a set of governance guidelines that are designed to ensure the effective implementation of plans and Member States collaboration to achieve this target. It is worth noting that of the proposed energy and climate targets for 2020, only the one related to EE was at risk of not being achieved, the realistic adjusted potential identified by the EC (European Commission, 2014c) stood at 17%, which represented a positive evolution from previous estimations in 2009 that pointed at the possibility of only reaching 9% of energy consumption reduction instead of the 20% goal defined (European Commission, 2011a). These figures present a positive evolution towards the policy targets. However a small gap on EE represents a risk that has to be mitigated.

The aim of this paper is to present an EE governance analysis based on a framework proposed by Jollands & Ellis, (2009a), as well as to provide information on individual Member States EE governance performance. This research is designed to provide insightful indications on current strengths and weaknesses of EE governance in the EU, which are important from a policy makers' perspective when designing, developing or evaluating policy agendas and possible targets for future implementation. The organisation of the paper is as follows. Section 2, presents and discusses the concepts of governance and its application on the EE field, as well as EU–wide actions, Section 3, presents the selected indicators and analysis methods, Section 4, unveils the results of the governance analysis and Member States performance, and Section 5 concludes and presents pathways for future work.

2. Energy efficiency governance

The concept of governance applied to EE was presented by the International Energy Agency (IEA, 2010) as well as by Jollands & Ellis (2009a, 2009b), and associated with the use of political influence, organisations and resources by decision making agents to deliver greater EE. This concept can be easily understood when looking at the EE policy context in the EU, which engages EU-wide institutions, Member States Governments, citizens and private sector players into a set of strategies and roadmaps towards enhanced EE. To do so the EU implements strategies and directives at different levels in order to create the required motivation to deliver the planned targets.

2.1. EE actions in the EU

The different levels of EE actions in the EU are part of an interconnected governance framework which aims at ensuring the necessary mechanisms for delivering the energy and climate related targets.

2.1.1. Strategic level

At strategic level the EU has been implementing a range of policy roadmaps and communications that create a framework for action, and enable new directives and regulations as well as support actions to be designed as supporting mechanisms for these strategies. Table 1, below, presents the existing pipeline of strategic orientations from 2020 until 2050 for climate and energy.

Table 1: Strategic level action of EE EU governance adapted from (Pereira, 2014).

Strategic level action	Main goals /Ambitions	References
Europe 2020 20-20-20	20% reduction of primary energy consumption; 20% increase in renewable energy; 20% reduction of GHG emissions.	(European Commission, 2010, 2011b; European Parliament, 2012a)
2030 framework for climate and energy policies	This action sets the standards to build the 2030 European strategy for energy and climate, taking into account the learning points from Europe 2020 20-20-20 and the Europe 2050 Roadmap; 40% reduction of GHG emissions: 27% increase in renewable energy; 27% reduction of primary energy consumption.	(European Commission, 2013a, 2013b, 2014a)
Europe 2050 Roadmap	A secure, competitive and decarbonised energy system; 80-95% reduction of GHG emissions (indicative).	(European Commission, 2011c; European Parliament, 2013; Faber et al., 2012)

The strategic level actions in Table 1, serve as guidelines for what should be achieved in the future in terms of climate and energy goals within 5 (i.e.: in 2020), 15 (i.e.: in 2030) and 35 (i.e.: in 2050) years. These high-level strategies enable the creation of laws that enforce actions and the implementation of policies in a defined timeframe and consistent with a set of requirements.

2.1.2. Legislative level

In terms of the laws created, the EED is the key legislative instrument, enacted in 2012 which is far more ambitious, when compared to the previous Energy Services Directive (European Parliament, 2006). The new instrument goes beyond energy services and end-uses, including actions concerning the EE in the generation, transmission and distribution of energy in the EU. The key measures implemented by this Directive are further described by the European Commission (2011d, 2011e), and include measures for the public sector, residential consumers, Small and Medium Enterprises (SMEs) and energy sales companies.

The legislative actions at the EU level go beyond the EED, which is an overarching instrument that incorporates other EU level laws. A more detailed analysis of the existing legislations can be found in Pereira (2014). Additionally to the legislative and strategic actions mentioned the EU develops a range actions that support specific parts of the policy roadmaps.

2.1.3. Support level

Support actions for EE represent efforts to remove market barriers and support the delivery of planned energy savings. These actions are designed to boost information sharing, training activities and financial support enabling a broader reach of the necessary evolution on practices and technologies across the EU that contribute to improvements on EE. An example of this support actions is the EU ENERGY STAR Programme, for labelling energy efficient office equipment, based on an agreement between the USA government and the EU (European Commission, 2013a).

These initiatives consist on actions that beyond their direct contribution to EE development, represent measures to improve energy security and environmental quality, such as Carbon dioxide (CO_2) emissions decrease and renewable energy sources (RES) integration and deployment.

2.2. Governance analysis methodology

The framework designed by the EU in terms of EE actions defines a set of measurable indicators to understand the level of compliance form Member States level in regards to EU-wide actions. These, in combination with other governance dimensions, will be used to conduct an analysis on the current status of governance of EE in the EU. The methodology that the authors will follow was developed by Jollands & Ellis (2009b), to enable the assessment of the level of compliance in different governance areas and at different levels (i.e.: local, national, international). The authors of the methodology defined the foundations for governance, as the necessary resources required to establish a governance system, consisting of: (i) institutional structures, (ii) human and financial resources, (iii) human capacity and (iv) political support. These dimensions provide the direction on the analysis to be performed in the following section.

Jollands & Ellis (2009b), represent their methodology through a radial chart as presented in the following diagram (Figure 1).

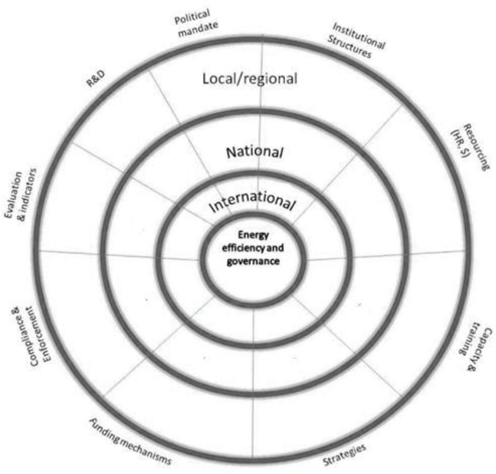


Figure 1: Schematic of governance dimensions.
Adapted from Jollands & Ellis (2009b)

The governance analysis to be presented for the EU Member States is based on an adaptation of the above presented methodology to be able to aggregate relevant indicators into an analysis that provides relevant outputs on the status of EE governance in the EU. The level of the analysis is international (i.e.: EU-wide), as indicators form the 28 Member States will be selected, collected and analysed. The following section presents the main analysis conducted for the selected dimensions within the defined framework. The analysis consists on an update and improvement of the governance analysis presented in Pereira (2014).

3. Data collection and analysis

The data gathered for the analysis is based on the World Energy Council Database for Energy Efficiency (World Energy Council, 2015), and publications released by the EC (European Commission, 2014d, 2015). The EE governance indicators collected enable the development of an analysis covering the four foundation dimensions of governance presented by Jollands & Ellis (2009b). The following table aggregates the indicators included within each dimension for analysis.

Table 2: EE governance indicators collected and analysed.

Dimension	Indicator
Institutional Structures	National energy agency
	Ministry department for EE
	Local energy agencies
Financial Capacity	Dedicated EE funds
Human Capacity	National energy agency
	Ministry department for EE
	Local energy agencies
Political Support	National EE programmes with targets
	Energy law with EE targets
	Dedicated EE law
	EPBD Transposition *(on 22/07/2014), due 09/07/2012*
	EPBD Cost-optimal calculation report *(on 22/07/2014), due 21/03/2013*
	EPBD Near Zero Energy Buildings report *(on 22/07/2014), due 04/03/2014*
	EED Targets presented *(on 16/07/2014), due 30/04/2013*
	EED Building renovation strategy *(on 16/07/2014), due 30/04/2013*
	EED EE obligations/alternative programmes *(on 16/07/2014), due 05/12/2013*
	3rd Round of National EE Action Plans (NEEAPs) delivered *(on 16/07/2014), due 30/04/2013*
	EED Transposition *(on 16/07/2014), due 05/06/2014*

The data analysis process presented in this paper is two staged. Firstly, the information obtained is categorized considering the level of compliance with the specific governance indicator being analysed. The analysis of the indicators within the institutional structures, financial capacity and political support dimensions for each Member State are categorized according to the following criteria: (*i*) Complies with the indicator; (*ii*) Does not comply with the indicator; (*iii*) No information available on the indicator; and (*iv*) Ongoing effort to comply with the indicator.

For the indicators within the Human Capacity/People dimension the categorisation is as follows: (*i*) Human resources allocated; (*ii*) No human resources allocated; (*iii*) No information available on the indicator.

The second stage of the governance analysis consists on the development of a global ranking of the 28 Member States, which complements the EU-wide analysis presented through the application of the Jollands & Ellis (2009b) framework on the first stage. The ranking method followed is based on the allocation of points, which are assigned according to the level of compliance attributed on the first stage of the analysis. The points given to obtain the global EE governance ranking of the 28 Member States are presented in the following table (Table 3).

Table 3: Governance ranking point assignment method.

Ranking Criteria	Points assigned	Ranking Criteria	Points assigned
Institutional, Legislative and Financial Dimensions		***Human capacity Dimension***	
Complies with the indicator;	1	Human resources allocated	1
Does not comply with the indicator;	-1	No human resources allocated	-1
No information available on the indicator;	0	No information available on the indicator	0
Ongoing effort to comply with the indicator.	0,5		

The main findings obtained from the aggregation of the indicator set collected are now presented, providing an overview on the aggregate performance of EU-28 Member State[4] for the analysed governance dimensions[5].

3.1. Institutional structures

In terms of institutional structures the Energy Agencies at National and Local Level are well implemented throughout the EU. In contrast

[4] The results presented include acronyms for each Member State as follows: Austria: AT; Belgium: BE; Bulgaria: BG; Croatia: HR; Cyprus: CY; Czech Republic: CZ; Denmark: DK; Estonia: ET; Finland: FI; France: FR; Germany: DE; Greece: EL; Hungary: HU; Ireland: IE; Italy: IT; Latvia: LV; Lithuania: LT; Luxembourg: LU; Malta: MT; Netherlands: NL; Poland: PL; Portugal: PT; Romania: RO: Slovakia: SK; Slovenia: SI; Spain: ES; Sweden: SE; United Kingdom: UK.

When presenting results "n.a." stands for: no information available.

[5] A detailed table in Annex A, presents the Member States included in each group of the analyses presented for the four dimensions analysed through this research.

with this only 53% of the Member States have a Ministry Department focused on EE issues.

3.2. Financial capacity

The financial capacity dimension was analysed by searching for the existence of dedicated EE funding at the Member State level as a positive point to promote and support EE improvements. The financial capacity analysis presents that no specific funding for EE exists in Belgium (BE), Estonia (ET), Greece (EL), Finland (FI), France (FR), Ireland (IE), Luxembourg (LU), Latvia (LV), Malta (MT) and Sweden (SE).

3.3. Human capacity/People

The dimension associated with human capacity and people is analysed through the data available on the allocation of human resources in the different types of organisations analysed in the institutional structures dimension.

Human resources allocated to the organizations engaged on promoting EE at national or regional level, through energy agencies, or at governmental level through ministry departments for EE vary widely. Therefore, in order to avoid a misleading analysis the human capacity compliance indicator was organized by categories, as described previously. Through this approach the authors considered that Member States were complying with the governance indicator when any number human resources where allocated to the organisation under analysis.

From the Member states allocation of human resources it is possible to observe a greater work force on EE related institutions by Central Europe and Nordic Countries. For National Energy Agencies, the Netherlands (NL) ranks highest with a total workforce of 1250 people. The case of the Ministry Departments for EE is led by Denmark (DK) with a workforce of 40 people. Sweden (SE) is the Member State with more staff allocated to Local Energy Agencies with a workforce of 303 people.

3.4. Political support

Regarding political support, the analysis is based on the level of compliance with the latest EU-wide EE related directives, the EED (European Parliament, 2012b) and the Energy Performance in Buildings Directive (EPBD) (European Parliament, 2010).

From the outputs obtained the most critical is the current status of implementation of the EED by Member States, as available data shows that only Cyprus (CY), Denmark (DK), Italy (IT), Malta (MT), and Sweden (SE) have the EE legislative instrument transposed into national legislation, the deadline was due on June, 2014 as outlined in Table 2. However, despite the existing failure to transpose the EED on time all Member States have presented their EE targets to meet by 2020, as one of the obligations stipulated, as well as have all presented the EE obligations schemes or alternative initiatives to foster the improvement of EE levels.

The outputs herein presented are aggregated and discussed in the following section providing the main results of the research process. The research framework for EE governance analysis obtainable in Figure 1 is presented in an adapted version according to the governance dimensions and indicators outlined in Table 2. In addition the Member States EE governance ranking is disclosed, providing a basis to better understand the regions in the EU where more effort is necessary to ensure the necessary contribution to achieve the set EE targets for 2020, which will dictate the potential to achieve the more ambitious targets towards 2030, of 27% on EE.

4. Results and Discussion

Understanding the status of EE governance in the EU is an underpinning issue considering the existing climate and energy policies, which are based on ambitious energy consumption reduction targets. The two stage analysis developed is beneficial in terms of the insights it provides at different levels, informing policy and decision makers.

The first stage of the analysis, based on the application of the governance framework of Jollands & Ellis (2009b) adapted to the available indicators, provides EU-wide information on the current status of the analysed dimensions. This is relevant to provide policy makers with the areas where more action is needed, or to enable ex-post analyses of the implemented laws, programmes and support initiatives to be promoted.

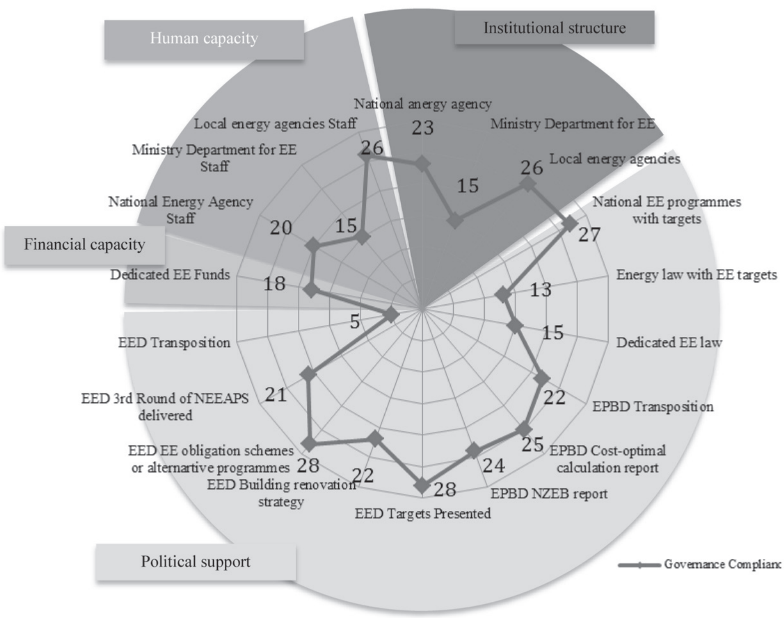
The second stage of the research based on the application of a method to rank Member States according to their EE governance performance (see Table 3 for ranking method) that provides complementary information. From the results obtained policy markers' can access a strengthened perspective on the status of EE governance in each Member State, this enables the design of EE actions that target particular challenges. In addition the ranking enables Member States to have a clear view of their positioning in the EU-wide framework. Beyond the information it provides to a Member State regarding their own situation, it is also a tool to assist the identification of best case practices, as those with a lower performance can analyse the governance practices of better ranked countries in search of successful actions and programmes. The result of these processes is presented below.

4.1. Governance analysis framework results

The governance analysis framework provides a global overview on all the indicators analysed, presented through the following graph (Graph 1).

The results (Graph 1) are presented in terms of number of Member States of the EU-28 that are complying with the governance indicator in each dimension (e.g.: In the financial capacity dimension, for the dedicated EE funds 18 Member States have a dedicated fund, therefore comply with the indicator).

In terms of institutional structure and human capacity the EU presents a significant potential for improvements. Despite the existing structures the results show that only 82% of the Member States have National Energy Agencies, 53% have Ministry Departments for EE. Local Energy Agencies present the best outcome with 92% Member States complying with the indicator. This analysis of the existing structures reflects a similar situation in the human capacity dimension as the data for human resources allocated to EE organizations is for the same institutions analysed in the institutional structures dimension.



Graph 1: EU governance analysis.

In terms of financial capacity 64% of the EU has dedicated EE financing schemes, which are either funded through the national public budget, through specific taxes for climate and energy related action programmes.

The political support dimension presents positive results associated with the EPBD transposition, with 79% of the Member States with the directive transposed into their national law. However, the situation associated with the EED is significantly weaker, with only 18% of the Member

States with the directive transposed. Considering that the transposition was due on June, 2014 and the data used for the analysis is from July, 2014 date by which the majority of the Member States failed to comply with the target date. The EED was designed to increase EU-wide efforts towards a more energy efficient EU, the failure to adopt the measures proposed will have a direct impact on the ability to deliver the planned goals, as well as on ensuring a wide-ranging collaborative effort for a more efficient use of energy resources.

From the indicator set collected the area of political support dimension provides a wider vision given a greater availability of information. Further indicators should be analysed in future work in terms of financial capacity, possibly looking for Research and Development related funding at each Member States developed through national funding programmes. For the human capacity dimension the analysis should evolve to include governance indicators related to available skills at each Member State in the area of EE. The institutional structure dimension herein presented was based on support structures; the governance analysis can be complemented by including indicators from the private sector (e.g.: indicators on the number of companies working in the field of EE).

To ensure that governance can be evaluated over time, allowing the identification of improvements and declines on governance performance, a set of common EU-wide indicators should be developed and collected over time providing periodic insights from the Member State level and from the EU overall situation.

The EC (European Commission, 2014b) is aware of the need of a stronger governance framework and as part of the 2030 energy and climate agenda has presented the ambition to improve the EE governance process existing in the 2020 energy and climate framework, towards a more consolidated and supportive system. Through this the EC envisions to: (*i*) deliver the EU energy and climate agenda goals (*ii*) improve Member States' approaches to EU policies implementation; (*iii*) stimulate the competition and integration of energy markets and (*iv*) reduce the uncertainties related to the set targets to drive more investment. The proposed governance framework improvement is supported by a three-step process (European Commission, 2014b), as presented in Figure 2.

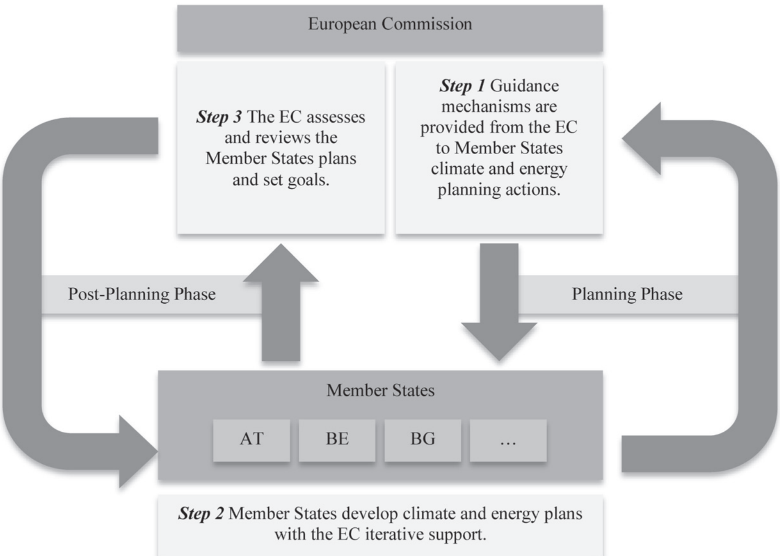


Figure 2: Governance improvement methodology, adapted from (European Commission, 2014b).

The main feature of the proposed improvement methodology (Figure 2) is the overarching nature of the process covering the whole planning phase and providing further support in the post-planning period, ensuring the adequacy and attainability of the developed plans.

4.2. Member States governance performance ranking results

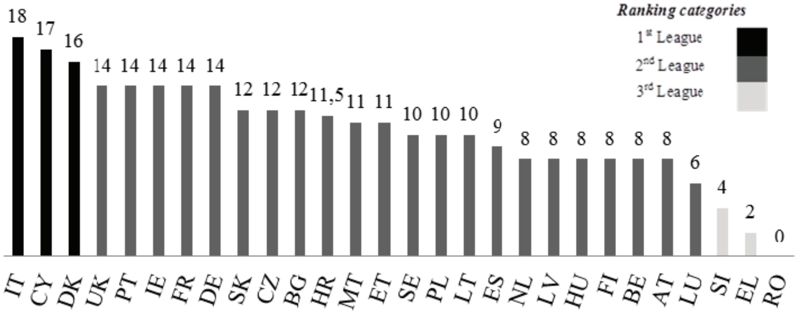
The development of a governance performance ranking was based on the goal to map the best performing Member States, according to the selected governance analysis criteria, whilst providing also information on Member States where stronger efforts have to be implemented to ensure a good level of EE governance.

To allow some comparison three different levels where created to aggregate the Member States into larger groups. From the raking method presented (see Table 3) and considering the governance criteria used for analysis any Member State can score a maximum of 18 points (i.e.: when complying with all the indicators), and a minimum of -18 points (i.e.: when failing to comply with all the indicators. The intervals used to group the Member States and a brief description of the categories is presented on the table below (Table 4). The definition of the categories proposed hereafter is based on the aim to classify Member States, according to three categories: 1st League, 2nd League and 3rd League in terms of EE governance performance. The performance point intervals established for each category were defined to enable the identification of a small group of top performing Member States (i.e.: 1st League), which corresponds to high levels of compliance with the EE governance indicators assessed in this analysis. The second category (i.e.: 2nd League) enables the identification of Member States with an average performance, which could benefit from best case practices from Member States included in the 1st League Member States group, whilst the third category (i.e.: 3rd League) aggregates all the Member States with low performance on EE governance. The authors propose point assignment interval thresholds that expand through the defined categories from a very limited 1st League category, from 15 to 18 points; to a more extended 2nd League, from 6 to 14 points; and a broader 3rd League, from - 18 to 5 points. Defining the point intervals through this method enables the creation of meaningful insights regarding the identification of governance practices among EU Member States, serving as guidance for policy makers. From the authors' point of view, this approach adds more value to the presented analysis than the definition of performance point intervals following the definition of equal ranges across the possible scores to be obtained by Member States.

Table 4: Ranking Categories.

Category	Performance Point Interval	Brief description
1st League	>14 points	Top performing Member States, contributing to the EE development and serving as potential sources of best case practices.
2nd League	<= 14 and >5 points	Member States with a medium-low governance performance.
3rd League	<= 5 points	Member States with a low performance, which can benefit from best case practices observed in 1st League Member States.

The point assignment process enabled the development of the overall Member States ranking. It is important to note that the point assignment process and final results are based on the available information at the time of development of this research. Some Member States performance may be depicted better or worse than it is in reality, given the impossibility to access the necessary data to point a particular indicator. Despite the effort of the authors that based the data collection process in reliable databases, some information was not available. This barrier encountered in the data collection analysis serves also as a suggestion for Member States and the EC to communicate this data and other of similar nature in a more organised manner, reducing the information access barrier. The following graph (Graph 2) presents the outputs of the performance analysis.



Graph 2: EU Member States performance analysis.

Considering the categories defined Italy (IT) is at the forefront of governance performance complying with all the analysed dimensions ranking in number one on the 1st League group followed by Cyprus (CY) and Denmark (DK). The group representing the 2nd League has as best performers Germany (DE), France (FR), Ireland (IE), Portugal (PT) and the United Kingdom (UK). The 3rd league is composed by Slovenia (SI), Greece (EL) and Romania (RO), all of which fail to comply with a significant number of the governance performance indicators selected. The results obtained indicate significant discrepancies throughout the EU in terms of governance. This analysis further validates the need to implement the governance improvement framework presented in Figure 2, to align the Member States governance levels to a highly-collaborative standard based on timely implementation of policies and effective national plans for EE.

5. Conclusion

The research conducted on the current status of EE governance is crucial considering the proximity to the 2020 deadline to deliver a 20% energy consumption reduction in the EU. This analysis gains greater relevance considering the already presented goals towards 2030 of 27% EE increase, with an ambition to raise the goal to 30%. Considering these set targets, the need of an aligned effort to govern and foster EE is well justified.

In line with this need for an aligned and collaborative environment towards reducing energy consumption, a first analysis was conducted applying a governance analysis model presented by Jollands & Ellis (2009b). From the dimensions analysed, the political support is the main source of concerns. Member States are failing to comply with the transposition of EU Directives, designed to drive impetus towards higher levels of EE. The EED as the EU framework legislation that embraces the energy system from generation to end-use on pursuing energy savings has only been implemented by 18 % of the Member States, in contrast with 79% for the EPBD.

The financial resources dimension has also potential for improvement considering that from this research only 64% of the Member States have dedicated EE funding instruments. Shifting towards an energy efficient EU requires political, social and financial stimulus in order to enable the market for EE to expand and reach a good level of development. Regarding the institutional structures and human capacity the analysis performed was based on public entities and agencies, these indicators present the existence of a good network of National and Local energy Agencies, and fewer Ministry Departments for EE.

These EU-wide governance status results provide the necessary information to structure strategic plans that tackle the identified gaps through different policies, potentially through support actions.

The complementary Member States governance performance ranking (see Graph 2) provided country specific information, which demonstrated a misalignment between the 28 Member States. For instance, through the set of selected indicators and available data, Italy (IT) was able to comply with all governance indicators whilst Romania (RO) failed in most of the dimensions, ranking last on the list. This information has to be taken into consideration when devising national, regional and EU-wide strategies, funding programmes and EE targets, to ensure that these are realistic and match the Member States ability to contribute. The combination of the two analyses conducted should be considered when implementing the governance improvement methodology proposed by the EC (see Figure 2). The analysis was based on publicly available information on the governance indicators selected. The results and Member State ranking are based on this data. To the best knowledge of the authors no database is yet available for the analysis of EE governance, the aggregation of this information on a single source would support better analysis and more robust recommendations.

Future work can include analysis of the interrelations, between the various governance dimensions (e.g.: the impact of the institutional structure on the timely transposition of directives). The analysis of these interlinks will contribute further to prioritising and informing decision makers on where to act, in order to trigger greater levels of EE. Furthermore the analysis herein disclosed shows the EU-wide situation and Member States performance, complementary works can focus on Member States individually to understand the regional priorities to foster EE, for this application the indicators must be adapted accordingly to yield relevant results.

Acknowledgments: The first author would like to acknowledge *Fundação para a Ciência e Tecnologia* (FCT) for supporting this work through the Doctoral Grant PD/BD/105841/2014, awarded on the framework of the MIT Portugal Program funded through the POPH/FSE. The authors would like to acknowledge that this work has been partially supported by the *Fundação para a Ciência e a Tecnologia* (FCT) under project grant UID/MULTI/00308/2013, and the Energy for Sustainability Initiative of the University of Coimbra.

References

EUROPEAN COMMISSION, Europe 2020 . A strategy for smart, sustainable and inclusive growth. - COM(2010) 2020 final. Brussels, 2010.

EUROPEAN COMMISSION, Background on energy in Europe. Brussels, 2011a.

EUROPEAN COMMISSION, Energy 2020 - A strategy for competitive, sustainable and secure energy. Brussels. 2011b.

EUROPEAN COMMISSION, Energy Roadmap 2050 - COM(2011) 885/2. Brussels, 2011c.

EUROPEAN COMMISSION, Press Release - 22 June 2011 - Strong impetus to energy savings and energy efficiency. Brussels, 2011d.

EUROPEAN COMMISSION, The Commission ' s new Energy Efficiency Directive - MEMO/11/440 - 22 June 2011 (pp. 1–6). Brussels, 2011e.

EUROPEAN COMMISSION, European Comission Energy Star Programme. Retrieved November 20, 2013 from http://www.eu-energystar.org/en/en_016.shtml, 2013a.

EUROPEAN COMMISSION, GREEN PAPER - A 2030 framework for climate and energy policies - COM(2013) 169 final. Brussels, 2013b.

EUROPEAN PARLIAMENT, Report on the Energy roadmap 2050, a future with energy (2012/2103(INI)) Committee on Industry, Research and Energy. Brussels, 2013.

EUROPEAN PARLIAMENT,. Directive 2006/32/EC of the European Parliament and the Council of the European Union of 5 April 2006 (2006). European Union: Official Journal of the European Union.

EUROPEAN PARLIAMENT, Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast), Pub. L. No. OJ L 153 (2010). Official Journal of the European Union.

EUROPEAN PARLIAMENT, Delivering on the Europe 2020 Strategy - Handbook for Local and Regional Authorities. Brussels, 2012a.

EUROPEAN PARLIAMENT, Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC, Pub. L. No. Volume

55 14 November 2012 (2012). European Union: Official Journal of the European Union, 2012b.

EUROPEAN COMMISSION. 2030 climate and energy goals for a competitive , secure and low-carbon EU economy, IP/14/54. Brussels, 2014a.

EUROPEAN COMMISSION, A policy framework for climate and energy in the period from 2020 to 2030, COM(2014) 15 Final. Brussels, 2014b.

EUROPEAN COMMISSION, Climate and energy priorities for Europe: the way forward. Brussels, 2014c.

EUROPEAN COMMISSION, Energy Efficiency and its contribution to energy security and the 2030 Framework for climate and energy policy, Annexes 1 of 3, COM(2014) 520 final. Brussels, 2014d.

EUROPEAN COMMISSION, Energy Efficiency Directive Target Reporting. Retrieved January 15, 2015 from http://ec.europa.eu/energy/efficiency/eed/eed_en.htm

FABER, J., SCHROTEN, A., BIES, M., SEVENSTER, M., MARKOSWKA, A., SMIT, M., RIET, J. (2012). Behavioural Climate Change Mitigation Options and Their Appropriate Inclusion in Quantitative Longer Term Policy Scenarios. Delft, April 2012.IEA, Energy Efficiency Governance - Handbook - Second Edition. Paris, 2010. Retrieved December 12, 2014 from http://www.iea.org/publications/freepublications/publication/name,3931,en.html

JOLLANDS, N., & ELLIS, M., Energy efficiency governance – an emerging priority - ECEEE Summer Studies 2009, 91–100, 2009a. Retrieved December 11, 2014 from http://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2009/Panel_1/1.086/paper

JOLLANDS, N., & ELLIS, M., Energy efficiency governance – an emerging priority - ECEEE Summer Studies 2009. In N. Jollands & M. Ellis (Eds.), . European Council for an Energy Efficient Economy, 2009b. Retrieved December 10, 2014 from http://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2009/Panel_1/1.086/presentation

PEREIRA, G. I. (2014). Connecting energy efficiency progress and job creation potential, Master Degree Dissertation, University of Coimbra, Coimbra, Portugal.

WORLD ENERGY COUNCIL. (2015). Energy Efficiency Policies and Measures Database. Retrieved January 28, 2015, from http://www.wec-indicators.enerdata.eu/world.php

ANNEX A

	Financial Capacity	Human Capacity/Poeple		
Governance analysis framework	Dedicated EE Funds	National Energy Agency Staff	Ministry Department for EE Staff	Local energy agencies Staff
Country	Status 1: Yes; 2: No; 3: n.a.	Staff count	Staff count	Staff count
AT	1	74	n.a.	12
BE	2	n.a.	6	19
BG	1	70	n.a.	6
CY	1	n.a.	40	2
CZ	1	n.a.	10	44
DE	1	100	10	59
DK	1	41	40	12
EE	2	n.a.	3	1
EL	2	150	n.a.	10
ES	1	141	0	42
FI	2	33	5	7
FR	2	1006	15	57
HR	1	100	n.a.	6
HU	1	60	n.a.	151
IE	2	43	6	16
IT	1	150	10	41
LT	1	29	n.a.	1
LU	2	8	n.a.	1
LV	2	n.a.	n.a.	3
MT	2	49	4	n.a.
NL	1	1250	0	4
PL	1	26	4	10
PT	1	40	10	23
RO	1	n.a.	0	n.a.
SE	2	35	0	303
SI	1	n.a.	6	6
SK	1	118	0	5
UK	1	n.a.	1600	34

	Institutional strucutres			Legislative compliance										
Governance analysis framework	National anergy agency	Ministry Department for EE	Local energy agencies	National EE programmes with targets	Energy law with EE targets	Dedicated EE law	EPBD Transposition	EPBD Cost-optimal calculation report	EPBD NZEB report	EED Targets Presented	EED Building renovation strategy	EED EE obligation schemes or alternartive programmes	EED 3rd Round of NEEAPS delivered	EED Transposition
Country	Status 1: Yes; 2: No; 3: n.a.	Status 1: Yes; 2: No; 3: n.a.	Status 1: Yes; 2: No; 3: n.a.	Status 1: Yes; 2: No; 3: n.a.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.	Status 1: Yes; 2: No; 3: n.a.; 4: under dev.
AT	1	3	1	1	2	2	2	1	1	1	1	1	1	2
BE	3	1	1	1	3	3	2	1	1	1	1	1	1	2
BG	1	3	1	1	2	1	1	1	1	1	1	1	1	2
CY	1	1	1	1	1	1	1	1	1	1	1	1	1	1
CZ	3	1	1	1	2	1	1	1	1	1	1	1	1	2
DE	1	1	1	1	2	1	1	1	1	1	1	1	1	2
DK	1	1	1	1	2	1	1	1	1	1	1	1	1	1
EE	1	1	1	1	1	2	1	1	1	1	1	1	1	2
EL	1	3	1	1	3	3	1	2	2	1	2	1	2	2
ES	1	2	1	1	4	4	1	1	2	1	1	1	1	2
FI	1	1	1	1	2	2	2	1	1	1	1	1	1	2
FR	1	1	1	1	1	1	1	1	1	1	1	1	1	2
HR	1	3	1	1	1	1	1	4	1	1	1	1	2	2
HU	1	3	1	1	1	2	1	1	1	1	2	1	2	2
IE	1	1	1	1	1	1	1	1	1	1	1	1	1	2
IT	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LT	1	3	1	1	2	2	1	1	1	1	1	1	1	2
LU	1	3	1	1	3	3	1	1	1	1	2	1	2	2
LV	3	3	1	1	2	1	1	1	1	1	1	1	1	2
MT	1	1	3	2	3	1	1	1	1	1	1	1	1	1
NL	1	2	1	1	3	3	2	1	1	1	1	1	1	2
PL	1	1	1	1	1	1	2	1	1	1	2	1	2	2
PT	1	1	1	1	1	1	1	1	1	1	2	1	1	2
RO	3	2	3	1	2	1	1	2	2	1	1	1	2	2
SE	1	2	1	1	1	2	1	1	1	1	1	1	1	1
SI	3	1	1	1	1	2	2	1	2	1	2	1	2	2
SK	1	2	1	1	1	1	1	1	1	1	1	1	1	2
UK	1	1	1	1	1	3	1	1	1	1	1	1	1	2

STRATEGIES FOR DEVELOPING BATTERIES FOR ELECTRIC VEHICLES: A REAL OPTIONS MODEL

Joana Fialho[1], Pedro Godinho[2] and João Paulo Costa[3]

Abstract

In this article we consider the task of developing batteries for electric vehicles, and we use a tool that incorporates a real option model and Monte Carlo simulation in order to define the best strategy for managing it. We assume that different levels of resources can be used to undertake the task, leading to different average advancement speeds and different costs, and that the level of resources being used may be changed according to the way the task is developing. We present a procedure that aims at identifying the strategies that maximize the net present value of the task. The approach is used in several different scenarios, in order to define some general rules that can help managers have an idea of the resources that may be available in the future, without endangering the maximization of the net present value of the crucial tasks.

Keywords: Real options, battery development, electric vehicles, Monte Carlo simulation

[1] CI&DETS – Instituto Politécnico de Viseu, Escola Superior de Tecnologia de Viseu and INESC-Coimbra (e-mail: jfialho@estv.ipv.pt)

[2] University of Coimbra, Faculty of Economics (e-mail: pgodinho@fe.uc.pt)

[3] University of Coimbra, Faculty of Economics and INESC-Coimbra (e-mail: jpaulo@fe.uc.pt)

http://dx.doi.org/10.14195/978-989-26-1039-9_5

1. Introduction

Public opinion has been increasingly concerned with the harmful effects of carbon dioxide gas on the environment. Such concerns extend to several different industries and technologies that resort to fossil fuels, like the automotive industry. So, automobile manufacturers have been trying to develop vehicles based on cleaner energies, which have a lesser impact on the environment. Electric Vehicles (EVs) are such an alternative to fossil fuel powered cars.

To improve the autonomy of EVs and increase the demand for such vehicles, the development of efficient batteries is crucial. The batteries are so important to the EVs that the evolution of these vehicles seems to have been driven by the developments achieved in the batteries (Magalhães, 2013).

In this work, we propose to use EV battery development as an application case for a financial valuation tool based on real options. We assume that the development of batteries is a uniform task, in the sense that it can be split into portions with identical characteristics, and that it is integrated into a major project consisting of developing a new EV. Such a task is subject to uncertainty, due to technical difficulties, market uncertainty and competitor actions, among others. Additionally, it is important to adapt the plan of action whenever unexpected events occur.

We assume that several different modes may be used to undertake this task, corresponding to the usage of different resources (e.g., different development teams), and leading to different costs and execution speeds. Managers have to decide which resources should be used to start the battery development, and in which circumstances it would be better to change the allocated resources. Each different combination of resources that can be used in the development process will be hereafter referred to as a "level of resources", and it is characterized by a cost per unit of time and a stochastic speed of task advancement.

The tool we use in this article is based on real options theory and allows the definition of a resource allocation strategy. We define a strategy as a set of rules that determine which level of resources shall be chosen, at each moment, and we aim to find the strategy that maximizes the value of the task of battery development. We assume that the advancement of the task is stochastic and that the project manager may change the level of resources allocated to the task while it is in progress. In order to identify the best strategy, the tool uses a method that allows the incorporation of operational flexibility and uncertainty in simulation-based valuation of projects or tasks: Least Squares Monte Carlo (originally proposed by Longstaff and Schwartz, 2001).

We present an evaluation exercise applied to a battery development task. We define several different scenarios, and extract some general conclusions by comparing the results achieved in these scenarios. We assume that a firm can either allocate a base team (Level 1) to this task or an enlarged team (Level 2) to undertake the task. These two alternatives lead to different costs and different development speeds. We build alternative scenarios to illustrate the type of results that can be obtained, and the circumstances in which each level of resources will be used. In particular, we are able to identify several cases in which it seems to be optimal to keep using the same level of resources until the end of the task, and others in which it seems best to adapt the level of resources to the way the task is developing. These rules may help managers forecasting the resources they will need in the future.

This paper is structured as follows: section 2 reviews some literature about evaluating and managing projects in the presence of uncertainty. Section 3 presents a brief review of the characteristics of batteries used in EVs. Section 4 presents a summary of the model and the evaluation procedure underlying the evaluation tool. Section 5 shows an application of the evaluation tool and presents some results. Section 6 concludes.

2. Evaluating projects in the presence of uncertainty

Most projects are characterized by the possibility of adapting the plan of action, according to the way the project is developing or to the arrival of new information, that is, they are characterized by operational flexibility. Adapting the actions in response to altered future conditions expands an investiment opportunity's value, improving the upside potential of the project and at the same time limiting the downside losses (Trigeorgis, 1993).

Traditional evaluation methods, such as the ones based on discounted cash flows, are not adequate to evaluate projects that have operational flexibility, because they assume a pre-determined and fixed plan (Yeo and Qiu, 2003). Furthermore, the realized cash flows will probably be different from what was expected, due to the uncertainty, the competitive interactions and the change that characterizes the actual marketplace (Trigeorgis, 1993). Real options methods take into account both operational flexibility and uncertainty, aiding the identification of the best decisions. A nice introduction to real options theory can be found in Dixit and Pindyck (1993).

Real options theory allows us to determine the best sequence of decisions to make in an uncertain environment, and provides the proper way to evaluate a project when such flexibility is present. The decisions are made according to the opportunities that appear along the project lifetime, which means that the optimal decision-path is chosen step by step, switching paths as events take place and opportunities appear (Cortazar *et al.*, 2008).

The models to evaluate real options can present some difficulties, like determining the model inputs, defining relations among the input parameters or being able to solve the option pricing algorithm (Santiago and Bifano, 2005). In spite of these difficulties, following a real options perspective may have a significantly positive impact in the financial performance of a firm.

Determining real option value depends on the type of option and on the type model that is considered but, in general, determining its exact value is a complex process. Often, numerical techniques, analytical approximations or simulation are used to make the valuation process tractable. Simulation is an attractive tool to evaluate real options, because it allows us to consider the state variables as stochastic processes and, nowadays, the simulation techniques are simple, transparent and flexible (Longstaff and Schwartz, 2001).

3. Characteristics of batteries for electric vehicles

The development of batteries has been slower than other areas of electronics (Zeng *et al.*, 2014). To increase the efficiency of an EV, it is necessary to increase the number of batteries and, consequently, to increase its weight and required power.

There are many kinds of batteries, like lead-acid, nickel-cadmium (Ni-Cd), nickel-metal hydride (NiMH) and lithium-ion batteries (LIBs) (Zeng *et al.*, 2014). The last ones, LIBs, have presented some advantages, like higher energy density, higher cell voltage, less memory effect, low self--discharge, and very good life cycle, and are environmentally sound as well as simple to charge and maintain. Thus, these batteries are widely used in laptops, smartphones and other types of electronic devices. However, the usage of these kinds of batteries in electric vehicles still presents some difficulties, such as the need to increase their energy and power densities, improve their safety, and lower the cost (Su *et al.*, 2014). Moreover, if there is no proper disposal for spent LIBs, the human health and environment can be endangered (Zeng *et al.*, 2014). Several different recycling technologies have been developed for spent LIBs, but most of them are still in pilot state.

Battery development tasks are therefore crucial processes for the makers of electric vehicles. In this paper, we consider a big industrial company striving to develop a new battery, but also having other tasks to which its Research and Development (R&D) staff may be allocated. The analysis of the better battery development strategies will result into a set of rules that may provide some general guidelines for the resource allocation policies of the company.

4. A model and an evaluation procedure for battery development

We assume that the development of the batteries is a homogeneous task that can be divided into several portions, termed "work units", with identical characteristics, which must be undertaken sequentially. We assume that the firm can allocate different levels of resources to the task. Different levels of resources may be related, for example, to different sizes of the team undertaking the task. Different levels of resources will lead to different costs and different development speeds. The benefits and costs resulting from completing the task will be related to the time it takes to undertake it.

We will now introduce the broad lines of the model. The detailed model can be found in Fialho (2013).

4.1 Time to complete one work unit

Randomness is introduced in the task advancement process by assuming that the time it takes to complete each work unit follows an exponential distribution (other authors, like Folta and Miller, 2002, use similar assumptions). Thus, the time to complete one work unit using a given level of resources, k, is a random variable following exponential distribution, $t^{(k)}$. The average of this random variable is defined as the reciprocal of the average number of units completed in a unit of time, when using the level of resources k.

4.2 Cost of using a level of resources

Costs are related to the usage of different levels of resources. The costs per unit of time are deterministic, and assumed to grow at a constant rate ρ. Considering a specific level of resources k, and defining $C_x^{(k)}$ as the instantaneous cost at time x, we have $dC_x^{(k)} = \rho C_x^{(k)} dx$. If level k is used at time x, we can say that the cost of using it in the next instant dx is $C_x^{(k)} dx$. Each time the level of resources changes, we assume that there is a "setup cost" incurred by the company.

4.3 Task worth

The concepts of task worth and instantaneous task worth are used to summarize the financial impacts derived from concluding the task. The task worth is the present value of the cash flows resulting from the task, at the moment of its completion, and the instantaneous task worth is the potential value of the task worth, if it were completed at a given moment. We model the evolution of the instantaneous task worth and, when the task is concluded, we have the effective task worth. Instantaneous task worth grows at a fixed rate, and it can have stochastic jumps due to "shocks" in the value of potential cash flows.

The instantaneous change in instantaneous task worth (R), in an instant of length dx is denoted as dR, and it can be modelled as:

$$dR = \alpha R dx + R dq$$

Parameter α represents the growth rate of the instantaneous worth and dq represents a jump process. The occurrence of these jumps in instantaneous task worth is assumed to follow a Poisson process with parameter p. So, the value of dq can be defined as:

$$dq = \begin{cases} 0, & \text{with probability } 1 - p \cdot dx \\ u, & \text{with probability } p \cdot dx \end{cases}$$

Any probability distribution can be defined for u and, in fact, other models for the task worth could also be used since the evaluation procedure is based on simulation. For example, a stochastic term based on a Brownian motion could easily be included, as is common in many real options models. However, in this kind of projects, expectations about future revenues and costs do not usually change in a continuous fashion, since we are not dealing with goods traded in markets in which prices change continuously (like capital markets or commodity markets). Instead, stochastic changes will usually be due to discrete events, like the entry of a competitor, shifts in the expected demand or technological changes, which may cause a significant instantaneous change in the expected cash flows.

If task completion takes longer, other risks and negative effects may affect the cash flows obtained by the company. For example, if a competitor is able to introduce, earlier, a similar product in the market, like new batteries, or even new electric cars, there may be a significant negative impact in the company cash flows. These effects are incorporated in the model through a penalty. This penalty is expressed by a function *g(x)*, where *x* denotes the time. This function is positive, decreasing, and it takes values from the interval [0,1]. If the task finishes at time *x*, the value of the instantaneous task worth, as defined before, is multiplied by *g(x)*, in order to get the task worth. Notice that the absence of such penalty can be taken into account by defining *g(x)=1*, for all *x*.

4.4 Net present value

The net present value of the task is calculated by discounting all the costs regarding the use of the resource levels, as well as the final task worth, to the moment the task starts, using an appropriate discount rate. Given the stochastic nature of the model, we have to use the expected value for the net present value. Concretely, and since we use simulation, a set of paths is simulated, the best strategy and the corresponding net present value are determined for each one, and finally and average of these net present values is calculated.

4.5 Procedure for determining the best strategy

The procedure we use aims to define the optimal strategy to execute a task, in particular the development of batteries for an electric car. The objective of the procedure is to choose which level of resources should be used at each moment, taking into account the way the task is developing, in order to maximize the task value. In order to find the best levels of resources, we use a method similar to the Least Squares Monte Carlo (LSMC, Longstaff and Schwartz, 2001).

LSMC can be applied to estimate the value of real options. It constructs regression functions to explain the payoffs for the continuation of an option through the values of the state variables. A set of simulated paths of the state variables is generated. With the simulated paths, the optimal decisions are defined for the last period. From these decisions, a conditional function is built, for the penultimate period, which defines the expected value taking into account the optimal decisions of the last period and the values of the state variables. With this function, optimal decisions are defined for the penultimate period. The process continues by backward induction until the first period is reached. The use of simulation allows integrating different state variables in an easy way.

The procedure we follow is based on LSMC, but some adaptations were necessary. We start by building many paths, with different strategies. The strategies used to build the paths include executing all work units with the same level of resources or using different levels of resources to finish the task. For each strategy, and for each path, we simulate the values of the time to execute the task, using the model presented in the previous section. With the time elapsed and the levels of resources utilized, we can determine the costs, and through the model for the instantaneous task worth, we also simulate the values for the task worth. Finally, we determine the net present value of the task, for each path and for each work unit.

In this procedure, we build, by backward induction and for all work units, regression functions for the values previously calculated for the paths. These functions explain the net present value of the task as a function of different state variables: the elapsed time, the instantaneous task worth, the number of work units already finished and the level of resources being used. For the first work unit, by definition, there is no path-dependent conditional information. In order to determine the best resource level to use in the first work unit, a specific level of resources (any level) is initially assumed. With the regression functions estimated for the following work units, the best strategy and the corresponding net present value at the beginning of the task are defined for all paths, and the average net present value obtained by starting to use that resource level is calculated. The process is repeated, assuming the task starts with each possible resource level, and the level leading to the largest net present value is selected as the best one. After this procedure, the regression functions allow defining rules which can guide management in deciding which strategy to use at each moment, under each set of conditions (as defined by the state variables). More details about this procedure can be found in Fialho (2013).

5. Application of the model to battery development

This evaluation exercise intends to illustrate the type of results that a real options model, based on Monte Carlo simulation, can provide when it is applied to a battery development task. Using the previously outlined model and procedure, we will analyse the patterns that can be found in the selection of the best levels of resources to use at each moment, considering the objective of maximizing the net present value of the benefits and costs related to the task. In the simulations, we consider that the task is divided into 20 work units, and we use 2100 paths. We also assume that the firm can either allocate a base team to this task (we denote this possibility as level 1 of resource usage) or an enlarged team (we denote this possibility as level 2 of resource usage). These two alternatives lead to different costs and different development speeds.

We define several different scenarios, and extract some general conclusions by comparing the results achieved in these scenarios. The different scenarios are built around a base case (or base scenario), which we will now describe. We consider a time unit of one fortnight (two weeks) and assume that, for each level of resources, the costs per unit of time are deterministic. In order to simplify the presentation of results, we consider a monetary unit (m.u.) of $5 000.

Level 1 of resource usage is assumed to have a cost that initially amounts to 10 m.u. per unit of time, and to allow the conclusion of an average of 1.5 work units per unit of time. Level 2 allows faster, but more expensive, development of the batteries: its cost initially amounts to 25 m.u. per unit of time, and it allows the conclusion of an average of 2.5 work units per unit of time. The setup cost is equal to the cost of using the new level of resources for one unit of time.

Costs are assumed to grow at a rate of 0.06% per fortnight (approximately the United States inflation rate in 2013). The discount rate, *r*, is assumed to be 8% per year (approximately the weighted average cost of capital of the auto & truck industry in the United States in the beginning of 2014, according to Damodaran, 2015). This leads to a discount rate of about 0.31% per fortnight.

We assume the instantaneous task worth to be 2 000 m.u. at the beginning of the task (that is, $10 000 000). The occurrence of stochastic jumps is modelled according to a Poisson process with a rate of 0.4, and each jump changes the instantaneous task work by a percentage given by a uniform distribution, with a minimum of 20% and a maximum of +20%. The task worth penalty (function *g(x)*, which penalizes the task worth as a function of the conclusion time) is defined in the following way: in the initial 8 units of time (16 weeks), the penalty is 0.625% per unit of time; for the subsequent 7 units of time (14 weeks), there is a penalty of 0.714% per unit of time; if the task takes more than 15 units of time (30 weeks) to be completed, we assume a fixed penalty of 10%.

We have chosen to build alternative scenarios to understand the type of results that will be obtained, and the circumstances in which each level of resources will be used. Scenarios were built by considering two alternative levels (beyond the base case value) for each of four parameters. These parameters are the costs of resource level 2 per unit of time (C2), the task advancement per time unit in level 2 (A2), the instantaneous task worth at the beginning of the task (R0) and the task worth penalty (Pen). For each of these parameters, three levels were considered: a "high level" (H), a "medium level" (M), which is the parameter value in the base scenario, and a "low level" (L). This led to a total of 81 scenarios.

For the low level (L) of these parameters, we use the following values: for a low C2, we use costs of 15 u.m. per unit of time; for a low A2, we use an average advancement rate of 1.8 work units per unit of time; for a low R0, we use 500 m.u.; and, for a low penalty (Pen), we use a function defining a penalty of 0.250% per time unit in the initial 8 units of time, 0.429% per time unit for the subsequent 7 units and a fixed penalty of 5% if the task takes more than 15 units of time. For the high level (H) of these parameters, we use the following values: for a high C2, we use costs of 40 m.u. per unit of time; for a high A2, we use an average advancement rate of 4 work units per unit of time; for a high R0, we use 10 000 m.u.; and, for a high penalty (Pen), we use a function defining a penalty of 1.25% per time unit in the initial 8 units of time, 2.86% per time unit for the subsequent 7 units and a fixed penalty of 30% if the task takes more than 15 units of time.

We started by analysing the base scenario, and we initially focused on resource usage as a function of the percentage of work that is completed. The results we obtained, after applying the evaluation procedure and obtaining the best strategy, show that level 2 should be used at the beginning of the task and, while the task is progressing, level 2 is often replaced by level 1 (Fig. 1).

The probability of choosing level 1, when 25%, 50%, 75% e 90% of the work is completed, is 54.57%, 58.86%, 59.29% and 59.57%, respectively, when the best strategy is used.

Next we focused on the level of resources that should be being used at each time. We concluded that the average time to finish the task is 10.9 units of time and, under the best strategy, the probability of finishing the task after 5, 10 and 15 units of time is 0.57%, 4.43% and 89.43%, respectively. Figure 2 summarizes the probability of using the different levels of resources as a function of elapsed time.

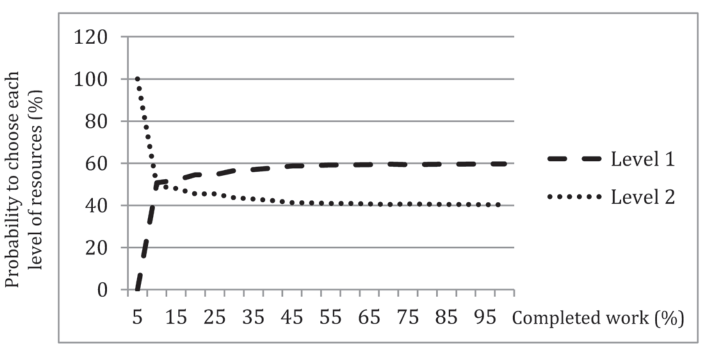


Figure 1: Probability of using each level of resources, for various levels of task completion, in the base scenario.

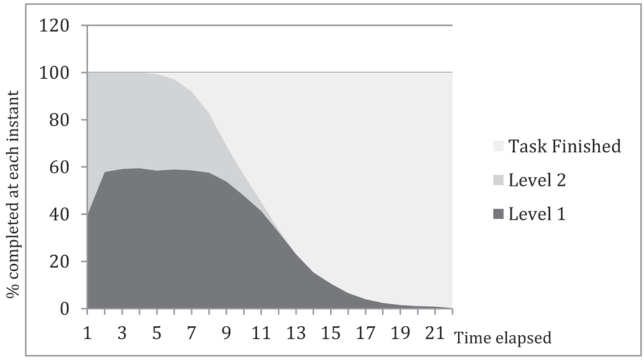


Figure 2: Probabilty of using each level of resources and of finishing the task, as a function of the elapsed time.

We then analysed what would happen in the alternative scenarios defined by considering the alternative parameter values (L and H), defined before, in order to derive some general rules that may provide managers an idea of what to expect concerning resource usage in the optimal strategy, which maximizes the expected net present value of the task.

At the outset, a larger usage of resource level 2 is expected to lead to a higher probability of completing the task sooner, since level 2 allows a faster average advancement rate of the task. We expect that higher costs of using resource level 2 and lower advancement when using this resource level will lead to a larger utilization of level 1, since they will make level 1 more attractive when compared to level 2. A lower initial task worth and a lower penalty for late completion are also expected to favour usage of resource level 1, since they will reduce the the financial benefit of concluding the task earlier.

In order to verify whether our initial expectations would hold, we analysed which level of resources was used in the beginning of the task and the probability of each level of resources being used when different portions of the task were completed. We also analysed what level of resources was being used, or if the task was completed, at different times.

We can point out some scenarios in which level 1 is used exclusively, or almost exclusively. These scenarios are shown in Table 1, in the appendix. Notice that, as expected, the probability of task completion is similar in all these scenarios, since the average task advancement with level 1 is not changed among different scenarios, and these scenarios use almost exclusively this level.

There is an interpretation for all the scenarios in Table 1: they correspond to the cases in which the additional benefit of using level 2 does not cover the additional costs. This is usually the case when the costs of level 2 (C2) are high and the average task advancement (A2) is low in this level - level 1 is almost exclusively used in these scenarios, except when both the initial instantaneous task worth (R0) and the penalty (Pen) are high. When C2 is high and A2 is medium, level 1 is almost exclusively used both when R0 is low and when R0 is medium and Pen is not high. When both C2 and A2 are high, level 1 is almost exclusively used only in the case in which both R0 and Pen are low.

When C2 is medium, level 1 is almost exclusively used when: both A2 and R0 are low; A2 is low and R0 is medium; A2 is medium and R0 is low; both A2 and Pen are low. Finally, when C2 is low, level 1 is almost exclusively used only if A2 is low and either R0 is low or Pen is low with a medium R0.

Summarizing these results, we can say that a low A2 (a low benefit in the task advancement speed) seems to be the main driver for using level 1 almost exclusively, occurring in 19 out of the 29 scenarios. The following most important factors seems to be low R0, which takes place in 17 scenarios, and high cost C2, which occurs in 15 scenarios. A lower penalty also has some influence in using level 1 almost exclusively, but such influence is smaller, with just 12 scenarios showing it. Notice that these results confirm the prior expectations that low A2, R0 and Pen, and high C2 favour the utilization of resource level 1.

Considering now the scenarios in which level 2 is used almost exclusively, we find 41 such scenarios, which are shown in Table 2, in the appendix. They correspond to the cases in which there is a large enough benefit of using level 2 to cover the additional costs of using this resource level.

Making a summarized analysis of the results, high average task advancement in level 2 seems to be the main driver for using level 2 almost exclusively, occurring in 22 out of the 41 scenarios. A low cost in level 2 and a high penalty occur in 20 scenarios, and are the following most important factors. A high initial instantaneous task worth also plays some role in the almost exclusive use of level 2, taking place in 17 scenarios. These results are also in accordance to the initial expectations.

There are 5 scenarios in which level 1 is used at the outset of the task, and level 2 has, subsequently, a significant probability of being selected. These scenarios are shown in Table 3 (in the appendix). In Table 4 (in the appendix) we show the 6 scenarios (including the base scenario) in which level 2 is used at the beginning of the task, and level 1 has, subsequently, a significant probability of being used. All these scenarios correspond to intermediate situations in which there is some balance in the costs and benefits of using level 2.

The results we obtained confirm the initial expectations, and also show the relative importance of the factors driving the choice of resource levels, and what can be expected in several different situations. In general, the benefit of level 2 in relation to level 1, in terms of advancement speed, seems to be the main driver for the choice of resource level. These results may help managers understanding what they can expect regarding optimal resource utilization, making it easier to formulate expectations concerning future resource availability, without hampering the maximization of the net present value of crucial tasks.

6. Conclusions and future research

In this article we consider an approach for defining the best strategy for managing a task, taking into account the maximization of the financial net present value, and apply it to an example of battery development for electric vehicles. This approach assumes that several different processes can be used to develop the task, characterized by different levels of resources, and leading to different average task advancement speed and different costs. The approach allows the definition of rules that allow the identification of the best level of resources to be used under each set of circumstances, using a procedure based in LSMC.

The approach is used in several different scenarios, in order to understand the type of management rules that would define the best levels of resources in a battery development project. In particular, we are able to identify several cases in which it seems to be optimal to keep using the same level of resources until the end of the task, and others in which it seems best to adapt the level of resources to the way the task is developing. These rules may help managers forecasting the resources that they will need in the future, without endangering the maximization of the net present value of the most important tasks.

In the future, it would be useful to apply this approach to case study data from a producer of batteries for electric vehicles. This would also allow us to assess the difficulties in calibrating the model and to apply the optimal decisions identified by the model.

It will also be important to incorporate other options that may be relevant for these kinds of tasks, like the abandonment option and the option to delay the beginning of the task. Finally, it is of great relevance to integrate a set of interdependent tasks with this approach, in order to define the best decisions for complete projects, like the project of developing a new electric vehicle.

Acknowledgments: This research was supported by Project EMSURE—Energy and Mobility for Sustainable Regions (CENTRO-07-0224-FEDER-002004)

References

Cortazar, G., Gravet, M. and Urzua, J. (2008). The valuation of multidimensional American real options using the LSM simulation method. Computers & Operations Research, pp. 35: 113-129.

Damodaran, A. (2015). Damodaran online - Data: Archives. http://pages.stern.nyu.edu/~adamodar/, accessed February 12th 2015.

Dixit, A. and Pindyck, R. (1993). Investment under uncertainty. New Jersey: Princeton University Press.

Fialho, J. (2013). Avaliação de Projetos de Investigação e Desenvolvimento na Área das Telecomunicações. Coimbra: Faculdade de Economia da Universidade de Coimbra.

Folta, T. and Miller, K. (2002). Real options in equity partnerships. Strategic Management Journal, 23(1):77-88.

Longstaff, F. and Schwartz, E. (2001). Valuing American Options by Simulation: A Simple Least-Square Approach. The Review of Financial Studies, 14: 1113-1147.

Magalhães, D. (2013). Modelo de Baterias com aplicação em sistemas de gestão de baterias (BMS) de Veículos Eléctrivcos (EVs). Porto: Faculdade de Engenharia da Universidade do Porto.

Santiago, L. P. and Bifano, T. (2005). Management of R&D projects under uncertainty: a multidimensional approach to managerial flexibility. IEEE Trans. on Eng. Man., 52(2): 269-280.

Su, X. et al. (2014). Silicon-Based Nanomaterials for Lithium-Ion Batteries: A Review. Advanced Energy Materials, 4 (1): 1-23.

Trigeorgis, L. (1993). Real Options and Interactions with Financial Flexibility. Financial Management, 22 (3): 202-224.

Yeo, K. T. and Qiu, F. (2003). The value of management flexibility - a real option approach to investment evaluation. International Journal of Project Management, (21):243-250.

Zeng, X., Li, J. and Singh, N. (2014). Recycling of Spent Lithium-Ion Battery: A Critical Review. Critical Reviews in Environmental Science and Technology, 44(10):1129-1165.

Appendix

Table 1: Most relevant results concerning the scenarios in which level 1 is used at the beginning of the task, and it is also used exclusively (or almost exclusively) throughout the task.

				Probability of level 1 being used when a percentage *x* of the task is completed (in %)			Probability of the task being completed for different moments *T* (in %)		
C2	A2	R0	Pen	*x*=25%	*x*=50%	*x*=90%	*T*=5	*T*=10	*T*=15
L	L	L	L	100.0	100.0	100.0	0.0	12.3	74.3
L	L	L	M	100.0	100.0	100.0	0.0	12.1	73.6
L	L	L	H	100.0	100.0	100.0	0.0	14.6	73.0
L	L	M	L	99.9	98.6	98.9	0.0	13.3	73.3
M	L	L	L	100.0	100.0	100.0	0.0	11.3	73.9
M	L	L	M	100.0	100.0	100.0	0.0	9.7	72.1
M	L	L	H	100.0	100.0	100.0	0.0	11.3	73.9
M	L	M	L	100.0	100.0	100.0	0.0	11.3	73.9
M	L	M	M	100.0	100.0	99.7	0.1	14.0	71.9
M	L	M	H	100.0	99.6	99.4	0.0	13.1	70.4
M	L	H	L	98.6	96.3	99.4	0.0	12.7	73.2
M	M	L	L	100.0	100.0	100.0	0.0	13.7	72.9
M	M	L	M	100.0	100.0	100.0	0.0	13.7	72.9
M	M	L	H	98.9	98.6	98.4	0.0	13.0	70.1
H	L	L	L	100.0	100.0	100.0	0.0	14.1	75.0
H	L	L	M	100.0	100.0	100.0	0.0	9.7	72.1
H	L	L	H	100.0	100.0	100.0	0.0	13.9	71.7
H	L	M	L	100.0	100.0	100.0	0.0	11.7	72.0
H	L	M	M	100.0	100.0	100.0	0.0	13.7	72.4
H	L	M	H	100.0	100.0	100.0	0.0	14.3	76.0
H	L	H	L	100.0	99.6	99.9	0.0	11.7	72.0
H	L	H	M	97.3	100.0	99.7	0.1	11.9	72.0
H	M	L	L	100.0	100.0	100.0	0.0	12.6	75.1
H	M	L	M	100.0	100.0	100.0	0.0	13.4	74.9
H	M	L	H	100.0	100.0	100.0	0.0	12.7	73.9
H	M	M	L	100.0	99.7	100.0	0.0	13.4	70.4
H	M	M	M	99.7	100.0	99.6	0.0	13.4	73.9
H	H	L	L	100.0	100.0	100.0	0.0	13.9	73.0
H	H	L	M	100.0	99.7	99.7	0.0	11.00	73.3

Table 2: Most relevant results concerning the scenarios in which level 2 is used at the beginning of the task, and it is also used exclusively (or almost exclusively) throughout the task.

				Probability of level 2 being used when a percentage x of the task is completed (in %)			Probability of the task being completed for different moments T (in %)		
C2	A2	R0	Pen	x=25%	x=50%	x=90%	T=5	T=10	T=15
H	H	H	H	99.4	95.3	99.7	49.2	99.9	100.0
H	H	H	M	96.7	98.9	98.0	47.6	99.6	99.9
H	H	M	H	99.7	100.0	98.6	53.1	100.0	100.0
H	H	M	M	99.0	99.1	94.0	51.2	99.0	99.9
H	H	L	H	100.0	97.4	93.7	47.2	98.6	100.0
H	M	H	H	99.7	99.7	99.4	2.4	86.7	99.9
H	M	M	H	99.6	97.9	92.6	1.4	84.7	99.6
H	L	H	H	98.1	94.9	94.0	0.0	35.4	93.0
M	H	H	H	100.0	100.0	99.7	53.4	100.0	100.0
M	H	H	M	100.0	100.0	100.0	52.3	100.0	100.0
M	H	M	H	100.0	100.0	100.0	52.1	100.0	100.0
M	H	M	M	99.7	100.0	99.4	49.4	99.9	100.0
M	H	M	L	99.4	99.4	96.0	52.9	99.3	99.7
M	H	L	H	100.0	100.0	100.0	56.0	100.0	100.0
M	H	L	M	100.0	100.0	100.0	55.6	99.0	100.0
M	H	L	L	100.0	100.0	100.0	56.6	100.0	100.0
M	M	H	H	100.0	100.0	100.0	1.6	85.4	100.0
M	M	H	M	94.6	100.0	98.6	2.3	88.1	99.9
M	M	H	L	97.0	98.1	95.4	1.6	85.0	99.6
M	M	M	H	100.0	100.0	100.0	2.4	88.4	99.9
M	L	H	H	99.9	97.7	96.6	0.1	33.9	93.4
L	H	H	H	100.0	100.0	97.7	50.0	100.0	100.0
L	H	H	M	99.9	100.0	100.0	49.9	100.0	100.0
L	H	H	L	97.0	98.9	100.0	48.4	99.0	100.0
L	H	M	H	99.9	100.0	100.0	53.3	100.0	100.0
L	H	M	M	100.0	100.0	100.0	50.9	100.0	100.0
L	H	M	L	99.9	99.9	99.6	54.4	100.0	100.0
L	H	L	H	100.0	100.0	100.0	53.0	100.0	100.0
L	H	L	M	100.0	100.0	100.0	54.3	99.9	100.0
L	H	L	L	100.0	100.0	100.0	53.0	100.0	100.0
L	M	H	H	100.0	99.9	99.9	2.3	88.0	100.0
L	M	H	M	91.1	96.6	97.1	2.8	86.0	98.9
L	M	H	L	82.7	96.9	100.0	2.0	79.7	99.4
L	M	M	H	100.0	100.0	100.0	3.9	88.0	100.0
L	M	M	M	100.0	100.0	100.0	3.4	86.6	99.9
L	M	M	L	97.0	99.0	97.1	3.1	86.4	99.3
L	M	L	H	100.0	100.0	100.0	4.1	87.3	99.7
L	M	L	M	100.0	99.7	99.7	4.2	87.3	99.6
L	M	L	L	100.0	100.0	99.9	2.7	87.1	99.9
L	L	M	H	98.4	99.1	98.7	0.3	51.2	94.3
L	L	H	H	100.0	100.0	99.7	0	20.6	92.6

Table 3: Most relevant results concerning the scenarios in which level 1 is used at the beginning of the task, and there is a significant probability of level 2 being used throughout the task.

				Probability of level 1 being used when a percentage ***x*** of the task is completed (in %)					
C2	A2	R0	Pen	x=25%	x=50%	x=75%	x=90%		
L	L	M	M	94.00	93.57	93.71	93.71		
L	L	H	L	55.57	49.71	39.57	35.57		
M	L	H	M	98.86	88.71	87.57	88.29		
H	M	H	L	79.57	54.86	49.43	55.14		
H	H	M	L	91.71	94.57	94.57	99.00		
				Probability of level 1 being in use in different moments ***T*** (in %)			Probability of level 2 being in use in different moments ***T*** (in %)		
C2	A2	R0	Pen	T=5	T=10	T=15	T=5	T=10	T=15
L	L	M	M	92.43	82.14	24.57	5.00	7.43	0.71
L	L	H	L	57.00	28.00	3.00	51.71	43.00	10.86
M	L	H	M	88.71	80.86	25.57	3.57	11.29	1.71
H	M	H	L	66.71	47.43	14.29	13.57	33.14	0.00
H	H	M	L	95.28	84.14	24.00	0.71	4.71	0.29

Table 4: Most relevant results concerning the scenarios in which level 2 is used at the beginning of the task, and there is a significant probability of level 1 being used throughout the task.

				Probability of level 1 being used when a percentage ***x*** of the task is completed (in %)					
C2	A2	R0	Pen	x=25%	x=50%	x=75%	x=90%		
L	L	H	M	12.29	23.43	22.29	16.43		
M	M	M	L	2.14	29.86	46.14	44.14		
M	M	M	M	54.57	58.86	59.29	59.57		
M	H	H	L	23.29	8.29	15.43	9.43		
H	M	H	M	20.43	12.86	13.00	16.43		
H	H	H	L	32.71	32.71	2.29	3.14		
				Probability of level 1 being in use in different moments ***T*** (in %)			Probability of level 2 being in use in different moments ***T*** (in %)		
C2	A2	R0	Pen	T=5	T=10	T=15	T=5	T=10	T=15
L	L	H	M	20.29	20.71	8.00	79.71	46.29	3.00
M	M	M	L	46.43	13.86	0.00	53.57	22.29	0.43
M	M	M	M	58.43	47.86	10.57	41.00	8.71	0.00
M	H	H	L	16.00	0.71	0.14	53.00	1.29	0.83
H	M	H	M	19.14	13.29	4.00	79.14	10.71	0.14
H	H	H	L	30.29	3.14	2.28	36.71	5.29	0.00

Microsimulation for traffic calming schemes assessment – a case study

Joana F. Dourado and Ana M. C. Bastos Silva[1]

Abstract

The accidents' data in Portugal show an urgent need to improve the safety of vulnerable users in urban areas. In 2013 73% of the total pedestrians fatalities registered in urban areas occurred in local roads (ANSR, 2014). Many urban roads are multifunctional, used by vulnerable users and motorized traffic, which results in substantial differences in speed and degree of protection, producing a drastic asymmetry between the mobility of motor vehicles and the safety of vulnerable users. Traffic calming appears as a useful tool to mitigate this problem.

In Portugal, recent legislation allows the implementation of urban traffic calming schemes particularly adapted for residential areas: 30 km/h zones and shared spaces (woonerf zones or home zones). With this in regard the existence of a minimum standardized methodology for the implementation of these schemes is essential to ensure an adequate selection of the urban area and a proper choice of the physical measures and road environmental changes to implement.

"Before and after" studies are essential evaluation approaches of the transformed urban areas. Nevertheless, before the implementation of these schemes it would be helpful to have an idea on the impact of these measures without having to disturb the real system. Using microsimulation models it is possible to represent the reality with high detail, simulating the real network conditions and developing a perform analysis with computer representations.

[1] Civil Engineering Department, University of Coimbra
e-mails: joana.dourado@uc.pt, abastos@dec.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_6

The feasibility and applicability of these concepts were tested in the development of a real case study, where changes were proposed to some areas of the neighbourhood Norton de Matos in Coimbra and a micro-simulation model was built to evaluate the local and peripheral effects of such traffic calming schemes. A methodology for implementing such interventions in residential areas is presented. For that purpose a resident/visitor survey was also elaborated.

The results were particularly interesting. Population showed to be aware of the speed impacts and open to the implementation of such measures, defending the speed control and the safeguard of the pedestrians in residential areas. In a general analysis the evaluation of the results showed that the application of traffic calming measures would not generate a significant increase of the inter-zonal travel times, achieving however significant reductions in through traffic volumes and speeds within the study case area.

Keywords: Microsimulation, Traffic Calming, 30 km/h Zone, Home Zones

1. Introduction

The accidents' data in Portugal show the urgent need to improve the safety for vulnerable users in urban areas. When dealing with multifunctional areas one must take in consideration the following: 1. the appropriate speed is usually the lowest of the speeds adequate to the individual functions (Avenoso and Beckmann, 2005); 2. In collisions with pedestrians, if the collision speed is less than 30 km/h, more than 90% of those struck survive (SWOV, 2012). With this in regard, traffic calming appears as a useful tool to achieve the required safety conditions (Ewing, 1999; van Schagen, 2003).

Traffic calming can be translated in a set of infrastructural physical measures that change the driver's behavior inducing lower speeds and greater awareness of vulnerable road users.

The principles of traffic calming date back to the beginning of the 1960's. Since then different solutions and measures have been tested: segregation of traffic and pedestrians, shared space solutions, 30 km/h speed limit complemented with measures like: speed humps, chicanes, raised intersections, and others. See authors: Hass-Klau (1992), Danish Road Directorate (1993), Ewing (1999), Schepel (2005), TSO (2007), SOWV (2010), IMTT (2010).

In Portugal, recent legislation allowed the implementation of urban traffic calming schemes: 30 km/h zones and shared spaces (*woonerf* zones). The first 30 km/h zones in Portugal were implemented recently in Lisbon in the end of 2013 (Nunes da Silva and Lajas Custódio, 2013). With this in regard the existence of a minimum standardized methodology for the implementation of these schemes is essential to guarantee not only that they are well implemented but also that the chosen areas are suitable.

"Before and after studies" are essential elements for the evaluation of the transformed urban areas. Nevertheless before the implementation of these schemes it would be helpful to have an idea on how these changes will affect the urban area without having to disturb the real system. Recently some research works have been using microsimulation to assess traffic calming impacts (Yousif *et al.*, 2013; Ghafghazi and Hatzopoulou, 2014).

The present article focuses on traffic calming measures, such as Woonerf zones and 30 km/h zones, which have potential application in multifunctional residential areas. The applicability of these concepts is tested by developing a case study in the neighborhood Norton de Matos, in Coimbra. Then to assess the impacts of the proposed measures a microsimulation model is proposed as a useful tool.

2. Traffic Calming Schemes in Residential Areas

2.1. The Woonerf Zones

In woonerf zones pedestrians, cyclists and motorized traffic share the same area, without the separation between carriageway and sidewalks.

Concerns with urban design are present and the layouts of the streets may include areas for children to play, seats, parking places and vegetation providing not only environmental improvements but also creating obstacles for higher speeds. It should be impossible for cars to drive very much faster than walking pace (in Portugal, speed limit is 20 km/h) otherwise, the scheme may give a false feeling of safety for pedestrians and cyclists (Hass-Klau, 1992).

The main function of the woonerf zone should be residential and the traffic terminating or originating from the area should not exceed the 100 vehicles/h (Hass-Klau, 1992; de Witt and Tallens, 2001). The targeted areas can be individual streets or a set of small streets. The entrances and exits of the woonerf zone shall be recognizable with the specific sign and changes in infrastructure and the landscape.

Many studies have been done to evaluate the benefits and costs of these schemes. Although the construction costs can be high, there are many benefits (Hass-Klau, 1992; van Schagen, 2003). Study cases made in the UK and in the Netherlands (Biddulph, 2010; SOWV, 2003) pointed out reduction of vehicles speed, increase in the feeling of safety and actual reduction in reported injury accidents, some of these findings even in more affordable projects.

2.2. The 30 km/h zones

The 30 km/h zones appeared as an alternative to the woonerf. The traditional separation between the carriageway and sidewalks is established, but it is imposed a 30 km/h speed limit and the streets are treated with traffic calming measures like: gateway treatment, speed humps, chicanes, raised intersections, raised crosswalks and changing surface pavement texture and colors (TSO, 2007; Hass-Klau, 1992).

These zones are well fitted for both access and distributor urban streets (TSO, 2007) especially in residential areas and roads passing in front of schools and shopping areas (ETSC, 2008).

In terms of effectiveness, according to studies of the Institute for Road Safety Research, SWOV (2010), Netherlands, the average number of crashes decreases by about 25% when 50 km/h speed limit is redesign as 30 km/h Zone. In the UK, the Transport Research Laboratory (TRL) reviewed results from 250 zones in England, Wales and Scotland (Webster and Mackie, 1996). The main findings indicated that average speeds reduced by about 15 km/h, annual accident frequency fell by 60% and traffic flow in the zones was reduced by 27%, causing the increase of 12% in the surrounding boundary roads.

2.3. Combining the two schemes

In terms of rehabilitating the street life and activities and discourage car use, the woonerf zones are more efficient than simple traffic calmed streets (Biddulph, 2012) imposed by vertical signing. However its implementation conditions are restrictive. With that in regard, combining 30km/h zones with smaller woonerf areas inside them may be a good approach.

Public participation is an essential key for the acceptance of the traffic calming measures implementations (van Schagen, 2003). The integration of the residents in the elaboration of the projects allows unheard concerns to be raised, creating more adequate and acceptable solutions (Biddulph, 2003) and stimulating the relations between authorities and population, enforcing the community feeling.

3. Methodology Approach

This study focused on the definition and evaluation of a proposal for the transformation of an existing neighborhood through the use of traffic calming measures. The methodology was based on five key stages (Figure 1): (i) selection of the intervention area; (ii) diagnosis; (iii) development of the solution; (iv) impact assessment using a microsimulation model; (v) analysis of the results.

3.1. Case Study Area

The selected case study concerns the neighborhood Norton de Matos and the surrounding streets, located in Coimbra, Portugal (Figure 2). This is a multifunctional residential area: it has not only housing, but also one basic school, a cultural center and many markets, coffee shops, and other facilities;

It is crossed by some local and major distributor roads, which creates conflicts between local activities and through traffic.

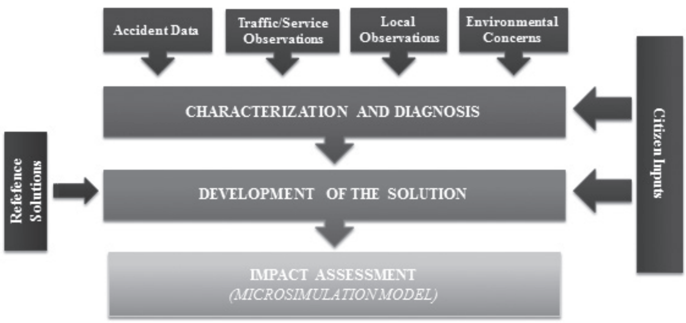


Figure 1: Methodological approach (from Dourado, 2013)

Figure 2: Study area, with the highlight of the intersections studied with more detail (from Dourado, 2013)

Due to lack of resources it was not possible to study the entire area with detail. Instead it was chosen the most interesting and problematic zone to analyze with more detail and more generic analyses and solutions for the remaining area. The zone chosen was the set of intersections between the distributor streets "Vasco da Gama", "Macau", "Mouzinho de Albuquerque" and "Daniel de Matos". This zone was chosen due to the presence of many facilities (the school, the cultural center and many cafes) and high through traffic volumes (Figure 2).

3.2. Characterization and Diagnosis

In order to develop an adequate solution for this residential area a diagnosis was first made (Figure 1). This diagnosis was based on five categories (Hoyle, 1995 quoting Smith *et al.*, 1980):

1. *Citizen inputs* based on residents and visitors surveys: a survey based on a 5-point scale was made to ascertain residents and visitors' satisfaction and perceived safety as pedestrians, easy of parking, use of the street by children to play (Biddulph, 2010) and degree of acceptance of possible approaches to the space occupancy changes.

2. *Traffic/Service observations:* road hierarchy, traffic volumes, traffic speed, parking conditions, public transport supply.

3. *Records: accident data*. The accident data was provided by official entities (ANSR) and included accidents between vehicles and vehicles and pedestrian.

4. *Local observations of resident and visitor activities;*

5. *Environmental concerns:* access for pedestrians, visual quality and noise.

3.3. Development of the solution

Given the non-existence of support manuals with neither built examples of 30 km/h zones nor woonerf zones in Portugal, the development of the solution was based on international best practices, particularly in northern/central Europe.

3.4. Microsimulation Model development

Microsimulation is a dynamic and stochastic modeling of individual vehicle movements within a transport system/network. These models simulate vehicle-by-vehicle, updating their position, speed, acceleration, lane position and other state variables on time steps as they interact with each other and with the environment (traffic signals, road geometry, etc.) (Gettman and Head, 2003). These interactions are modeled using complex algorithms (microscopic models) that describe some maneuvers and behaviors. The three fundamental sub-models are: lane changing, car following and gap acceptance.

In traffic analysis, the microscopic traffic simulation models are being increasingly used (Habtemichael, 2013). Simulation-based studies allow investigating hypothesis on safety and efficiency of traffic operations under controlled environment and different traffic conditions (Lee *et al.*, 2004 and Abdel-Aty *et al.*, 2006).

The development of a simulation model involves five steps (Vilarinho and Tavares, 2012): model codification, verification, calibration, validation (it should be used different data or performance indicators from the calibration) and scenario testing.

4. Characterization and Diagnosis Results

4.1. Data collection

Accident data: 14 accidents were recorded from 2007 to 2011, being 6 collisions with pedestrians and 8 collisions between vehicles, all resulting in light injuries. The majority of these accidents occurred in the streets "Vasco da Gama", "Macau", "Daniel de Matos". This data refers only to accidents with injured people or situations that required the police presence, so there may have been other less severe accidents not reported.

Traffic volumes: traffic volumes during the morning rush hour were globally the most higher. The most frequented street was "Rua Mouzinho de Albuquerque" with approximately 600 vehicles per hour arriving to the intersection. In both intersections, the traffic circulated fluently, so no capacity problems were recorded. During the afternoon, although the global volumes were a bit lower, the congestion level of the intersections was higher, due to illegal parking near the facilities.

Speed: The average speeds measured in three sections: "Daniel de Matos", "Mouzinho de Albuquerque" and "Macau" in free flow speed conditions, with no stopping or decelerating at the intersections or crosswalks, were between 38 km/h and 45 km/h and the percentile 85 was between 42 km/h and 51 km/h. However the range of speeds recorded in the same locals could go from 25 km/h to 60 km/h or more.

Parking conditions: Just outside the facilities there is not enough parking during the peak hours. However, there are free spaces in almost every street in the area during all day. Most of the parking places are not formalized.

Access for pedestrians: The minimum sidewalk width of 1,5 m is not respected everywhere and some pedestrian paths are degraded.

Local activities, visual quality and noise: The study area offers a good social environment. Inside the neighborhood Norton de Matos, some of the squares have suffered recent improvements, offering places for socialization. The discomfort caused by the vehicles noise, it is more significant in the distributor streets.

Citizen inputs: the sample included young people, adults and older people. The majority of residents interviewed were retired and in the case of the visitors they were young or active people. Nevertheless their opinion did not differ radically in most of the answers. The majority of people feel that cars pass too fast in the residential area, causing unsafety feelings. When asked about possible changes in the neighborhood, both residents and visitors showed good acceptance: almost 80% agreed with the 30 km/h speed limit in that area, 77% agreed with the implementation of traffic calming measures in the major streets and 82% answered affirmatively when asked if they agreed with giving the priority to pedestrians in local access streets.

4.2. Discussion of the Diagnosis Results

The characterization work and diagnostic allowed confirming that although there are no serious problems of road safety, the quality of life of the neighborhood has been deteriorating in time with the car invasion, thus justifying a global intervention in order to defend and promote the presence of vulnerable users and the improvement of local life quality.

Registered accidents reflect the existence of conflicts between vehicles and pedestrians on the major distributor roads. The speed measurements reinforce these findings. Although the average values are not very high, they may be sufficient to cause accidents with serious consequences, considering that this is an area with many facilities and strong presence of pedestrians (including children). Also the high range of speed values in the same street suggests that the road design is permissive, allowing driving behavior depending almost exclusively from driving personal style, which can vary drastically and also be unpredictable.

The through traffic tends to be another problem. Although it was not possible to assess the weight associated with this traffic component (which would require making origin/destination surveys), it is expected to be significant. Although the measured traffic volumes may not be very high in terms of road capacity, they tend to be high for a residential area. The priority given to the major distributor roads that cross this area, allowing high speeds, may encourage the presence of through traffic.

The illegal parking near the facilities (school and cultural center) creates conflicts in the intersections justifying an intervention. Also reorganizing the parking offer in the neighborhood is a way of not only maximizing the offer but also of improving the public space quality.

5. Development of the Solution Proposal

The proposed solution seeks to respond to three major objectives:

- Eliminate/decrease the through traffic;
- Speed control and homogenization;
- Decrease conflicts and accidents.

The basic principle of the solution concerned the protection of the vulnerable users and urban experience over vehicle mobility, without affecting the accessibility. With this in regard the global solution includes two strategies of intervention: (i) Creation of a 30 km/h zone covering the entire neighborhood; (ii) Identification of restrictive areas inside the neighborhood to transform into woonerf zones.

5.1. Redefinition of the road hierarchy

The first task focused on the redefinition of the road hierarchy in the study area, downgrading some of the current major distributor roads that cross the center of the neighborhood - Streets "Vasco da Gama", "Daniel de Matos", "Macau" and "Mouzinho de Albuquerque". This changes aims to affect the traffic conditions (increasing travel times) on these streets to discourage the through traffic, sending it to the peripheral major distributor roads.

5.2. New circulation plan

The definition of a new circulation plan (Figure 3) was made, which included the alteration of some traffic direction changes and creation of one way roads. The aim was to create a group of closed circuits inside the area, in order to eliminate the long rectilinear trajectory that links street Mouzinho de Albuquerque and Macau (the major crossing axis).

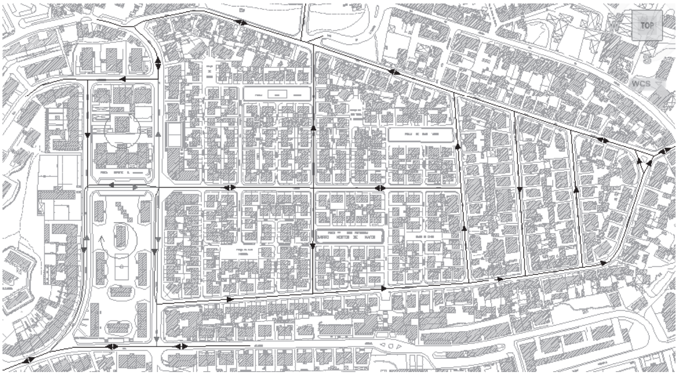


Figure 3: New circulation plan (from Dourado, 2013)

5.3. The 30 km/h Zone

With the definition of the 30 km/h zone, a new legal speed limitation is imposed, which contributes to the increase of safety without high prejudice of the motor vehicles circulation. This includes the application of the road signal and the implementation of traffic calming measures in the distributor roads, such as: chicanes by reorganizing the parking spaces alternately on each side of the street, raised platforms in the intersection and raised crosswalks. A detail solution was made for a group of intersections in the neighborhood (Figure 4): pavement elevation until the sidewalks height, colorful and textured pavements and enlargement and safeguard of the sidewalks, reposition of crosswalks, lighting improvement.

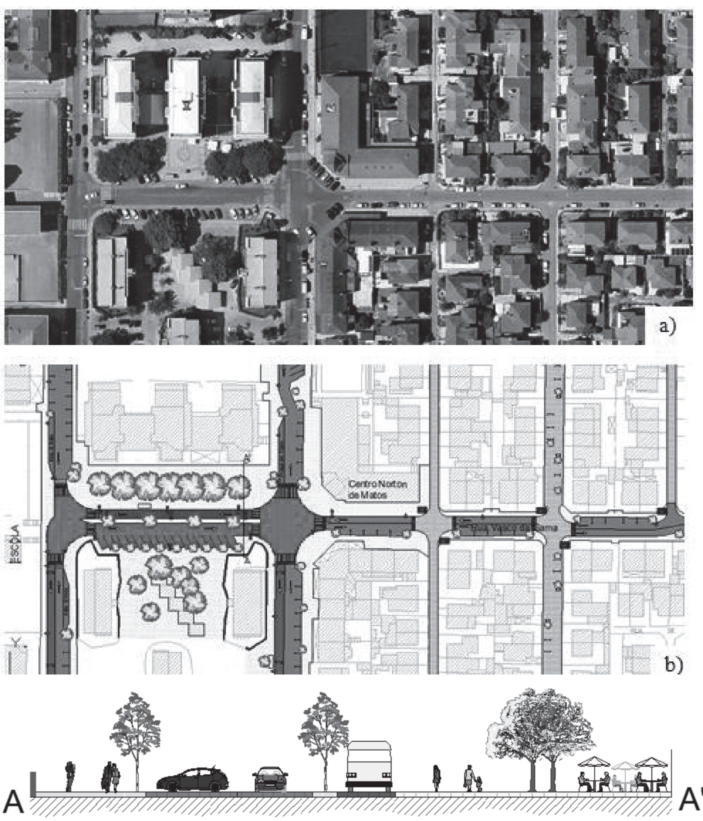


Figure 4: Solution proposed for the study area: a) present situation; b) proposed project plant and traverse cutting of the proposal for the intersections. (from Dourado, 2013)

5.4. Woonerf Zones

Creation of woonerf zones inside the 30 km/h zone, in local access streets and squares inside the neighborhood (smaller streets in figure 4, b)). Delimitating the parking places in these streets and adding some street furniture, without blocking the circulation of the resident's cars or emergency vehicles, creating good spaces for children to play and neighbors to socialize. The entries in these streets should be signalized with the respective road signal and raised platforms in the intersections can contribute to a gate effect. Parking inside these zones should be restricted to the residents.

5.5. Complementary measures

To respond to the parking supply problems one suggests some parking management measures, like: *kiss&ride* near the school, charging tariffs and time limitations, spaces reserved to residents.

For the public transport in this particular area, one proposes the creation of bus lanes for the existing public transport routes.

6. Impact Assessment using a Microsimulation Model

Given the impossibility of making before and after analysis, one opted to build a simulation model using the software *Aimsun*, from *Transport Simulation Systems*® in order to assess the possible effects of the proposed measures.

Thus, two scenarios were created:

- The Current Situation - reference scenario;
- Proposed Solution - incorporating amendments to traffic directions and the resulting impedances from maximum speed changing.

6.1. Building the model

In order to assess the impacts of the proposed solution for the study area one decided to expand the codified area to some surrounding spaces. Thus, for the simulation the study area became delimited by "Rua do Brasil", "Estrada da Beira", "Av. Mendes da Silva" and "Av. Cónego Urbano Duarte".

The codification of the area was based on air photographs and cartography, courtesy of the Municipality of Coimbra. Additionally centroids were defined to represent traffic origin/destination in the network. Also, bus circuits were established.

The O/D matrix was deduced from the collected data sample of traffic volumes. Due to the lack of data, the calibration of the model was made in a qualitative way, based on local observations, known dynamics of the area and using for quality control purposes the traffic volumes recorded in the studied intersections. As this exercise was mainly a comparative study to evaluate the magnitude of the effects resulting from the application of the proposed solution compared to a reference situation, it was considered not particularly relevant to reliably simulate the real circulation conditions. The software does not allow a detailed representation of all traffic calming measures (i.e., pavement elevation). To overcome this, one represented impedances in speed in the sections were the measures were supposed to be implemented.

6.2. Performance Indicators

To answer the objectives of the study one selected the following indicators: total travel time, total travel distance; traffic volumes assigned to each arc of the network; average speed in the network; average density of the network.

At the same time, the selection of the routes by the model for each O/D pair was analyzed.

6.3. Simulation results and discussion

As previously referred there were selected two scenarios: a reference scenario and the proposed solution. Even taking into account possible mismatches with the real situation, the simulation results (see Figure 5) corresponded generally to the expectations:

- General decrease of traffic volumes inside the neighborhood Norton de Matos, especially in the street Vasco da Gama (interior segment of the neighborhood);
- Significant increase of traffic volumes in some surrounding streets of the neighborhood;
- Significant increase of traffic volumes in the street Daniel de Matos (targeted street in the solution) and in the segment of the Street Vasco da Gama, that links the two targeted intersections.

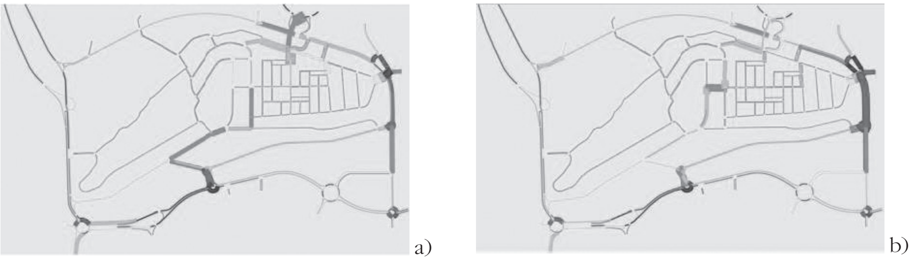


Figure 5: Network traffic assignment: a) reference situation; b) after changes. Traffic volumes increase with the darkness of the lines and density (veh./ Km) increases with thickness (from Dourado, 2013)

In terms of routes choice, there were registered alterations. As an example, the O/D pair “Portela/Solum”, which corresponds to south/north centroids, shows a clear reduction in the slice of traffic that crosses the neighborhood - from 31.87% to 12.39% (Figure 6). It is also possible to see the adoption of a new route using the street “Paulo Quintela”, which has suffered from traffic volumes increase after the creation of the 30 km/h zone.

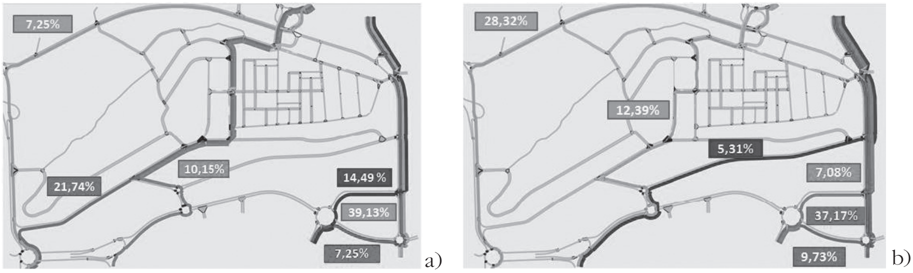


Figure 6: Routes choice for Portela/Solum O/D pair. a) reference situation, b) after alterations (from Dourado, 2013)

The global performance indicators show interesting results: the total travel time increased 2,35% (from 202,65 h to 207,41 h), the total travel distance remained approximately constant, the average speed in the network decreased from 50.01 km/h to 49.08 km/h and the average density of the network slightly increased from 4.36 vehicles/km to 4.48 vehicles/km.

The simulation results allowed concluding that the measures introduced in the neighborhood would not cause drastic changes in the average global speed neither in the average travel time correspondent to the area of influence of the intervened zone. Although these results cannot be generalized for other situations, these conclusions are especially important to help mitigate the idea that traffic calming measures penalize significantly the general community circulation. This means that it may be possible to delimit areas subject to traffic calming measures without causing significant disturbances in the local mobility and accessibility.

Nevertheless it is important to highlight the possible less positive effects of this proposal: the migration of traffic and conflicts/accidents to other peripheral streets inside the neighborhood. It is essential to guarantee that there are conditions to accommodate this traffic and that the speed control measures are efficient. In this case, the proposed measures to these streets are expected to respond to these demands.

7. Conclusions

Through the study of foreign experience, namely north and center European countries, it was possible to determine extremely positive results associated with the implementation of traffic calming in residential areas.

In order to test the applicability of these concepts a study case was developed. A set of integrated measures was proposed to the neighborhood Norton de Matos. The solution definition was preceded by a diagnosis based on population surveys complemented with local observations and collection of traffic and speed data. To assess the impacts of the proposed measures to the target area and to its influential zones, a microsimulation model was built with Aimsun Software.

There are still some aspects of this study that should be taken into account in the analysis and assessment of the results:

- The isolated study of the target area without contemplating the adjacent areas may be a limitation in terms of results interpretation and in the definition of the solutions;

- The difficulties in the data collection for the diagnosis showed to be a limitation not only for the proposal definition but also for the construction of the simulation model. The impossibility to use an O/D matrix representative of the real traffic demand conditioned significantly the quality and robustness of the results analysis.

Despite the difficulties, the microsimulation showed to be an effective tool for evaluating the potential effects of proposed amendments. The results are in line with the expected, indicating the effectiveness of traffic calming solutions and identifying some less desirable effects.

This work is not a finished product. One can identify a set of actions that should be further developed: the extension of the zone (in terms of detailed solution) and an extensive data collection to build, calibrate and validate the microsimulation model. These two actions are essential to assess the overall effects associated with the implementation of the 30 km/h zone and other measures.

Acknowledgements: This work has been carried out in the framework of projects EMSURE - Energy and Mobility for Sustainable Regions (CENTRO-07--0224-FEDER-002004).

References

Abdel-Aty, M., Dilmore, J. and Hsia, L., Applying Variable Speed Limits and the Potential for Crash Migration. Transportation Research Record, 1953, pp. 21–30, 2006.

ANSR. (2014). Relatório Anual de Vítimas a 30 dias- 2013. ANSR.

Avenoso, A. and Beckmann, J., The Safety of Vulnerable Road Users in the Southern, Eastern and Central European Countries (The “SEC Belt”), European Transport Safety Council, Brussels, 2005.

Bastos Silva, A., Vasconcelos, L., Correia, G. and Santos,S., Microssimulação Aplicada aos Estudos de Acessibilidade. 7º Congresso Rodoviário Português (CRP), 10 a 12 de Abril, Lisboa, 2013.

Biddulph, M., Evaluating the English Home Zones Initiatives. Journal of the American Planning Association, 76:2, 199-218, 2010.

Biddulph, M., Street Design and Street Use: Comparing Traffic Calmed and Home Zones. Journal of Urban Design, 17:2, 213-232, 2012.

Biddulph, M., Towards Successful Home Zones in the UK. Journal of Urban Design, 8:3, 217-241, 2003.

Danish Road Directorate, An Improved Traffic Environment – A Catalogue of ideas. Ministry of Transport, Denmark, 1993.

De Wit, T. and Talens, H., Traffic Calming in the Netherlands, CROW, the Dutch National Information and Technology Centre for Transport and Infrastructure, 2001.

Dourado, J.F., Soluções de Acalmia de Tráfego com Potencial de Aplicação em Áreas Residenciais Multifuncionais, Master Dissertation on Civil Engineering under the suppervision of Professor Ana Maria César Bastos Silva, University of Coimbra, 2013. Written in portuguese.

ETSC-European Transport Safety Council, Shlow! Show me How Slow- Reducing Excessive and Inappropriate Speed Now: a Toolkit. European Transport Safety Council. Brussels, 2008.

Ewing, R., Traffic Calming: State of the Practice, Institute of Transportation Engineers, Washington. 1999.

Gettman, D. and Head, L., Surrogate Safety Measures from Traffic Simulation Models. Transportation Research Record, 1840, pp. 104-115, 2003.

Ghafghazi, G. and Hatzopoulou, M., Simulating the air quality impacts of traffic calming schemes in a dense urban neighborhood. Transportation Research Part D 35 (2015) 11–22. 2014.

Gomes, G., May, A. and Horowitz, R., Congested Freeway Microsimulation Model Using VISSIM. Transportation Research Record, 1876, pp. 71–81, 2004.

Habtemichael, F.G. Improved Active Traffic Management System for Motorway Safety and Efficiency: Benefits of Reducing the Driving Task Difficulty, Thesis specifically prepared to obtain the PhD Degree in Transportation Systems under supervision of Professor

Luís Guilherme de Picado Santos, Universidade Técnica de Lisboa, Instituto Superior Técnico, Lisboa, 2013.

Hass-Klau, C. et al., Civilised Streets- A Guide to Traffic Calming, Environmental & Transport Planning, Brighton, 1992.

Hoyle, C. L., Traffic Calming. American Planning Association. Michigan, 1995.

IMTT- Instituto de Infra-Estruturas Rodoviárias, Disposições Normativas. Medidas de Acalmia de Tráfego, Instituto de Infra-Estruturas Rodoviárias, 2010.

Lee, C., Hellinga, B. and Saccomanno, F., Assessing Safety Benefits of Variable Speed Limits. Transportation Research Record, 1897, pp. 183–190, 2004.

Nunes da Silva, F. and Lajas Custódio, R. Zonas 30 – Segurança Rodoviária, Vida e Vitalidade para os Bairros da Cidade de Lisboa, 7º Congresso Rodoviário Português (CRP), 10 a 12 de Abril, Lisboa, 2013.

Schepel, S., Woonerf revisited - Delft as an example, International Childstreet2005 Conference, Delft, 2005.

Smith, D. et al., State of the Art Report: Residential Traffic Management, U.S. Department of Transportation Federal Highway Administration, Report No. FHWA/RD-80/092, 1980.

SWOV- Institute for Road Safety Research, SWOV Fact Sheet- Zones 30: Urban Residential Areas. Leidschendam. 2010.

SWOV- Institute for Road Safety Research. SWOV Fact Sheet-Vulnerable Users, SWOV, Leidschendam, 2012.

TSO - The Stationery Office, Local Transport Note 01/07- Traffic Calming. Department for Transport, The Stationery Office, London, 2007.

van Schagen, I., Traffic Calming schemes - Opportunities and Implementation Strategies, SWOV, Leidschendam, 2003.

Vilarinho, C. and Tavares, J.P., Traffic Model Calibration: A Sensitivity Analysis. Compendium of papers of the 15th Edition of the Euro Working Group of Transportation. International Scientific Conference. 10 to 13 December, 2012, Paris. 2012.

Webster, D. and Mackie, A., Review of traffic calming schemes in 20 mph zones. TRL Report 215. Transport Research Laboratory, Crowthorne, 1996.

Yousif, S., Alterawi, M. and Henso, R., Effect of Road Narrowing on Junction Capacity Using Microsimulation, Journal of Transportation Engineering, ASCE/June 2013.

Página deixada propositadamente em branco.

Reliability Metrics for the evaluation of the Schedule Plan in Public Transportation

Jorge Freire de Sousa[1,2], **João Mendes-Moreira**[3,4],
Luís Moreira-Matias[5] **and João Gama**[3,6]

Abstract

Nowadays, the major Public Transportation Companies around the world use intelligent transportation systems based on automated data collection frameworks. The existence of these data has driven to the development of new approaches to the operational planning of public transportation. These approaches, commonly known as ADC-based operational planning strategies (ADC from Automated Data Collection), to improve public transportation reliability consist of adjusting the definitions made on the initial steps of the operational planning process by using real-world data. This type of changes concentrates mainly on restructuring routes and adjusting the existing schedule plan (SP). However, the usefulness of such tunings from a company point-of-view is often of difficult evaluation.

This paper starts by presenting a brief review on improving the network definition based on historical location-based data. Then, it presents a broad review on ADC-based evaluation techniques of the schedule plan reliability, discussing the existing metrics.

[1] CEGI-INESC TEC, 4200-465 Porto, Portugal

[2] DEGI-FEUP, U. Porto, 4200-465 Porto, Portugal

[3] LIAAD-INESC TEC, 4200-465 Porto, Portugal

[4] DEI-FEUP, U. Porto, 4200-465 Porto, Portugal

[5] NEC Laboratories Europe, D-69115 Heidelberg, Germany

[6] Faculdade de Economia, U. Porto 4200-465 Porto, Portugal

e-mails: jfsousa@fe.up.pt, joao.mendes.moreira@inescporto.pt, luis.matias@neclab.eu, jgama@inescporto.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_7

The purpose of this paper is to critically describe the performance indicators used in the evaluation of the SP reliability, following the aforementioned bibliographic reviews. They will be certainly useful to shape the approaches developed by the research community for improving the quality of public road transportation operations based on data collected by ADC systems.

This paper focuses on two different, yet highly related, approaches: 1) changing the network definition; 2) evaluating and adjusting the SP in place. The automatic control strategies and the different actions to improve the SP remain out of the scope of this paper.

Keywords: Automated Data Collection (ADC); Operational Planning (OP); Public Transportation (PT); Network Design; Schedule Plan (SP); Reliability Metrics.

1. Introduction

In the last three decades, following a growing demand for fast transportation services in urban areas and clear advances both in real-time communications and in vehicle location technologies, public road transport companies have made important investments in information systems mostly dedicated to their operations (Furth *et al.*, 2003; Barabino *et al.*, 2013; Mazloumi *et al.*, 2010; Hounsell *et al.*, 2012). Automatic Vehicle Location (AVL), Automatic Passenger Counting (APC), Automatic Fare Collection (AFC), and multimodal traveler information systems are just some examples of this kind of answer from the operators to a major concern of reliability and service quality level from passengers. As a consequence of this effort in Advanced Public Transportation Systems, these companies have been able to collect massive data, indeed real continuous flows of data (Furth *et al.*, 2003).

The existence of these new data has driven to the development of new approaches for planning the operations of public transport companies, and many researchers have highlighted their potential to offer insights on new ways to evaluate and improve service reliability by means of both operational planning (OP) and control (Strathman *et al.*, 1999; Strathman, 2002; Strathman *et al.*, 2003). Some of these approaches imply changes in the initial phases of OP. However, the usefulness of such changes from a company view-point is not always easy to assess (Mendes-Moreira & de Sousa, 2014).

The aim of this paper is precisely to discuss works on the improvement of quality of public road transport operations based on OP strategies, mainly by using AVL and AFC data, and to improve the metrics and the approaches used in the evaluation of public transport reliability.

This paper starts with the presentation of how operational planning is usually done in public transport companies. Then, the next two sections are assigned to the review of literature; section 3 is dedicated to the alterations in the network design and section 4 to the evaluation of the schedule plan, both concerned with the reliability of the public transport service. Following the discussion presented in these two sections, challenges, opportunities and research ideas are proposed in section 5. Finally, in section 6, some highlights and thoughts on future research trends are synthesized.

2. Service and Operational Planning

Service and OP at public transport companies include network and route design, frequency determination, and vehicle and crew scheduling (Ceder, 2002; Wilson *et al.*, 2009):

a) *Network definition:* It consists of defining the *lines*, *routes*, and bus stops. Here, a route is an ordered sequence of directed road stretches between an origin and a destination, passing by multiple bus stops. A line is defined as a set of routes – typically two – with very similar paths, but inversely ordered.

b) *Trips definition:* The most common method involves firstly the definition of the set of bus stops for which schedule time-points will be set, necessarily including the origin and the destination ones. Then, timestamps are assigned to the previously defined schedule time-points. In high-frequency routes, however, this timetabling may also be defined by setting the time between two consecutive trips in the same route (this is called headway-based). The set of resulting trips is often defined as the *schedule plan* (SP).

c) *Definition of duties:* A duty is the work that a driver and/or a bus must perform in a day. The definition of the drivers' duties has much more constraints than the definition of bus duties (for instance, governmental legislation, union agreements and company rules). The logical definition of bus duties is commonly done prior to the drivers' duties.

d) *Assignment of duties:* It consists of physically assigning the previously defined logical duties to the companies' drivers and buses. The assignment of driver duties to drivers is called rostering.

Some authors (Vuchic, 2005; Ceder, 2007) also call this sequential or hierarchical process tactical planning. They claim that in the first stage of the process – the network design – decisions are less frequent; service considerations, judgement, and manual analysis tend to dominate. Going down the list, the dominance of these factors change up to the point where crew scheduling decisions are constantly made using cost considerations as the main driver and where computer-based analysis dominates in the optimization of the system (Aguerrebere, 2012).

Anyway, the data needed for this set of functions are extensive, encompassing all the inputs required for travel demand forecasting, as well as information on usage of the system and current route and network performance. And AVL and AFC data are most likely to be effective in characterizing this usage and the performance of the existing system (Wilson *et al.*, 2009).

There are internal and external causes for reliability problems arising in PT networks. The former are usually associated to persistent problems as they include factors such as route configuration, inappropriate scheduling, times for the entry and exit of passengers at bus stops. The external ones are more sporadic. They include traffic congestion and accidents, raining days or badly parked cars. Only the internal causes are addressed by means of OP strategies, trying to avoid unreliability on a long-term perspective. The sporadic problems are lessened by control strategies, ensuring just corrective actions for a specific situation in a given moment (Abkowitz & Tozzi, 1987).

In this paper we will only concentrate our attention in the internal causes and in the ADC-based OP strategies to improve the reliability of PT operations by using real-world data, since they facilitate the analysis of the supply and demand, while allowing new network proposals. This type of approaches focuses on (a) altering routes or restructuring the network or on (b) adjusting the existing SP.

3. On Modifying the Network Design

As previously mentioned in section 2, the PT network design consists mainly of defining the number of routes to build, along with their paths and bus stops.

The importance of the design and planning of the network structure for PT success is usually undervalued, and surprisingly the topic is more or less neglected in standard texts on public transport or transport policy (Nielsen & Lange, 2010). Maybe the infrequent nature of the network design can explain this underrating but getting the network right is usually more important than the often debated and studied mode selection.

Also, during the design stages of PT, little attention is paid to operational reliability, even if many design choices have a great influence on schedule observance. During the network design, reliability should be taken into account as a design parameter (Oort & Nes, 2009a).

It is clear that the data obtained from ADC systems can have wide-ranging applications within public transport. One of them, of particular interest, is the opportunity to make use of these databases to develop a better picture of how public transport systems are performing and being used (Wilson *et al.*, 2009). Better estimates of performance measures and usage attributes may be made at lower cost than by conventional methods, and, for the first time, it is possible to evaluate important attributes such as those concerned with reliability and its effects, until recently virtually impossible to quantify due to scarcity of data. Indeed, while there is an extensive bibliography about demand variations due to fare changes, findings on demand variations due to the level of service are occasional.

There are a few AVL-based works focused on improving the PT reliability by adjusting the route definition. The most common approach is changing the location of the bus stops. The work in Yan *et al.* (2006) used historical GPS data to mine human mobility patterns in a major Taiwan intercity bus operation to find an optimal compromise between the location of bus stops and operational costs. The researchers did so by discovering the demand patterns using both a stochastic demand scheduling model and heuristic-based methods to solve the models. A passenger wait cost-based model is developed in Li and Bertini (2008) to find the optimal bus stop spacing based on historical AVL data.

More recently, a couple of works were published showing that many cost-effective opportunities to improve the level of service reliability, together with the application of operational instruments, both related to network structure, lead to highly reliable services.

In Oort and Nes (2009a), a tool is developed to calculate the additional waiting time due to variability and transfers based on actual journey and passenger data. A case study in The Hague shows that in the case of long lines with large variability, splitting the line could result in less additional travel time because of improved reliability. This benefit compensates for the additional transfer time, provided that the transfer point is well chosen.

Public transport network planners often propose network structures that either assume a certain level of regularity or are even especially focused on improving service reliability, such as networks in which parts of lines share a common route or the introduction of short-turn services. The key idea is that travelers on that route will have a more frequent transit service. The impact of such network designs on service regularity is rarely analyzed in a quantitative way.

Oort and Nes (2009b) present a tool that can be used to assess the impact of network changes on the regularity on a transit route and on the level of transit demand. The tool can use actual data on the punctuality of the transit system. The application of such a tool is illustrated in two ways. A case study on introducing coordinated services shows that the use of such a tool leads to more realistic estimates than the traditional approach. Second, a set of graphs is developed which can be used for a quick scan when considering network changes. These graphs can be used to assess the effect of coordinating the schedules and of improving the punctuality.

Nielsen and Lange (2010) try to demonstrate the importance of network planning and design for the success of public transport. They present proposals for the structuring of multi-modal public travel networks in different types of urban and rural districts, and they also give examples of “good” practices from different regions and countries.

Also, Oort *et al.* (2013) describe the state of publicly available transit data, with an emphasis on the Dutch situation. The value of insights from Automatic Vehicle Location data is demonstrated by examples. A software tool, that makes comprehensive operational analysis possible for operators and public transport authorities, was able to identify several bottlenecks when applied in practice.

4. On Evaluating SP Reliability

In this section, the authors decided to follow quite closely the terms of the survey presented in Moreira-Matias *et al.* (2015).

The SP reliability is a vital component for service quality. Improvements on reliability may increase the service demand and, consequently, the companies’ profitability. Low reliability levels lead to a limited growth in the number of passengers and to a decreased perceived comfort (Strathman *et al.*, 1999). It is possible to establish three distinct axes on evaluating SP reliability (Oort, 2011): 1) the unexpected increases on the waiting time on bus stops; 2) the time spent in crowded situations caused by transport overloading; and 3) delays on the passengers’ arrivals due to travel time variability (TTV). The first two axes are mainly related with passengers’ *comfort* and *experience* criteria. The value of such extra time consumptions varies from the passenger condition (seated or standing). However, these two aspects are mainly *satisfiers*: additional aspects that the passengers like to have but are not essential factors to abandon the services provided by a certain PT company. On the other hand, the last one is a fundamental issue by the disturbances that it does on the passengers' daily activities (Oort, 2011). By directly affecting the *convenience* and the *speed* of transportation, it is crucial to maintain the travelers' confidence on the PT network (i.e., a *dissatisfier*).

For the aforementioned reasons, this survey is focused on carrying the SP reliability evaluation by the existing TTV. Once established, it is expected that an SP meets the passengers' demand by *following* their mobility needs (namely, their daily routines). Typically, service unreliability is originated by one (or many) of the following causes (Fattouche, 2007; Cham, 2006): schedule deviations at the terminals, passenger load variability, running time variability, meteorological factors, and driver behavior.

Today's urban areas are characterized by a constant evolution of road networks, services provided, and location (for instance, new commercial and/or leisure facilities). Therefore, it is highly important to automatically assess how the SP suits the needs of an urban area. An efficient evaluation can lead to important changes in an SP. These changes will lead to a reduction in operational costs (for instance, by reducing the number of daily trips in a given route) and/or a reliability improvement in the entire transportation network, which will increase the quality of the passengers' experience and, therefore, the number of customers.

An SP consists of a set of *k* schedules, which provide detailed information about every trip running on previously defined routes. Each schedule contains a timetable. Different routes may have different timetables. Nevertheless, they share the number *k* of schedules and the daily *coverage* of each schedule.

A schedule planning process for a given route relies on three distinct steps: the first step is defining the number *k* of schedules and their individual coverage; second, the *schedule time-points* are chosen among all bus stops in the route; and finally, the third step is defining timetables for each route schedule containing the time the buses pass at each scheduled time-point (per trip). This process is done for all routes.

From the aforementioned definition of SP, it is possible to divide the SP evaluation into two different dimensions: the suitability of the number of schedules k and of the set of their daily coverages and the reliability of their timetables (to test whether the real arrival times of each vehicle at each bus stop are meeting the previously defined timetable). Although there is an obvious impact on the definition of the timetable, to the authors' best knowledge, until very recently there was no research reported in the literature addressing the evaluation of the number of schedules and their daily coverage (Mendes-Moreira *et al.*, 2015).

This section defines and reviews evaluation methodologies with regard to the reliability of timetables.

Evaluation Metrics

When evaluating an SP, it is important to differentiate *low-frequency services* and *high-frequency services* (Turnquist, 1982): in low frequency services, passengers arrive at the bus stops shortly before the bus's scheduled services, whereas in high-frequency services, the customers tend to arrive at the stops *randomly* (Jolliffe & Hutchinson, 1975; Turnquist, 1978; Bowman & Turnquist, 1981; Ceder & Marguier, 1985). In the first scenario, *punctuality* is the main metric, whereas the service *regularity* is the most important metric in high-frequency routes. There is no exact boundary between these two scenarios. Fan and Machemehl (2000) conducted a data-driven experiment in Austin, Texas (USA), where they identified a 10-min threshold. Recent studies have also used 10–12 min as a threshold between low and high frequency services (Oort, 2011; Trompet *et al.*, 2011).

Polus (1979) presented a landmark paper proposing four measures of performance for evaluating SP reliability on arterial routes: overall TT, congestion index, overall travel speed, and delay. All these measures were route based and highly focused on the operational perspective. The first three are mainly variations of the remaining ones – which are based on ratios between the actual and expected run times. Delay was a more sophisticated measure, defined as all the time consumed while traffic is impeded in its movement somehow – but also reported as hard to obtain by then.

The AVL data enabled the possibility to extend this analysis to other granularities than route based such as segment based or stop based. Following such advances, four main indicators were first proposed by Nakanishi (1997) and followed by other similar studies (Strathman *et al.*, 1999; Barabino *et al.*, 2013). These indicators are outlined as follows: 1) *on-time performance* (OTP); 2) *run time variation* (RVT); 3) *headway variation (HV)*; and 4) *excess waiting time (EWT)*. The first two indicators are more applicable to low-frequency routes, whereas the last two focus on the high-frequency routes (Turnquist, 1982; Strathman, 1998; Strathman *et al.*, 1999). This set of indicators is the most widely known formulations of these metrics, which have been used on multiple studies in the last decade. They are formally presented here.

OTP indicates the probability that buses will be where the schedule says they are supposed to be. It is possible to represent this metric by an *arrival delay* (AD) in a given trip i, i.e., AD_i as function of both the *scheduled arrival time*, i.e., SAT_i, and the *actual arrival time*, i.e., AAT_i. Therefore, it can be defined as follows (Strathman, 1998):

$AD_i = AAT_i - SAT_i$. (1)

The RTV represents the variation on the run times performed by each trip. Some introductory concepts on this subject will be presented below. Typically, the TT reports the trip duration, from terminal to terminal, and is often referred to as *round-trip time*. TT is often used to define the time required to go from one point of interest to the other. This last definition is used in this survey. One of the factors that mostly affect the RTV is the *dwell time*, which is the total time the bus has to stay at a given bus stop for passenger boarding and alighting. From the passenger perspective, a larger variation can mean a longer waiting time in some stops and/or missed transfers. From the operational planners' perspective, greater RTV translates into higher costs as a result of the extra hours that must be added to accommodate passenger load (Strathman, 1998). This indicator is more appropriate for routes that cover long distances, facing many traffic lights and regular traffic delays (Sterman & Schofer, 1976). Given a set of *n* trips of interest, it is possible to compute the RTV as follows (Strathman *et al.*, 1999):

$$\mathrm{RTV} = n^{-1} \times / \sum_{i=1}^{n} |\mathrm{SAT}_i - \mathrm{AAT}_i| \,/\, \mathrm{AAT}_i \quad (2)$$

In high-frequency routes, where the trips start within very short headways, the OTP is not that relevant (Hounsell & McLeod, 1998). The HV represents the probability that controllers are able to maintain a regular and stable headway between each pair of vehicles running in the same routes.

Let $f_{i,j}$ be the frequency (i.e., scheduled headway) established between a given pair of trips (i, j), whereas $H_{i,j}^{b}$ represents the observed headway on such pair of trips at a bus stop of interest, i.e., b. The *headway ratio* on the bus stop b, i.e., $\mathrm{Hr}_{i,j}^{b}$, is defined as follows (Strathman *et al.*, 1999; Strathman, 1998):

$$\mathrm{Hr}_{i,j}^{b} = \left(\frac{H_{i,j}^{b}}{f_{i,j}^{b}}\right) \times 100 \quad (3)$$

where the value 100 represents a perfect SP matching. Given a set of *n* trips of interest, it is possible to compute the standard deviation and the mean value of Hr (σ_{Hr}^{b} and μ_{Hr}^{b}, respectively). We can do it by calculating every possible $\mathrm{Hr}_{i,i+1}$: $i \in \{1, \ldots, n-1\}$ at a bus stop *b*. Then, it is possible to obtain the HV at bus stop b throughout these *n* trips as follows (Lesley, 1975):

$$\mathrm{HV}^{b} = \left(\frac{\sigma_{\mathrm{Hr}}^{b}}{\mu_{\mathrm{Hr}}^{b}}\right). \quad (4)$$

The EWT is an estimation of the excessive waiting time that passengers experience as a consequence of unreliable service. It is possible to calculate the EWT at a bus stop *b*, i.e., EWT^{b}, as a function of HV^{b}. A possible way to do so is presented as follows (Welding, 1957):

$$\mathrm{EWT}^{b} = \left(\frac{{\sigma_{\mathrm{Hr}}^{b}}^{2}}{2 \times \mu_{\mathrm{Hr}}^{b}}\right). \quad (5)$$

The bus stop *b* used to compute statistics on the first two indicators is the destination bus stop. For the last two indicators, any bus stop can be considered a reference if it has a frequency scheduled to it, i.e., $f_{i,j}^{b}$. Commonly, such statistics are computed by the transit companies aggregating its values to a fixed time granularity (typically, 1-h periods) (Barabino *et al.*, 2013), but they can be also computed according to the trip.

SP Evaluation Studies

Many works have evaluated schedule reliability by measuring the aforementioned indicators on historical AVL data sets. Strathman (1998) and Strathman *et al.* (1999) evaluated schedule reliability on the Tri-Met by measuring indicators (1–4), whereas the work by Bertini *et al.* (2003) solely focuses on the first two ratios. Traditionally, the HV was often disregarded by the transit planners due to the intrinsic *chaos* assumed (as the schedule time-points on the timetables are not the central variable to confirm service reliability). Nevertheless, recent advances have changed this reality: in Strathman *et al.* (2003), AVL/APC data were considered to evaluate the impact of the HV on the operational control. Another perspective of the Tri-Met data is presented in Berkow *et al.* (2007), where an analysis of indicators (2–4) demonstrated the feasibility of using AVL data along with other data sources to better accomplish their evaluation. Lin and Ruan (2009) formulated probabilistic headway regularity metrics (HV). Then, the authors tested their approach using AVL data from Chicago. In Bellei and Gkoumas (2010), relations between transit assignment, bus bunching events, and operation models are mined from the location-based data. This study aimed to identify irregularities in HV's distribution function caused by an inadequate SP. The reliability of an express service implemented in Montreal, Canada, is evaluated in El-Geneidy and Surprenant-Legault (2010) by employing the first two indicators. A large-scale evaluation was performed by Hounsell *et al.* (2012), where the data acquired through the *iBus* (an AVL/APC framework installed on a bus fleet running in the city of London, U.K.) were used to evaluate all the four main indicators of schedule reliability.

Another approach to evaluate schedule reliability on a route is the segment-based one. It consists of identifying segments/parts of a route where there are greater schedule deviations, and therefore, the SP should be adjusted by changing the timetable or by introducing bus priority lanes and/or traffic signals in intersections. One of the first authors to carry out such work was Horbury (1999) based on the HV. In Mandelzys and Hellinga (2010), it is proposed to measure indicators (1 and 2) using stop-based metrics and to identify the causes for larger deviations through an empirical framework. The work of El-Geneidy *et al.* (2011) proposes a way of identifying *where* the schedule is unreliable by evaluating the first two indicators on the schedule time-points.

Recently, the methodological approach to evaluate SP reliability has evolved from the key indicators to using nonparametric deterministic methods such as data envelopment analysis (DEA), as described in Mendes-Moreira and de Sousa (2014). The main advantage enabled by employing such a method is the possibility of directly comparing metrics from distinct dimensions by introducing decision-making units. Lin *et al.* (2008) used AVL data to establish confidence intervals for the DEA scores based on the four indicators previously introduced. Despite its usefulness in identifying cost-based relationships between the resources used and the service produced, the DEA models are not addressed in this survey as they usually use a wider scope of data on the companies' management than we do.

Many of the aforementioned works have often employed the four traditional transit measures to evaluate schedule reliability. Nevertheless, few works have been successful in identifying the factors behind poor performance measurements. Such measurements focus mainly on the passengers' perspective about the service. Recently, innovative approaches have emerged on this research topic, such as the day-to-day variability. Mazloumi *et al.* (2010) proposed to determine the nature and shape of TT distributions for different departure time windows at different times of the day, using data from a route running in Melbourne, Australia. Factors causing TTV in public transport are also explored using regression methods. A method for finding interesting contexts to justify RTV is proposed in Jorge *et al.* (2012): Distribution rules are employed to identify particular conditions that lead to systematic bus delays. The HV is explored using a sequence mining approach in Moreira-Matias *et al.* (2012). The goal is to highlight sequences of bus stops where a failure to meet the schedule systematically leads to bus bunching situations further stops ahead in the route. Recently, an innovative study was presented by Chen *et al.* (2009), where three novel metrics were proposed to address three distinct granularities, namely, stop, route, and network levels. This approach seems promising. However, it also fails to deliver a unique indicator on the SP reliability.

The contribution of this ongoing generator of historical trip data to evaluate SP reliability is that it replaces the old estimations on TTV with real values (Bertini & El-Geneidy, 2003). The findings of the evaluations previously described consisted of identifying unreliable schedule time--points (El-Geneidy *et al.*, 2011; Jorge *et al.*, 2012) or badly designed bus priority lanes (Hounsell *et al.*, 2012). In this work, some evaluations also build dwell time models that help to understand how this variable changes from trip to trip and throughout the day.

The four metrics are well established in the literature. However, they focus mainly on the passengers' perception of service quality, particularly the EWT. The OTP can help the planners identify the *exact* schedule time-points to be changed, whereas the RTV shows a more general perspective on network service, which can lead to more profound studies on the drivers' behavior, terminal dispatching policies, or on the current schedule's slack. The HV is the most used metric. Even so, it is possible to observe that the company's perspective on such RTV is not addressed as a primary goal of these evaluation studies.

Nevertheless, even if it is possible to identify *what* is happening and *where* changes must be performed to improve SP reliability, it is not easy to identify *how* it is possible to improve it.

5. Research Challenges

Two main issues may be identified where further AVL-based research should be employed to improve the evaluation of SP reliability: (a) creating a unique evaluation indicator, considering the company's perspective on the evaluation by including external factors in the evaluations or by developing cost-related evaluations and to (b) evaluate the reliability of the current schedule's number and coverage.

The aforementioned four evaluation metrics are classical but widely used in evaluation studies. However, distinct metrics (which are highly correlated to the main ones) are continuously emerging. It is known that the importance of each one of these indicators *depends* on the frequency established in the route. However, to the author's best knowledge, there is no consensual, individual, and integrated reliability ratio. This gap in the literature leads to an important research question: *Is it possible to build a consensual frequency-dependent reliability ratio based on these four main indicators?*

The first step in building an SP is defining both the schedule's number and day coverage. Then, a timetable is assigned to each schedule in a stepwise process already discussed. This definition has an explicit impact on the definition of timetables. However, to the authors' best knowledge, no research addresses the evaluation of whether the schedule's number and coverage still suit the current demand patterns and network behavior. Consequently, a question arises: *Is it possible to assess whether the schedule's number and coverage are suitable for the network needs based on historical AVL data?*

Another topic that can be a challenge is to rethink the OP. Section 2 briefly revises the steps of the traditional OP. Although AVL-based research has recently emerged on improving route definition, most AVL-based works on OP focus on the SP. The state of the art relies on deterministic and cost-based models. The AVL data make it possible to perform a bottom-up OP evaluation, namely, correctly exploring the available resources or even reducing them if possible to meet the current demand. A complete AVL-based framework to redesign all the steps of the OP is an intelligent transportation system (ITS) that could be a research goal on this topic for the medium-term future.

6. Conclusion and Future Trends

Over the last decade, various relevant contributions have emerged on location-based ITS applications for improving the OP of mass transit transportation networks. The spatiotemporal features of this type of data provided novel opportunities to reveal underlying patterns on unexpected behaviors that are deteriorating the quality of the service. These data are now affordable and widely available as a standard in every medium/large-sized mass transit company.

Such innovation revolutionized the way to improve both operational planning and control in these networks. The theoretical traffic models, which were the state of the art for improving OP during the 1980s and 1990s are now being progressively replaced by complex yet efficient statistical and machine learning models.

It is also important to provide real-time information to the passengers about what is *happening* in the network (i.e., on-the-spot information on arrival times). More than building an exact but time-consuming prediction on arrival time, the researchers have focused on building simple frameworks capable of learning from location-based data streams and of providing predictions with low uncertainty.

The AVL-based improvements to planning and control are becoming increasingly mature, but the existing evaluation studies are still mainly proofs of concept focused on the passengers' perspective. Some challenges have been addressed in the previous section.

The high availability of reliable data reporting the vehicle operations in real time pushes up the will of researchers in this field. It is expectable that data driven models will prove themselves as state-of-the-art methods for improving PT reliability. More than ever, the AVL data are a real-time stream. Such availability, along with the expansion of urban centers, can progressively change the traditional focus on planning to an autonomous data-driven real-time control, which may reduce the manpower required for those tasks.

Acknowledgements: This work is financed by the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project UID/EEA/50014/2013

References

Abkowitz M. & Tozzi, J., "Research contributions to managing transit service reliability," J. Adv. Transp., vol. 21, no. 1, 1987, pp. 47–65.

Aguerrebere, Carlos G. Q., Redesigning and improving the efficiency of bus transit networks using automatically collected data: the case of Gipuzkoa, Spain, MSc Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2012.

Barabino, B., Di Francesco, M. & Mozzoni, S., “Regularity diagnosis by automatic vehicle location raw data,” Public Transp., vol. 4, no. 3, 2013, pp. 187–208.

Bellei, G. & Gkoumas, K., “Transit vehicles’ headway distribution and service irregularity,” Public Transp., vol. 2, no. 4, 2010, pp. 269–289.

Berkow, M., Chee, J., Bertini, R. & Monsere, C., Transit performance measurement and arterial travel time estimation using archived AVL data, presented at the ITE District 6th Annual Meeting, Portland, OR, USA, 2007.

Bertini, A. & El-Geneidy, R., “Generating transit performance measures with archived data,” Transp. Res. Rec.: J. Transp. Res. Board, vol. 1841, no. 1, 2003, pp. 109–119.

Bowman, L. & Turnquist, M., Service frequency, schedule reliability and passenger wait times at transit stops,” Transp. Res. Part A: Gen., vol. 15, no. 6, 1981, pp. 465–471.

Ceder, A. & Marguier, P., “Passenger waiting time at transit stops,” Traffic Eng. & Control, vol. 26, no. 6, 1985, pp. 327–329.

Ceder, A., “Urban transit scheduling: Framework, review and examples,” J. Urban Planning Develop., vol. 128, no. 4, 2002, pp. 225–244.

Ceder, A., Public Transit Planning and Operation: Theory, Modeling and Practice. Oxford, U.K.: Butterworth-Heinemann, 2007.

Cham, L., Understanding Bus Service Reliability: A Practical Framework Using AVL/APC Data, Ph.D. dissertation, Massachusetts Institute of Technoogy, Cambridge, MA, USA, 2006.

El-Geneidy, A. & Surprenant-Legault, J., “Limited-stop bus service: An evaluation of an implementation strategy,” Public Transp., vol. 2, no. 4, 2010, pp. 291–306.

El-Geneidy, A., Horning, J. & Krizek, K., “Analyzing transit service reliability using detailed data from automatic vehicular locator systems,” J. Adv. Transp., vol. 45, no. 1, 2011¸ pp. 66–79.

Fan, W. & Machemehl, R., “Characterizing bus transit passenger waiting times,” in Proc. 2nd Mater. Specialty Conf. Can. Soc. Civil Eng., Montreal, QC, Canada, 2002, pp. 1–10.

Fattouche, G., Improving High-Frequency Bus Service Reliability through better Scheduling, Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2007.

Furth, P., Hemily, B., Muller, T. & Strathman, J., Uses of Archived AVLAPC Data to Improve Transit Performance and Management: Review and Potential, Transportation Research Board, Ed. Washington, D.C., USA, 2003.

Horbury, A., “Using non-real-time automatic vehicle location data to improve bus services,” Transp. Res. Part B: Methodol., vol. 33, no. 8, 1999, pp. 559–579.

Hounsell, N. & McLeod, F., “Automatic vehicle location: Implementation, application, and benefits in the United Kingdom,” Transp. Res. Rec.: J. Transp. Res. Board, vol. 1618, no. 1, 1998, pp. 155–162.

Hounsell, N., Shrestha, B. & Wong, A., “Data management and applications in a world-leading bus fleet,” Transp. Res. Part C, Emerging Technol., vol. 22, 2012, pp. 76–87.

Jolliffe, J. & Hutchinson, T., “A behavioural explanation of the association between bus and passenger arrivals at a bus stop,” Transp. Sci., vol. 9, no. 3, 1975, pp. 248–282.

Jorge, A., Mendes-Moreira, J., de Sousa, J. F., Soares, C. & Azevedo, P., “Finding interesting contexts for explaining deviations in bus trip duration using distribution rules,” in Proc. IDA, 2012, pp. 139–149.

Lesley, L., “The role of the timetable in maintaining bus service reliability,” in Proc. Operating Public Transp. Symp., 1975.

Li, H. & Bertini, R., "Optimal bus stop spacing for minimizing transit operation cost," in Proc. ASCE Traffic Transp. Studies, 2008, pp. 553–564.

Lin, J. & Ruan, M., "Probability-based bus headway regularity measure," IET Intell. Transp. Syst., vol. 3, no. 4, 2009, pp. 400–408.

Lin, J., Wang, P. & Barnum, D.,"A quality control framework for bus schedule reliability," Transp. Res. Part E: Logistics Transp. Rev., vol. 44, no. 6, 2008, pp. 1086–1098.

Mandelzys, M. & Hellinga, B., "Identifying causes of performance issues in bus schedule adherence with automatic vehicle location and passenger count data," Transp. Res. Rec.: J. Transp. Res. Board, vol. 2143, no. 1, 2010, pp. 9–15.

Mazloumi, E., Currie, G. & Rose, G., "Using GPS data to gain insight into public transport travel time variability," J. Transp. Eng., vol. 136, no. 7, 2010, pp. 623–631.

Mendes-Moreira J. & de Sousa, J.F., "Evaluating changes in the operational planning of public transportation", In Rossi, R. and de Sousa, J.F. (Ed.), Computer-based Modelling and Optimization in Transportation, vol. 262. New York: Springer-Verlag, 2014, pp. 57–68.

Mendes-Moreira, J., Moreira-Matias, L., Gama, J. & de Sousa, J.F., "Validating the Coverage of Bus Schedules: A Machine Learning Approach," Information Sciences, Vol.293, 2015, pp.299-313.

Moreira-Matias, L., Ferreira, C., Gama, J., Mendes-Moreira, J. & de Sousa, J. F., "Bus bunching detection by mining sequences of headway deviations," in Advances in Data Mining. Applications and Theoretical Aspects, vol. 7377. Berlin, Germany: Springer, 2012, ser. LNCS, pp. 77–91.

Moreira-Matias, L., Mendes-Moreira, J., de Sousa, J.F. & Gama, J., "Improving Mass Transit Operations by Using AVL-Based Systems: A Survey", IEEE Transactions on Intelligent Transportation Systems, pp.-, 2015.

Nakanishi, Y., "Bus performance indicators: On-time performance and service regularity," Transp. Res. Rec.: J. Transp. Res. Board, vol. 1571, 1997, pp. 1–13.

Nielsen, G. & Lange, T., Network Design for Public Transport Success: Theory and Examples, presented at the Thredbo Transport conference, 2010.

Oort, N. van & Nes, R. van, "Line length vs. Reliability: network design dilemma in urban public transport", Transportation Research Record: Journal of the Transportation Research Board, Vol. 2112, Issue 1, 2009a, pp. 104-110.

Oort, N. van & Nes, R. van, "Regularity analysis for optimizing urban transit network design", Public Transp., vol. 1, no. 2, 2009b, pp. 155-168.

Oort, N. van, Service reliability and urban public transport design, Ph.D. dissertation, Netherlands TRAIL Research School, Delft, The Netherlands, 2011.

Oort, N. van, Sparing, D., Brands, T. & Goverde, R., Optimizing Public Transport Planning and Operations Using Automatic Vehicle Location Data: The Dutch Example. In: 3rd International Conference on Models and Technologies for Intelligent Transport Systems, Dresden, Germany, 2013, pp. 291-300.

Polus, A., "A study of travel time and reliability on arterial routes," Transportation, vol. 8, no. 2, 1979, pp. 141–151.

Strathman, J., Automated Bus Dispatching, Operations Control and Service Reliability: Analysis of Tri-Met Baseline Service Date, University of Washington, Seattle, WA, USA, Tech. Rep., 1998.

Strathman, J., Kimpel, T. & Dueker, K., "Automated bus dispatching, operations control and service reliability," Transp. Res. Rec., vol. 1666, 1999, pp. 28–36.

Strathman, J., Kimpel, T., Callas, S. & Northwest, T., Headway deviation effects on bus passenger loads: Analysis of Tri-Met's archived AVLAPC data, Citeseer, State College, PA, USA, Tech. Rep., 2003.

Strathman, J., Tri-Met's experience with automatic passenger counter and automatic vehicle location systems, Center Urban Studies, Portland State Univ., Portland, OR, USA, 2002.

Trompet, M., Liu, X. and Graham, D., "Development of key performance indicator to compare regularity of service between urban bus operators," Transp. Res. Rec.: J. Transp. Res. Board, vol. 2216, no. 1, 2011, pp. 33–41.

Turnquist, M., "A model for investigating the effects of service frequency and reliability on bus passenger waiting times," Transp. Res.Rec., vol. 663, 1978, pp. 70–73.

Turnquist, M., Strategies for improving bus transit service reliability, Transportation Research Board, Washington DC, USA, Tech. Rep., 1982.

Vuchic, V.. Urban Transit. Wiley, 2005.

Welding, P., "The instability of a close-interval service," OR, vol. 8, no. 3, 1957, pp. 133–142.

Wilson, N., Zhao, J., & Rahbee, A., "The potential impact of automated data collection systems on urban public transport planning", In Wilson, N. and Nuzzolo, A. (Ed.), Schedule-Based Modeling of Transportation Networks: Theory and Applications. New York: Springer, 2009, pp. 75-99.

Yan, S., Chi, C. & Tang, C., "Inter-city bus routing and timetable setting under stochastic demands," Transp. Res. Part A: Policy Practice, vol. 40, no. 7, 2006, pp. 572–586.

Página deixada propositadamente em branco.

The impact of a crisis in a cost-benefit analysis: What has changed in the Douro Interior sub-concession economic evaluation?

Sandrina Filipe, Pedro Godinho and Joana Dias[1]

Abstract

In this paper we analyse the impact of a crisis in the economic evaluation of a road infrastructure. We consider the Douro Interior sub-concession, awarded by the Portuguese government to a private consortium in 2008, and concluded in 2012. The government decision was based on economic analysis made on 2008, which concluded that the economic benefits of the infrastructure would be larger than its economic costs. Today there are some doubts whether it was really worthwhile to build the infrastructure, but any analysis based on the current situation must take into account that Portugal underwent a severe crisis with significant impact on its macro-economic indicators. The exercise we undertake consists of analysing what would be changed in the 2008 analysis if it would have been possible to foresee the macro-economic changes in the Portuguese situation.

In order to perform this analysis, we use the economic evaluations published in 2008 and we update the macro-economic data according to the information available at the beginning of 2015. The resulting evaluation highlights the components of the analysis that would have been altered if the crisis and other macro-economic changes had been foreseen.

We conclude that the changes in the Portuguese macro-economic situation would be enough to justify the alteration of the recommendation produced by the economic evaluations. This highlights the importance of using methodologies that allow the incorporation of risk in the assessment of such projects, as well as the

[1] University of Coimbra, Faculty of Economics
E-mails: sandrina.filipe@fe.uc.pt, pgodinho@fe.uc.pt, joana@fe.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_8

explicit incorporation of the postponement option in the analysis. Additionally, it shows that some caution is necessary when performing ex-post analysis of economic studies: unforeseen changes in the macro-economic situation may be enough to alter the recommendation provided by such studies.

Keywords: Cost-Benefit Analysis; Highways and roads; Economic factors

1. Introduction

Over the last decades there has been a significant growth of the investment in road infrastructures in Portugal (Ferreira, 2013; Rosmaninho, 2010). Empirical studies show that such investment had positive impacts on employment, private investment and private output, and that it also contributed to increased labour productivity and purchasing power (Pereira & Andraz, 2007).

In order to carry out such investments within the constraints of the government budget, Portuguese authorities adopted a funding method based on Public–Private Partnerships (PPP) and "shadow tolls", that is, tolls which are paid by the government instead of the user. Under this scheme, private companies are responsible for the design, construction, financing, operation and maintenance of the roads. In return, they are paid an amount that depends on the traffic volume, usually for a period of 30 years (Pereira & Andraz, 2012).

The Douro Interior Sub-Concession was awarded on 2008 to "AENOR – Douro Interior S.A.", now "Ascendi Douro, Estradas do Douro Interior, S.A." (Ascendi, 2015), and it was concluded in 2012 (Ascendi, 2015). The decision to build the infrastructure was supported by economic studies made in 2008: KPMG/VTM (2008) and Reis *et al.* (2008). Despite these studies, some political sectors expressed their skepticism about the benefits of the infrastructures whose construction was planned to start in 2008 and 2009, and criticised the costs that they would entail to the taxpayers.

Nowadays, there seems to be some signs that traffic volume has been below what was expected in 2008. This probably means that the project benefits are also lower than previously expected. One justification for a lower traffic volume can possibly be found in the economic and financial crisis that has affected Portugal.

Godinho and Dias (2012) evaluated the postponement option concerning this infrastructure, based on the same data that was used by KPMG/VTM (2008) and Reis *et al.* (2008). A simulation analysis performed by the authors confirmed that the expected benefits were larger than the expected costs. However, by taking into account the uncertainty concerning GDP changes and fuel prices, and by using real options analysis, the authors concluded that the expected net social benefit would be maximized by waiting until some of the uncertainty was resolved. This result raises methodological questions concerning the way benefit-cost analysis has been conducted, and makes us wonder whether its conventional application is sufficient for achieving robust conclusions.

In this paper, we will try to give insights into some methodological issues by answering some questions:

- What influence did the crisis have in the economic evaluations, comparing the analysis made in 2008 with the analysis made in the light of what is known today?
- If the current macro-economic situation (early 2015) had been anticipated at the time, would the recommendation given by the studies be different?
- What methodological lessons can be learned from this case?

The rest of the paper is structured as follows. In the next section we outline the main components of cost-benefit analysis considered in this study. In sections 3 and 4, we present the changes that result from updating macro-economic variables, and their impact on the components of Cost-Benefit Analysis. In section 5, we report the results, and we present the conclusions in section 6.

2. The Components of Cost-Benefit Analysis (CBA)

CBA is a crucial tool for determining the project economic viability (Afonso, 2013; Reis *et al.*, 2008), and it was the methodology used in the economic analysis made in 2008, regarding the Douro Interior Infrastructure. In what follows, we will detail the main options that were considered in the CBA, describing the main benefits and costs, and the way they were quantified. The direct net benefits estimated in those analysis concerned mainly three groups of stakeholders:

- Infrastructure Managers;
- Users;
- Society.

The direct net benefits assumed for the infrastructure managers were, as it is usual in the case of road infrastructures:

- Construction costs (initial costs): initial investment related to pavement construction, including expropriation costs (Reis *et al.*, 2008);
- Maintenance and operation costs: all activities that occur throughout pavement life cycle in order to keep its quality above a required threshold;
- Residual value (considered as a negative cost or a benefit): present value of the road infrastructure considering its use beyond the horizon of analysis, based on its structural and functional value.

The direct net benefits considered for the users were:

- Value of travel time: travel time savings, quantified according to recommendations from the HEATCO project regarding Portugal (Bickel *et al.*, 2006);
- Vehicle operating costs: both fuel and nonfuel costs were considered. These costs depend essentially on travelled distance, driving speed and fuel prices, according to Department for Transport (DfT) (2009);
- Benefits associated to new users: some users will travel due to the availability of the new infrastructure.

The impact to society includes both direct and indirect effects:

- Accident casualties: costs of fatalities and injuries (serious and light injuries), and property damage, quantified according to recommendations from the HEATCO project regarding Portugal (Bickel *et al.*, 2006);
- Air pollution emissions: emission of pollutants that affect public health, the buildings, the ecosystem and risks associated with the environment, quantified according to the recommendations of HEATCO project regarding Portugal (Bickel *et al.*, 2006);
- Wider economic benefits: improved labour supply and increased output in imperfectly competitive markets, measured indirectly according to the recommendations of DfT (2005).

No tolls were applied in this infrastructure, so tolls were not considered in the CBA. According to Reis *et al.* (2008), the net impact of noise was negligible, so it was not included in the analysis. The specific quantitative information regarding the infrastructure was gathered from KPMG/VTM (2008), KPMG (2008) and Reis *et al.* (2008).

3. Incorporation of the impact of macro-economic changes

The original studies concerning this infrastructure are KPMG/VTM (2008) and Reis *et al.* (2008). Both studies assumed a time horizon of 30 years for the analysis, from 2009 to 2038, and the macro-economic assumptions were based on the information available in 2008 and forecasts produced by national and international institutions at that time. These studies considered the same data regarding the construction, operation and use of the infrastructure (traffic volume and construction costs, for instance), and similar macro-economic settings, presenting some differences in the methodological choices made. This results in different but similar estimated Economic Net Present Value (ENPV): KPMG/VTM (2008) estimate an ENPV of 261 million euros and Reis *et al.* (2008) estimate an ENPV of 225 million euros. In this paper we follow mostly the methodological choices of Reis *et al* (2008).

The economic environment turned out to be very different from what was expected at the time. We want to assess how these differences impact the results of the previous studies. In order to do so, we will consider recent macro-economic data (as of January 2015) and assess its impact on the ENPV. We will now briefly describe the variables whose values are adjusted, and outline their impact.

Gross Domestic Product (GDP)

GDP influences directly and significantly several components of the CBA: value of travel time, environmental costs, accident costs and wider economic benefits. Apart from this direct influence, changes in GDP will also be reflected in other components, probably even more important than the ones already cited: GDP will have an impact on traffic, which is an essential driver of the value of many benefits. The evolution of GDP since 2008 has been much worse than what was expected at the time, contributing to a decrease in the ENPV.

To use updated GDP data, values from 2009 to 2013 were gathered from PORDATA (2014). International Monetary Fund (IMF) forecasts were used for the 2013 to 2019 period (IMF, 2015). As there were no forecasts available for the rest of the time horizon, a constant GDP growth rate was considered, equal to the one defined by the IMF for 2019.

GDP disparity

GDP disparity for the region in which the infrastructure is located (Douro Interior) in relation to the whole country is taken into account in the analysis. The average per capita GDP for the Douro Interior region is lower than the national average, leading to smaller economic benefits of the infrastructure.

In Reis *et al.* (2008) it was considered that the Douro Interior *per capita* GDP was 68,0% of the national average. The updated available information from the *Instituto Nacional de Estatística* (INE, 2014a) leads to values of 70,0% and 71,9% for 2012 and 2013, respectively. In the present study we assumed an average *per capita* GDP of the Douro Interior region remaining at 71,9% of the national average for the rest of the time horizon.

Changes in GDP disparity have an impact on the value of travel time and on the wider economic benefits. Since disparity was reduced and, therefore, the *per capita* GDP of the region is now a larger percentage of the corresponding national value, this will contribute to an increase of the ENPV due to a larger valuation of time savings and wider economic benefits.

Fuel prices and non-fuel operation costs

Petrol and diesel are the types of fuel used by almost all road vehicles in Portugal, and so they were the only two type of fuels considered in the analysis. Fuel prices have a direct impact on users' operation costs and also an indirect impact on different components of the CBA, through their impact on traffic (if fuel costs increase, some users will stop using the infrastructure).

Reis *et al.* (2008) assumed constant fuel prices before taxes, equal to the average of their values in the 12 months previous to the analysis (from October 2007 to September 2008), as reported by the *Direcção Geral de Energia e Geologia* (DGEG). KPMG/VTM (2008) do not report the assumed fuel cost, but they use constant vehicle operation costs per unit of distance, implying that constant fuel prices were assumed. In this paper, we will update these values using annual data available until 2014 (DGEG, 2015). Beyond 2014 we assume a constant price, equal to the 2014 average price.

Value-added tax (*Imposto sobre o Valor Acrescentado*, IVA) and tax on oil products (*Imposto Sobre Produtos Petrolíferos*, ISP) are indirect taxes, so their direct impact is removed from the economic analysis. However, these taxes are paid by users, so they impact their decisions concerning the use of the infrastructure, and also provide some indications in relation to the value of the perceived benefits. For IVA, previous analysis considered a constant rate of 20% and for ISP the values prevailing in 2008 were also assumed to be kept constant throughout the analysis.

IVA and ISP will only have an impact after the infrastructure starts being used. This means that their values will only be relevant for time periods when at least some road sections are already opened to traffic (2011 and beyond). In 2011, IVA has changed from 20% to 23%, so in this paper a constant rate of 23% is considered. We notice that, apart from its impact in fuel prices, IVA is also considered in the non-fuel operation costs.

ISP values were also updated, according to the information from DGEG (DGEG, 2015). Values of ISP are available up to 2014, and they were assumed to remain constant afterwards.

Both fuel prices and taxes on both fuel and non-fuel costs increase in comparison with the initial analysis. This will lead to a direct increase in operating prices and, more significantly, a decrease in traffic. Both these aspects will contribute to a decrease of the ENPV.

Wages as percentage of GDP and total tax revenue as a percentage of GDP

Wages as percentage of GDP, and total tax revenue as a percentage of GDP, have an impact on the wider economic benefits of the infrastructure.

Wages as percentage of GDP were updated, up to 2013, using values from INE (2014b). After 2013, we assumed a constant value, equal to the value corresponding to 2013. Tax revenue as a percentage of GDP was gathered from OECD (2014). Values up to 2012 were available and, for later years, we assumed values equal to the percentage corresponding to 2012.

Both these variables had lower values than the ones assumed in the initial studies. The specific impact of these variables, as measured in the wider economic benefits, concerns the additional entries to the labour market, and while a reduction in the wage to GDP ratio will tend to increase the benefit, a reduction in the tax to GDP ratio will tend to decrease it. The resulting impact is, therefore, undefined.

Average load of heavy vehicles

The benefits of the infrastructure for heavy vehicles will depend on the average load of such vehicles. We updated the value of this average load, using data from INE for the years 2012 (INE, 2013) and 2013 (INE, 2014c). For the following years, we assumed a constant value equal to 2013.

The values we gathered were greater than those that had been assumed in the previous studies, leading to an increased benefit of travel time savings for heavy vehicles, and also to an increase in wider economic benefits. Consequently, this will contribute to an increase in ENPV. However, some caution is necessary in considering this impact, since we did not consider that a larger load may mean that fewer heavy vehicles will use the infrastructure. In a sensitivity analysis we consider the case in which the average load is unchanged.

Wages, wage taxes and unemployment rate

Wages have an impact on both the construction costs and the costs of maintaining and operating the infrastructure. Wage taxes and unemployment rate have an impact on the determination of the costs of labour, according to the methodology proposed by the European Commission (EC, 2008).

In the analysis of Reis *et al.* (2008), a different treatment was given to labour costs, depending on whether the corresponding jobs were created by the project or not. In the case of the labour used for building the infrastructure, it was considered that those jobs were not created by the project, and the cost of the corresponding wages was not adjusted: the underlying assumption is that those workers would be working somewhere else if they were not in the project.

In the case of jobs created by the project - the labour costs for maintaining and operating the infrastructure - Reis *et al.* (2008) assumed that those jobs would not exist without the infrastructure, and the approach presented by EC (2008) for calculating shadow wages was followed. According to this approach, taxes and social security payments should be used to adjust the cost of labour and, when there is high involuntary unemployment (more than 15%-20%), the unemployment rate should also be used to adjust the cost of labour. In this approach, the shadow wage rate is calculated as:

$SWR=W\ (1-u)(1-t)$ (1)

Where SWR is the shadow wage, W is the market wage, u is the unemployment rate and t is the rate of social security payments and relevant taxes. In EC (2008) it is also noticed that, unless there is high involuntary unemployment, the above expression will understate the cost of labour. Taking this into account, and considering that the unemployment rates in 2008 were well below 15% (and a significant rise was not expected), Reis *et al.* (2008) only considered a conversion factor regarding the rate of social security payments and relevant taxes (*u* was considered equal to zero).

In this paper, we follow an approach similar to Reis *et al.* (2008). However, we take into account that there was a significant rise in unemployment in Portugal, exceeding 15% in some years, so we also include the unemployment rate in the determination of the shadow wage rate for the labour created by the project.

Both in the case of labour used to build the infrastructure and labour used for its operation and maintenance, we make a correction to the wages, in order to take into account the changes in real wages occurred in Portugal. A zero growth rate in real wages had been assumed in the previous studies, meaning that wages were expected to increase at the same rate as the inflation. In this paper, we use the nominal wages reported by PORDATA (2015) to calculate the growth of real wages until 2012. In some of the years, we can observe a negative growth in real wages. For future years, we assumed that the real wage would be kept constant (that is, that it would rise at a rate equal to inflation).

For the labour used to operate and maintain the infrastructure, we updated the rate of social security payments and relevant taxes, and we incorporated unemployment in the analysis. For wage taxes and social security payments, we used data until 2013 available from OECD (OECD, 2014) and we assumed that the wage percentage corresponding to such contributions would remain constant thereafter. For unemployment, we used the national unemployment rates available from 2009 to 2014 from INE (2014b, 2015), and forecasts from 2015 to 2019 from the IMF (2015). Beyond 2019, we assumed a constant unemployment rate equal to the 2019 forecast.

Comparing the updated values to the forecasts used in the previous studies, wages are lower and both the unemployment rate and the wage taxes and social security payments are higher. This leads to a reduction of the cost of labour, contributing to an increase of the ENPV. In order to get a better picture of the impact of the unemployment rate on the results, we will also perform a scenario analysis, considering the cases in which the unemployment rate is not considered in the shadow wage rate, and also the case in which a different evolution is considered for this variable.

4. Adjustments to traffic forecasts

Original traffic forecasts were made for a given scenario of GDP and fuel prices. Since there were important changes in these variables, it was necessary to analyse how these changes would impact the forecasts. This analysis is particularly important since traffic is generally the most important variable determining the benefits in a CBA of a road infrastructure (Godinho & Dias, 2012), due to its impact on the most important components of the CBA.

In order to re-estimate traffic volume, we took into consideration the traffic estimates used by Reis *et al.* (2008) and KPMG/VTM (2008) and also the original GDP values and fuel prices assumed by traffic consultants, as gathered from (KPMG/VTM, 2008). The traffic was then calculated based on the updated GDP and fuel costs (as described in the previous section). We also used estimates of traffic elasticity in response to the GDP and fuel prices, as reported by Silva and Caetano (2013). These authors estimate different elasticities for different NUTS III national regions and, for the Douro Interior, they report traffic elasticities of 0,865 in response to GDP and -0,587 in response to fuel prices.

The estimation of traffic changes was made separately for each type of fuel and for the light and heavy vehicles. This means that we ended up having also a new distribution of vehicles in the infrastructure.

Taking into account the revised GDP and fuel price values, as well as the previous elasticities, we obtained a reduction in the estimated traffic. This will contribute to a substantial decrease in the ENPV.

5. Results and analysis

After making all the adjustments to the model used in previous studies, and taking the methodology of Reis *et al.* (2008) as the main reference, we will now analyse the impact of the changes in the CBA results. We will analyse this impact for each CBA component, and compare our results with those of KPMG/VTM (2008) and Reis *et al* (2008), as shown in Table 1.

Similarly to the previous studies, we used a discount rate of 4% (defined by *Despacho nº13 208/2003, 25 June, Ministry of Finance*). We can argue that, at least at the peak of the crisis, a larger rate would be justifiable. However, this rate is defined by the Government and remained unchanged for the whole period. We obtain a negative ENPV: -43 million euros. The Economic Rate of Return (ERR), the rate that produces a zero value for the ENPV, is 3,5%, less than half the value obtained by KPMG/VTM (2008) and well below the value of 6,3% obtained by Reis *et al* (2008).

Table 1: Comparison of the results obtained with the updated macro-economic values with those of the previous studies.

Unit: million euros

Components of ENPV	**(KPMG/VTM, 2008)**	**(Reis *et al.*, 2008)**	**Updated values**
Travel time savings	470,5	590,2	400,6
Vehicle operating costs	23,5	1,0	0,1
Accident costs	112,2	159,0	99,7
Environmental externalities	-1,0	-3,2	-2,2
Benefits related to new users	126,2	34,3	22,5
Wider economic benefits	NC	45,9	31,4
Construction, maintenance and operation costs	-531,5	-688,6	-680,2
Residual Value	51,1	86,1	85,4
NPV of economic benefits	**731,3**	**913,4**	**637,4**
NPV of economic costs	**480,4**	**688,6**	**680,2**
ENPV	**260,9**	**224,9**	**-42,8**
Economic Rate of Return **(ERR)**	**7,2%**	**6,3%**	**3,5%**
Benefit/Cost Ratio **(B/C ratio)**	**1,47**	**1,33**	**0,94**
Legend	NC: Not taken into account		

The reduction of GDP plays an important role in the reduction of ENPV, since the crisis led to GDP values that are lower than expected in 2008 and GDP has an impact on several components of the CBA. The reduction of estimated traffic due to the GDP reduction and to the increase in fuel prices (in several years, traffic was reduced by more than 20%), also had a clear impact on the reduction of the ENPV.

Separately analysing the different components of the CBA, and comparing the values now obtained with those of Reis *et al* (2008) (which has a methodology closer to ours), we notice very significant decreases in the net benefits associated to travel time savings and accident costs and also some important decreases in the benefits related to new users and in the wider economic benefits, induced by the decrease in traffic. In the case of accident costs we must notice that, since most of the traffic is base traffic, and a decrease in such traffic occurs both with and without the infrastructure, it should be expected that such a decrease would also lower the benefits associated with a reduction in accident costs.

The reduction of traffic also leads to a very small decrease in the net costs of environmental externalities, but the impact is only marginal. The changes related to labour market, namely the reduction in wages and the increase in unemployment, lead to a reduction in construction, maintenance and operation costs, but the impact is small.

We must also notice that, although the two studies made in 2008 have performed a sensitivity analysis on the variables that were considered critical, they did not foresee a negative ENPV as calculated in the present analysis. The macro-economic impacts of the crisis that occurred in Portugal were too deep to be captured by the analysis made at the time.

We conclude that changes in macro-economic variables have as net effect a very significant reduction in the ENPV, which becomes negative - therefore, the change in the macro-economic situation is enough to turn an infrastructure that was expected to provide significant net economic and social benefits into an economic and social liability.

We must acknowledge that the incorporation of some updates in the analysis might have been made in different ways - for example, different assumptions might have been considered for the future values of fuel prices and unemployment rate, and different assumptions might have been considered for the impact of changing the average heavy vehicle load. In order to take this into account, we made a sensitivity analysis for the fuel prices, and considered different scenarios for the unemployment rate and heavy vehicle load.

In the case of fuel prices, we considered values 10% above and 10% below the base scenario of our analysis (Table 2). For the unemployment rate, we analysed the impact of excluding it from the analysis (considering full-employment as in the previous studies) and we also analysed the impact of assuming that from 2019 on it will decrease, at a rate of 10% a year, until it reaches 7% - the average value of the unemployment rate for the period of 2000-2010 (see Table 3). Finally, for heavy vehicle load, we considered an alternative scenario in which its value is the one assumed by Reis *et al.* (2008) (Table 4).

Table 2: Sensitivity analysis for fuel prices

CBA Components Unit: ENPV in million euros	**Base scenario**	**Fuel price** **+10%**	 **-10%**
Travel time savings	400,6	388,6	413,8
Vehicle operating costs	0,1	-0,8	1,0
Accident costs	99,7	96,7	103,1
Environmental externalities	-2,2	-2,2	-2,3
Benefits related to new users	22,5	22,6	22,6
Wider economic benefits	31,4	30,4	32,4
Residual Value	85,4	85,4	85,4
Total benefits	**637,4**	**620,7**	**655,8**
Construction, maintenance and operation costs	-680,2	-680,2	-680,2
Total Costs	**-680,2**	**-680,2**	**-680,2**
ENPV	**-42,8**	**-59,5**	**-24,3**
ERR	**3,5%**	**3,3%**	**3,7%**
B/C ratio	**0,94**	**0,91**	**0,96**

In the case of changes in fuel prices, there are some changes in the ENPV, but it always remains negative. In the cases of the unemployment rate and heavy vehicle load, the changes in ENPV are negligible.

We can reach the conclusion that the recent changes in the macro-economic situation in Portugal lead to an ENPV of the Douro Interior sub-concession that is negative. Godinho and Dias (2012) had already concluded that a significant positive ENPV, as provided by previous analysis, could not be enough to ensure that the best option would be to start building the infrastructure right away. These authors use the same data of Reis *et al.* (2008) and apply a model in which GDP and fuel prices follow stochastic processes estimated from historical data. The authors use Monte Carlo simulation and reach a sizeable expected ENPV of 232 million euros, in line with the studies made in 2008. Using real options analysis the authors conclude that, in spite of this large ENPV, the expected ENPV would be maximized if the decision to postpone the project was made, in order to resolve some uncertainty, instead of starting building right away. The authors only considered uncertainty in GDP and fuel prices - it is possible that the inclusion of other sources of uncertainty would reinforce this conclusion.

Table 3: Scenario analysis for unemployment rate

CBA Components Unit: ENPV in million euros	**Base scenario**	**Unemployment rate**	
		Decreases until it reaches 7%	**Excluded**
Travel time savings	400,6	400,6	400,6
Vehicle operating costs	0,1	0,1	0,1
Accident costs	99,7	99,7	99,7
Environmental externalities	-2,2	-2,2	-2,2
Benefits related to new users	22,5	22,5	22,5
Wider economic benefits	31,4	31,4	31,4
Residual Value	85,4	85,4	85,4
Total benefits	**637,4**	**637,4**	**637,4**
Construction, maintenance and operation costs	-680,2	-680,3	-680,8
Total Costs	**-680,2**	**-680,3**	**-680,8**
ENPV	**-42,8**	**-42,9**	**-43,5**
ERR	**3,5%**	**3,5%**	**3,5%**
B/C ratio	**0,94**	**0,94**	**0,94**

The results of this paper, and those of Godinho and Dias (2012), call into question the methodologies used in the evaluation of public projects, particularly those requiring sizeable investments. The identification and analysis of critical variables, as well as the use of Monte Carlo simulation, as prescribed by EC (2008), are clearly steps in the right direction. However, we must acknowledge that although real options theory has been widely researched in a corporate finance context, it has not been widely used for public projects evaluation. Real options theory has the potential to provide governments with the tools that lead to more robust decisions and, particularly, to the definition of contracts that contribute to maximize the social benefits while minimizing social risks. In the case of the Douro Interior sub-concession, a contract that structured the investment by stages, allowing the government to abandon later stages if the situation evolved differently from what was expected, might have minimized the risk associated with a downturn while allowing the construction to proceed if conditions were good.

Table 4: Scenario analysis for heavy vehicle load

CBA Components Unit: ENPV in million euros	**Base scenario**	**Heavy vehicle load** **Values of *Reis et al.* (2008)**
Travel time savings	400,6	397,7
Vehicle operating costs	0,1	0,1
Accident costs	99,7	99,7
Environmental externalities	-2,2	-2,2
Benefits related to new users	22,5	22,5
Wider economic benefits	31,4	31,1
Residual Value	85,4	85,4
Total benefits	**637,4**	**634,1**
Construction, maintenance and operation costs	-680,2	-680,2
Total Costs	**-680,2**	**-680,2**
ENPV	**-42,8**	**-46,0**
ERR	**3,5%**	**3,5%**
B/C ratio	**0,94**	**0,93**

6. Conclusions and future research

In this paper, we analysed the impact of a crisis in the economic evaluation of the Douro Interior sub-concession. In order to perform such analysis, we plugged the recent macro-economic data into the original studies performed in 2008. We conclude that the inclusion of these data would be enough to change the recommendation provided by the studies. So, if it would have been possible to forecast, in 2008, the current macro-economic situation, the decision would probably have been different.

Our results, and also the results of previous studies like Godinho and Dias (2012), call into question the methodologies used for assessing large public projects. We show that unexpected changes in macro-economic conditions may turn a sizeable positive ENPV into a negative one, even if the original analysis was based in the best forecasts of macro-economic conditions, as the ones published by reputed international institutions. Additionally, since economic forecasts usually are not available for all the period of analysis, it is advisable to adopt a conservative attitude concerning the assumptions regarding macro-economic conditions.

We believe it is time to give real options models a more relevant role in the public sector. Real options models may help governments not only make better decisions but also define contracts that mitigate public risk. We intend to use real options models, similarly to Godinho and Dias (2012), to analyse the impact of considering other options in this infrastructure - namely the abandonment option and the options that would be available if the investment would have been staged as series of outlays.

It will also be interesting to make a qualitative ex-post analysis of the regional impacts of the infrastructure. Particularly, it may be useful to assess if the new infrastructure contributed for positive economic effects in its region. The objective would be to find indicators capable of evaluating the improvement of conditions for residents and regional development.

A further doubt stems from the evaluations of other concessions that were decided at the same time of the Douro Interior infrastructure. Would the conclusions reached in this paper be similar in the case of those other concessions?

Acknowledgments: This research was supported by Project EMSURE—Energy and Mobility for Sustainable Regions (CENTRO-07-0224-FEDER-002004)

References

Afonso, I. T. (2013). Análise Custo Benefício de Projectos Rodoviários. In *Proceedings of 7º Congresso Rodoviário Português. n. 136*.

Ascendi. (2015). Ascendi - Subconcessão Douro Interior. *Ascendi*. Retrieved January 14, 2015, from http://www.ascendi.pt/gca/?id=102

Bickel, P., Friedrich, R., Burgess, A., Fagiani, P., Hunt, A., Jong, G. D., … Tavasszy, L. (2006). *Developing harmonised European approaches for transport costing and project assessment (HEATCO). Deliverable 5: Proposal for harmonised guidelines. Institut für Energiewissenschaft und Rationelle Energieanwendung*. Stuttgart, Germany.

Dft. (2005). *Transport, wider economic benefits and impacts on GDP*. London, UK.

DfT. (2009). *Value of Time Operating Costs - TAG Unit 3.5.6*. London, UK.

DGEG. (2015). Preços dos combustíveis líquidos. Retrieved January 7, 2015, from http://www.dgeg.pt/?cn=6891700270037129AAAAAAAA

EC. (2008). *Guide to cost-benefit analysis of investment projects: Structural funds, cohesion fund and instrument for pre-accession — Final report. Directorate General Regional Policy*. Brussels, Belgium.

Ferreira, S. (2013). *Parcerias Público-Privadas: A estimação da elasticidade preço da procura da A28*. Tese de Mestrado em Economia e Gestão das Cidades, Universidade do Porto.

Godinho, P., & Dias, J. (2012). Cost-Benefit Analysis and the Optimal Timing of Road Infrastructures. *Journal of Infrastructure Systems*, *18*(4), 261–269. doi:10.1061/(ASCE)IS.1943-555X.0000105

IMF. (2015). World Economic Outlook Database, October 2014. Retrieved January 8, 2015, from http://www.imf.org/external/pubs/ft/weo/2014/02/weodata/index.aspx

INE. (2013). Estatísticas dos transportes e comunicações 2012. (Instituto Nacional de Estatística, Ed.). Lisboa: INE.

INE. (2014a). Anuário Estatístico da Região Norte 2013. Lisboa.

INE. (2014b). *Anuário Estatístico de Portugal 2013*. Lisboa.

INE. (2014c). Estatísticas dos transportes e comunicações 2013. Lisboa.

INE. (2015). Estatísticas do Emprego - 4.º Trimestre de 2014. Lisboa.

KPMG. (2008). *Estudo Económico a 75 anos com Actualização após Adjudicação — Projecto Douro Interior*.

KPMG/VTM. (2008). *KPMG/VTM. (2008). Estudo de Impacto Económico Global—Subconcessão Douro Interior*.

OECD. (2014). OECD Factbook 2014: Economic, Environmental and Social Statistics. Paris: OECD Publishing.

Pereira, A. M., & Andraz, J. M. (2007). *Public investment in transportation infrastructures and industry performance in Portugal* (No. 45). *Journal of Economic Development* (Vol. 32). The Economic Research Institute Chung-Ang University.

Pereira, A. M., & Andraz, J. M. (2012). On the economic and budgetary effects of investments in SCUTS: the Portuguese toll-free highways. *The Annals of Regional Science*, *48*(1), 321–338. doi:10.1007/s00168-010-0404-6

PORDATA. (2014). PIB e PIB per capita a preços constantes (base=2011) em Portugal. Retrieved December 18, 2014, from http://www.pordata.pt/Portugal/PIB+e+PIB+per+capita+a+precos+constantes+(base+2011)-933

PORDATA. (2015). Salário médio mensal dos trabalhadores por conta de outrem da Construção: remuneração base e ganho por sexo - Portugal. Retrieved January 2, 2015, from http://www.pordata.pt/Portugal/Salario+medio+mensal+dos+trabalhadores+por+conta+de+outrem+da+Construcao+remuneracao+base+e+ganho+por+sexo-899

Reis, J., Godinho, P., Barata, E., & Cruz, L. (2008). *Estudo Integrado dos Impactes Económicos Globais associados às Concessões da AE Transmontana, Túnel do Marão e Douro Interior* (pp. 1–45). Coimbra.

Rosmaninho, G. P. (2010). *A regulação dos operadores de infra-estruturas rodoviárias*. Mestrado em Engenharia Civil, Universidade Técnica de Lisboa.

Silva, J. de A. e, & Caetano, L. (2013). Determinantes da Evolução do Tráfego Rodoviário. Análise da sua Evolução em Portugal Continental. In *Proceeding of 7º Congresso Rodoviário Português. n. 128*.

Preference Based Decisions in Health

Patrícia Antunes[1,4], Pedro Lopes Ferreira[2,4] and Lara Noronha Ferreira[3,4]

Abstract

The decision making process within a context of limited resources has important consequences on the health of individuals. The criteria by which health technology assessment and health interventions are decided, as well as the review of current procedures represent a challenge for decision makers in the face of the current economic slowdown. The decision making process requires a transparent and rational approach to ensure the maximization of health care and the reduction of inequalities. Economic evaluation techniques may provide a valid solution to achieve these goals. Preference based measures are part of economic evaluation techniques. These approaches incorporate utilities for health outcomes and can be used in cost-benefit analyses to aid resource allocation decisions. They may also meet the needs of citizens. However, the methodological challenges regarding particular situations and the preference based measures currently in use require further research in order to improve cost-benefit analysis and explore alternative preference elicitation methods for decision making. The comparative studies concerning the application of different techniques to obtain utility values show that techniques such as Discrete Choice Experiment are promising with regard to the use of the more traditional Visual Analogue Scale, the Time Trade-Off and the Standard Gamble. However, this technique still presents challenges and requires improvements. Further applications of this technique along with the

[1] Regional Health Authority, Coimbra

[2] Faculty of Economics, University of Coimbra

[3] University of the Algarve-ESGHT

[4] Centre for Health Studies & Research, University of Coimbra
E-mails: patriciaferreira@mail.telepac.pt, pedrof@fe.uc.pt, lnferrei@ualg.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_9

implementation of hybrid methods are needed to support the decision making process. This paper provides a brief description of each of these techniques and illustrates some of their applications.

Keywords: Preference based decisions, Standard Gamble, Time Trade-Off, Discrete Choice Experiment, Economic Evaluation.

1. Introduction

In health policies as in other areas of the economy, decisions are based on the existence of limited resources. In any of the public, private or social components of a health system these decisions have clear consequences on the health of populations.

Economic evaluation is a process to prioritize and decide on efficient allocation of human, financial and technological resources. Its main purpose is to support the decision makers on the choice of efficient and equitable options, when comparing costs and consequences of alternative programmes and health interventions (Brazier *et al.*, 2007; Drummond *et al.*, 2005). The decision making process is rather complex as it is not limited to the decision on different alternatives according to the health gains originated. In this process multiple criteria is taken into account such as social equity, the quality of patient experience, impacts on the wider economy and the quality of evidence upon which to base a decision (Devlin & Sussex, 2011).

The process of choosing among technologies and among health interventions has contributed to a greater efficiency in the allocation of resources (Garrido *et al.*, 2008). The multiple challenges that public health systems embrace, such as the maximization of health, the reduction of inequalities, the current economic context and the substantial pressure on state budgets, assume particular complexity to the decision-makers regarding the choice of new technologies or health programs and on the appraisal of existing procedures. Baltussen and Niessen (2006) state that problems of this nature are handled with some difficulty by the decision makers, using heuristic or intuitive approaches to reduce its complexity. Decision makers may also be influenced by individual motivations and political pressures that will constitute a disturbing factor. The risk of arbitrariness is therefore real. The variability of the decision in seemingly similar situations may sometimes be explained by contextual issues such as personal values, different risk perceptions, different priorities and variable budgets. For this reason, the choice of the most suitable method to decide on the technology, program or intervention in health is of the utmost importance, in particular with regard to the independence, transparency and impartiality of the definition and measurement of value.

The purpose of economic evaluation is to identify, measure, assess and compare costs and consequences when choosing various alternatives that lead to different use of resources (Brazier *et al.*, 2007). When looking for an informed decision, Drummond *et al.* (2005) emphasize the need to perform a systematic analysis for a clear identification of all the relevant alternatives and the importance of knowing the different perspectives of the stakeholders involved as well as determining the associated opportunity costs, comparing them with the consequences obtained in each alternative. Devlin and Sussex (2011) argue that in theory if it was possible to determine the cost-effectiveness of all health services and technologies, it would be possible to maximize the benefit of the population constrained to a specific budget. Regardless of the scenario involved it is necessary to determine the cost-effectiveness analysis with the greatest accuracy possible and accomplish the preferences of stakeholders. Given the methodological challenges in assessing particular situations (children, elderly, specific diseases) it is necessary to deepen the knowledge and improve the various existing techniques for the determination of cost-effectiveness as well as to explore alternative techniques of elicitation of preferences to support decision making. In this sense, it is intended to present and discuss the most common techniques of elicitation of preferences in health.

2. Preference based measures in health

The decision on allocation of health resources can be hampered if just considering the physiological impact of an intervention. This limited approach contradicts the comprehensive concept of health, adopted by the World Health Organization since 1946, which defines health as a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (World Health Organization, 1946).

The Quality-Adjusted Life Year (QALY) is an indicator based on preferences for support decision making in public institutions when applying techniques of economic evaluation. This indicator allows us to optimise the available resources and may be used in different situations (e.g., in different diseases) and in individual and population scopes (Kobelt, 2013; Rowen *et al.*, 2014). QALY represents a real innovation in health economics and decision making process, when combining the quality of life with length of life (Brazier *et al.*, 2007). Quality of life is measured on a full health-dead 1-0 scale, where one equals full health and zero is equal to being dead, with negative values considered to be worse than death. In practice for any individual the prospect of living Y years in less than full health may be equated to a prospect of living X years in full health where $X<Y$ (Brazier *et al.*, 2007).

The instruments used for obtaining health state values have two components: a descriptive system for describing health or its impact on quality of life and an algorithm for assigning values to each state described by the system (Tsevat *et al.*, 1994; Brazier *et al.*, 2007). While the measurement of health status objectively assess physiological results or inquire the individual about their functional capacity and limitations, the measurement of value (value, preferences, utility) asks the individual in relation to the value he assigns to a specific health condition, assuming, for example, that individuals with similar limitations may score them distinctly. Drummond *et al.* (2005) explain that although the terms “value”, “preferences” and “utility” are sometimes used indistinctly, there are differences in their meaning: the term “preference” applies to the broader concept and the terms “value” and “utility” fall as types of preferences.

There are several instruments or generic preference based measures of health including the Quality of Well-Being or QWB, the 15-D, the Health Utilities Index or HUI, the EQ-5D and the SF-6D. The QWB, developed in late 1970, is the oldest of the QALY instruments (Kaplan *et al.*, 1976; Brazier *et al.*, 2007). This instrument has three multilevel dimensions relating to function (mobility, physical activity and social activity) and a list of symptom and problem complexes. Weights have been estimated from a sample of health states using a Visual Analogue Scale. The aim is to produce a value (and utilities) on a scale of preferences that ranges between zero and one, or death and full health (Drummond *et al.*, 2005; Seiber *et al.*, 2008). The 15-D is a generic instrument, consisting of 15 dimensions that cover most of the dimensions of a preference based measure. The valuation of 15-D has been assessed through the combined use of rating scales and magnitude estimation (Sintonen, 1994, 1995). The HUI is a family of generic health measurement instruments: HUI1, HUI2 and HUI3, being HUI3 the most widely used. The first one was developed for use in an economic evaluation of neonatal intensive care, the second one is now used as a generic based-preference for children and the third one is for adults (Horsman *et al.*, 2003; Brazier *et al.*, 2007). HUI3 has eight dimensions and has been valued using the Standard Gamble technique (Horsman *et al.*, 2003; Drummond *et al.*, 2005). The EQ-5D is an instrument developed by a group of researchers (EuroQol Group) for use in economic evaluation. This index derives from an initial descriptive system with six attributes (mobility, self-care, usual activities, social relationships, pain/discomfort and anxiety/depression), subsequently revised to include only five attributes (social relationships were excluded). Each attribute has three levels – no problems, some problems and extreme problems – that will enable the definition of 243 ($=3^5$) possible health states. These are valued using a Visual Analogue Scale and the Time Trade-Off technique. The values will be situated on a scale that ranges between zero and one, or death and full health. Negative values may be assumed for health states considered worse than death (Drummond *et al.*, 2005; Ferreira, 2010; Ferreira *et al.*, 2013a). The Portuguese version of the EQ-5D was completed in 1998 and published in 2013, by Ferreira and his collaborators (Ferreira *et al.*, 2013a). The SF-6D was developed from the SF-36 instrument and has six dimensions (physical functioning, role limitation, social functioning, pain, mental health and vitality), each with four to six levels, allowing the generation of 18.000 distinct health states (Brazier *et al.*, 2002). In any analysis, the use of the SF-6D cannot be made independently, and must be preceded by the implementation of the SF-36 questionnaire (Ferreira *et al.*, 2013b). The health states obtained with the SF-6D are valued using the Standard Gamble technique.

Figure 1 schematically represents the process of measuring value of a specific health condition.

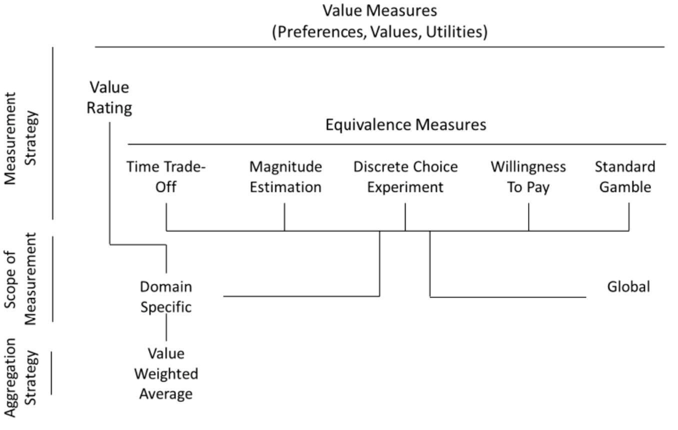


Figure 1: Ascertain a value for a health state[5]

Tsevat *et al.* (1994) point to the advantage of measuring value instead of measuring a health state being the former a method that best fits the commitments of the decision making process.

In 1944, von Neumann and Morgenstern during their research on Theory of Games developed a rational theory on decision making in a context of uncertainty, which became known as the expected utility theory or theory of von Neumann-Morgenstern utility. The normative model developed by these researchers determines the behavior of a rational

[5] Adapted from Tsevat *et al.* (1994). Using Health-related Quality-of-life Information. *Journal of General Internal Medicine, 9*(10), 576–582, Figure 1.

individual in decision making, under a context of uncertainty. This model considers that decision agents have preferences that dominate a set of possible choices, satisfying the axioms of transitivity, independence and continuity (Drummond *et al.*, 2005).

Generally speaking, the term "utility" can be understood as the more an individual prefers a certain result the more utility will be associated with it. However, this concept reveals itself complex when it attempts a more specific definition and, especially, when looking for its measurement (Drummond *et al.*, 2005).

The methods used in the process of measuring preferences vary in several aspects: the type of preferences (value or utility), the context of certainty/uncertainty, and the response of individuals (scaling or choice). Schematically, the techniques for measuring preferences can be represented as follows (table 1).

Table 1: Methods of measuring preferences[6]

Response method	Question framing	
	Certainty (values)	**Uncertainty (utilities)**
Scaling	Rating scale Category scaling Visual analogue scale Ratio scale	
Choice	Time trade-off Paired comparison Equivalence Person trade-off	Standard gamble Discrete choice

According to Brazier *et al.* (2007) different techniques can generate distinct health values. Thus, it is important to reflect on the advantages and disadvantages of techniques more commonly chosen to integrate the instruments for measuring quality of life. The three major techniques for valuing health states are the Visual Analogue Scale (VAS), the Standard Gamble (SG) and the Time Trade-Off (TTO). More recently a new technique

[6] Adapted from Drummond *et al.* (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 143, Table 6.1.

has been introduced: the Discrete Choice Experiment (DCE). However, which is the best technique?

The next section provides a brief description of each of these techniques.

3. Visual Analogue Scale

The VAS is a graduated ruler, usually drawn horizontally with 10 cm or vertically with 20 cm, with well defined end-points, on which respondents are able to indicate their judgement about a health state. The distances between intervals on a VAS represent the same difference and correspond to the difference in terms of perceived preference by individuals (Brazier *et al.*, 2007). The VAS has been used in economic evaluation for three decades and is the most common technique taking part in the QWB, HUI, 15-D and EQ-5D instruments for valuing health states. This technique can be used to assess chronic states considered better than death, states worse than death and temporary health states. In order to use VAS in economic evaluation it is necessary to ensure comparability among respondents. This requires clear and unambiguous end-points (full health, death) and a clear definition of the concept of full health in order to minimize the risk of different interpretation among respondents and even researchers (Brazier *et al.*, 2007).

To use this technique it is desirable to ensure that health states valuation can be placed on a zero to one scale, where zero is for states equivalent to death and one represents a state of full health. However, it is also necessary to allow for states valued worse than being dead. After obtaining a value for death and based on the interval properties of the scale, then all health states are calculated according to the following formula:

$$A_i = \frac{R_i - R\,(dead)}{R\,(full\,health) - R\,(dead)}$$

where A_i represents the adjusted VAS rating for health state h_i; R $(dead)$ is the raw rate given to being dead; R_i is the raw rate given to health state h_i and R $(full\ health)$ represents the rate attributed to the full health state. The value of A_i would lie between one (full health state) and zero (death) although it may assume negative values for states valued as worse than being dead (Brazier *et al.*, 2007).

Preferences for temporary health states may be measured using a VAS as long as respondents are fully acknowledged that the health states will last for a specific period of time (weeks, months or years; a period of time less than life expectancy) after which the person will return to full health (Brazier *et al.*, 2007).

4. Standard Gamble

The SG offers the respondent two alternatives, a certain intermediate outcome (state *i*) and the uncertainty of a gamble with two possible outcomes, one of which is better than the certain intermediate outcome - full health - and the other is worse - death (Torrance, 1986). This technique is used for measuring temporary health states, chronic health states and chronic health states better or worse than death. The method changes its format according to the situation and for a chronic health state preferred to death can be illustrated as shown in Figure 2.

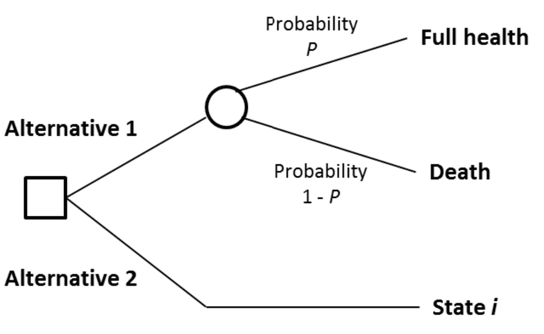


Figure 2: Illustration of a standard gamble [7]

[7] Adapted from Drummond *et al.* (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 150, Figure 6.2.

Alternative 1 offers two possible outcomes, with distinct probabilities: either the individual is returned to normal health and lives for an additional *t* years (probability *P*), or the individual dies immediately (probability 1-*P*). Alternative 2 has the certain outcome of chronic state h_i for life (*t* years). In this game, the probability *P* of the best outcome is varied until the individual is indifferent between the certain intermediate outcome (alternative 2) and the gamble (alternative 1). In this case, the value of health state *i* (h_i) is equal to the *P* value measured on the utility scale, in which full health for *t* years corresponds to one and immediate death to zero (Ferreira, 2003; Drummond *et al.*, 2005). The technique used for the variation of value *P* is called ping-pong, since lower values are toggled with higher values until reaching indifference.

5. Time Trade-Off

The TTO is a less complex technique developed as an alternative to the SG. While the SG assumes a choice between a certain outcome and a game based on uncertainty, the TTO requires a simple choice between two certain alternatives. It is a less complex technique because respondents are not required to understand probabilities (Brazier *et al.*, 2007). The TTO is a technique that can be used in chronic health states considered better or worse than death and temporary health states (*mutatis mutandis* to the technique). Figure 3 illustrates the alternatives for a chronic health condition preferred to death.

For a chronic health state preferred to death, alternative 1 involves living in a less than full health state (h_i) for a period *t* (e.g., life expectancy of a particular chronic condition), followed by death. Alternative 2 involves full health for a period *x* where $x<t$, followed by death. In this technique, time *x* will be varied until the individual is indifferent between the two alternatives. In this case, the value of the health state *i* (h_i) is equal to x/t (Ferreira, 2003; Drummond *et al.*, 2005).

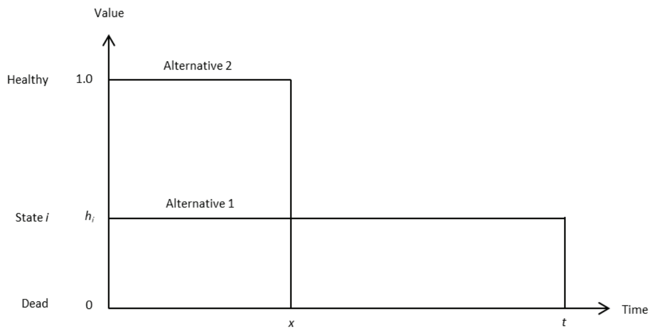


Figure 3: Time trade-off for a chronic health state preferred to death[8]

6. Limitations of the most common techniques

As mentioned before, VAS, SG and TTO are the standard techniques for eliciting preferences in health. However, there has been an intense debate about which is the best technique. Although the TTO and SG are recognizably superior techniques when compared to VAS, since they are based on choices that involve sacrifice, the values obtained from each of these techniques generate little consensus (Rowen *et al.*, 2014). Van Osch *et al.* (2004) reported that the utility values obtained by these techniques can be distorted by biases due to loss aversion (TTO, SG), scale compatibility (TTO), utility curvature for life duration (TTO) and probability weighting (SG). These concerns were also shared by other authors (Brazier *et al.*, 2012). Generally speaking, the cumulative effect of these limitations on obtaining utility values is not known when using TTO. It is also assumed that the utility values are higher when obtained through the SG. Even though props and visual aids have been developed to address literacy concerns, the TTO and SG have been criticized for their complexity towards the respondents and also because they exclude particular groups such as children, the elderly or different cultures (Brazier

[8] Adapted from Drummond *et al.* (2005). *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press, 152, Figure 6.4.

et al., 2012; Rowen *et al.*, 2014). On the other hand, VAS is a technique that is not strictly based on preferences as it does not imply a choice between options (Brazier *et al.* 2007).

The limitations of these techniques have led to a growing interest in others such as the Magnitude Estimation, Person Trade-Off and Discrete Choice Experiment (DCE). Although these techniques are less frequent, DCE has been getting increasingly featured in studies when eliciting preferences in health.

7. The discrete choice experiment for elicitation of preferences

DCE is a method for eliciting preferences that has been used since 1960. It is based on mathematical psychology and was originally applied in the marketing area. Since then, DCE has been used for eliciting preferences of individuals in other academic areas such as transportation and environmental economics and more recently in health (Earl & Kemp, 1999; Kenny *et al.*, 2003; Lancsar & Louviere, 2008; Norman *et al.*, 2013). The use of this method of elicitation is particularly relevant because preferences play a key role in priority setting and resource allocation in health (Lancsar & Louviere, 2008). The use of DCE is also justified by its simplicity of implementation and for being a method easier to understand (Kjær, 2005).

Discrete choice occurs when an individual is faced with a choice between a finite set of alternatives, mutually exclusive and that contemplate all the possibilities. In DCE the respondents have to choose one alternative out of a given number of alternatives (two or more). DCE is usually implemented with questionnaires that describe the good/service by a number of attributes and according to the most relevant attributes, the respondent associates a utility.

The purpose of this method is to observe the response of the individual to the change of these attributes throughout levels to which they are assigned. Each combination, built with the various attributes and associated levels, will constitute a scenario (*i.e.* a different package of good/service) whose ultimate aim is to allow the individual to feel able to make exchanges between scenarios according to what matters most to him/her.

Thus, it becomes possible to identify the level that influences the choice of each attribute, i.e., estimating the marginal rate of substitution of the attributes (Kjær, 2005; Lancsar & Louviere, 2008). The cost attribute (*e.g.*, transportation costs, wages and taxes) plays an extremely important role in DCE as it allows to elicit preferences concerning the willingness to pay. This does not mean that the respondents are directly asked about their willingness to pay. Instead they are asked to carry out monetary exchanges for improvements in attributes (Kjær, 2005). Similarly, the life expectancy is also a way of measuring the value that is attributed to a health state.

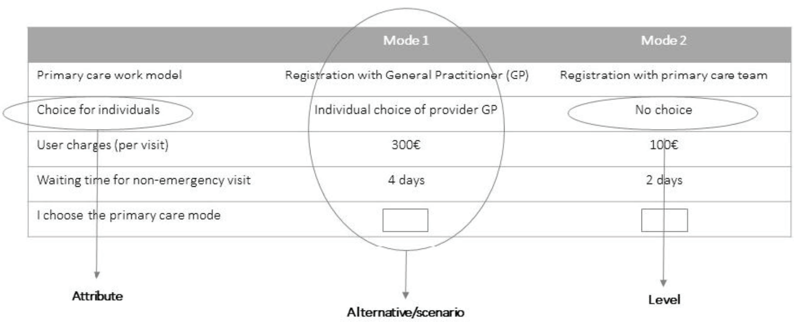


Figure 4: A choice set[9]

In order to analyse the results of DCE, each respondent choices converges into a single utility value, assigning weights to each attribute as a linear combination (Earl & Kemp, 1999; Kjær, 2005).

Choices made in the DCE are analysed using random utility theory, i.e., for an individual i conditioned on choice j, the utility can be decomposed into an explainable component V_{ij} and a non-explainable or random component ε_{ij}:

$U_{ij} = V_{ij} + \varepsilon_{ij}, j=1,\ldots,J$

[9] Adapted from Hjelmgren & Anell (2007). Population preferences and choice of primary care models: A discrete choice experiment in Sweden. *Health Policy, 83,* 314–322.

The non-explainable component ε_{ij} may be due to unobservable or unobserved attributes, unobserved variations in preferences, specification error and measurement error.

It is assumed that an individual chooses a particular option 1 if, and only if, its utility is higher than the utility of any other option on the set of J alternatives. Whereas Y_i is a random variable that denotes the choice outcome and assuming a joint probability distribution for ε_i, the probability P that utility is maximized by choosing option 1 is given by:

$$P(Y_i{=}1) = P(U_{i1} > U_{ij})$$
$$= P(V_{i1} + \varepsilon_{i1} > V_{ij} + \varepsilon_{ij})$$
$$= P(V_{i1} - V_{ij} > \varepsilon_{ij} - \varepsilon_{i1}), \forall j \neq 1$$

To move from a probabilistic choice model to an econometric choice model, the observable or systematic component of utility must be specified:

$$V_{ij} = X'_{ij}\beta$$

where X'_{ij} is a vector of variables representing observed attributes of option j and β is a vector of preferences parameters to be estimated (McIntosh *et al.*, 2010). V_{ij} can also depend on the characteristics of individual decision makers (Z'_i – vector of characteristics of individual), interacted with attributes of the good or service (X'_{ij}) but this would be ignored to simplify the discussion:

$$V_{ij} = X'_{ij}\beta + Z'_i\gamma$$

It is not wise to apply DCE without fully understanding the method. Therefore, it is important to study more deeply the theory, the design of the method and the results that can be obtained in order to build a valid model for research studies (Lancsar & Louviere, 2008).

Each design of DCE takes into account the specific context of the ongoing study and divides the several distinct phases in order to simplify the use of the method. Several authors argue for similar designs yet with slight differences. However, there is some consensus to split the design of DCE in five main stages: i) identification of attributes; ii) identification of levels; iii) experimental design; iv) data collection; and v) analysis of data (Kjær, 2005; Lancsar & Louviere, 2008). For the successful implementation of the method it is crucial that researchers have a broad notion of what is being studied and that the respondents are fully aware of the good/service under review (Lancsar & Louviere, 2006).

The first stage of the design involves defining the attributes of interest. These must be comprehensive and measurable, and can be qualitative or quantitative. For the identification of the most appropriate attributes in the characterization of a good/service, there are several possible sources of information including the existing literature, focus groups, interviews with key persons (e.g., decision makers) and experts (Kjær, 2005; Lancsar & Louviere, 2008). The main objective in the definition and identification of attributes is that they are relevant to the decision makers and that, simultaneously, are meaningful for the majority of the respondents. There is no maximum limit on the number of attributes but some researchers, particularly in the area of health economics, suggest not to exceed twelve attributes (Lancsar & Louviere, 2008). In identifying the attributes there are several aspects to be taken into consideration in particular if these are irrelevant or not, i.e., to what extent their exclusion/inclusion interferes with the results obtained and if the attributes are, among themselves, mutually dependent or causality related. These characteristics are particularly important because they may influence individuals by inducing a decrease in utility or even the transfer of utility between attributes (Kjær, 2005).

The second stage involves assigning levels to the identified attributes. Thus, for each of the attributes identified in the previous step reasonable and relevant levels are assigned. The range of these levels should be such as to enable the respondents to carry out exchanges between the various combinations of attributes. The number of levels of an attribute influence the significance of the attribute (considering equal variation intervals). However, the greater the number and levels of attributes the more complex will be the design (Kjær, 2005; Lancsar & Louviere, 2008).

With regard to obtaining values for utility measuring instruments already validated, these two phases will be naturally overcome.

The experimental design stage involves the design of the hypothetical choice sets including the formation and pairing alternatives. The main objective is to create the DCE so that the number of alternatives is as small as possible without questioning the ability of the respondent to infer utility on all the alternatives presented (Kjær, 2005). When pairing the alternatives it is also important that the differences in attribute levels for each choice set are not multi-correlated. The design will allow the estimation of a matrix, resulting from all possible combinations of levels of the different attributes of alternatives (choice set). The number of alternatives significantly increases as the number of attributes and the number of associated levels increase:

$a=l^b$

where a represents the number of alternatives; l is the number of levels and b represents the number of attributes (Lancsar & Louviere, 2008). Full factorial design refers to a design in which all the possible alternatives are represented. However, the full factorial design can only be easily applied to small experiments, with a very limited number of attributes and levels. In studies involving a larger number of attributes and levels it is necessary to reduce the size of the design. In this case the fractional factorial design is the most suitable option. Fractional factorial design involves a selection or a subset maintaining the properties of the original full factorial design. Although some loss of statistical information is involved, the fractional factorial design allows the estimation of effects of interest as efficiently as possible.

In order to ensure the optimisation of the design four principles need to be simultaneously considered: level balance (the levels of an attribute occur with equal frequency in the design); orthogonality (the difference in the levels of each attribute varies independently over choice sets); minimal overlap (a level does not repeat itself in a choice set); and utility balance (the utilities of the alternatives within each choice set are approximately equal). Full factorial design and fractional factorial design can be obtained through manual classifications or preferably using specific software (Kjær, 2005; Lancsar & Louviere, 2008).

The next stage involves the collection of data. Studies on the design of DCE recommend the inclusion in the model of an additional alternative on the choice set, particularly in circumstances where it is realistic to consider that the good/service is not consumed. This alternative is considered a no choice and prevents the individual from being forced to choose something that is not important to him/her (Kjær, 2005). With the inclusion of this alternative the respondent may "opt out", choose "no", "not to participate" or keep the existing situation (*status quo*). In some complex models the limitation of this choice resides in the fact that the individual can choose the easiest answer, i.e., can choose to apply a heuristic to avoid making a choice deemed difficult. Another perspective on the inclusion of an additional alternative is the possibility that this reflects a situation of indifference to the individual (Kjær, 2005; Lancsar & Louviere, 2008). For the success of data collection it is essential to ensure that the respondent understands what is expected of him/her which requires a thorough explanation on the part of researchers with regard to the context and objectives of the study, the description of the attributes and the contextualization of each of the scenarios (choice sets). At this stage, researchers may also consider important to perform a validation using rationality tests. Therefore, they need to include additional choice sets to test the various axioms, such as completeness, transitivity and continuity (Lancsar & Louviere, 2006). Finally, it is also important to include questions regarding socio-demographic data (*e.g.*, age, gender, education, income level, occupation). The collection of data can be done using interviews, questionnaires, or a combination of these methods, and using the Internet.

With regard to the sample, it should reflect the population for which the results will be generalized and to whom the opportunity costs matter. The sample size will depend on the number of questions that each individual responds, the size of the population, the response rate that is expected to obtain in the study and the statistical power intended to obtain from the model (Kjær, 2005; Lancsar & Louviere, 2008).

The last stage involves data analysis and includes the choice of the probability model and the interpretation of data. Several models are available for the estimation of DCE whose choice depends, among other factors, on the design chosen for the method (Kjær, 2005). If the DCE happens to be dichotomous (yes or no answer for a given alternative) or present a choice set consisting of two alternatives, then binomial discrete choice models (logit or probit) are used. When the choice set is made up of three or more alternatives then it is suitable to use multinomial (logit or probit), mixed logit, nested logit, and heteroscedastic extreme-value models. By increasing the number of alternatives in the choice set, these models put a series of challenges that do not exist in the binary choice (Kjær, 2005; Ryan *et al.*, 2008). Generally, the estimation of discrete choice models is based on the maximum likelihood method, although other methods may also be considered (Kjær, 2005). While DCE values are estimated on an unobserved anchored latent scale, they can be anchored on the health utility scale (zero to one scale) by incorporating duration (time) as an attribute. This method is referred to DCE_{TTO} (Bansback *et al.*, 2014).

8. Discussion

There are many concerns regarding the standard methods used to provide utilities. First, different utility values are obtained with SG and TTO and this raises the question of which method to use. Also, when a respondent has experienced a certain outcome (*e.g.,* disabled person, person with a chronic disease or a surviving cancer patient), he/she will approach the valuation task differently than a person who has not experienced the consequences of a disease or treatment. It is not unusual for a patient to refuse to sacrifice life-expectancy in order to be relieved of his/her health problem.This may be due to the time frame being longer than the time he/she expects to live. The use of a valuation task that offers an alternative like “full heath” to a person suffering from a disease also raises main ethical concerns. This would not be an issue if the task was to be applied for the general population. Nevertheless, the potential limitations of standard utility measurement techniques in particular groups need to be carefully considered.

Although DCE is not free from bias either, this problem can be minimized by the experiment design. Another main advantage on the use of DCE is the ability to control the experiment. The possibility of constructing a DCE offers the opportunity to focus to the precise issue of interest, like a group of persons suffering from a specific disease. This also allows for a range of attributes and levels to be valued within the same survey and therefore provides a richer set of information. Several other benefits are identified in this technique: ease of understanding (mainly by respondents), wider application (great number of people) and the fact that preferences are not influenced by time. Several studies have also begun to use DCE in particular groups, such as children and the elderly (Rowen *et al.*, 2014). Still, there are many differences between these valuation techniques and Table 2 resumes the main ones.

Table 2: Comparison between VAS, TTO, SG and DCE

Technique	Summary of differences
VAS	• Choiceless method
TTO	• Utilities biased upwards by loss aversion and scale compatibility • Tasks involved are too complex for certain populations • Exclude particular groups • Typically conducted with an interviewer (more time and cost consuming)
SG	• Utilities biased upwards by probability weighting and loss aversion. • SG utilities are systematically higher than TTO utilities • Tasks involved are too complex for certain populations • Exclude particular groups • Typically conducted with an interviewer (more time and cost consuming)
DCE	• Allow a range of attributes and levels to be valued within the same survey (provide a richer set of information) • Task cognitively simple • Typically conducted without an interviewer (less time and cost consuming) • Wider application

The comparison evidences that techniques such as DCE are promising with regard to the use of the TTO and SG (Braziet *et al.*, 2012; Krabbe *et al.*, 2014). But, apart from all the benefits identified DCE still presents challenges regarding the anchorage on the health utility scale with one for full health and zero for dead, which require further analysis (Rowen *et al.*, 2014).

The future will pass by carrying out additional studies using this technique and for assessing the implementation of hybrid methods (Bansback *et al.*, 2014).

9. Conclusion

In a context of scarce resources and particularly in health, the decision making process has important consequences on individuals. Whether in choosing new technologies or health programs, whether in the re-evaluation of existing procedures, the decision making process is a real challenge under the current economic context. The concern about the use of rational, exempt, replicable and transparent approaches clearly stands in the decision making process. On the other hand, to meet the concerns of citizens it is important to make a rigorous use of preference based measurement instruments in health and economic appraisal in particular. The review of currently available techniques and the conduction of additional studies with more promising methods are essential in supporting the decision making process.

References

Baltussen, R. & Niessen, L. (2006). Priority setting of health interventions: the need for multi-criteria decision analysis. Cost Effectiveness and Resource Allocation, 4, 14.

Bansback, N., Hole, A. R., Mulhern, B. & Tsuchiya, A. (2014). Testing a discrete choice experiment including duration to value health states for large descriptive systems: addressing design and sampling issues. Social Science & Medicine, 114, 38-48

Brazier, J., Roberts, J., Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics, 21(2), 271-292.

Brazier, J., Ratcliffe, J., Salomon, J. A. & Tsuchiya, A. (2007). Measuring and Valuing Health Benefits for Economic Evaluation. Oxford: Oxford University Press.

Brazier, J., Rowen, D., Yang, Y. & Tsuchiya, A. (2012). Comparison of health state utility values derived using time trade-off, rank and discrete choice data anchored on the full health-dead scale. European Journal of Health Economics, 21(2), 271-292.

Devlin, N. J. & Sussex, J. (2011). Incorporating Multiple Criteria in HTA: Methods and Processes. London: Office of Health Economics.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J. & Stoddart, G. L. (2005). Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press.

Earl, P. E. & Kemp, S. (1999). The Elgar companion to consumer research and economic psychology. Cheltenham; Northampton, MA, USA: Edward Elgar.

Ferreira, L. (2003). [Utilities, QALYs and quality of life measurement]. Revista Portuguesa de Saúde Pública, 3, 51-63.

Ferreira, P. (2010). [Measurement of health status and quality of life. In: 30 years of the National Health Service: a commented path] (461–484). Coimbra: Almedina.

Ferreira, P. L., Ferreira, L. N., Pereira, L. N. (2013a). [Contribution for the validation of the Portuguese version of EQ-5D]. Acta Médica Portuguesa, 26(6), 664–676.

Ferreira, L. N., Ferreira, P. L., Pereira, L. N., Rowen, D., Brazier, J. E. (2013b). Exploring the consistency of the SF-6D. Value in Health, 16(6), 1023–31.

Garrido, M. V., Kristensen, F. B., Nielsen, C. P. & Busse, R. (2008). Health technology assessment and health policy-making in Europe. Copenhagen: WHO European Observatory on Health Systems and Politics.

Hjelmgren, J. & Anell, A. (2007). Population preferences and choice of primary care models: A discrete choice experiment in Sweden. Health Policy, 83, 314–322.

Horsman, J., Furlong, W., Feeny, D., Torrance, G. (2003). The Health Utilities Index (HUI): concepts, measurement properties and applications. Health and Quality of Life Outcomes, 1, 54.

Kaplan, R. M., Bush, J. W., Berry, C. C. (1976). Health status: types of validity and the index of well-being. Health Services Research, 11(4), 478–507.

Kenny, P., Hall, J., Viney, R. & Haas, M. (2003). Do participants understand a stated preference health survey? A qualitative approach to assessing validity. International Journal of Technology Assessment in Health Care, 19(4), 664–81.

Kjær, T. (2005). A review of the discrete choice experiment - with emphasis on its application in health care. Health Economics Papers, 1.

Kobelt, G. (2013). Health Economics: An Introduction to Economic Evaluation (3rd ed.). London: Office of Health Economics.

Krabbe, P. F. M., Devlin, N. J., Stolk, E. A., Shah, K. K., Oppe, M., Van Hout, B., … Xie, F. (2014). Multinational evidence of the applicability and robustness of discrete choice modeling for deriving EQ-5D-5L health-states values. Medical Care, 52(11), 935-943

Lancsar, E. & Louviere, J. (2006). Deleting "irrational" responses from discrete choice experiments: a case of investigating or imposing preferences? Health Economics, 15(8), 797–811.

Lancsar, E. & Louviere, J. (2008). Conducting Discrete Choice Experiments to Inform Healthcare Decision Making A User's Guide. Pharmacoeconomics, 26(8), 661–677.

McIntosh, E., Clarke, P., Frew, E. & Louviere, J. (2010). Applied Methods of Cost-Benefit Analysis in Health Care. Oxford: Oxford University Press.

Norman, R., Viney, R., Brazier, J., Burgess, L., Cronin, P., King, M. & Street, D. (2013). Valuing SF-6D Health States Using a Discrete Choice Experiment. Medical Decision Making. [published online before print] doi:10.1177/0272989X13503499

Rowen, D., Brazier, J. & Van Hout, B. (2014). A comparison of methods for converting DCE values onto the full health-dead QALY scale. Medical Decision Making. [published online before print] doi: 10.1177/0272989X14559542

Ryan, M., Gerard, K., Amaya-Amaya, M. (2008). Using discrete choice experiments to value health and health care. Dordrecht: Springer.

Seiber, W. J., Groessl, E. J., David, K. M., Ganiats, T. G. & Kaplan, R. M. (2008). Quality of Well Being Self-Administered (QWB-SA) Scale User's Manual. San Diego.

Sintonen, H. (1994). The 15-D Measure of Health Related Quality of Life : Reliability, Validity and Sensitivity of its Health State Descriptive System. Melbourne: Centre for Health Program Evaluation.

Sintonen, H. (1995). The 15-D Measure of Health Related Quality of Life. II Feasibility, Reliability and Validity of its Valuation System. Melbourne: Centre for Health Program Evaluation.

Torrance, G. (1986). Measurement of Health State Utilities for Economic Appraisal. Journal of Health Economics, 5(1), 1-30.

Tsevat, J., Weeks, J. C., Guadagnoll, E., Tosteson, A., Mangione, C. M., Pliskin, J. S., Cleary, P. D. (1994). Using Health-related Quality-of-life Information. Journal of General Internal Medicine, 9(10), 576–582.

Van Osch, S. M. C., Wakker, P. P., Van den Hout, W. B., Stiggelbout, A. M. (2004). Correcting biases in standard gamble and time tradeoff utilities. Medical Decision Making, 24, 511-517.

World Health Organization. (1946). The Preamble of the Constitution of the World Health Organization. New York. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q =intitle: PREAMBLE+TO+THE+CONSTITUTION+OF+THE+WORLD+HEALTH+ORGANIZATION#0.

Assessing the effect of education on subjective well-being in Portugal: a study of mediating effects

Maria da Conceição Pereira[1,2] and Daniel Martins[1]

Abstract

Research on happiness has shown that individuals derive utility from material as well as from non-material factors, from which some useful policy implications have been derived. Since education has not received the attention we believe it deserves within this literature, we aim at contributing to fill in this gap by assessing the effects of education on subjective well-being, thus raising awareness for the importance of investing in education. Accordingly, in this study we conduct an analysis of the mechanisms that transmit the effect of education into subjective well-being, focusing on Portugal, to take into account country specificities, and using data from the European Social Survey. In order to test such mechanisms, we add to a baseline regression, which includes the education level, a large set of potential mediating variables to test whether education affects SWB through the following channels: 1. Higher lifetime earnings; 2. Higher professional status; 3. Less risk of unemployment; 4. Higher social capital, and 5. Better health.

The analysis shows that most of the considered variables contribute to carry the effects of education into subjective well-being. This is evidenced by a reduction of the coefficients of the education variables following the introduction of each mediator in the regression, thus confirming the hypothesized channels of transmission. Moreover, we find that education does not exert a direct effect on well-being, that secondary education provides a wider range of benefits than higher education, and that the human capital theory is not enough to account

[1] University of Coimbra, Faculty of Economics

[2] GEMF

E-mails: mpereira@fe.uc.pt, uc2010131267@student.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_10

for all the mechanisms transmitting the effect of education into subjective well--being in Portugal.

Keywords: Education, Subjective well-being, Happiness, Life satisfaction, Portugal

1. Introduction

Several authors have argued that results of research on the economics of happiness should be used to support public policy (e.g. Diener and Seligman, 2004; Frey, 2008; Helliwell, 2003; Layard, 2005). Some go as far as to call for a national system of well-being indicators. For instance, Diener and Seligman (2004: 2) wrote that such a system "would supplement and enhance [economic or other current social indicators'] value by placing them within an over-arching framework of well-being, underscoring the shortcomings of economic indicators. [...] The purpose of the production of goods and services and of policies in areas such as education, health, the environment, and welfare is to increase well-being. Therefore, well-being is the common desired outcome, and it follows directly that society should measure this outcome to provide a common metric for evaluating policies".

In fact, emphasis on well-being is important because there seems to be a pronounced gap between the information contained in traditional indicators of economic performance like GDP, and what matters for common people's well-being (Stiglitz *et al.*, 2009). The economics of happiness challenges the traditional economic thinking by relying on the principle that the best way to find out what is really important for individuals is to ask them, thus using survey-based indicators of subjective well-being (hereafter SWB). These indicators reveal peoples' affective (pleasant and unpleasant feelings) and cognitive (satisfaction with life) evaluations of their lives. In so doing, this research has been able to foster the debate on the determinants of individual's well-being, showing that individuals derive utility from material and non-material factors, including volunteering (e.g. Meier and Stutzer, 2006), social relations (e.g. Pugno, 2007; Powdthavee, 2008), religion (e.g. Clark and Lelkes, 2009), good governance (e.g. Frey and Stutzer, 2000) and the environment (e.g. Brereton *et al.*, 2008). These insights suggest a broader notion of utility than conventional economics, thereby raising awareness for the need to pursue various goals, beyond those typically related to economic performance, and reshaping individuals and policy makers' preferences towards welfare enhancing choices.

The economics of happiness research has been producing valuable contributions to help derive policy implications in several domains. For instance, Di Tella *et al.* (2001) have shown that a 1-percentage point increase in the unemployment rate is compensated for, in terms of well-being, by a 1.7-percentage-point decrease in the inflation rate, implying that the misery index wrongly ascribes the same importance to the two causes of economic discomfort. Since unemployment has a more devastating effect on SWB, this result suggests that, in the trade-off between unemployment and inflation, more emphasis should be put on unemployment reduction policies. Another interesting example of a policy insight obtained from happiness research results from finding that the costs of unemployment go well beyond the loss of income (Winkelmann and Winkelmann, 1998). This suggests that policies meant to help individuals find a job would provide advantages beyond the traditional unemployment benefit, which does not compensate unemployed individuals for the psychological costs of unemployment, such as the loss of self-esteem (cf. Frey *et al.*, 2002). Still another example, meaningful for our study, was provided by Oreopoulos (2005), showing how compulsory school laws increase lifetime wealth, health and happiness. Other findings with relevance for policy making are reviewed in Diener and Seligman (2004).

Education is one of the determinants usually included in happiness regressions[3]. However, there is no consensus about its net effect. This puts the economics of happiness in a fragile position in terms of policy inferences concerning investment in education. Some studies find a

---

[3] Dolan *et al.* (2008) and Diener *et al.* (1999) provide extended surveys concerning the determinants of SWB, including education. Others, like Frey and Stutzer (2002), Myers and Diener (1995) and Diener and Seligman (2004) do not cover education as a main determinant of SWB.

positive association between education and subjective well-being (e.g. Blanchflower and Oswald, 2004), whilst others find a negative relationship (e.g. Caporale *et al.*, 2009). Moreover, the sign and strength of this link is frequently associated with a country's stage of development (e.g. Veenhoven, 1994), or with characteristics of the population, like income (Diener *et al.* 1993), or job status (Clark and Oswald, 1994).

Moreover, most of these studies have given education a secondary role, and studied it amongst many other variables. So far education has not been given the central stage role it deserves, most likely because researchers accept that its effects are mainly transmitted through income (Layard, 2005), thereby diverting their attention to the latter. Some exceptions are Botha (2014), Chen (2012), Ferrante (2009), and Salinas *et al.* (2011), which focus on mediating effects in the relationship between education and SWB. Even so, these papers take partial looks at those mechanisms and miss on undertaking a systematic effort to dissect all possible mechanisms.

Our purpose is then to contribute to fill in this gap in the literature, by undertaking a systematic analysis of the role of education in influencing well-being. Therefore, in contrast to previous studies that focus on a single or reduced set of mediating variables, we investigate the potential mediating role of a large set of variables that may help carry education's effects into SWB. Following the literature, we will test whether education affects SWB through the following channels: 1. Higher lifetime earnings; 2. Higher professional status; 3. Less risk of unemployment; 4. Higher social capital, and 5. Better health. These indirect effects stem from individuals apprehending education as an investment good, in the sense that by putting in both material and non-material efforts into education in the present, returns will emerge in the future, which again can be of a material and non-material nature. Alternatively, education can be seen as a consumption good, thus being enjoyed for its intrinsic value. Therefore, we also investigate whether education has a direct effect on SWB.

Many governments around the world have upheld a political agenda that recognizes the benefits of education (Michalos, 2008). Portugal is one of the OECD countries that has made more progress in improving the baseline qualifications of its population. This has been crucial since Portugal has had one of the worst records in terms of attainment rates for secondary and tertiary education. In 2012, the proportion of working-age adults (25-64 year-olds) whose highest level of attainment was upper secondary education was 19%, which contrasts with the OECD average of 44%, and the proportion of working-age adults with a tertiary diploma was 19% compared to the OECD average of 33% (OECD, 2014). Nevertheless, Portugal has made an impressive progress in recent years. Portugal spent 5.9% of its GDP on education in 2009, compared to 4.9% in 1995 (OECD, 2012)[4], and is one of the OECD countries that have shown more progress in improving the level of education of its population. The proportion of adults that have not attained upper secondary education decreased from 81% in 2000 to 62% in 2012 (OECD, 2014). In 2012 the proportion of 25-34 year-olds with at least an upper secondary education was 58% and the proportion of 55-64 year-olds with a similar education level was only 20%; equally, the proportion of 25-34 year-olds with tertiary education was 28%, while the proportion of 55-64 year-olds with that same level was only 11% (OECD, 2014). This marked cross-generational difference leads to the expectation that Portuguese educational attainment will significantly improve over the years. Given the tough financial constraints that threaten to jeopardize these efforts and achievements, it becomes important to discuss which dividends education brings about, namely in terms of well-being, and in particular to inform the Portuguese authorities about the education's role.

We thus undertake this analysis for Portugal using data from the European Social Survey (ESS). Our analysis and empirical strategy can be extended /replicated to all of the countries included in the used data-set. Nevertheless, our focus on Portugal enables us to take into account country specificities. Given that the institutional context conditions the results, country-specific analysis are more suited to substantiate country specific policy recommendations.

---

[4] This figure decreased to 5.5% in 2011. See Santiago (2012) for a list of measures that lead to a reduction in resources available to education in Portugal due to the economic crisis.

The remainder of the paper is structured in a straightforward way. In section 2 we provide a brief survey of the literature focusing on the channels through which education exerts its effects on SWB; in section 3 we present the data and empirical strategy; in section 4 we present and discuss the estimation results; and in section 5 we conclude.

2. Previous work/research

Although education is seen as vital for societies, its role on well-being is not undisputed from an individual's point of view. A positive relationship between education and individual SWB has been documented by many authors (e.g. Blanchflower and Oswald, 2004; Botha, 2014; Chen, 2012; Delhey, 2004; Di Tella, 2003; Ferrer-i-Carbonell, 2005). Nonetheless, others find a negative effect of education on SWB (e.g. Caporale *et al.* 2009; Clark and Oswald, 1994, 1996). More education is commonly associated with longer working schedules, more stressful professions, higher dispersion of income and higher job expectations, which can upset the positive effects of education. More education is further responsible for distress in case of job loss (Clark and Oswald, 1994). Even the process of acquiring education is stressful and costly. Authors such as Veenhoven (2010) and Ferrante (2009) believe that beyond a certain threshold level, education may lose its significance, possibly because getting an education might involve costs that outweigh its benefits. Ferrante (2009) considers that real life opportunities commonly fall short of peoples' aspirations and finds that, beyond a certain level of education (secondary education for Italy), further education fuels socioeconomic aspirations (education rises people's earnings and job expectations, for example), in such a way as to depress SWB.

That education is more strongly related to well-being in poor countries is supported by Ferrer-i-Carbonell (2005), Veenhoven (1994) and Hartog and Oosterbeek (1997), and for individuals with lower income is evidenced by Diener *et al.* (1993).

Many of those who find a positive effect of education on SWB posit that such effect is exerted through indirect channels. From these different studies it is possible to identify mechanisms such as higher income, higher employability, rapid promotions, more secure jobs, social integration and increased health. For example Diener (1993) finds that once income is controlled for the effect of education becomes insignificant. Similarly, Helliwell (2002) found a small effect of education on SWB, after controlling for participation in social activities, health, trust in people, and higher income.

The more popular channels are grounded on the theory of human capital (Becker, 1964; Schultz, 1960). According to this theory, education adds to the stock of human capital, thereby promoting productivity. This leads to a higher income earning capacity and employability, which, in turn, leads to high-wage and high benefit jobs. Other than contributing to individuals' future returns by increasing their stock of capital, education may promote future earnings and employability by acting as a signal to the labour market, informing potential employers about workers' abilities, which are unobserved by them. This alternative idea has become known as signalling theory (Spence, 1973). In addition, this signal helps matching the most capable workers to the jobs that are more adequate to them, thereby promoting their well-being. Likewise, improved information and knowledge acquired through schooling increase labour market search efficiency (Haveman and Wolfe, 1984).

Education has played a central role in explaining growth. Within economic growth theory, the benefits of education are seen as deriving from an accumulation of human capital, which acts as a fundamental input to production and, thus, to output growth in the long run, as well as from a boost in total factor productivity, which operates as a driving force for growth (Teixeira, 2014). The increased labour market search efficiency, fostered by education, can also be seen as offering a growth driver. Matching workers to jobs where their productivity is the highest enhances their productivity, and this also serves the general interest of society, since a better allocation of resources leads to growth. Growth feeds back into the individual, since living in more developed environments boosts individual happiness (cf. Stevenson and Wolfers, 2008). For a long time many researchers and policy makers conformed to the impotence of public policy to foster happiness by promoting living standards, due to the observation that in many industrialized countries real per capita income, which is the main proxy for material well-being, has risen markedly since World War II, while average well-being has not. This finding has become known as the Easterlin paradox (Easterlin, 1974, 1995). Recently, however, a reassessment of the Easterlin paradox by Stevenson and Wolfers (2008) established a very strong relationship between SWB and growth.

The above perspectives support the role of education in facilitating the access to better quality jobs. These have multifaceted characteristics, in the sense that they are associated with higher wages, higher job status, and less vulnerability to unemployment, amongst others. Earnings' potential has been particularly emphasized, together with occupational status (e.g. Diener, 1999; Teixeira, 2014). In this context it is interesting to note that Portugal is the OECD country that exhibits the highest wage gain from attaining university-level degree (70% as compared to the OECD average of 59%, according to the latest data, 2012). Furthermore, there is a large body of evidence that a higher level of education leads to a lower risk of unemployment, from academic (e.g. Cuñado and de Gracia, 2012), as well as from international institutions. For instance, the OECD Education at a Glance (2013) reports that the unemployment rate in Portugal, between 2008 and 2011, among 25-64 year-olds without an upper secondary education, raised from 7.6% to 13.3% (the OECD average increased from 8.8% to 12.6%); while among 25-64 year-olds with tertiary education it increased from 5.8% to 8.0% (the OECD average increased from 3.3% to 4.8%).

Recently, attention has also been given to the role of social capital in channelling the benefits of education into well-being. Social capital has been defined as "networks together with shared norms, values and understandings that facilitate cooperation within or among groups" (Côté and Healy, 2001: 41). It is common to group the social resources that enable such a cooperation in the following dimensions: social networks, participation in social activities, involvement in civic activities and trust (Nieminen *et al.*, 2008). A vast number of studies has documented the effect of education on social capital. For example Helliwell and Putman (2007) go so far as to say that education is the most important predictor of some forms of social capital. The effect of social capital on well-being is also well known (e.g. Han *et al.*, 2013; Powdthavee, 2008; Pugno, 2007; Winkelmann, 2009). The indirect effect of education on SWB via social capital is dealt with by Helliwell (2002) and Chen (2012). Böhnke and Kohler (2008: 15) also recognizes that "educational skills are not only an (increasingly) important prerequisite for labour market integration, but also form the basis for social integration and participation in a modernized world".

For instance, through education individuals gain access to a vast network of people, which broadens their minds, thereby promoting well-being (cf. Chen, 2012). One of the indirect effects of education on SWB postulated by Helliwell (2002) flows through social connectedness, proxied by participation in social activities, namely church attendance, and membership in voluntary associations. According to Helliwell (2002), trust in people is another benefit that can carry the effects of education into SWB. Chen (2012) argues that education increases happiness by enhancing ones capacity to get involved with the world, not only family and neighbourhood, but the outside world as well, thereby becoming more open-minded about other cultures, thus broadening ones' horizons, which brings positive feelings. Chen (2012) focuses on specific forms of social capital, which are social networks (proxied by the number of acquaintances and contacts with family members during last New Year's vacation) and cosmopolitanism (captured by international travel experience and frequency with which one talks about international issues). He finds that these mediating factors are more important in explaining the relationship between education and SWB than income, in some Asian countries, except for China, for which income plays a more important role, supposedly due to China's relative low level of personal income.

The effect of education on health is also quite popular (e.g. Hartog and Oosterbeek, 1997; Oreopoulos, 2007). The rationale is that more educated individuals adopt more healthy lifestyles. In other words, they "are assumed to be more efficient producers of health; they have less unhealthy habits and visit their doctor when required" (Hartog and Oosterbeek, 1997: 2). In parallel, education contributes to the acquisition of information, and provides an advantage in the choice of less hazardous occupations and locations (Haveman and Wolfe, 1984). The effect of education on SWB through its effect on health is tackled by Gerdtham and Johannesson (2001).

In contrast to these indirect effects, education can provide utility per se. There are individuals who value the acquisition of knowledge independently of leveraging it into future benefits, and others obtain utility from attending school, as a form of entertainment (Lazear, 1977). While some studies do not find a direct effect of education on SWB (e.g. Helliwell, 2002), others detect such an effect (e.g. Blanchflower and Oswald, 2004; Cuñado and de Gracia, 2012; Salinas *et al.*, 2011). Cuñado and de Gracia (2012) interpret the effect of education on SWB, after controlling for income, labour status, and health, as evidence of a self-confidence or self-esteem effect of education. We note, however, that the differences across studies often derive from the authors' subjective appreciation of the size of the coefficients of the education variables, as well as from omitted factors.

Most of the studies on the benefits of education focus on marketed effects of schooling, but Haveman and Wolfe (1984) identify a vast number of non-marketed effects, such as better marriage, child quality, consumer choice efficiency, crime reduction, social cohesion, amongst others. Likewise, Diener (1999: 293) recognizes that "education may contribute to SWB by allowing individuals to make progress toward their goals or to adapt to changes in the world around them", and Botha (2014) point to the role of education in providing life skills to avoid public shame, in rising social-status, and prestige. Sabates and Hammond (2008) provide a good review of the impact of education on well-being, covering less popular effects, such as self-esteem, self-efficacy, and risk of depression.

Education is also viewed as a positional good (e.g. Botha, 2014; Helliwell and Putman, 2007; Salinas *et al.*, 2011), since it provides a differentiation device, thereby conferring social status and advantage in job competition. No matter whom people compare themselves with, be it region-wise (Helliwell and Putman, 2007), income-wise (Botha, 2014) or other, acquiring this differentiation provides a stimulus to individual's well-being.

3. Methodology / Data and empirical strategy

The dataset that we use to study the relationship between education and SWB is from the European Social Survey. This is a cross-national survey that has been conducted every two years in more than 30 European countries since 2001. As mentioned previously, only Portuguese data will be used. The study will consider the most recent three rounds, which refer to 2008, 2010 and 2012.

We adopt the following model for our baseline specification:

$$WB_i^{*}=\alpha+\beta \cdot EDU_i+\gamma \cdot X_i+\delta_t+\varepsilon_i$$

where WB_i^{*} is the well-being level reported by individual i, which acts as a proxy for the true individual well-being WB and is measured by the answer to the question "All things considered, how satisfied are you with your life as a whole nowadays?". Answers range from 0 to 10, where 0 means extremely dissatisfied and 10 means extremely satisfied[5]. EDU is education in the formal sense of the term, thus measured by the level of education attainment[6].

X is a vector of control variables, including gender (men being the reference category), age, and marital status[7], δ is a time effect (a dummy

[5] This measure was chosen because life satisfaction is less responsive to short-term circumstances (Helliwell, 2002) than happiness. However, experiments using the question "Taking all things together, how happy would you say you are?", to proxy well-being, converged to similar results.

[6] Seventeen categories were combined into three: basic education (reference category), comprising no education level, basic - levels 1, 2 and 3, and vocational qualifications diploma courses of level 2; secondary education, which also includes vocational qualifications diploma courses of level 3 and technological specialization courses; and higher education, including polytechnic and university programmes at the bachelor's level or equivalent, post graduate programmes, including masters and doctoral degree.

[7] The category "divorced" also includes individuals legally separated. The reference category is single.

variable for each round), and ε_i is the error term, also capturing non--available factors. Since our dependent variable is measured on an ordinal scale, the model is estimated as ordered probit, in which the true individual well-being is the latent variable.

Our empirical strategy consists of adding to this baseline specification factors expected to mediate the effect of education on SWB, and thus estimate the following model:

$$WB_i^{*}=\alpha+\beta \cdot EDU_i+\gamma \cdot X_i+\theta \cdot H_i+\delta_t+\varepsilon_i$$

where *H* may stand for any one of such mediating variables. We will test a considerable number of mediators and predict the following:

Hypothesis 1: Education enhances SWB by raising earnings.

Hypothesis 2: Education enhances SWB by improving job status.

Hypothesis 3: Education enhances SWB by reducing the risk of unemployment.

Hypothesis 4: Education enhances SWB by boosting social capital.

Hypothesis 5: Education enhances SWB through its positive effect on health.

Table A1 in the appendix presents a detailed description of the explanatory variables, arranged according to the hypothesis being tested[8]. Here we briefly summarize the proxies used for each mediator. Earnings is proxied by the individual's household total income, computed in equivalent terms. That is, an adjustment is undertaken to consider economies of scale in consumption within the household.

Job status, referring to the position in which a worker is placed in the professional ladder, will be proxied by a set of variables, including workers having a supervisory role, autonomy in organizing own daily work, and capacity to influence management decision.

As to the third hypothesis, both current and past unemployment will be tested. Current unemployment includes those who are both actively looking for a job and those who are not. Past unemployment is measured considering two time horizons, a short-term unemployment spell, referring

[8] Table A2 displays summary statistics of all variables and Tables A3 to A6 present correlation coefficients between the different proxies considered in hypothesis 2 to 5.

to a past experience of unemployment between 3 and 12 months long, and a long-term unemployment spell, referring to a past unemployment experience lasting 12 months or more.

As discussed in the previous section, social capital is a multifaceted concept, which does not have a universal measurement. Despite this, three aspects of social capital are commonly identified. Drawing on prior research, we will proxy such established aspects of social capital as follows. Social participation and social networks will be measured by the frequency of social meetings, by having or not someone to discuss intimate and personal matters, by social activity in general, and by religious activity. Civic engagement considers whether individuals voted in the last election, other civic participation (such as having worked in a political party), and perceptions of safety in the neighbourhood. Trust will focus in people and will encapsulate a perception not only of trust but of fairness and helpfulness as well, as detailed in the appendix.

The fifth hypothesis will be tested using self-reported health and disabling health problems.

By adding these potential mediators to estimations we expect the education coefficients to decrease. This will be evidence that part of the effect of education on SWB is mediated by such factors.

One last procedure will consist of analysing more comprehensive models. Some of these are meant to inspect whether there exists a direct effect of education on SWB, or whether the effect is exclusively exerted through indirect factors. In order to do that, we run a regression comprising variables from all hypothesis.

After deleting individuals that have missing values in any of the used variables we are left with 2785 observations.

4. Empirical Results

Tables 1 to 5 present the estimation results for different specifications, meant to test hypothesis 1 to 5, respectively. The results of the estimation of the baseline specification are repeatedly shown in column 1 of all tables for comparison purposes.

Model 1 shows that gender has no effect on life satisfaction in Portugal, which contrasts with most countries where women seems to be happier than men. A possible explanation is that Portuguese women have a higher participation in the labour market than in other developed countries, and this may deteriorate their well-being, especially around motherhood period. As in most studies, there is a U-shaped relationship between life satisfaction and age. Compared to being single, being married positively contributes to life satisfaction[9]; whereas being divorced or separated, as well as being widowed, negatively impacts on life satisfaction, results which conform to expectations. As to the education variables, we observe that they are both statistically significant, and in particular that individuals with a higher education level are the ones reaching higher satisfaction.

Model 2 reports the results of testing hypothesis 1. We accomplish this by regressing a model with H=income. Expectedly, income has a positive effect on SWB. Adding income to the baseline regression reduces the education variables coefficients in a remarkable way. What is even more striking is the reduction of the higher education coefficient, which suggests that the advantage from having a higher education comes exclusively from attaining a higher income. This is in accordance with the fact reported in section 2 that the net income gain from attaining a university-level degree is indeed very high in Portugal (the highest in OECD). Hence, we probably would not have found such a huge effect of income on the education role in influencing SWB in other OECD countries.

[9] In some of the coming models we note that the statistical significance of this variable is not upheld. This may be due to mediation effects, but we will not conjecture about it since it is not the focus of our study.

Table 1: Hypothesis 1

	Model 1	Model 2
Gender (man)	0.0221 [0.0405]	-0.0150 [0.0409]
Age	-0.0067*** [0.0014]	-0.0056*** [0.0014]
Age squared	0.0003*** [0.0000]	0.0003*** [0.000]
Married	0.1140* [0.0597]	0.0521 [0.0603]
Divorced	-0.2684*** [0.0832]	-0.2635*** [0.0832]
Widowed	-0.1824** [0.0823]	-0.1515* [0.0825]
Secondary Educ.	0.2511*** [0.0590]	0.1698*** [0.0600]
Higher Educ.	0.3534*** [0.0697]	0.1476* [0.0753]
Income		0.2386*** [0.0331]
Pseudo R^2	0.0194	0.0237

Estimations also include round dummies. Standard errors are shown in parenthesis. Single, double and triple asterisks indicate significant coefficients at the 10%, 5% and 1% levels, respectively.

Table 2 presents the estimation results for the models that include the variables regarding job status. Model (3) shows that individuals who are responsible for supervising the work of some employees report higher levels of satisfaction with life. Interestingly, a higher degree of responsibility in this respect seems to be neutral to life satisfaction. Higher responsibility in supervising workers brings more decision latitude, but expanding options bring about psychological decision costs, namely in gaining information to support decisions and in confronting one's own bad decisions (Ferrante, 2009). The negative feelings thus gained may cause stress and worries that neutralize any eventual positive effect that professional status could confer. Furthermore, the education coefficients decrease with respect to model 1, suggesting a mediation effect, in particular for higher education. Models (4) and (5) inform us that autonomy in deciding how individuals' daily work is organized and capacity to influence management decision exert a positive effect on well-being and decrease the education coefficients relative to model (1) – these effects are more pronounced than in model (3). We can see that, bringing together these three proxies, in model (6), does not lead to a much steeper decline in education coefficients, and only the highest degree of autonomy and the proxy "influence" keep its statistical significance. This may result, at least in part, from these variables capturing overlapping features of job status (Table A3, in the appendix, shows high correlation coefficients between autonomy and influence). Thus, we conclude that there is some evidence that, in Portugal, education promotes well-being by providing a higher job status.

Table 3 reports estimations that test education as a risk of unemployment reduction device. Model (7), which adds current unemployment to the baseline regression, confirms the usual result that unemployment is detrimental to life satisfaction. By observing a modest reduction in the education coefficients we also learn that education somehow boosts SWB by reducing the risk of unemployment. Models (8) and (9) consider past unemployment, of a short and long term nature, respectively. We can see that past unemployment experience only depresses life satisfaction scores and mediates the relationship between education and well-being in the case of a long unemployment spell. Model (10) combines current with past long unemployment to find out that they both seem relevant in explaining SWB, but they do not seem to carry different effects from education into SWB, since the drop in the coefficients of education is not much different from the ones observed in models (8) and (9) (this is probably related to the fact that these variables are correlated; Table A5, in the appendix, shows a correlation coefficient of 0.55 between current unemployment and long-term past unemployment). We believe that, overall, these results endorse hypothesis 3, although in a mild manner.

In Table 4, models (11) to (19), we include social capital to test whether education acts through it to improve SWB in Portugal. We can see that individuals more socially integrated score higher in satisfaction with life, no matter what dimension of social capital we test for. In addition, we see that the coefficients of the education variables slightly decrease in most models. In models (11) through (14) we test for the effect of social networks and participation, by using as proxy for this dimension of social capital the frequency of social meetings, having someone to discuss intimate and personal matters, social activity and religious activity, respectively. We find a significantly positive effect on SWB of all such proxies for social networks and participation. This dimension of social capital also seems to play a mediating role in the SWB-education relation, although attending religious services seems to be an exception. The strongest reduction in the education coefficients occurs when participation in social activities is introduced in model (13).

Table 2: Hypothesis 2

	Model 1	Model 3	Model 4	Model 5	Model 6
Gender (man)	0.0221 [0.0405]	0.0057 [0.0410	-0.0179 [0.0410]	-0.0106 [0.0408]	-0.0218 [0.022]
Age	-0.007*** [0.0014]	-0.0069*** [0.0014]	-0.0068*** [0.0014]	-0.0071*** [0.0014]	-0.0071*** [0.007]
Age squared	0.0003*** [0.0000]	0.0003*** [0.0000]	0.0003*** [0.0000]	0.0003*** [0.0000]	0.0003*** [0.000]
Married	0.1140* [0.0597]	0.1123* [0.0597]	0.1073* [0.0597]	0.1019* [0.0597]	0.1035* [0.060]
Divorced	-0.268*** [0.0832]	-0.266*** [0.0832]	-0.2796*** [0.0833]	-0.2854*** [0.0833]	-0.2828*** [0.083]
Widowed	-0.1824** [0.0823]	-0.1775** [0.0824]	-0.1738** [0.0824]	-0.1696** [0.0824]	-0.1670** [0.082]
Secondary Educ.	0.2511*** [0.0590]	0.2326*** [0.0594]	0.2084*** [0.0593]	0.2100*** [0.0592]	0.2055*** [0.060]
Higher Educ.	0.3534*** [0.0697]	0.3109*** [0.0714]	0.2682*** [0.0711]	0.2705*** [0.0710]	0.2418*** [0.072]
Workers sup. - some		0.1944*** [0.0703]			0.0906 [0.074]
Workers sup. - many		0.1114 [0.0981]			0.0084 [0.101]
Autonomy - low			0.2158*** [0.0606]		0.0350 [0.083]
Autonomy - medium			0.2975*** [0.0579]		0.0895 [0.080]
Autonomy - high			0.3745*** [0.0577]		0.1775** [0.083]
Influence - low				0.2515*** [0.0514]	0.2234*** [0.071]
Influence - medium				0.3225*** [0.0544]	0.2508*** [0.074]
Influence - high				0.3488*** [0.0568]	0.1983** [0.082]
Pseudo R^2	0.0194	0.0201	0.0231	0.0238	0.0245

Estimations also include round dummies. Standard errors are shown in parenthesis. Single, double and triple asterisks indicate significant coefficients at the 10%, 5% and 1% levels, respectively.

Table 3: Hypothesis 3

	Model 1	Model 7	Model 8	Model 9	Model 10
Gender (man)	0.0221 [0.041]	0.0130 [0.041]	0.0234 [0.041]	0.0149 [0.041]	0.0112 [0.041]
Age	-0.0067*** [0.001]	-0.0081*** [0.001]	-0.0070*** [0.001]	-0.0080*** [0.0014]	-0.0085*** [0.000]
Age squared	0.0003*** [0.000]	0.0002*** [0.000]	0.0003*** [0.000]	0.0002*** [0.000]	0.0002*** [0.000]
Married	0.1140* [0.060]	0.1034* [0.06]	0.1134* [0.060]	0.0959 [0.060]	0.0960 [0.060]
Divorced	-0.2684*** [0.083]	-0.2614*** [0.083]	-0.2656*** [0.083]	-0.2765*** [0.083]	-0.2675*** [0.083]
Widowed	-0.1824** [0.082]	-0.1993** [0.082]	-0.1827** [0.082]	-0.2040** [0.083]	-0.2071*** [0.083]
Secondary Education	0.2511*** [0.059]	0.2285*** [0.059]	0.2543*** [0.059]	0.2142*** [0.060]	0.2137*** [0.060]
Higher Education	0.3534*** [0.070]	0.3185*** [0.070]	0.3598*** [0.070]	0.3119*** [0.071]	0.3040*** [0.070]
Current unemployment		-0.3528*** [0.066]			-0.2697*** [0.079]
Past unempl. - short			-0.1098 [0.069]		
Past unempl. - long				-0.2610*** [0.058]	-0.1425** [0.069]
Pseudo R^2	0.0194	0.0218	0.0196	0.0211	0.0221

Estimations also include round dummies. Standard errors are shown in parenthesis. Single, double and triple asterisks indicate significant coefficients at the 10%, 5% and 1% levels, respectively.

Although this reduction in coefficients is not comparable with that observed in model (2), the increase in the pseudo-R^2 in model (13) is more expressive than in model (2), suggesting that adding social participation improves the explanatory power of the model more than income. A possible explanation for this is that education is already capturing the contribution of income to SWB, whereas social capital adds new dimensions of individuals' needs, and this agrees with the literature that finds a more significant role for post-materialistic concerns than for materialist concerns on SWB (e.g. Delhey, 2004; Böhnke and Kohler, 2008).

Models (15) through (17), which test for civic engagement, give evidence of a positive contribution to SWB for all the proxies considered, as well as a mediating role for voting and civic participation, while no significant changes are detected in the education coefficients when security in the neighbourhood is added to the baseline regression. In Model (18) we observe that trust exerts a mediating role in the SWB-education relation. In model (19) we include all dimensions of social capital (including one proxy of each dimension), and find that the drop in the education coefficients is quite remarkable[10]. Nevertheless, when compared to model (2), we find that the mediating effect of income is still more powerful than that of social capital.

Finally, regression results from testing hypothesis five are displayed in Table 5, models (20), (21) and (22). In model (20) we can see that if individuals have a better perception about their health, their SWB scores are significantly higher. We also note that the explanatory power of this regression increases considerably. Additionally, education coefficients decrease significantly when this variable is added. Similarly, those who have problems that hamper daily activity are less satisfied with their lives. The mediating role of the latter, however, is smaller than that of perceived health. The combination of the two variables adds little explanation power to the estimation and in terms of mediation, the latter variable does not seem to add any new channel of influence of education on SWB.

Table 4: Hypothesis 4 (social networks and participation)

	Model 1	Model 11	Model 12	Model 13	Model 14
Gender (man)	0.0221 [0.041]	0.0063 [0.041]	0.0275 [0.041]	0.0025 [0.041]	0.0449 [0.041]
Age	-0.0067*** [0.001]	-0.0064*** [0.001]	-0.0061*** [0.001]	-0.0059*** [0.001]	-0.0072*** [0.001]
Age squared	0.0003*** [0.0000]	0.0003*** [0.0000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]
Married	0.1140* [0.060]	0.1262** [0.060]	0.0896 [0.060]	0.1086* [0.060]	0.1071* [0.060]
Divorced	-0.2684*** [0.083]	-0.2689*** [0.083]	-0.2643*** [0.083]	-0.2693*** [0.083]	-0.2694*** [0.083]
Widowed	-0.1824** [0.082]	-0.1608* [0.082]	-0.1656** [0.082]	-0.1677** [0.082]	-0.1981** [0.083]
Secondary Educ.	0.2511*** [0.059]	0.2267*** [0.059]	0.2333*** [0.059]	0.2060*** [0.059]	0.2620*** [0.059]
Higher Educ.	0.3534*** [0.070]	0.3283*** [0.070]	0.3235*** [0.070]	0.2974*** [0.070]	0.3646*** [0.070]
Social meeting - medium		0.1948*** [0.066]			
Social meeting - high		0.3787*** [0.060]			
Someone to talk to			0.4421*** [0.065]		
Social act. - same				0.3170*** [0.041]	
Social act. - more				0.3402*** [0.080]	
Relig. act. - sometimes					0.1012** [0.043]
Relig. act. - often					0.1736** [0.081]
Pseudo R^2	0.0194	0.0233	0.0232	0.0246	0.0201

[10] Including all variables used to measure social capital leads the secondary and higher education variables to drop to 0.1619 and 0.1947, respectively, and the R^2 to rise to 0.0355.

Table 4 (continued): Hypothesis 4 (civic engagement; trust)

	Model 15	Model 16	Model 17	Model 18	Model 19
Gender (man)	0.0160 [0.041]	0.0173 [0.041]	0.0044 [0.041]	0.0100 [0.041]	-0.0126 [0.041]
Age	-0.0078*** [0.001]	-0.0065*** [0.001]	-0.0059*** [0.001]	-0.0066*** [0.001]	-0.0067*** [0.001]
Age squared	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]
Married	0.0978 [0.060]	0.1138* [0.060]	0.1201** [0.060]	0.1080* [0.060]	0.0898 [0.060]
Divorced	-0.2747*** [0.083]	-0.2725*** [0.083]	-0.2493*** [0.083]	-0.2563*** [0.083]	-0.2631*** [0.083]
Widowed	-0.1906** [0.082]	-0.1833** [0.082]	-0.1651** [0.082]	-0.1539* [0.082]	-0.1487* [0.083]
Secondary educ.	0.2369*** [0.059]	0.2367*** [0.060]	0.2515*** [0.059]	0.2268*** [0.059]	0.1748*** [0.060]
High education	0.3110*** [0.071]	0.3245*** [0.071]	0.3497*** [0.070]	0.2973*** [0.070]	0.2137*** [0.071]
Social act. - same					0.3035*** [0.041]
Social act. - more					0.3054*** [0.080]
Vote	0.1815*** [0.044]				0.1441*** [0.044]
Civic participation		0.1113* [0.058]			
Neighbourhood - safe			0.2133*** [0.045]		
Neighbourhood - very safe			0.2894*** [0.072]		
Trust				0.0862*** [0.011]	0.0816*** [0.011]
Pseudo R^2	0.0209	0.0197	0.0217	0.0244	0.0302

Estimations also include round dummies. Standard errors are shown in parenthesis. Single, double and triple asterisks indicate significant coefficients at the 10%, 5% and 1% levels, respectively.

Table 5: Hypothesis 5

	Model 1	Model 20	Model 21	Model 22
Gender (man)	0.0221 [0.041]	-0.0354 [0.041]	-0.0016 [0.041]	-0.0352 [0.041]
Age	-0.0067*** [0.001]	-0.0000 [0.002]	-0.0035** [0.001]	0.0002 [0.001]
Age squared	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]
Married	0.1140* [0.060]	0.0770 [0.060]	0.0872 [0.060]	0.0722 [0.060]
Divorced	-0.2684*** [0.083]	-0.2991*** [0.083]	-0.3131*** [0.083]	-0.3138*** [0.083]
Widowed	-0.1824** [0.082]	-0.2005** [0.082]	-0.1765** [0.082]	-0.1948** [0.082]
Second. educ.	0.2511*** [0.059]	0.1787*** [0.059]	0.2105*** [0.059]	0.1735*** [0.059]
Higher educ.	0.3534*** [0.070]	0.2673*** [0.070]	0.3154*** [0.070]	0.2656*** [0.070]
Health - fair		0.5109*** [0.059]		0.3946*** [0.067]
Health - good		0.7930*** [0.064]		0.6354*** [0.076]
D. H. P - some			-0.4408*** [0.055]	-0.2000*** [0.062]
D. H. P - a lot			-0.6703*** [0.094]	-0.3117*** [0.105]
Pseudo R^2	0.0194	0.0322	0.0276	0.0334

Estimations also include round dummies. Standard errors are shown in parenthesis. Single, double and triple asterisks indicate significant coefficients at the 10%, 5% and 1% levels, respectively.

As a whole, these findings support the five hypothesis we put forward. More specifically, we conclude that the first hypothesis is supported to a greater extent, seeing that income is the mediating factor leading to a higher reduction in the education coefficients. Since separately adding variables from each of the five hypothesis reduces the education coefficients, but do not annul their significance, we went further to investigate whether education effects are totally or only partially mediated. In addition, since we have seen in model (2) that the monetary factor is not enough to exhaust the education effects on SWB, we aim to check the role of non-monetary factors in channelling such affects. Table 6 presents several comprehensive models that enables us to investigate these issues.

Model (23) assembles all five hypothesis. From each set of measures associated with each of the five hypothesis we retained the one with the highest mediation effect[11]. When we consider all transmission mechanisms, we observe that the effect of education on SWB is no longer significant, which indicates that there is no direct effect of education on well-being. Other experiments considering all hypothesis, using different variables or using all proxies examined in this study, produced similar results. Considering different sets of mediators and using data for 18 OECD countries, Helliwell (2002: 24) also found that the coefficients of education have a low magnitude, inferring that the benefits of education "flow primarily through [...] participation, health, perceived trust, and higher income". However he did not conduct a proper examination of mediation. In contrast, Cuñado and de Gracia (2012), using ESS data for Spain and controlling for income, main activity, and health status, found that a very low level of education (not completed primary) keeps its significance, and interpret this as a direct effect. However, in view of our results, we believe that this effect could be picking up other omitted mediators.

[11] The exception to this refers to unemployment. The variable with the highest mediating effect was long-term past unemployment, but we retained instead current unemployment, because this is the typical measure included in standard happiness estimations, and the difference in mediating effects for the two variables, measured by the drop in education coefficients, is quite small.

Models (24) through (27) include income and the set of significant variables from each of the hypothesis 2 to 5. In model (24) we can see that, after controlling for income, adding job status is not enough to drive secondary education to non-significance, but it is sufficient to annul higher education's significance. In models (25), (26) and (27) we can see a similar pattern when separately including unemployment, social capital and health, respectively. These results suggest that secondary education provides a wider variety of benefits than tertiary education. Finally, we realized that adding income, job status and unemployment is sufficient to drive the higher education variable to non-significance, but it is not enough to annul the secondary education's significance. This proves that the theory of human capital is not sufficient to account for all mechanisms transmitting the effect of education into SWB, given that the statistical significance of the secondary education coefficient indicates that this level of education provides benefits other than higher income, higher job status and less vulnerability to unemployment. Furthermore, this result is also evidence of the higher variety of benefits enabled by secondary education, as compared to higher education, suggesting that secondary education can provide benefits that go beyond materialistic aspects, namely by providing people with social resources and/or healthier lives, which promote well-being. Indeed, when adding perceived health or social capital to the latter specification the effect of secondary education disappears[12]. Using European Quality of Life survey data, Böhnke and Kohler (2008: 29) found a similar result when health was included in the regression, concluding that an important role of attaining secondary education is related to living in healthier conditions.

5. Concluding Remarks

Education has been left with an undeserved secondary role in the economics of happiness. Thus, a more thorough analysis of the role of

[12] We do not report the results of such estimations for parsimony reasons.

education is in need. We aimed in this study to contribute to fill in this void in the happiness literature. We have focused on Portugal, which is one of the OECD countries that has made more progress in improving the baseline qualifications of its population, though still lagging behind in terms of educational attainment. This occurs despite Portugal spending on education almost as much as the OECD average (5.9% compares to the OECD average of 6.2% in 2009).

We have found that education has a positive impact on SWB and, more specifically, that having a higher education seems to pay off in terms of SWB. By observing the changes in the coefficients of the education variables we have been able to confirm all postulated hypothesis. I.e. that education enhances SWB by contributing to attain higher earnings, higher job status, less vulnerability to unemployment, more social resources, and better health.

In particular, we have realized that income is the factor that reduces education coefficients to a larger extent. Furthermore, we have realized that the drop in the coefficient of higher education, after controlling for income, is much larger than that of secondary education, almost levelling the latter. This result is in accordance with the outstanding earnings advantage for those with university-degree in Portugal.

Table 6: Testing the direct and indirect effects of education on SWB

	Model 1	Model 23	Model 24	Model 25	Model 26	Model 27	Model 28
Gender (man)	0.0221 [0.041]	-0.1133*** [0.042]	-0.0485 [0.042]	-0.0208 [0.041]	-0.0464 [0.041]	-0.0604 [0.041]	-0.0552 [0.042]
Age	-0.0067*** [0.001]	-0.001 [0.002]	-0.0059*** [0.001]	-0.0072*** [0.001]	-0.0056*** [0.001]	0.0007 [0.002]	-0.007*** [0.001]
Age squared	0.0003*** [0.000]	0.0004*** [0.000]	0.0004*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]
Married	0.1140* [0.060]	0.0234 [0.060]	0.0481 [0.060]	0.0422 [0.060]	0.0321 [0.061]	0.0271 [0.060]	0.0388 [0.060]
Divorced	-0.2684*** [0.083]	-0.2966*** [0.083]	-0.0277*** [0.083]	-0.2636*** [0.083]	-0.2585*** [0.083]	-0.3073*** [0.083]	-0.2767*** [0.083]
Widowed	-0.1824** [0.082]	-0.1740** [0.083]	-0.1389* [0.083]	-0.1754** [0.083]	-0.1207 [0.083]	-0.1704** [0.083]	-0.1627** [0.083]
Secondary educ.	0.2511*** [0.059]	0.0401 [0.061]	0.1391** [0.061]	0.1449** [0.061]	0.1015* [0.061]	0.1155* [0.060]	0.1140* [0.061]
Higher educ.	0.3534*** [0.070]	-0.0035 [0.077]	0.0683 [0.077]	0.1241 [0.076]	0.0249 [0.076]	0.1121 [0.076]	0.0458 [0.077]
Income		0.1480*** [0.034]	0.2220*** [0.034]	0.2169*** [0.033]	0.2238*** [0.033]	0.1839*** [0.033]	0.2004*** [0.034]
Workers sup. - some			0.0718 [0.074]				0.0677 [0.074]
Workers sup. - many			-0.0814 [0.102]				-0.0655 [0.102]
Autonomy - low		0.2134*** [0.061]	0.0219 [0.083]				0.0377 [0.083]
Autonomy - medium		0.25702*** [0.058]	0.0867 [0.080]				0.0969 [0.080]
Autonomy - high		0.3570*** [0.058]	0.1875** [0.084]				0.1978** [0.084]
Influence - low			0.2229*** [0.071]				0.2177*** [0.071]
Influence - medium			0.2225*** [0.074]				0.2206*** [0.074]
Influence - high			0.1611** [0.082]				0.1433* [0.082]
Current unempl.		-0.3108*** [0.067]		-0.2190*** [0.077]			-0.2169*** [0.078]
Past unempl.- long				-0.1292* [0.067]			-0.1274* [0.067]
Social act. - same		0.2528*** [0.042]			0.2964*** [0.041]		
Social act. - more		0.2893*** [0.080]			0.2815*** [0.080]		
Vote					0.1367*** [0.011]		
Trust					0.0807*** [0.0437]		
Health - fair		0.4741*** [0.060]				0.3739*** [0.067]	
Health - good		0.7132*** [0.065]				0.6006*** [0.076]	
D. H. P - some						-0.1845*** [0.062]	
D. H. P - a lot						-0.2954*** [0.105]	
Pseudo R2	0.0194	0.0432	0.0281	0.0256	0.0340	0.0359	0.0270

Estimations also include round dummies. Standard errors are shown in parenthesis. Single, double and triple asterisks indicate significant coefficients at the 10%, 5% and 1% levels, respectively.

We note that the superior role of material factors is not a symptom of a developed country. As suggested by Chen (2012), monetary factors are likely to be more relevant in countries where material achievements have not been fulfilled for the general population, while non-monetary factors are likely to play a more crucial role in abundant societies. Chen (2012) showed that non-monetary factors, namely the number of interpersonal networks and the degree of cosmopolitanism, are much more important than income for some countries in East Asia. Still, this same study finds a superior contribution of income for China, and rationalizes it on the basis of being a country where personal income is generally low, thus playing a more important role for SWB.

However, we also found that social capital adds more explanatory power to the model than income, which barely alters the regression's pseudo-R^2. Perceived health also gives a stronger contribution to explain SWB than income. In this respect Portugal shows evidence of a high income country. This is more in accordance with studies that account for a stronger relationship between income and life satisfaction in poor countries than in rich ones (e.g. Ferrer-i-Carbonell. 2005). In this vein, Delhey (2010) and Böhnke and Kohler (2008) found that individuals living in poor countries derive more satisfaction from material concerns, whereas individuals living in more affluent countries derive more satisfaction from post-materialistic concerns. In a study conducted using the ESS second round data (covering 20 European countries), Lima and Novo (2006) found that Portugal was the European country where objective variables (gender, age, household income, education level, and marital status) mattered the most to the prediction of SWB, namely when compared to subjective variables (perceived income, perceived health, and collectivist and individualistic values). We are thus confronted with mixed signals about Portugal's profile in terms of development stage.

In addition, given that we found that none of the considered factors separately exhausts the effects of education on SWB, we investigated whether the complete set accomplished that. I.e., we checked whether the education effects are totally or only partially mediated, and found that the effect of education on well-being in Portugal is exclusively exerted through indirect channels. Individuals value education for what it brings them, in terms of material and non-material rewards, not for itself.

Moreover, since the monetary factor alone is not sufficient to drive the education's coefficients to non-significance, we tested for each hypothesis after controlling for income. In this manner, when we added the set of variables used to test for each hypothesis, we learned that higher education loses its significance, while secondary education still keeps it. This suggests that secondary education can provide a wider variety of benefits than higher education. Even when we included all variables that can be associated with human capital theory, in model (28), that same pattern emerges. This proves that such a theory is not enough to account for the mediating effects of education in SWB in Portugal, and is further evidence of the higher variety of benefits enabled by secondary education, as compared to higher education. This result is quite remarkable, given that apart from validating the traditional focus on education, it reinforces its role by evidencing that individuals can derive utility from education in ways that are not usually emphasized in mainstream economics.

Our results support the focus of recent Portuguese governments in investing in education. In particular, in view of the results that show that the secondary level of education enables a wider variety of benefits, we feel inclined to say that public policy should focus its efforts in assuring that the population at least attains this level of education. A big step has already been taken by having extended compulsory school to twelve years since 2009. If young people "ignore or heavily discount future consequences when deciding to drop out of school [...] making school compulsory or offering incentives to stay in school may help improve lifetime outcomes" (Oreopoulos, 2007). However, much remains to be done. For instance, investment in the secondary level of education for the adult population needs to be reinforced. The program of education for adults has suffered with the financial problems that the country has gone through. A program of adult education called New Opportunities was established in 2005, but its restructuring since 2012 has been leading to a decrease in the offer of education for adults. The new centres for qualification and vocational education that replaced the New Opportunities' centres are still awaiting for financing and do not seem to be responding adequately to the needs of the adult population. Santiago *et al.* (2012) provide a list of policy recommendations directed at improving school outcomes in Portugal, within an OECD Report reviewing the implementation and use of assessment and evaluation procedures.

References

Becker, Gary S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. University of Chicago Press, Ltd London. 1993, 3rd ed.5 for NBER.

Blanchflower, David G. and Oswald, Andrew J. (2004). Well-being over time in Britain and the USA. Journal of Public Economics, 88, 1359–1386.

Böhnke, Petra and Kohler, Ulrich (2008). Well-being and Inequality. WZB Discussion Paper. Social Science Research Center Berlin (WZB).

Botha, Ferdi (2013). Life Satisfaction and Education in South Africa: Investigating the Role of Attainment and the Likelihood of Education as a Positional Good. Social Indicators Research, 118(2), 555–578.

Brereton, Finbarr; Clinch, J. Peter and Ferreira, Susana (2008). Happiness, geography and the environment. Ecological Economics, 65 (2), 386-396.

Caporale, Guglielmo Maria; Georgellis, Yannis: Tsitsianis, Nicholas and Yin, Ya Ping (2009). Income and happiness across Europe: do reference values matter? Journal of Economic Psychology, 30(1), 42-51.

Chen, Wan-chi (2012). How Education Enhances Happiness: Comparison of Mediating Factors in Four East Asian Countries. Social Indicators Research, 106(1), 117–131.

Clark, Andrew E. and Oswald, Andrew (1994). Unhappiness and Unemployment. The Economic Journal, 104 (424), 648-659.

Clark, Andrew E. and Oswald, Andrew (1996). Satisfaction and Comparison Income. Journal of Public Economics, 61, 359-381.

Clark, Andrew and Lelkes, Orsolya (2009). Let us Pray: Religious Interactions in Life Satisfaction. Working Paper No. 2009-01, Paris School of Economics.

Côté, Sylvain and Healy, Tom (2001). The Well-being of nations. The role of human and social capital. Organization for Economic Co-operation and Development, Paris.

Cuñado, Juncal and de Gracia, Fernando Pérez (2011). Does Education Affect Happiness? Evidence for Spain. Social Indicators Research, 108(1), 185–196.

Delhey, Jan (2010). From Materialist to Post-Materialist Happiness? National Affluence and Determinants of Life Satisfaction in Cross-National Perspective. Social Indicators Research, 97, 65-84.

Delhey, Jan (2004). Life satisfaction in an enlarged Europe. Report for the European Foundation for the Improvement of Living and Working Conditions. Luxembourg: Office for Official Publications of the European Communities.

Di Tella, Rafael; MacCulloch, Robert and Oswald, Andrew (2001). Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. American Economic Review 91(1), 335-341.

Diener, Ed; Sandvik, Ed; Seidlitz, Larry and Diener, Marissa (1993). The relationship between income and subjective well-being: Relative or absolute? Social Indicators Research, 28, 205-215.

Diener, Ed; Suh, Eunkook; Lucas, Richard and Smith, Heidi (1999). Subjective Well-being: Three Decades of Progress. Psychological Bulletin, 125 (2), 276-302.

Diener, Ed and Seligman, Martin (2004). Beyond Money: Toward an Economy of Well-Being. Psychological Science in the Public Interest, 5 (1), 1-31.

Di Tella, Rafael; MacCulloch, Robert and Oswald, Andrew J. (2003). The Macroeconomics of Happiness. Review of Economics and Statistics, 85(4), 809-27.

Dolan, Paul; Peasgood, Tessa and White, Mathew (2008). Do We Really Know What Makes us Happy? A Review of the Economic Literature on the Factors Associated with Subjective Well-being. Journal of Economic Psychology, 29 (1), 94-122.

Easterlin, Richard A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz, Paul A. David and Melvin W. Reder (Eds.), New York: Academic Press, 89-125.

Easterlin, Richard A. (1995). Will Raising the Incomes of all Increase the Happiness of all?, Journal of Economic Behavior and Organization, 27 (1), 35-47.

Ferrante, Francesco (2009). Education, Aspirations and Life Satisfaction. Dipartimento di Scienze Economiche Università di Cassino. Working Paper 3/2009.

Ferrer-i-Carbonell, Ada (2005). Income and Well-being: an Empirical Analysis of the Comparison Income Effect. Journal of Public Economics, 89 (5-6), 997-1019.

Frey, Bruno S. and Stutzer, Alois (2002). What Can Economists Learn from Happiness Research? Journal of Economic Literature, 40 (2), 402-435.

Frey, Bruno S. and Stutzer, Alois (2000). Happiness, Economy and Institutions. The Economic Journal, 110(15), 918–938.

Frey, Bruno S. (2008). Felicidade: uma Revolução na Economia. Gradiva: Lisboa.

Gerdtham, Ulf-G. and Johannesson, Magnus (2001). The relationship between happiness, health, and socio-economic factors: Results based on Swedish microdata. Journal of Socio-Economics, 30, 553–557.

Han, Sehee; Kim, Heaseung and Lee, Hee-Sun (2013). A Multilevel Analysis of the Compositional and Contextual Association of Social Capital and Subjective Well-Being in Seoul, South Korea. Social Indicators Research, 111(1), 185-202.

Haveman, Robert H. and Wolfe, Barbara L. (1984). Schooling and Economic Well-Being: the role of nonmarket effects. The Journal of Human Resources, 19(3), 377-407.

Helliwell, John F. (2002/3?). How's life? Combining Individual and National Variables to Explain Subjective Well-being. Working Paper 9065, National Bureau of Economic Research.

Helliwell, John F. and Putman, Robert D. (2007). Education and Social Capital. Eastern European Journal, 33(1), 1-19.

Layard, Richard (2005). Happiness: lessons from a new science. Penguin Books: London.

Lazear, Edward P. (1977). Education: Consumption or Production? Journal of Political Economy, 85, 569-98.

Lima, Maria Luísa and Novo, Rosa (2006). So far so good? Subjective and social well-being in Portugal and Europe. Portuguese Journal of Social Science, 5(1), 5-33.

Meier, Stephan and Stutzer, Alois (2006). Is Volunteering Rewarding in Itself?. Economica, 75 (297), 39-59.

Myers, David G. and Diener, Ed (1995). Who Is Happy? Psychological Science, 6(1), 10-19.

Nieminen, Tarja; Martelin, Tuija; Koskinen, Seppo; Simpura, Jussi; Alanen, Erkki; Härkänen, Tommi and Aromaa, Arpo (2008). Measurement and Socio-demographic Variation of Social Capital in a Large Population-based Survey. Social Indicators Research, 85, 405-423.

OECD (2012). Education at a Glance 2012 – Country Note. OECD Publishing. http://dx.doi.org/10.1787/eag-2012-en. Accessed in 23/01/2015.

OECD (2014). Education at a Glance 2014: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2014-en. Accessed in 23/01/2015.

Powdthavee, Nattavudh (2008). Putting a Price Tag on Friends, Relatives, and Neighbours: Using Surveys of Life-Satisfaction to Value Social Relationships. Journal of Socio-Economics, 37(4), 1459-1480.

Pugno, Maurizio (2007). The subjective well-being paradox: a suggested solution based on relational goods. In Handbook on the Economics of Happiness, Luigino Bruni and Pier Porta (coord.), Elgar, 263-289.

Sabates, Ricardo and Hammond, Cathie (2008). The Impact of Lifelong Learning on Happiness and Well-being. Institute of Education, London.

Salinas-Jiménez, Maria del Mar; Artés, Joaquín and Salinas-Jiménez, Javier. 2010. Education as a Positional Good: A Life Satisfaction Approach. Social Indicators Research, 103(3), 409–426.

Santiago, Paulo; Donaldson, Graham; Looney, Anne and Nusche, Deborah (2012). OECD Reviews of Evaluation and Assessment in Education: Portugal 2012. OECD Publishing. http://dx.doi.org/10.1787/9789264117020-en. Accessed in 23/01/2015.

Schultz, Theodore W. (1961). Investment in Human Capital. American Economic Review, 51, 1-17.

Spence, Michael (1973). Job Market Signaling. The Quarterly Journal of Economics. 87(3), 355-374.

Table A1: Description of the variables tested under the five hypothesis

H1	**Income** Q: Please tell me which letter describes your household's total income, after tax and compulsory deductions, from all sources? If you don't know the exact figure, please give an estimate. Cat: Since intervals differ throughout the three rounds, the mid-term point of the income interval bracket was standardized. In order to consider income in equivalent terms, it was also divided by the square-root of the number of household members. This adjustment is meant to consider economies of scale in consumption within the household.
H2	**Workers supervision** Q1: In your main job, do/did you have any responsibility for supervising the work of other employees? Q2: How many people are/were you responsible for? Cat.: No (omitted category); some workers (1 to 9); many workers (10 or more).
	Autonomy Q: How much the management at your work allows/allowed you to decide how your own daily work is/was organised? Cat: No autonomy (omitted category); low autonomy (1 to 4); medium autonomy (5 to 7); high autonomy (8 to 10).
	Influence on management decision Q: How much the management at your work allows/allowed you to influence policy decisions about the activities of the organisation? Cat.: No influence (omitted category); low influence (1 to 4); medium influence (5 to 7); high influence (8 to 10).
H3	**Current unemployment** Q: Which of these descriptions applies to what you have been doing for the last 7 days? In paid work; in education; unemployed and actively looking for a job; unemployed, wanting a job but not actively looking for a job; permanently sick or disabled; retired; in community or military service; doing housework, looking after children or other persons. Cat.: In paid work or other categories (reference category); unemployed (either looking for a job or not).
	Past unemployment (short and long-term) Q1: Have you ever been unemployed and seeking work for a period of more than 3 months? Q2: Have any of these periods lasted 12 months or more? Cat.: Past short-term unemployment (yes to Q1 and no to Q2); past long-term unemployment (yes to Q1 and yes to Q2).

H4	**Social networks and participation** (frequency of social meetings, having someone to discuss intimate and personal matters, social activity, religious activity) Q1: How often do you meet socially with friends, relatives or work colleagues? Cat.: Never, less than once a month, or once a month (reference category); medium (several times a month, once a week); high (several times a week, every day). Q2: Do you have anyone with whom you can discuss intimate and personal matters? Cat.: No (reference); yes. Q3: Compared to other people of your age, how often would you say you take part in social activities? Cat.: Much less than most or less than most (reference); about the same; more than most or much more than most. Q4: Apart from special occasions such as weddings and funerals, about how often do you attend religious services nowadays? Cat.: Only on special holy days, less often or never (reference); sometimes (once a week or at least once a month); often (every day or more than once a week).
	Civic Engagement (vote, other civic participation and perceptions of safety in the neighbourhood) Q1: Did you vote in the last national election? Cat.: No (reference); yes. Q2: During the last 12 months, have you done any of the following? Contacted a politician, government or local government official; worked in a political party or action group; worked in another organisation or association; worn or displayed a campaign badge/sticker; signed a petition; taken part in a lawful public demonstration; boycotted certain products? Cat.: No to all question (reference); yes to any one of the above questions. Q3: How safe do you – or would you - feel walking alone in this area after dark? Cat.: Unsafe or very unsafe (reference category); safe; very safe.
	Trust Q1: Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people? Q2: Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair? Q3: Would you say that most of the time people try to be helpful or that they are mostly looking out for themselves? A: Average of answers (ranging from 0 to 10) to these three questions.
H5	**Perceived health** Q: How is your health in general? Would you say it is… Cat.: Bad or very bad (reference); fair; very good or good.
	Disabling health problem Q: Are you hampered in your daily activities in any way by any longstanding illness, or disability, infirmity or mental health problem? Cat.: No (reference); yes to some extent; yes a lot.

Table A2: Descriptive statistics (average/percentages)

Life satisfaction	5.6	**Life satisfaction per education level**		**Education Level**	
		Basic	5.41	Basic	76%
		Secondary	6.33	Secondary	15%
		Higher	6.39	Higher	9%

Gender		**Age**	55	**Marital status**		**Rounds**	
Male	39%			Single	20%	2008	32%
Female	61%			Married	55%	2010	36%
				Divorced/separated	9%	2012	32%
				Widowed	16%		

H1

Annual Income - Round 4		**Annual Income - Round 5**		**Annual Income - Round 6**	
J: Less than 5000	3.86%	J: Less than 1800	1.57%	J: Less than 5500	15.91%
R: 5501 to under 7500	11.8%	R: 1800 to under 3600	6.48%	R: [5501, 7500]	20.54%
C: 7501 to under 10000	16.23%	C: 3600 to under 6000	16.4%	C: [7501, 10000]	21.33%
M: 10001 to under 12000	25.77%	M: 6000 to under 12000	30.35%	M: [10001, 12000]	17.61%
F: 12001 to under 14000	18.05%	F: 12000 to under 18000	21.71%	F: [12001, 14000]	11.63%
S: 14001 to under 17000	10.56%	S: 18000 to under 24000	10.12%	S: [14001, 17000]	5.87%
K: 17001 to under 20000	5.79%	K: 24000 to under 30000	5.99%	K: [17001, 20000]	3.05%
P: 20001 to under 25000	4.54%	P: 30000 to under 36000	3.54%	P: [20001, 25000]	1.58%
D: 25001 to under 35000	2.38%	D: 36000 to under 60000	2.95%	D: [25001, 35000]	1.81%
H: 35000 or more	1.02%	H: 60000 to under 90000	0.49%	H: above 35000	0.68%
		U: 90000 to under 120000	0.1%		
		N: 120000 or more	0.29%		

H2

Workers' supervision		**Autonomy**		**Influence**	
None	87%	None	21%	None	33%
Some	9%	Low	21%	Low	26%
Many	4%	Medium	28%	Medium	22%
		High	30%	High	19%

Table A2 (continued): Descriptive statistics (average/percentages)

H3							
Current Unemployment		**Past unemployment - short**	9%	**Past unemployment - long**	15%		
Unemployed	10%						
Other job status	90%						
H4							
Freq. of social meetings		**Having someone to talk to**	90%	**Social activity**		**Religious activity**	
Low	13%			Less than most	42%	Rarely	55%
Medium	25%			About the same	51%	Sometimes	38%
High	62%			More than most	7%	Often	7%
H4 (continued)							
Vote (has voted in the last national elections)	68%	**Perceptions of safety in the neighbourhood**		**Trust in people** (average)	4,2		
		Unsafe	28%				
		Safe	62%				
		Very safe	10%				
H5							
Perceived health		Disabling health problem					
Bad	17%	None	78%				
Fair	33%	Some	17%				
Good	50%	A lot	5%				

Table A3: Correlation coefficients - variables tested in H2

	W. sup. - none	W. sup. - some	W. sup. - many	Aut. - none	Aut. - low	Aut. - medium	Aut. - high	Infl. - none	Infl. - low	Infl. - medium	Infl. - high
Workers sup. - none	1.000										
Workers sup. - some	-0.802	1.000									
Workers sup. - many	-0.542	-0.065	1.000								
Autonomy - none	0.184	-0.148	-0.101	1.000							
Autonomy - low	0.163	-0.132	-0.088	-0.270	1.000						
Autonomy - medium	0.031	-0.017	-0.028	-0.324	-0.325	1.000					
Autonomy - high	-0.342	0.268	0.197	-0.335	-0.336	-0.403	1.000				
Influence - none	0.205	-0.170	-0.104	0.737	-0.156	-0.246	-0.279	1.000			
Influence - low	0.151	-0.124	-0.079	-0.304	0.610	-0.060	-0.217	-0.409	1.000		
Influence - medium	-0.078	0.051	0.058	-0.271	-0.228	0.552	-0.097	-0.373	-0.317	1.000	
Influence - high	-0.328	0.285	0.150	-0.251	-0.250	-0.226	0.673	-0.339	-0.288	-0.263	1.000

Table A4: Correlation coefficients - variables tested in H4

	S. M. - low	S. M. - med.	S. M. - high	Som. to talk	S. A. - less	S. A. - same	S. A. - more	R. A. - rarely	R. A. - somet.	R. A. - often	Vote	Civic part.	Neig. -unsafe	Neig. -safe	Neig. -v. safe	Trust
Social meeting - low	1.000															
Social meeting - medium	-0.224	1.000														
Social meeting - high	-0.502	-0.731	1.000													
Someone to talk to	-0.198	-0.018	0.154	1.000												
Social act. - less	0.274	0.030	-0.218	-0.178	1.000											
Social act. - same	-0.228	-0.016	0.174	0.161	-0.868	1.000										
Social act. - more	-0.083	-0.027	0.082	0.030	-0.236	-0.279	1.000									
Relig. act - rarely	-0.018	-0.042	0. 050	0.044	-0.052	0.050	0.003	1.000								
Relig. act. - sometimes	0.013	0.058	-0.061	-0.022	0.038	-0.023	-0.028	-0.871	1.000							
Relig. act. - often	0.009	-0.029	0.019	-0.045	0.030	-0.055	0.050	-0.297	-0.210	1.000						
Vote	-0.028	0.013	0.009	0.024	-0.040	0.037	0.005	-0.095	0.104	-0.014	1.000					
Civic participation	-0.034	0.014	0.012	0.058	-0.138	0.082	0.106	0.053	-0.057	0.006	0.110	1.000				
Neighbourh. -unsafe	0.086	0.002	-0.062	-0.034	0.074	-0.086	0.024	-0.026	0.035	-0.017	0.003	-0.046	1.000			
Neighbourh. - safe	-0.048	0.040	-0.002	0.017	-0.025	0.046	-0.041	0.018	-0.030	0.024	-0.008	0.002	-0.797	1.000		
Neighbourh. - very safe	-0.050	-0.068	0.095	0.024	-0.070	0.054	0.030	0.010	-0.004	-0.013	0.009	0.068	-0.209	-0.424	1.000	
Trust	-0.059	0.006	0.036	0.065	-0.086	0.060	0.049	0.002	-0.004	0.003	0.050	0.072	-0.201	0.113	0.118	1.000

Table A5: Correlation coefficients - variables tested in H3

	Current unemp.	Past. un.- short	Past. un.- long
Current unemp.	1.000		
Past unemp. - short	0.143	1.000	
Past unemp. - long	0.551	-0.130	1.000

Table A6: Correlation coefficients - variables tested in H5

	Health - bad	Health - fair	Health - good	D. H. P - none	D. H. P - some	D. H. P - a lot
Health - bad	1.000					
Health - fair	-0.317	1.000				
Health - good	-0.452	-0.703	1.000			
D. H. P - none	-0.572	-0.053	0.478	1.000		
D. H. P - some	0.415	0.111	-0.416	-0.863	1.000	
D. H. P - a lot	0.379	-0.097	-0.193	-0.415	-0.101	1.000

Adaption of the Profitability estimation focused on benefits due to personal affection

Stephan Printz[1], René Vossen[1] and Sabina Jeschke[1]

Abstract:

Strategic investment decisions are characterized by high innovation potential and long term effects on the competitiveness of enterprises. Due to uncertainty and risk in this complex decision-making situation, the need for well-structured support activities arises. A method that considers cost and the long-term added value is the cost-benefit effectiveness estimation. One of those methods is the "**p**rofitability **e**stimation **f**ocused on **b**enefits – PEFB"-method (trans. Savelsberg 2008) developed at the institute of management cybernetics at RWTH Aachen University in 2000, by WEYDANDT in the course of his PhD. This method copes with the challenge of strategic investment decision by integrating long term non-monetary aspects as well as by mapping the chronological sequence of an investment within the organisation´s target system. For that reason, this method is characterised as a holistic approach for the evaluation of costs and benefits of an investment. In applying this participation-oriented method in business environment, the overall interdisciplinary investment teams were formed by more than 500 participants. Due to feedback of these participants within accomplished assessments, a survey with 110 participants was conducted focussing on the statements of verbal probability. The results of the survey identified a structural deficit in the conception of the methodology regarding the estimation of probabilities due to personal affection. This paper describes the well-structured PEFB-method to reduce risk and uncertainty in decision making progress. Moreover, the results of the survey concerning the bias due to personal affection within the evaluation process are

[1] Institute for Management Cybernetics e.V. assoc. Institute of RWTH Aachen University
E-mails: stephan.printz@ima-zlw-ifu.rwth-aachen.de, rene.vossen@ima-zlw-ifu.rwth-aachen.de, sabina.jeschke@ima-zlw-ifu.rwth-aachen.de

http://dx.doi.org/10.14195/978-989-26-1039-9_11

addressed. Finally, to achieve more accurate PEFB-method results, recommendations for an adaption of the PEFB are given.

Keywords: Multi-Criteria Decision, Cost-Benefit Analysis, Risk and Uncertainty Analysis, Profitability Estimation Focused on Benefits, PEFB-method.

1. Introduction

Literature just like practice shows that capital budgeting has not always distinguished strategic types of investment (Graham & Harvey, 2001). Latest advances in research give evidence that there is substantial need for distinction in strategic investment decisions (Alkaraann & Northcott, 2006). Strategic investment decisions (SID) are defined as

"(...) substantial investments that involve high levels of risk, produce hard-to-quantify (or intangible) outcomes and have a significant long-term impact on corporate performance." (Alkaraan & Northcott, 2006)

In addition, Carr & Tomkins point out that there is also an effect on the whole organisation regarding the competitiveness (Carr et al., 2010). Typical examples for SID are companies' mergers and acquisitions, e.g. a new assembly line, advanced technology of manufacturing or the launch of new software (Van Cauwenbergh et al., 1996; Slagmulder et al., 1995; Accola 1994). Controlling the complexity and uncertainty surrounding strategic investment decisions presents particular challenges to the management (Alkaraan & Northcott, 2006; Dempsey, 2003; Slagmulder, 1997). For this reason, an efficient information search and evaluation is necessary (Preißler, 1991). Especially for the evaluation of accounting SID, there is a need to pay more attention to scenario-based techniques (Cornelius et al., 2005; Alessandri et al., 2004; Miller & Waller, 2003). Thus, SID have a greater need to pay attention to probability of occurrence (Martzoukos & Trigeorgis, 2002). Field studies give evidence that traditional profitability analysis to assess SID is supplanted by substantial sophisticated techniques in terms of linking qualitative and financial aspects (Alkaraan & Northcott, 2006; Adler, 2000; Slagmulder et al., 1995). While the quantification and assignment of cost is examined extensively, there are less methods for the quantification and assignment of long term benefits (Schönheit, 1996; Zangemeister, 1994). New sophisticated techniques face this challenge for the quantification of benefits and the estimation of their impacts within an efficient information system.

2. Methodology Review

Indeed, most research has aimed at facing the question as to which analyses are being used to assess SID (Sandahl & Sjörgren, 2003). One method commonly used for SID is the Discounted Cash Flow (DCF)-Analysis (Pike, 1996) (1). The DCF, also known as present value (PV) uses risk-adjusted discount rates (r) to monetise the time frame of SID (Allemann, 2002; Adler, 2000; Slagmulder et al., 1995). The cash flows (c_t) are separated by their period of time (t) (for a detailed overview refer to Brealey et al., 2011). To evaluate an investment, the initial negative cash flow (C_0) is added and leads to the Net Present Value (NPV) (2).

$$PV = DCF = \sum_{t=1}^{T} \frac{C_t}{(1+r)^t} \quad (1)$$

$$NPV = C_0 + PV = C_0 + \sum_{t=1}^{T} \frac{C_0}{(1+r)^t} \quad (2)$$

Accurate forecast of long-term measures cannot be made by comprehensively monetising (Ansoff, 1988). For instance, risk rises by involving new technology (Demchek, 1992), dynamic feedback loops in the SID (Paitch & Sterman, 1993) and a long forecast horizon (Webb, 1994). Above all, DCF has been examined by using excessively high discount rates compared to the market standard to assess SID (Adler, 2000). Due to missing flexibility in monetising option values, DCF is most appropriate in assessing non-strategic investment, where intangible elements and risks are low (Pike, 1996). Finally, DCF is supposed to overemphasise on the short term, exclude the non-financial benefits and neglect the assessment of cross-functional and cross-departmental benefits (Adler, 2000).

According to SID, neglecting long term effects may lead to wrong decisions (Nitzsch, 2006). The issue is the lack in considering qualitative aspects.

An established method developed for decision support in the field of economics, finance and psychology considering also qualitative aspects is the utility analysis (UA) (Cascio, 2000). In terms of system technology, UA is characterised by "organizing a multidimensional target system with a set of complex alternatives in order to set preferences according to the decision maker" (Zangemeister, 1976 - translation from German). This approach ensures that decision making process is performed systematically and comprehensible by a person or team (Kaplan & Rubak, 1994). The value in use of an option (U_i) for all potential circumstances (s) is defined by a weighting (w_z) of the partial utility values (u_{iz}) (3), provided that the sum of the weights (w_z) equals to 1 (4). However, the so-called value synthesis of the part worth and the overall benefit requires a consistent cardinal scale and implies independent benefits from the target system (Zangemeister, 2000). Although it is important to estimate benefits, UA uses non-monetary dimensions and may lead to standard errors due to uncertainty (Alexander & Cronshaw, 1984; Alexander et al., 1986).

$$U_i = \sum_{z=1}^{s} u_{iz} w_z \quad (3)$$

$$\sum_{z=1}^{s} w_z = 1 \quad (4)$$

Since UA has become an established method (Holling, 1998), the challenge is to combine the precise measures of TPA and the advantages of UA (Reichwald et al., 1996). In particular, the investment decisions are related to monetary assessable measures, but SID need to consider non-monetary aspects as well. Hence, the need is to use classification of direct, indirect and uncertain impact classes in a team assessment (Pittermann, 1998). With respect to UA, an integration of a target system with low complexity and mapping scenarios through the use of probabilities increases transparency of the SID (Zangemeister, 2000). By the same token, the support in different kinds of applications with little temporal and personal expense is required (Weydandt, 2000).

However, a method assimilating the precise measures DCF and the assets of the Utility Analysis is required. In addition, a holistic perception of the SID in order to quantify non-monetary aspects is needed. With regard to the UA, the implementation of weightings in order to prioritise the sequence of investment is required. In compliance with these requirements, the method has to use the accuracy of DCF as well as the monetary evaluation dimension. Finally, the method has to be non-complex and easy to perform, leading to less expenditure of time and less effort.

3. Profitability estimation focused on benefits (PEFB)-method

The PEFB-method (Savelsberg, 2008) is a holistic, participation-oriented approach for the evaluation of costs and benefits of an investment (Unger, 1998). This approach combines precise measures of financial accounting with the usage of an interdisciplinary investment team to reduce bias in decision making process for non-monetary aspects. By assessing non-monetary aspects in an interdisciplinary investment team with different kinds of hierarchical levels, the transparency through SID is increased. Furthermore, the method requires less expense regarding financial and temporal effort and is also transferable to other application cases.

Based upon the utility analysis by Nagel (Nagel, 1988) and the profitability analysis of IT-investment by Ott (1992), Weydandt (2000) expands this approach for technical investment. Not only measurable monetary factors are considered, such as costs and revenues, but also a quantification of non-monetary aspects is included, such as time, quality, flexibility or the enterprise environment (Strina et al., 2003). With regard to the problem-solving process (Sell & Schimweg, 2002), Figure 1 shows an overview of the PEFB-method (Jursch et al., 2010; Strina et al., 2003).

The PEFB-methods consists of seven steps, which are in detail:

1. constitution of the interdisciplinary investment team,
2. current situation analysis,
3. target situation analysis,

4. compilation of measures,
5. investment evaluation,
6. planning of actions and
7. reflection.

Step 1 is concerned with building a representative interdisciplinary investment team consisting of involved executives and employees. Hence, not only the affected departments are represented, but also the various levels of hierarchy. This team will be responsible to conduct the whole evaluation, implementation and reflection process. By building up an interdisciplinary investment team and with the support of a facilitator, a holistic view on the SID is ensured. Within the current state analysis (step 2), skills and competencies inside the enterprise are identified. Furthermore, by means of the target state analysis (step 3), the strategically, tactical and operational objectives of the project are identified and noted in a specification sheet. Step 4 executes a compilation of concrete measures to achieve the target state.

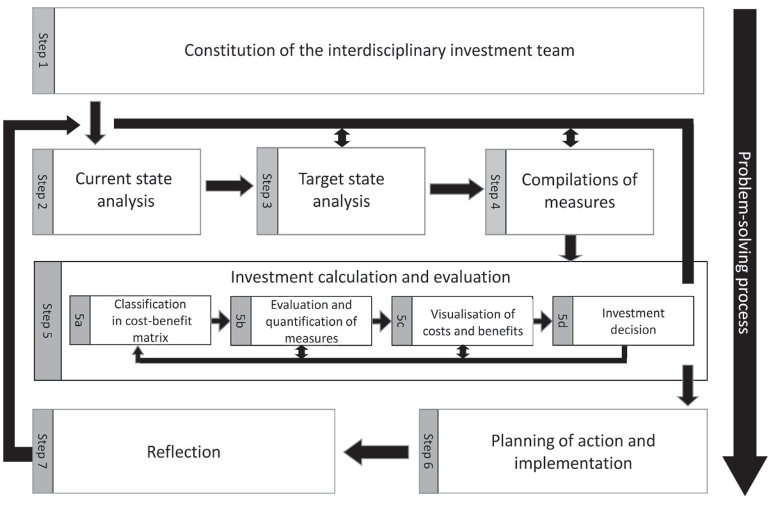


Figure 1: The PEFB-method (Source: Uribe, K. Henning, G. Strina: Measurement of Participation-oriented Evaluation: NOWS. In: Knowledge is Orange - Industry-Focused Applications of Knowledge Management across Germany. 2004)

In step 5, the investment calculation and evaluation is carried out through four sub tasks (Strina et al., 2003). As shown in Figure 1, it encompasses the classification of costs and benefits, the evaluation and quantification of measures, visualisation of cost and benefits as well as the investment decision. The method uses two different portfolios to classify the measures. One portfolio for the benefit classification (see Figure 2) and another for the cost classification (see Figure 3). Costs and benefits of the investment are defined and assigned in terms of their impact on the project as "direct", "indirect" or "difficult to ascertain".

		Probability of occurrence		
		High	Medium	Low
Benefit matrix	Direct	**1**	**3**	**6**
	Indirect	**2**	**5**	**8**
	Difficult to ascertain	**4**	**7**	**9**

Figure 2: Benefit portfoli

		Probability of occurrence		
		High	Medium	Low
Cost matrix	Direct	**9**	**7**	**4**
	Indirect	**8**	**5**	**2**
	Difficult to ascertain	**6**	**3**	**1**

Figure 3: Cost portfolio

Direct costs or benefits are those which are related to the investment so that the impact can be measured directly, e.g. acquisition costs or increase of productivity. Accordingly, indirect costs or benefits are a derivate from direct effects, for instance maintenance cost or increased quality. Finally, "difficult to ascertain" costs or benefits contain effects which can only be presumed, like demotivation of employees or improving the image of the company. Moreover, the measures are classified regarding their probabilities of occurrence into the classes high, medium and low. Each cell of the matrix contains a so called risk level, reaching from 1 to 9. Meanwhile direct benefits with high probability are assigned with the risk level 1 (refer to Figure 2), the risk levels in the cost portfolio (see Figure 3) are arranged contrarily. Hence, direct costs with high probabilities are assigned to the risk level 9.

In the framework of the visualisation, the risk levels define a ranking scale of measures (Nagel, 1988). After the classification and quantification of the aspects, the filled cells of the matrix are aligned in two numerical series. The overall benefits (B_n) for a risk level (j) are computed from the summation of the single benefits (b_j) (5). The calculation of the overall costs (C_m) for a risk level (j) of the individual costs (c_j) is done accordingly (6).

$$B_n = \sum_{j=1}^{9} b_j \qquad (5)$$

$$C_m = \sum_{j=1}^{(9-m)+1} c_j \qquad (6)$$

The cumulated values are added according to their risk level scale (Jursch et al., 2010; Strina et al. 2003). The value of each risk level is recorded in a risk oriented chart (see Figure 4). In relation to the possible intersection situations, four general cases can be distinguished. On the one hand, when the cost function is beneath the benefit function in all nine levels, the investment is economically evaluated as recommendable without restrictions. On the other hand, if the cost function is always above the benefit function, the investment is evaluated as definitely uneconomical. Finally, in the case of an intersection of both curves, an interpretation of the risk-level is required.

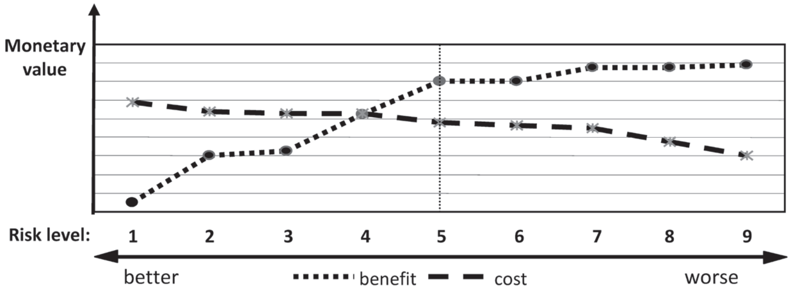


Figure 4: Risk levels visualisation based on costs-benefits evaluation

Investment scenarios with an intersection of both graphs in the risk level 1 are the most economically reasonable (better). In terms of the descriptive interpretation, the overall costs are exceeded by direct and highly probable benefits. According to the alignment, if the intersection point is located to a greater extent to the left of risk level 5, the investment is more recommendable. Contrarily, the worst economically reasonable investment scenario is a location of the intersection point at the risk level 9 (worse). Hence, direct and highly probable costs go beyond all possible benefits. Thus, the investment is getting worse exceeding benefit risk level 5.

In step 6 (planning of action or implementation – depending on the results of the assessment), the next actions towards the investment decision are determined. In particular, the advice for the evaluation of strategies is directly incorporated into the transfer process. Finally, the interdisciplinary investment team appraises the experiences gathered during the process step 7. Accordingly, in the reflection phase a review of the whole investment evaluation process is carried out and appropriate recommendations are made.

4. The experimental design

In general, there is a substantial need to assess also long-term effects for the evaluation of SID, thus an estimation of the probabilities of occurrence is used. This estimation is realised in PEFB-method by an interdisciplinary investment team. The team consists of different hierarchical levels with various attitudes towards the SID. Due to the various attitudes, an effect on quality of the results is supposed. Thus, in order to investigate the effect of verbal probabilities in the context of personal affection, a survey was conducted. According to the used probability terms of low, medium and high in the PEFB-method, the definition of the boundaries regarding the used probability terms has to be analysed.

H1: The perception of risks in investments is influenced by personal affection in group decisions which leads to a bias in probability evaluation with regard to the low and medium boundary.

H2: The perception of risks in investments is influenced by personal affection in group decisions which leads to a bias in probability evaluation with regard to the medium and high boundary.

The Participants were recruited randomly and the duration of the survey was about 10 minutes. Furthermore, the participants' social demographical background, like age and profession were assessed. Within the 24 items of the survey, 14 items were related to investigate the bias in probability assessment due to personal affection. The risk attitude was assessed by a point scale (*1=risk-averse, 5=risk seeking*). Furthermore, personal affection was assessed by four items on an open scale. The following questions are examples of items to determine the personal affection.

Q1 (investment_small): Imagine you buy an electric pineapple peeler for kitchen use. With the help of this device, you expect to peal fast and easy. Moneywise, you take no financial risks. For you, what would be a low, medium respectively high probability to satisfy your demands or to be dissatisfied with your investment?

Q2 (investment_big): Imagine you buy a large tenement with the goal of economic success in the long-run. The rents received ought to contribute essentially to your financial security in your retirement. Financially, you take a risk by buying. For you, what would be a low, medium respectively high probability to regard the investments not profitable (e.g. by building an incineration plant in striking distance to rental property, high vacancy, high repair costs etc.)?

5. Results

Table 1 shows detailed information regarding the number of participants (N), mean value (M) and standard Deviation (SD) of the survey. The majority of the participants are employees (40.9%) or students with a master degree (31.8%) followed by executives (14.5%). The average age of the participants was 32.55 years (SD = 11.62).

Most of the participants stated that they are risk-averse in general, although their attitude towards risks of everyday life is associated with risk seeking. An exception is formed by the attitude towards risks of natural forces by high risk aversion. Average measures in sample for investment without personal affection show boundary values for low to medium of 32.26% and 68.75% for medium high probabilities. This results decrease for personal affection to the boundary value of low to medium to 26.49% and for the boundary value of medium to high to 57.51%.

Table 1: Descriptive data of the study variables

variable	*N*	*M*	*SD*
Attitude towards risks	110	2.23	1.17
Attitude towards risks of everyday life	110	3.48	1.12
Attitude towards risks of natural forces	110	2.07	1.12
Investment boundary low to medium (low personal affection)	110	32.26	13.85
Investment boundary medium to high (low personal affection)	110	68.75	16.02
Investment boundary low to medium (high personal affection)	110	26.49	15.48
Investment boundary medium to high (high personal affection)	110	57.51	19.96

To test the hypothesis, an independent-sample *t-test* was conducted. Table 2 shows the results of the analysis with regard to the boundaries of mean value due to personal affection (PA) and the t-test (t).

Table 2: Results of independent-sample t-test for investment risk at low to medium and medium to high level in dependence on personal affection

variable	N	M (PA =low)	M (PA = high)	t
Investment risk low to medium level	110	32.257	26.494	2.911**
Investment risk medium to high level	110	68.751	57.5055	4.609**

Significant differences in perception of investments risk at low to medium (t = 32.257, p <.001) as well as medium to high level (t = 26.494, p <.001) were found. As proposed in hypothesis 1, people that are personally affected differ in their estimation of investment risks from low to medium boundary in comparison to participants that are not personally affected. In fact, people also differ in the estimation of investment risks from medium to high boundary in comparison to not personally affected participants. To that effect, applying PEFB-method entails some risks regarding the validity of the results. These risks have to be considered and eliminated by conceptual adaption.

6. Conclusion

Current research addressing the assessment of SID consistently demonstrates that there is a need for more comprehensible methods. The focus of these methods should rely on theoretically and adaptive approaches. Identifying and evaluating the long–term effects under uncertainty and risk is the main research challenge. Without knowledge about the effects, forecasting and evaluation of SID will be hazardous. One way to meet this challenge is applying the PEFB-method.

The contribution of this paper is the examination of the PEFB-method regarding the conceptual design and the challenges in applying for SID. In particular, the findings of the survey give evidence to a bias due to personal affection regarding the classification of occurrence probabilities. Within the scope of the PEFB-method, wrong alignment of probabilities under the assumption of fixed cost benefit components lead to an average bias of 2.6 risk levels. That implies a movement of the cost or benefit graph by the same amount of risk levels. This results in movement of the intersection point and the results of the method as well.

The findings of this study induce some critical reflections. These critical reflections are related to a superordinate and a method-orientated frame of reference. Within the scope of critical reflections to superordinate aspects, the interdisciplinary investment team are just interacting components. Owing to the psychological phenomena of "Groupthink" (Mullen et al., 1994) the conformity of the group is weighted higher than the result of the assessment. Besides these facts, there are also interrelations among the interdisciplinary team, like the impact of hierarchies, know-how and practical knowledge. In particular, personal commitment related to the company and the SID as well as professional experience may promote bias in probability estimation by group decision.

Furthermore, within the method-oriented frame of reference unambiguous definitions of the SID result in inconsistent considerations. Hence a fine grasp of every step of the PEFB-method just like a clear instruction of the computational rule of the investment decision has to be given. From that point of view, group discussion has to be managed sensitively.

Further research should focus on setting up a new survey in order to investigate the effects of facilitating the PEFB-method, the influence of hierarchies and the impact of different know-how among the interdisciplinary investment team. To ensure the validity of the results of the PEFB-method, an examination of the reliability is necessary. Currently, validation workshops provide evidence that the proved bias due to personal affection is a systematic error. Therefore, future research should focus on methods to eliminate this systematic bias through personal affection. Possible methods might be algorithms to support group process assessment and to ensure a common understanding of probability terms. In particular, artificial intelligence (AI)-based methods like Markov chain, Bayesian networks or a fuzzy-logic might be promising.

References

ACCOLA, W.L. (1994): Assessing risk and uncertainty in new technology investments. Accounting Horizons 8, pp. 19–35.

ADLER, R.W. (2000): Strategic investment decision appraisal techniques: the old and the new. Business Horizons 43 (6), pp. 15–22.

ALESSANDRI, T.; Ford, D.N.; Lander, D.M.; Leggio, K.B.; Taylor, M. (2004): Managing risk and uncertainty in complex capital projects. Quarterly Review of Economics and Finance 44, pp. 751–767.

ALEXANDER, R.A.; Cronshaw S.F. (1984): The utility of selection programs: A finance-based perspective. 92nd annual Convention of the American Psychological Associations. Toronto, Canada.

ALEXANDER, R.A.; Cronshaw S.F.; Barrick, M.R. (1986): Extending the managerial finance model of utility analysis to deal with uncertainty in parameter estimates. 1st annual conference of the Society for Industrial and Organizational Psychology, Chicago.

ALKARAAN, F.; Northcott, D. (2006): Strategic capital investment decision-making: a role for emergent analysis tool? A study of practise in large UK manufacturing companies. British Accounting review 38 (2), pp. 603-626.

ALLEMANN, J. (2002): A new view of telecommunications technologies. Telecommunications Policies 26, pp. 87-92.

ANSOFF, H. I.: (1988): The new Corporate Strategy. John Wiley and Sons, New York, p.30.

BREALEY, R.A.; Myers, S.C.; Allen, F. (2007): Fundamentals of Corporate Finance. 5th Edition. New York: Mc Graw-Hill/Irwin, Chapter 8, pp. 222-248.

CARR, C.; Kolehmainen, K.; Mitchell, F. (2010): Strategic investment decision making practise: A contextual approach. Management Accounting Research 21, pp. 167-184.

CASCIO, W.F. (2000): Costing Human Resources: The financial impact of Behaviour in organizations, fourth edition. Southwestern College Publishing, Cincinnati.

CORNELIUS, P.; Van de Putte, A.; Romani M. (2005): Three decades of scenario planning in shell. California Management Review 48, pp. 92–111.

DEMCHEK, C.C. (1992): Complexity, rogue outcomes and weapons systems. Public Administration Review 52 (4), pp. 347-355.

DEMPSEY, M.J. (2003): A multidisciplinary perspective on the evaluation of corporate investment decision making. Accounting, Accountability & Performance 9 (1), pp. 1–33.

GRAHAM, J.R.; Harvey, C.R. (2001): The theory and practice of corporate finance: evidence from the field. Journal of Financial Economics 60, pp. 187–243.

HOLLING, H. (1998): Utility Analysis Of Personnel Selection. An Overview And Empirical Study Based On Objective Performance Measures. Methods of Psychological Research Online 3 (1).

JURSCH, S.; Bischoff, S.; Hauck, E.; Flachskampf, K.H.; Jeschke, S. (2010): Value Oriented Cost-effectiveness Estimation of Innovative Intermodal Loading Unit. International Journal of Trade, Economics and Finance 1 (3), pp. 271-276.

KAPLAN, S.N.; Ruback, R.S. (1994): The Valuation of Cash Flow forecast: An empirical analysis. Working Paper No. 4724. National Bureau of economic research. Massachusetts, Cambridge, MA 02138.

MARTZOUKOS, S.H.; Trigeorgis, L. (2002): Real (investment) options with multiple source of rare events. European Journal of Operational Research 136, pp. 196–206.

MILLER, K.D.; Waller, H.G. (2003): Scenarios, real options and integrated risk management. Long Range Planning 36, pp.93–107.

MULLEN, B.; Anthony, T.; Salas, E.; Driskell, J.E. (1994): Group Cohesiveness and quality of decision making. An integration of tests of the Groupthink Hypothesis. Small Group Research, 25 (2), pp. 189-204.

NAGEL, K. (1988): Nutzen der Informationsverarbeitung: Methoden zur Bewertung von strategischen Wettbewerbsvorteilen, Produktivitätsverbesserung und Kosteneinsparung. Oldenburg.

NITZSCH, R.v. (2006): Entscheidungslehre. 1. Auflage, Aachen, Mainz GmbH.

OTT, H.J. (1993): Wirtschaftlichkeitsanalyse von EDV-Investitionen mit dem WARS-Modell am Beispiel einer Einführung von CASE. Wirtschaftsinformatik, 35 (6), pp. 522-531.

PAITCH, M.; Sterman, J.D. (1993): Boom, bust, and failures to learn in experimental markets. Management Science 39 (12), pp. 1439-1458.

PIKE, R. (1996): A longitudinal survey of capital budgeting practises. Journal of Business Finance and Accounting 23 (1), pp. 79-92.

PITTERMANN, P. (1998): Erfolgspotentialerfassung: Betriebswirtschaftliche Bewertung ganzheitlicher Veränderungsprozesse dargestellt an Fallbeispielen aus dem Qualitätsmanagement. Eschborn.

PREIßLER, P. (1991): Führen mit Kennzahlensystemen: Unverzichtbares Instrumentarium, in: Gablers Magazin 3/1991, Wiesbaden.

REICHWALD, R.; Höfer, C., Weichselbaumer, J. (1996): Bewertung von Reorganisationsprozessen. Stuttgart.

SANDAHL, G.; Sjögren, S. (2003): Capital budgeting methods among Sweden's largest groups of companies. The state of the art and a comparison with earlier studies. International Journal of Production Economics 84 (1), pp. 51–69.

SLAGMULDER, R.; Bruggeman, W.; Van Wassenhove, L. (1995): An empirical study of capital budgeting practices for strategic investments in CIM technologies. International Journal of Production Economics 40 (2–3), pp. 121–152.

SLAGMULDER, R. (1997): Using management control systems to achieve alignment between strategic investment decisions and strategy. Management Accounting Research 8 (1), pp. 103–139.

STRINA, G.; Uribe, J.; Henning, K. (2003): NOWS Technique: Participation oriented Evaluation of further Education Strategies and Measures. Proceedings of E-Learning 2003. World Conference on E-learning in Corporate, Government, Healthcare & Higher Education, pp. 331-337.

SAVELSBERG, E. (2008): Innovation in European freight transportation – basic, methodology and case studies for the European markets. Berlin, Springer-Verlag.

SCHÖNHEIT, M. (1996): Wirtschaftliche Prozeßgestaltung – Entwicklung, Fertigung, Auftragsabwicklung. Berlin.

SELL, R.; Schimweg, R. (2002): Probleme lösen: In Komplexen Zusammenhängen Denken. 6. überarbeitete Auflage. Berlin, Springer-Verlag.

UNGER, H. (1998): Organisationales Lernen durch Teams. Rainer Hampp Verlag, München.

VAN CAUWENBERGH, A.V.; Durinck, E.; Martens, R.; Laveren, E.; Bogaert, I. (1996): On the role and function of formal analysis in strategic investment decision processes: results form an empirical study in Belgium. Management Accounting Research 7 (2), pp. 169–184.

WEBB, G.K. (1994): Electronic industry model: a report on two decades of implementation. International Journal of Forecasting 10 (4), pp. 583-596.

WEYDANDT, D. (2000): Beteiligungsorientierte wirtschaftliche Bewertung von technischen Investitionen für prozeßorientierte Fertigungsinseln. Aachen, Shaker Verlag.

ZANGEMEISTER, C. (1976): Nutzwertanalyse in der Systemtechnik – Eine Methodik zur multidimensionalen Bewertung und Auswahl von Projektalternativen. Diss. Techn. Univ. Berlin 1970, 4. Aufl., München: Wittemann.

ZANGEMEISTER, C. (1994): Erweiterte Wirtschaftlichkeitsanalyse (EWA), in: Fortschrittliche Betriebsführung und Industrial Engineering 43, S. 63-71.

ZANGEMEISTER, C. (2000): Erweiterte Wirtschaftlichkeitsanalyse: (EWA) ; Grundlagen, Leitfaden und PC-gestützte Arbeitshilfen für ein "3-Stufen-Verfahren" zur Arbeitssystembewertung. 2. überarbeitete Auflage. Bremerhaven: Wirtschaftsverlag NW, Verlag. für Neue Wissenschaft.

Página deixada propositadamente em branco.

A BI-REGIONAL (RECTANGULAR) INPUT-OUTPUT MODEL FOR PORTUGAL: *CENTRO AND REST OF THE COUNTRY*

Pedro Ramos[1], Luís Cruz[1], Eduardo Barata[1], André Parreiral[2] and João-Pedro Ferreira[1]

Abstract

Regional Input-Output models aim to quantify the impacts on industry's outputs, and other economic indicators, of different final demand vectors for goods and services produced in the same or in different regions. These models are well suited for regional economic analysis as they combine inter-industrial and inter-regional economic interdependencies. MULTI2C is a general flexible procedure, developed by a group of researchers from the University of Coimbra, Portugal, that allows for the construction of that kind of models for different geographic configurations.

This work explores the construction of a bi-regional input-output model for Portugal, based on the MULTI2C approach, considering two regions: the NUT II Centro of Portugal and the Rest of the Country. This model considers rectangular matrices with 431 products and 134 industries. Further, it considers different types of households according to their main source of income, i.e., labour earnings, capital income, real estate income, retirement benefits and other social transfers. This modelling framework may be closed with respect to the consumption of different household's types, but this paper considers as endogenous the labour earnings type. Besides the model structure and the methodological choices for its construction, this work focuses on estimating interregional trade.

[1] University of Coimbra, Faculty of Economics

[2] Energy for Sustainability - University of Coimbra, Faculty of Economics

E-mails: pnramos@fe.uc.pt, lmgcruz@fe.uc.pt, ebarata@fe.uc.pt, acruzparreiral@fe.uc.pt, joao.ferreira@fe.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_12

Finally, the model is used to assess the impacts in the Centro region of Portugal, and in the Rest of the Country, derived from a shift in income's distribution in the Centro region, consisting in a reduction of the labour share, compensated by an increase in business investment, which however do not confine to the NUT II Centro of Portugal but, into some extent, spillover to the Rest of the Country.

Keywords: Input–Output Models; Household Income; Regional Economics.

1. Introduction

Regional Input-Output models aim to quantify the impacts on industry's outputs, and other economic indicators, of different final demand vectors for goods and services produced in the same or in different regions. These models are well suited for regional economic analysis as they combine inter-industrial and interregional economic interdependencies.

The initial purpose of this work is to build an input-output model for the *Centro* region (*C*), Portugal. This model uses 2010 data and analyses the interactions between the *Centro* region and the rest of Portugal, hereafter designated as "Rest of the Country" (*RC*). The *Centro* region is a NUT II located in mainland Portugal, occupying the central part of its territory (between Lisbon and Oporto) and corresponding to 31% (28 405 Km^2) of the country's total area. This region has 2.3 millions of residents (22% of the country's total) and its GDP represents about 18.5% of the Portuguese GDP. Next, the bi-regional *Centro-Rest of the Country* input-output model is explored to assess the impacts in both regions, derived from a shift in the income's distribution in the *Centro* region, consisting in a reduction of the labour share in this region, compensated by an increase in business investment.

According to the scope and objectives considered, the analysis is organized as follows. Section 2 presents, in sub-section 2.1, the structure of the bi-regional *Centro - Rest of the Country* input-output model and the main topics considered to build it up; in sub-section 2.2 are explained the main procedures regarding the consideration of different household types, according to their main source of income; sub-section 2.3 proposes a discussion on the estimation of the inter-regional trade between the two regions. Section 3 presents an application of the model to assess the impacts in the *Centro* region, and in the *Rest of the Country*, derived from a shift in the income's distribution in the *Centro* region. Section 4 concludes.

2. The Modelling framework

The bi-regional input-output model proposed in this work is an application of the MULTI2C (multi-sectoral multi-regional Coimbra model) framework. MULTI2C is a general flexible approach, developed by a group of researchers, mainly from the University of Coimbra (Portugal) that allows for the construction of input-output tables for different geographic configurations[3]. The MULTI2C approach has a great level of detail concerning both the products (or groups of products) included and the industries that produce them. The bi-regional *Centro-Rest of the Country* input-output model uses 2010 data and is focused on the Portuguese NUT II *Centro* region. Accordingly, this model considers that the 431 products included in the MULTI2C approach are produced by the 134 industries in the two different regions or are being internationally imported, i.e., part of these products are produced outside the Portuguese territory.

2.1 The structure of the bi-regional Centro-Rest of the Country input-output model

The structure of the multi-sector bi-regional *Centro-Rest of the Country* input-output model is schematically represented in Table 1.

[3] As a rule, MULTI2C models are of the bi-regional kind, as the one used here, although multi-regional structures are also considered. Sargento *et al.* (2013) have already adopted a similar framework, dividing Portugal in the "interior" and the "coast" parts of the country. The interior-coast dichotomy is again considered in Ramos *et. al* (2014). For an example of a tri-regional application please see Ferreira *et al.* (2014).

Table 1: Structure of the multi-sector bi-regional C-RC input-output model

		Products				Industries				Other Final Demand		Residues	Total
		Centro (C)		Rest of the Country (RC)		Centro (C)		Rest of the Country (RC)		Centro (C)	Rest of the Country (RC)		
Products	Centro (C)	0				IC^{CC}	HC^{CC} (Lab)	IC^{CRC}	HC^{CRC} (Lab)	OFD^{CC}	OFD^{CRC}	R^{C}	TPO^{C}
						HI^{C} (Lab)	0	0	0	0	0		THI^{C} (Lab)
	Rest of the Country (RC)					IC^{RCC}	HC^{RCC} (Lab)	IC^{RCRC}	HC^{RCRC} (Lab)	OFD^{RCC}	OFD^{RCRC}	R^{RC}	TPO^{RC}
						0	0	HI^{RC} (Lab)	0	0	0		THI^{RC} (Lab)
Industries	Centro (C)	P^{CC}	0	0		0				0			TIO^{C}
		0	THI^{C} (Lab)										THI^{C} (Lab)
	Rest of the Country (RC)	0		P^{RCRC}	0								TIO^{RC}
				0	THI^{RC} (Lab)								THI^{RC} (Lab)
Taxes less subsidies on products, falling upon intermediate consumption or final demand		0				$T(IC)^{C}$	$T(HC)^{C}$ (Lab)	$T(IC)^{RC}$	$T(HC)^{RC}$ (Lab)	$T(OFD)^{C}$	$T(OFD)^{RC}$		TT
International Imports destined to intermediate consumption or final demand						$M(IC)^{C}$	$M(HC)^{C}$ (Lab)	$T(IC)^{RC}$	$M(HC)$ (Lab)	$M(OFD)^{C}$	$M(OFD)^{RC}$		TM
Total Intermediate Consumption / Final Demand, at purchasers' prices						TIC^{C}	THC^{C} (Lab)	TIC^{RC}	THC^{RC} (Lab)	OFD^{C}	OFD^{RC}		TIC + TFD
Gross Value Added which is not directly distributed to households						$NHVA^{C}$	0	$NHVA^{RC}$	0	0	0		TNHVA
Savings and net transfers to other institutional sectors of households living mainly from labour income							S^{C} (Lab)		S^{RC} (Lab)				TS
Total		TPO^{C}	THI^{C} (Lab)	TPO^{RC}	THI^{RC} (Lab)	TIO^{C}	THI^{C} (Lab)	TIO^{RC}	THI^{RC} (Lab)	OFD^{C}	OFD^{RC}		

Legend:

C - Centro

RC - Rest of the Country

IC^{ij}, *i,j = C, RC* - Intermediate consumption of i's regional products, used by j's industries

HC^{ij} (Lab), *i,j = C, RC* - Final consumption of i's regional products, consumed by households mainly dependent from labour income living in region j

OFD^{ij}, *i,j = C, RC* - Other final demand for i's regional products, used in region j

R^{i}, *i= C, RC* – Residues on the demand of products in region i

HI^i (Lab), $i = C, RC$ - Region i's households income distributed to the households that live mainly from their labour compensations (it includes the mixed income distributed to the self-employed workers)
TPO^i, $i = C, RC$ - Total output of products produced in region i, at basic prices
P^{ii}, $i = C, RC$ - i's regional products, according to their production industry (generic element of the i's supply table)
TIO^i, $i = C, RC$ - Region i's total industry output, at basic prices
$T(g)^i$, $i = C, RC; g = IC, HC, OFD$ - Taxes less subsidies on products, falling upon g, in region i
TT - Total taxes less subsidies on products
$M(g)^i$, $i = C, RC; g = IC, HC, OFD$ - International imports destined to use g, in region i
TM - Total International Imports
TIC^i, $i = C, RC$ - Total intermediate consumption by industries, in region i, at purchaser's prices
THC^i (Lab) , $i = C, RC$ - Total region i's consumption by households mainly dependent on labour income, at purchaser's prices
THI^i (Lab) , $i = C, RC$ - Total region i's households income distributed to the households that live mainly from their labour earnings
OFD^i, $i = C, RC$ - Other final demand in region i, at purchaser's prices
TIC + TFD - Total intermediate and final demand, at purchaser's prices
$NHVA^i$, $i = C, RC$ - Gross Value Added which is not directly distributed to households, in region i
TNHVA - Total Gross Value Added which is not directly distributed to households
S^i, $i = C, RC$ - Savings and net transfers to other institutional sectors of the households that live mainly from their labour income, in region i
TS - Total savings and net transfers to other institutional sectors of the households that live mainly from their labour income

The structure of the model in Table 1 is based on a set of characteristics and hypothesis, which leads to the classification of this modelling approach as a "closed rectangular bi-regional input-output model", using domestic flows at basic prices. Below are presented the key procedures and the main implications regarding (i) the use of domestic flows, expressed at basic prices; (ii) the rectangular model structure, *(iii)* the regional production technologies and, finally, *(iv)* the endogenization of households consumption, namely those that live mainly from labour income.

(i) Domestic flows, expressed at basic prices

The model considers primarily domestic flows (unlike the National Accounts Supply and the Use Tables, from where it is derived (INE, 2012a), which favour an accounting approach at "total flows"). This means that this model focuses on the products of the industries operating within the Portuguese economy, i.e., the rows and the columns for each of the 431 products include only the products actually produced in Portugal (in *C* and/or *RC*) excluding the internationally imported uses.

The model is at basic prices. Total products output (TPO) and total industries output (TIO) are evaluated at basic prices (nevertheless the industries output (TIO) includes, as usually, intermediate consumption at purchasers' prices - IC); total final demand is also considered at purchasers' prices, although the final demand of each product is consistently considered at basic prices; finally, trade and transport margins are considered as inputs provided by trade services (wholesale, retail or specifically motor vehicles or fuels) or transport services (by different types of freight).

(ii) Rectangular model (industry-technology assumption)

The model considers 431 products and 134 industries, therefore allowing each industry to produce more than one product, whether they are primary or secondary products (i.e., that are main products of other industries). The rows of matrices P^{CC} and P^{RCRC} describe the products produced by each industry in *C* and *RC*, respectively. Regarding primary products, whenever each industry produces more than one product, it is used more specific information from different sources about the actual structure and major dominant products of each industry in *C* and *RC*. Concerning secondary products, we assume the same weight in total industries production in each of the regions (note that those products represent a non-significant share of the total output).

Rows corresponding to products (431 products × 2 regions) describe their different destinations, which include: the intermediate consumption (IC) in each region (naturally, a product produced in *C* can be inter--regionally exported and used as intermediate consumption in *RC*); the final consumption of the different types of households in both regions; and other destinations in the "Other Final Demand".

Columns corresponding to industries describe their technologies in absolute values, i.e., each product intermediate consumption in each industry, according to the origin's region (*C* or *RC*); the intermediate inputs internationally imported (although in this case, the total inputs are not disaggregated by products); the (non-deductible) taxes less subsidies falling upon the purchased inputs (in order to assure that each industry IC is expressed at purchaser prices); the income generated in each industry and in each region, i.e., the gross value added (GVA), whether it is directly distributed to households living mainly from their labour income, or distributed to some other institutional sector (NHVA).

This rectangular bi-regional input-output model admits that each industry has its own technology, identically to the production of all its primary or secondary products (as explained above secondary products represent only a residual value in total industry production)[4].

(iii)Regional production technologies

Moreover, in general, it is assumed that each industry has the same production technology in both regions, i.e. each input has the same weight in the intermediate consumption regardless the production place.

The equal production technology assumption in both regions is not the most accurate for all the cases, because it does not take into account the disparities of activities within some industries and/or regional differences. E.g., the original information provided by the Portuguese National Accounts Supply and Use Tables (INE, 2012a) considers one column vector of technologies for the electricity industry. However, this industry is composed by different activities such as production and distribution. Further, electricity can be generated by a wide variety of sources, encompassing a set of activities with different production technologies. Thus, as the regions possess different structures of electricity production and distribution there is the need to develop procedures to reflect such regional disparities.

Accordingly, we consider a division of the original electricity industry into 10 different industries. The electricity industry vector of production and intermediate consumption, for the 431 products, provided by the Portuguese National Accounts Supply and Use Tables (INE, 2012a), was

[4] See Sargento *et al.* (2011) for a discussion on the (dis)advantages of rectangular input-output models. Deeper descriptions of this kind of structure can be seen in the pioneering contribution of Oosterhaven (1984) and in Miller and Blair (2009: Chapter 5).

taken as the primary data. Then, to disaggregate this vector by the different activities and regions, we proceed to the estimation of each activity's total production and intermediate consumption. The data sources used to estimate such values (for 2010) were: the Portuguese National Statistical Institute (INE, 2015a, 2015b), which supplies information on total electricity production, as well as on total electricity produced by cogeneration plants; the database *Quadros de Pessoal*, from the Statistical Department of the Ministry of Labour and Social Solidarity (GEP-MTSS, 2011), which provides information on the number of workers in each activity; and the reports published by companies that manage the thermoelectric power plants (EDP Produção, 2011, 2012, 2013; Turbogás/Portugen, 2011; Tejo Energia/Pegop, 2011; EEM, 2008, 2011) to estimate the electricity produced by conventional sources, as well as the energy inputs used in such production; lastly, additional data on the power plants in operation throughout the country was gathered from Pinto and Faria (2015).

The disaggregation of the electricity industry considers one industry of electricity distribution and 9 industries of electricity production according to the following sources: 1) wind; 2) geothermal; 3) hydro; 4) photovoltaic; 5) coal; 6) fuel oil; 7) natural gas; 8) diesel; 9) cogeneration. Each of these sub-industries has its own production technology. The total electricity sector, that merges the 10 sub-industries, produces two distinct products (in each region): produced electricity and distributed electricity. Further, a single column vector of technologies for the refined petroleum industry would hide the regional disparity of its input structure. Indeed, while some regions require more goods and services related to the management activities, others consume mainly inputs directly related to the production process of refined products. Accordingly, we assume diverse production technologies for the different regions, based on Regional Accounts' information (regarding remunerations, Gross Value Added and production of the different regions), taking into account our knowledge of the actual location of the refineries in Portugal.

(iv) Closed model

The model considers, both in *C* and *RC*, different household's types, according to their main source of income, namely: labour earnings, capital income, real estate income, pensions and other social transfers.

The model is "closed" regarding the consumption of households that live mainly from labour income (employees or self-employed workers), i.e., labour income endogenously determines this group's final consumption. The income generated in each region contributes only for the consumption of households living in the same region; commuting and other periodical or seasonal migrations between *C* and *RC* (that are negligible between these regions) were not considered. Consumption of other household's types (the non-labour income dependent ones) is considered exogenous, i.e., their consumption expenditures are independent of the generation of productive income (which we do not know how, where and when is distributed to these families), and therefore considered as part of the Other Final Demand.

The Other Final Demand includes the consumption of other household's types (the non-labour income dependent ones), the consumption expenditures of general government and non-profit institutions; the investment (i.e., demand for products used as investment goods, produced in the country, allocated to *C* or *RC* according to the place of production), the consumption of non-residents in Portugal that visit both regions and, finally, other international exports of goods and services.

Residential and business rents paid to households are not considered as benefiting those living mainly from labour income, but are instead included in the NHVA vector. Thus, an increase in these incomes does not automatically induce an increase in consumption (as the consumption of the other household's types is deemed to be exogenous).

Finally, it is also important to mention that the portion of Table 1 inside the bold border - a square matrix of dimension 1134 (431 products, 134 industries and 2 extra rows relating to household income, for each one of the two regions) - is the core of the input-output framework implemented. Indeed, one departs from this core to compute the inverse matrix, which comprises a set of multipliers that measure impacts of exogenous final demand changes on products and industries production. Also, this inverse matrix includes the impacts on the income of the households that live mainly from their labour earnings, caused by those shocks. Noteworthy, it is also possible to assume exogenous shocks on such income and compute their effects on products/industries outputs. Further, exogenous final demand shocks can be formulated either in terms of products, or by redirecting them to industries.

2.2 The consumption structure of the different household types

The multi-sector bi-regional *Centro-Rest of the Country* input-output model distinguishes among private consumption structures, by different household types, according to their main source of income, namely: *(i)* labour earnings (wages and mixed income), *(ii)* capital income, *(iii)* real estate income, *(iv)* retirement benefits (pensions) and *(v)* other social transfers. The technical details on the procedures used to derive the consumption structure for these different household types are briefly mentioned ahead.

The vector of national household's final consumption, for the 431 products, provided by the Portuguese National Accounts Supply and Use Tables (INE, 2012a), is taken as the prime data. Then, we proceed to its disaggregation, firstly by estimating the total households' consumption by each household type, in *C* and *RC*. The data sources used to estimate such values are: the "Household Budget Survey 2010-2011" (HBS) (INE, 2012b), which provides information on the consumption expenditure per household type; the "Census 2011" (INE, 2012c) for data concerning the number of individuals per region and per household type (in order to reweight the sample of the HBS); and the regional *per capita* "Local Purchasing Power Index" (INE, 2013), also referring to 2011, to decide upon the total relative consumption between the two regions. Secondly, we estimate the consumption of the 431 products by each household type, in *C* and *RC*. As the HBS considers data disaggregated only for 200 products, such information is used as a primary data source and then to estimate the corresponding disaggregation for the more detailed 431 products level (as considered by the National Accounts), it is assumed that the relative significance of several sub-products is the same in the 2 regions and for all the household types, namely identical to the corresponding shares in the National Accounts.

Accordingly, Table 2 synthesises the estimated consumption structures, by household type (including the different values of the actual residential rents paid by these groups and the consumption of residents outside the Portuguese territory), in *C* and *RC*. This table refers to resident household's consumption, at purchasers' prices, including also the consumption of internationally imported goods and services. Table 2 considers the aggregation of the 431 products into 10 groups of products.

Table 2: Consumption structure by household type (%)

Region	**Centro (C)**				**Rest of the Country (RC)**			
Household's main source of income Products	Wages	Capital & real estate income	Pensions	Other social transfers	Wages	Capital & real estate income	Pensions	Other social transfers
Agriculture, forestry and fishing products	3,86	3,93	6,54	4,98	3,60	3,97	6,09	4,95
Food industry products	15,61	13,42	20,52	25,15	17,20	13,54	20,88	24,99
Other products of industry and construction	33,70	27,12	28,32	23,91	30,80	27,36	25,42	24,08
Energy, water supply and sewerage	4,07	3,99	6,52	5,97	3,79	4,03	5,66	5,94
Accommodation and food services; Wholesale and retail trade, repair of motor vehicles and motorcycles	11,37	12,89	8,47	8,68	11,80	13,00	9,06	8,63
Transportation and storage; information and communication products	7,20	6,99	6,58	9,94	7,91	7,05	7,42	9,93
Financial, insurance and real estate services	7,63	6,23	5,21	5,08	6,71	6,29	5,82	5,05
Other services	12,92	21,40	14,06	11,22	13,56	21,26	15,55	11,11
Housing rents	1,68	1,24	1,61	4,59	2,43	1,25	2,02	4,68
Resident's expenditures abroad	1,95	2,80	2,17	0,47	2,20	2,26	2,07	0,64

These data confirms the importance of taking into consideration distinct consumption structures of the different households' types, as well of distinguishing it for the *C* and *RC* regions. Indeed, firstly, it is possible to identify that the highest proportion of consumption expenditure, for the majority of household types in both regions, concerns to "Other products of industry and construction". Further, in both regions, households mainly depending on income from "Capital and real estate" are leaders in the relative importance of "Other services", "Resident's expenditures abroad" and "Accommodation and food services; Wholesale and retail trade, repair of motor vehicles and motorcycles", while households depending predominantly from "Pensions" and from "Other social transfers" stand out for "Agriculture, forestry and fishing products" and for "Housing rents" and "Food industry products", respectively. Secondly, the products included in the categories "Agriculture, forestry and fishing products" and "Energy, water supply and sewerage" have higher relative values in the *C* than in the *RC* region, for the majority of the household types. Conversely, products included in the categories of "Food industry products", "Accommodation and food services; Wholesale and retail trade, repair of motor vehicles and motorcycles", "Transportation and storage; information and communication products" and "Housing rents" stand relatively higher in the families of the *RC*.

Finally, as mentioned in subsection 2.1, there is the need to convert the information in Table 2 from purchasers' prices to basic prices and from "total flows" to "domestic flows". Accordingly, from the matrix at purchasers' prices, and for each product, are successively removed: the Value Added Tax (VAT); other taxes less subsidies; the wholesale and retail margins; and international imports. Basically, the margins are reclassified to the trade and transport industries, being considered as household's consumption of those industries products. Taxes less subsidies are removed from each product and inserted at the bottom part of Table 1 ($T(g)^i$). A similar procedure is used regarding international imports ($M(g)^i$). Further, it is assumed equal propensity to (internationally) import in both regions and for each of the 431 products.

2.3 The Centro - Rest of the Country inter-regional trade

The estimation of inter-regional trade is perhaps the most critical issue in building up regional input-output models. This task is accomplished, in a first approach, by the "residual method"[5]. This method is based on the assumption that if there is more supply (including international imports) than demand (including international exports) of one product in a given region, this product is inter-regionally exported (in net terms) by this region. Thus, all the components of each product's supply and demand are estimated for the region and the difference gives us the interregional net trade. For bi-regional models net flows are symmetrical between the two regions (and, obviously, the export's destination region and the import's origin region are known for each case).

Further, it is critical to have also information on exports and imports gross values (at least regarding the endogenous part of the model)[6]. The approach here implemented firstly estimates gross imports. The method for estimating gross imports is based on a detailed product classification that relies on several hypotheses and the expertise of the MULTI2C team members, namely their effective knowledge of the Portuguese reality. Three different product types (A, B and C) and corresponding assumptions are considered regarding the determination of gross inter-regional imports, as follows:

Type A products, regionally non-tradable. These products have necessarily to be produced in the same region where they are consumed; the inter-regional imports of these products are zero in both regions, as there is no inter-regional trade; the residue between these products supply and demand is not significant in the majority of the products but it may not be

[5] This is the general methodology adopted in the MULTI2C models. The methodological issues and corresponding procedures are described in Barata *et al.* (2011), where the aim was to estimate inter-regional gross imports, for single region models (for small Portuguese NUTS III regions, in the interior part of the country), and developed in Ramos *et al.* (2013) already in a bi-regional model context.

[6] This is the well-known *crosshauling* problem, firstly approached by Robinson and Miller (1988).

exactly 0 (it has a positive value in one region and the symmetric value in the other). These residues are included in the R^i column of Table 1.

Type B products, fully internationally and inter-regionally tradable. These products move around the two regions at negligible costs; there is no reason for a local delivering preference; thus, demand is assumed to be satisfied by local supply or by imports from the other region according to the proportion of the regional products output.

Type C products, regionally tradable between specific locations. This "intermediate" category includes mainly products with high transportation costs or with a strong regional preference; they also include some products for which the demand is usually locally manifested, but satisfied by national companies - a significant part of these product's production process takes place in the company's (national or regional) headquarters, often located in *RC* – we call this the *"*headquarters effect*"* –, and this *"*headquarters*"* participation on the production process is considered equivalent to an inter-regional export of the headquarter region to its establishment region.

These procedures generate a first estimate of the gross inter-regional trade in both regions. Gross imports depend on the type of product classification, net inter-regional exports are previously known and gross exports are obtained residually. However, contrarily to what is observed in net inter-regional trade, after this, most of the product's gross imports do not match with the gross exports of the same products in the other region. It is therefore essential to consider a final adjustment, consisting on a simultaneous increase of one product's inter-regional exports and imports (respectively XIR and MIR) in one region, combined with a simultaneous reduction of the same product's exports and imports, in the other region, until the interregional trade gross flows are equal (i.e., XIR (i) = MIR (j), *i,j = C, RC*). The distribution of these two adjustment weights is associated with the product's relative output in both regions.

Table 3 shows the 10 main products (out of P88 aggregation of the 431 original products) regarding gross and net exports from the *C* region, which can be considered as forming its economic basis. It is important to note that in the cases where export's destination is investment or consumption expenditure of the general government or of non-profit institutions (which are not estimated by region of demand, but for the country as a whole) the gross exports do not include inter-regional exports to *RC*. International re-exportation of imported goods is equally excluded as they were not "distributed" through the two regions.

Table 3: Main products (international and inter-regional) exported from Centro region

Unit: 10^6 € (2010)	***Gross exports***	*share (%) of total C region gross exports*	***Net exports***
Production, distribution and trade of electricity	2240.91	9.2	394.52
Fabricated metal products, except machinery and equipment	1385.94	5.7	600.08
Pulp, paper and paperboard (excluding corrugated)	1245.77	5.1	997.26
Agriculture, farming of animals, hunting and related services	1137.82	4.7	129.89
Food and beverage services	1100.52	4.5	497.50
Freight transport by road and removal services	1028.43	4.2	798.08
Plastics products	805.09	3.3	414.05
Accessories for motor vehicles	783.74	3.2	208.81
Basic chemicals, fertilizers and nitrogen compounds, plastics and synthetic rubber in primary forms	777.10	3.2	-419.20
Refractory ceramic products; Ceramic building materials; Other porcelain and ceramic products, non-refractory	678.75	2.8	637.13
…	…	…	…
Total Exports from Centro region	24232.04	100.0	-1189.35

Table 3 data reveals that the three major gross exporting industries in the *C* region are "Production, distribution and trade of electricity", "Fabricated metal products, except machinery and equipment" and "Pulp, paper and paperboard". However, regarding net exports, only "Pulp, paper and paperboard" remains on the top three and it is relevant to note the role of "Freight transport by road and removal services" and "Refractory ceramic products; Ceramic building materials; Other porcelain and ceramic products, non-refractory" in the *C* region economic basis.

3. The impacts of income (re)distribution in the Centro region of Portugal

The bi-regional *Centro - Rest of the Country* input-output model is here considered to assess the impacts, in both regions, resulting from a shift in income's distribution in the *C* region. More specifically, it is assumed a 5% reduction in the labour earnings in *C*, with such amount being "relocated" in gross fixed capital formation (GFCF). Naturally, the GFCF increase is not confined to the *C* region, but it is expected to spread all over the country. This happens because, contrarily to what generally happens with labour income, the application of the capital remuneration is not tied to the place/region where the production occurs and the operating surplus is perceived. Actually, employees generally live in the place/region where they work, but the same may not happen with capital holders as they usually spend their income in investments through different parts of the country, regardless their living place. Further, in this modelling framework what is actually relevant is where the production of the investment goods takes place, which may happen anywhere in the Portuguese territory, or even abroad, and not necessarily at the actual location of the investment.

The results of this modelling exercise indicate that the shift in income's distribution would generate a net expansionist effect on the Portuguese economy. However, concerning *Centro* region's total production, the model foresees a reduction of approximately 116 million € (0.2% of total output at basic prices), which is expected to be balanced by an increase of approximately 1000 million € in the *Rest of the Country*'s total production (0.4% of total output). Table 4 shows how these effects are distributed in the *Centro* region, emphasising the products whose production would change the most, as a consequence of this shock.

Table 4: Centro region's products with higher (+ / -) changes in production

(Year: 2010)	**Absolute** (10^6 €)	**Relative** (%)
Higher increases		
Development of building projects; Construction of buildings	44.08	2.0
Civil engineering	38.51	2.2
Specialized construction	24.19	2.1
Fabricated metal products, except machinery and equipment	16.19	0.8
Cement, lime and plaster; Articles of concrete, cement and plaster	11.42	1.5
Total change in products whose production increased	177.26	0.8
Higher reductions		
Monetary intermediation	-13.57	-1.4
Wholesale and retail trade and repair of motor vehicles and motorcycles	-14.13	-1.6
Food and beverage services	-23.74	-1.2
Renting of own or leased real estate	-58.75	-2.4
Retail trade, except of motor vehicles and motorcycles	-61.19	-1.9
Total change in products whose production decreased	-293.59	-0.8
Total net change in Centro region's production	**-116.33**	**-0.2**

To analyse Table 4 results it is important to note that the output decrease in *Centro* region happens because the initial income reduction occurs for *Centro* region's households that depend mainly from their labour income, and that these households spend most of their money in the region where they live. Additionally, a portion of this income is spent on non-tradable (*type A*) products, which have to be produced in the *Centro* region itself. As expected, the products with higher production increase are those that are (directly or indirectly) linked to investment, while the higher reductions are mainly associated with household's consumption expenditure (that is deemed to depend on labour income). The overall (negative) impact of this shock in the *Centro* region shows that the (negative) effect on the consumption products' output outweighs the (positive) impact resulting from investment and related goods' output.

Further, it is worth to highlight that the household's saving rate and transfers to other institutional sectors are relatively high, such that only a portion of household's income is applied into consumption expenditures. Contrarily, when household's income is considered to be redistributed in benefit of firm's profits, it is assumed that this amount is fully invested, therefore generating a net expansionist effect for the country as a whole.

This scenario may be criticised and considered as extreme. Indeed, we assume that the labour compensations reduction is fully converted into more investment (except for the part corresponding to taxes on products falling upon the investment goods), but actually only a fraction of household's income is available for consumption – e.g., a part of the revenue retained by companies (NHVA) is used in income taxes or in more transfers to other institutional sectors. Accordingly, we have tested the sensitiveness of these modelling results to this assumption, by considering an additional scenario, where only 80% of the total labour income's reduction in *Centro* region is compensated by an increase in GFCF. Even under this new milder assumption, the national expansionist impact remained, namely because household's income that is not consumed (so, it is saved or transferred to other sectors) is about 40%, while we assume not to invest 20% of firm's profits. Further, this new scenario estimates an output decrease in *Centro* region of about 179 million € (0.30% reduction against the 0.20% in the initial scenario) strengthening the negative effect on *Centro* region's production, while the production growth in the *Rest of the Country* is estimated to be approximately of 711 million € (an increase of 0.27% comparing with 0.38% in the initial scenario).

Noteworthy, in general, these scenarios confirm that regions benefit from income distributions that favour workers (because they generally spent their income in the same region), while are negatively affected by income distributions that benefit capital holders (as it is not certain where the increased profits that benefit capital, after the income redistribution, are really invested, and where the corresponding investment goods are produced). These research results are in accordance with those proposed in Ramos *et al.* (2011), when estimating the impact on employment for different scenarios concerning the distribution of productivity gains among employees and capital owners, in Portuguese depressed peripheral regions.

4. Conclusions

The leading ambition of this paper is to establish the structure of a bi-regional input-output model for the *Centro* region of Portugal and the *Rest of the Country*. The characteristics of this modelling framework allow assessing the impacts of a shock occurring in the *Centro* region, not only on the region itself, but also the spillover effects that leak to the *Rest of the Country*. These spillovers return then to *Centro* region in the form of a feedback effect. These leakages and feedbacks are of two kinds: *(i)* interindustry effects, due to the fact that one region's intermediate consumption demand is often satisfied by supplies from other regions, and *(ii)* income effects, as this kind of interregional contamination also occurs with household's consumption (namely of those living mainly from labour income). In this modelling approach, the interregional effects are influenced not only by the private consumption structures in each region but also by the relative weight of different household types, according to their main source of income.

The major conclusion is that the shock effects do not confine to the place where they hit. This is why multi-regional models are vital in regional analysis. We cannot assume that all the benefits (or costs) of a regional occurrence are locked into that region. Sometimes major upshots befall somewhere else and they may even come about elsewhere in a different manner.

This research estimates the impact in the *Centro* region and nationwide of an income distribution shock hitting directly this region only. More specifically we analyse a 5% redistribution of the income generated in the *Centro* region, considering a reduction in the income distributed to the households that live mainly from their labour income, and transferring it to the firms, which are supposed to spend it in investment. The results indicate that this shift in income distribution has an expansionist effect in the Portuguese economy, namely because the share of this additional firm's profit invested is higher than the share of household's income that is consumed. However, remarkably, the *Centro* region should not benefit with this income redistribution towards firms. Actually, the *Rest* *of the Country* takes the gain, while the *Centro* holds the burden. The arguments to explain this result include the fact that household's consumption is largely concentrated on the region where those households live (and work), whereas investments are usually spread all over the country. Thus, overall, this analysis claims that regions themselves may benefit from income distributions that favour workers but are negatively affected by income distributions that benefit capital holders.

Acknowledgments: This work has been framed under the *Energy for Sustainability* Initiative of the University of Coimbra and supported by the R&D Project *EMSURE - Energy and Mobility for Sustainable Regions* (CENTRO-07-0224-FEDER-002004).

References

BARATA, E.; Cruz, L.; Sargento, A.; Ramos, P.; Ferreira, J.-P. (2011) "Deriving Regional Input-Output Matrices to Assess Impacts in Small Portuguese Peripheral Regions", in Cámara, A., Cardenete, M., Medina, A., Monrobel, J. (eds.), *Sectores estratégicos para un nuevo modelo económico*, IV Jornadas Españolas de Análisis Input-Output, Universidad Rey Juan Carlos, Madrid.

EDP PRODUÇÃO (2011) *Environmental Statement 2010 – Sines Thermoelectric Power Plant (Portuguese only)*, Directorate of Thermal Production - EDP Gestão da Produção de Energia, S. A., Sines, Portugal.

EDP PRODUÇÃO (2012) *Environmental Statement 2011 –Ribatejo Thermoelectric Power Plant (Portuguese only)*, Directorate of Thermal Production - EDP Gestão da Produção de Energia, S. A., Carregado, Portugal.

EDP PRODUÇÃO (2013) *Environmental Statement 2012 – Lares Thermoelectric Power Plant (Portuguese only)*, Directorate of Thermal Production - EDP Gestão da Produção de Energia, S. A., Figueira da Foz, Portugal.

EEM (2008*) Vitória Thermoelectric Power Plant (Portuguese only)*, Empresa de Eletricidade da Madeira, S.A. – Directorate of Production Services, Funchal, Portugal.

EEM (2011) *Annual Report 2010*, Empresa de Eletricidade da Madeira, Funchal, Portugal.

FERREIRA, J.; Ramos, P.; Cruz, L.; Barata, E. (2014) "Spill-over effects in the Portuguese economy: Lisbon Metropolitan Area vs. Rest of the Country", *22nd International Input-Output Conference,* Lisbon, Portugal.

GEP-MTSS (2011) *Quadros de Pessoal (Portuguese only)*, Database of the Office for Strategy and Planning – Ministry of Labour and Social Solidarity.

INE (2012a) *Portuguese National Accounts 2010 (Base 2006)*, Statistics Portugal – National Accounts, Lisbon, Portugal.

INE (2012b) *Household Budget Survey - 2010/2011 (Portuguese only)*, Statistics Portugal, Lisbon, Portugal.

INE (2012c) *Census 2011*, Database of the Official Statistics Portal, Statistics Portugal, Lisbon, Portugal.

INE (2013) *Study on the Local Purchasing Power 2011*, Statistics Portugal, Lisbon, Portugal.

INE (2015a) *Gross production of electricity (kWh) by Geographic localization (NUTS - 2002) and Type of electricity production,* Database of the Official Statistics Portal, Statistics Portugal, Lisbon, Portugal (available at http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0002106&contexto=bd&selTab=tab2; accessed in 9/1/2015).

INE (2015a) *Electricity production (kWh) in cogeneration plants by Geographic localization (NUTS - 2002)*, Database of the Official Statistics Portal, Statistics Portugal, Lisbon, Portugal (available at http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0002011&contexto=bd&selTab=tab2; accessed in 9/1/2015).

MILLER, R.; Blair, P. (2009) *Input-Output Analysis – Foundations and Extensions*, 2nd Ed., Cambridge University Press, Cambridge, UK.

OOSTERHAVEN, J. (1984) "A Family of Square and Rectangular Inter-Regional Input Output Tables and Models", *Regional Science and Urban Economics*, 14, 565-582.

PINTO, R.; Faria, F. (2015) *Power plants in operation (Portuguese only)*, Wikienergia – Electricity Museum/EDP, Lisbon, Portugal (available at http://wikienergia.com/~edp/index.php?title=Categoria:Centrais_el%C3%A9ctricas_em_actividade; accessed in 9/1/2015).

RAMOS, P.; Barata, E.; Pimentel, A. (2013) "Um Modelo Input-Output Bi-Regional Litoral-Interior para Portugal: metodologia de construção e alguns resultados sobre a estimativa de comércio inter-regional", in Santos, J., St-Aubyn, M., Lopes, J., Santos, S. (coord.), *Livro de Homenagem a João Ferreira do Amaral*, Almedina, Coimbra.

RAMOS, P.; Barata, E.; Cruz, L.; Sargento, A. (2014) "An Input-Output Model with Resources-Constrained Sectors: An Application to the Agri-Food Development Strategy in the Context of a Portuguese Bi-Regional Model", *61st Annual North American Meetings of the Regional Science Association International*, Washington D.C., USA..

RAMOS, P.; Castro, E.; Cruz, L. (2011) "Economically Sustainable Demography: Reversing Decline on Portuguese Peripheral Regions", *19th International Input-Output Conference,* Alexandria, USA.

ROBINSON, H.; Miller, J. (1988) "Cross-Hauling and Nonsurvey Input-Output Models: Some Lessons from Small Area Timber Economies", *Environment and Planning A*, 20, 1523-1530.

SARGENTO, A.; Ramos, P.; Barata, E.; Cruz, L. (2013) "Regional planning insights from a Portuguese bi-regional Input-Output model", International Conference on Economic Modelling - *EcoMod2013*, Prague.

SARGENTO, A.; Ramos, P.; Hewings, G. (2011) "Input-Output Modelling based on Total-Use Rectangular Tables: Is this a Better Way?" *Notas Económicas*, 34, 8-34.

TEJO ENERGIA/PEGOP (2011) *Environmental Statement 2011 – Pego Thermoelectric Power Plant (Portuguese only)*, Tejo Energia - Produção e Distribuição de Energia Elétrica, S.A., and PEGOP - Energia Elétrica, S.A., Abrantes, Portugal.

TURBOGÁS/PORTUGEN (2011) *Environmental Summary 2010 - Tapada do Outeiro Combined Cycle Power Plant*, Turbogás - Produtora Energética S.A. and Portugen - Energia, S.A., Gondomar, Portugal.

Página deixada propositadamente em branco.

Effects of sectoral aggregation on an input-output table

Carmen Ramos[1]

Abstract

The importance of the consequences of sectoral aggregation, from the perspective of input-output analysis is the foundation of this paper. This issue is addressed from two perspectives: firstly, the effects of aggregation on the amount of information contained in an input-output matrix, in this case the statistical theory of information is used (Theil, 1957).

Moreover, the changes experienced by the output multipliers due to different sectoral aggregations are analyzed. In this sense we have considered various proposals as the hypothetical extraction (Dietzenbacher and Van der Linden, 1997) and Pure Linkage (Sonis and others, 1995).

The theoretical results derived from this study will be applied to the Input-Output Table for Asturias (2010). We have also considered different standard aggregations as NACE or CIIU, among others.

Keywords: Input-Output Analysis, Sectoral Aggregation, Loss information, Pure linkage

[1] University of Oviedo (Spain)
E-mail: cramos@uniovi.es

http://dx.doi.org/10.14195/978-989-26-1039-9_13

1. Introduction

The subject of aggregation in the input-output analysis may seem obsolete and overcome, since computers and specific computer programs enable users to work with almost any level of disaggregation, hence, why worry about the aggregation when there is the possibility to carry out studies with a very large number of sectors? Nonetheless, this approach is a bit simplistic due to the fact that the performance of economic analyses often requires the use of different statistical sources, and they are usually presented with different aggregations, which makes it necessary to homogenize (aggregate) them. Input-output analysis is similar in this respect: it is very common to have input-output tables with a certain number of sectors whose information needs to be "combined" with other data sources.

On a separate issue, the construction of symmetric matrices requires the joint use of supply and use tables, which previously need to be "made" square, aggregating in an analogous way the products and the industries.

The problems of aggregation have been studied by different authors such as Leontief (1951) himself, Hatanaka (1952), Theil (1957), Ara (1959), Tilanus y Theil (1965), Morimoto (1970), Blin and Cohen (1977), Doblado (1988), Lauritzen (1989) and Russo (2001), among others.

In this study we are going to address the subject of sectoral aggregation from the perspective of the effects it causes, in a double area: focusing on the information itself contained in an input-output table, considering how this one varies when it has more or less aggregated matrices and, on the other hand, focusing on the results of the structural analysis derived from the calculation of the multipliers.

The theoretical results derived from this study will be applied to the Input-Output Table for Asturias (2010) published by SADEI[2].

---

[2] Asturian Society of Economic and Industrial Studies.

2. Statistical theory of information: some concepts

It seems reasonable to assume that the more disaggregated an input-output table is, the more information it provides, as a higher number of data about the economic reality of a country or a region is available. Therefore, using matrices with a smaller number of sectors implies a loss of information that can be more or less considerable depending on how this aggregation is being performed and on the initial distribution of the matrix. In order to analyse this loss of information we are going to use the so-called statistical theory of information.

The statistical theory of information was initiated, among others, by Shannon, who in 1948 proposed entropy as a measure of the amount of information, that is to say, of the disorder or the uncertainty in a certain system. The entropy associated with a random variable *X* is related to the probability distribution of the variable and it shows how predictable the process subject to uncertainty is.

Let *X* be a discrete random variable whose associated probability distribution is $P= (p_1, p_2,\ldots p_n)$. We call entropy of the random variable *X* or of the distribution $P= (p_1, p_2,\ldots p_n)$ the expression

$$H(X) = H(p_1, p_2 \ldots . p_n) = -\sum_{i=1}^{n} p_i \log p_i \quad (1)$$

This measure can be interpreted as the average information provided by each element ($\log p_i$). The entropy is non-negative, it will take the highest value in the event that the distribution is uniform, in this case, $p_i = 1/n$ and $H(X)=\log n$. In other words, if all the sectors have a very similar (the same) volume of sales or of purchases, the input-output table will be more informative, since the probability of a purchase (sale) taking place is similar for all the sectors, hence it is "difficult" to predict from which ones it comes from. Nevertheless, if the distribution happened to be degenerate[3] the entropy would reach its lowest possible value, 0, as in that case a single sector would be the one to make all the purchases

[3] A distribution is degenerate when all the values of a variable have null probability, except for one with unitary probability.

(sales), so the origin of the transactions would be easily identifiable (predictable) and the information contained in the table would be null.

Since we are going to work with input-output tables, and consequently with double-entry tables, it will be necessary to consider not a single random variable, but a two-dimensional magnitude (X,Y), where X represents the sectors that make purchases and Y, the sectors that sell. In that case, Shannon entropy will be formulated as follows

$$H(X,Y) = -\sum_i \sum_j p_{ij} \log p_{ij} \quad (2)$$

where p_{ij} represents the joint probabilities associated with the two-dimensional random variable (X,Y). The interpretation and bounds of this measure are analogous to those indicated for a one-dimensional variable.

3. Aggregation in input-output tables

As it has been indicated before, it is often necessary to aggregate the input-output tables. In this regard, consider situations in which it is essential to combine different sources of information where the sector classifications are different or when we use supply and use tables, which are rectangular since they have different classification in rows and columns.

3.1. Aggregation and amount of information

Consider the matrix **Z** of intermediate demand

$$\mathbf{Z} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{bmatrix} \quad (3)$$

where x_{ij} represents the purchases that the sector j makes with regard to the sector i.

In order to apply the Information Theory in the input-output context, we will consider X and Y as two qualitative variables: X would represent the buying sectors, while Y would represent the selling sectors. The values that appear in the intermediate demand matrix contain the purchases/ sales between the different sectors or the number of times that sectors i and j interact (through economic exchanges). From these values we can obtain a relative frequency table, by only dividing each cell between the total of the table, that is[4],

$$f_{ij} = \frac{x_{ij}}{\sum_i \sum_j x_{ij}} \quad (4)$$

Therefore, from the matrix **Z** presented before, we can obtain the following one

$$\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix} \quad (5)$$

where f_{ij} are the relative frequencies associated with the table **Z**[5].

The most frequent notion of probability states that the probability is an ideal number with wich the relative frecuency converges, when n tends to infinity. On the basis of this concept we will use relative frequencies as probabilities. If we calculated the entropy, we would obtain:

$$H_0(X,Y) = -\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \log p_{ij} \quad (6)$$

It is denoted by $H_0(X,Y)$ to indicate that it is the initial moment (before the aggregation).

Let us imagine that it is necessary to perform an aggregation in the sectors that are included in the table due to any of the previously indicated

---

[4] See Doblado (1988).

[5] Observe that f_{ij} is not exactly a technical coefficient, as its denominator is the sum of the intermediate consumption and not the total output. The use of this type of coefficients is of interest in the input-output applications since they allow to get to know the amount of intermediate purchases (sales) of the total of intermediate consumption, eliminating the influence of the primary inputs (final demand).

reasons. More specifically, let us suppose[6] that the sectors from 1 to m are aggregated, both in rows and in columns. The resulting aggregated matrix would be the following:

$$\begin{bmatrix} \sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} & \sum_{i=1}^{m} p_{i,im+1} & \cdots & \sum_{i=1}^{m} p_{in} \\ \sum_{j=1}^{m} p_{m+1,j} & p_{m+1,m+1} & \cdots & p_{m+1,n} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{j=1}^{m} p_{nj} & p_{n,m+1} & \cdots & p_{nn} \end{bmatrix} \quad (7)$$

In general terms, it seems logical to think that the aggregation is going to lead to a reduction in the amount of information contained in an input-output table; going forward in this point, we are going to quantify and compare the information before and after the aggregation. As it has been indicated before, the expression that represents the information contained in the table can be found in (6).

After performing referred operation, the entropy will take the value $H_1(X,Y)$ with regard to the final moment. That is, $H_1(X)$ takes the following value

$$-\left[(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij})\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right) + \left(\sum_{i=1}^{m} p_{i,m+1}\right)\log\left(\sum_{i=1}^{m} p_{i,m+1}\right) + \ldots \left(\sum_{i=1}^{m} p_{i,n}\right)\log\left(\sum_{i=1}^{m} p_{i,n}\right) + \left(\sum_{j=1}^{m} p_{m+1,j}\right)\log\left(\sum_{j=1}^{m} p_{m+1,j}\right) + \right.$$

$$\left. \ldots + \left(\sum_{j=1}^{m} p_{nj}\right)\log\left(\sum_{j=1}^{m} p_{nj}\right) + \sum_{i=m+1}^{n}\sum_{j=m+1}^{n} p_{ij}\log(p_{ij}) \right] \quad (8)$$

3.2. Empirical results

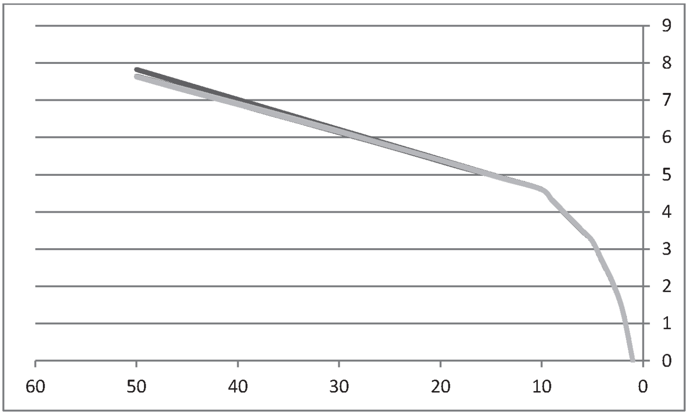
With the aim of illustrating this analysis, we have performed simulations on the basis of random samples of size 2500 (since economies formed by 50 sectors[7] have been assumed). The samples have been randomly generated and the probability distributions originated correspond to intermediate situations, that is to say, there are neither degenerate samples

[6] This supposition doesn't diminish the generality of the reasoning that has been used.

[7] Different analyses with different number of sectors have been carried out and the results have been analogous.

nor uniformly distributed samples. On the basis of each of these samples we have performed different aggregations and determined the amount of information in each level of grouping.

The graph which represents the different levels of information according to the performed aggregation is the following:



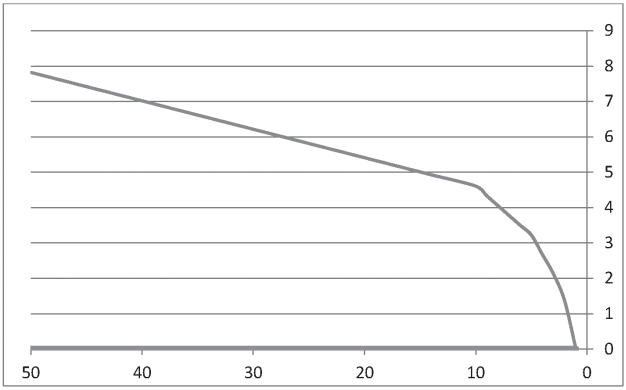
Graph 1: Information associated with an input-output table according to the performed aggregation

Since the aggregation is carried out starting from the highest number of sectors to the lowest, the ordinate axis is arranged in the reverse direction (from highest to lowest). In the vertical axis, the amount of information associated with each level of aggregation is presented.

As can be observed, the entropy progressively decreases and finally becomes null as tables are more aggregated. Furthermore, in the case of very aggregated tables the reduction in the amount of information is greater than when these tables are less aggregated.

It is necessary to point out that the results are very similar, despite the fact that the samples and the aggregations that have been performed differ. It is also true that only intermediate situations have been considered, in other words, the extreme cases haven't been taken into account: neither a uniform distribution of the flows between sectors, nor an economy formed by a single sector that carries out the totality of the transactions.

Below, and in order to complement the graphical analysis, we have considered the extreme situations and we have made a graph where we have included the entropy associated with a uniform distribution and with a degenerate distribution.



Graph 2: Amount of information associated with an input-output table. Extreme situations

The line which is situated over the ordinate axis (coinciding with it) represents the amount of information associated with a degenerate distribution, where, regardless of the level of aggregation, the information is zero. The other curve that appears in the graph represents the amount of information associated with a uniform distribution, for each of the levels of aggregation. Thus, all the intermediate cases, which are the usual ones, will be found between both situations.

Given the form of measurement of this amount of information, it can be observed that only in very extreme situations, close to the degenerate distribution, do the values move away from the upper bound to a great extent. On the basis of the previous results, we can indicate that the measure of entropy proposed by Theil is very sensitive to degenerate distributions, but less sensitive to other kinds of situations.

3.3. Loss of information

A concept that may be of interest is that of the loss of information associated with the sectoral aggregation of an input-output table. This loss can be defined on the basis of the difference in the entropy before and after the aggregation, and the result is the following:

$$\Delta(H) = H_0(X,Y) - H_1(X,Y) =$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\log p_{ij} +$$

$$\left[(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij})\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right) + \left(\sum_{i=1}^{m} p_{i,m+1}\right)\log\left(\sum_{i=1}^{m} p_{i,m+1}\right) + .. + .\left(\sum_{i=1}^{m} p_{i,n}\right)\log\left(\sum_{i=1}^{m} p_{i,n}\right) + \left(\sum_{j=1}^{m} p_{m+1,j}\right)\log\left(\sum_{j=1}^{m} p_{m+1,j}\right) + \ldots \right.$$

$$\left. + \left(\sum_{j=1}^{m} p_{nj}\right)\log\left(\sum_{j=1}^{m} p_{nj}\right) + \sum_{i=m+1}^{n}\sum_{j=m+1}^{n} p_{ij}\log(p_{ij}) \right] = -\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\log p_{ij} + \left[\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right)\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right) + \right.$$

$$\left. \sum_{j=m+1}^{n}\left(\sum_{i=1}^{m} p_{i,j}\right)\log\left(\sum_{i=1}^{m} p_{i,j}\right) + \sum_{i=m+1}^{n}\left(\sum_{j=1}^{m} p_{i,j}\right)\log\left(\sum_{j=1}^{m} p_{i,j}\right) + \sum_{i=m+1}^{n}\sum_{j=m+1}^{n} p_{ij}\log(p_{ij}) \right] \quad (9)$$

Observe that the first term is positive as it is affected by a negative sign. Likewise, the rest of the summands have a negative sign because, even though there are sums of probabilities, these sums always happen to be, as a result of their construction, less than one and, therefore, their logarithm is negative.

The first term of this expression ($\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\log p_{ij}$) can be decomposed in the way that is shown below

$$\sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}\log p_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\log p_{ij} + \sum_{i=m+1}^{n}\sum_{j=m+1}^{n} p_{ij}\log p_{ij} + \sum_{i=1}^{m}\sum_{j=m+1}^{n} p_{ij}\log p_{ij} + \sum_{i=m+1}^{n}\sum_{j=1}^{m} p_{ij}\log p_{ij}$$

If we denote by $\Delta(H) = H_0(X,Y) - H_1(X,Y)$, we obtain

$$\Delta(H) = \left[(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij})\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right)\right]^T + \left[\sum_{j=m+1}^{n}\left(\sum_{i=1}^{m} p_{i,m+1}\right)\log\left(\sum_{i=1}^{m} p_{i,m+1}\right)\right]^T + \left[\sum_{i=m+1}^{n}\left(\sum_{j=1}^{m} p_{i,j}\right)\log\left(\sum_{j=1}^{m} p_{i,j}\right)\right]^T$$

where

$$\left[(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij})\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right)\right]^T = (\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij})\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right) - \sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\log p_{ij} \quad (10)$$

Also

$$\left[\sum_{j=m+1}^{n}\left(\sum_{i=1}^{m} p_{i,m+1}\log\left(\sum_{i=1}^{m} p_{i,m+1}\right)\right)\right]^T = \left[\sum_{j=m+1}^{n}\left(\sum_{i=1}^{m} p_{i,m+1}\log\left(\sum_{i=1}^{m} p_{i,m+1}\right)\right)\right] - \sum_{i=m+1}^{n}\sum_{j=1}^{m} p_{ij}\log p_{ij} \quad (11)$$

Finally

$$\left[\sum_{i=m+1}^{n}\left(\sum_{j=1}^{m}p_{i,j}\log\left(\sum_{j=1}^{m}p_{i,j}\right)\right)\right]^{T}=\left[\sum_{i=m+1}^{n}\left(\sum_{j=1}^{m}p_{i,j}\log\left(\sum_{j=1}^{m}p_{i,j}\right)\right)\right]-\sum_{i=m+1}^{n}\sum_{j=1}^{m}p_{ij}\log p_{ij} \quad (12)$$

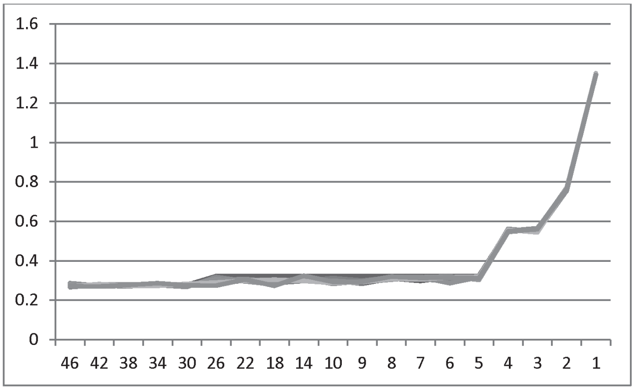
The expression (10) contains the changes in the amount of information caused by the aggregation in rows and columns, that is to say, the loss of information due to those sectors of activity that are aggregated. The expression (11) reflects the loss in the amount of information in the sectors which are not the ones intended to be directly joined, but those that are aggregated in a collateral way, as a result of the aggregation carried out with regard to other sectors. It reflects the loss of information caused by the aggregation carried out in columns. The expression (12) is analogous to (11) but it refers to the collateral aggregations in rows.

To conclude, the term $\sum_{i=m+1}^{n}\sum_{j=m+1}^{n}p_{ij}\log p_{ij}$ that is shown in (10) refers to those sectors that have not been aggregated and, therefore, it becomes null when it is subtracted from the corresponding terms of $H_0(X,Y)$.

Given the properties of the logarithmic function, we can deduce that the entropy before the aggregation is higher than the entropy after having carried out this operation. Thus, and considering what was previously indicated with regard to the signs of the different addends, we reach the conclusion that if we take all the terms into account, the aggregation leads to a reduction in the information that an input-output table provides. It is a decrease in absolute terms, that is, the indicator which quantifies the information contained in the matrix is lower.

If the distribution is intermediate, the loss will be lower when the aggregation does not contribute drastically to move the matrix coefficients away from a uniform distribution.

Below we present a graph derived from the performed simulations that reflects the loss of information associated with different aggregations with non-degenerate intermediate distributions.



Graph 3: Loss of information. Intermediate distributions

Once more, the horizontal axis of the graph is arranged from the highest to the lowest when it comes to the number of sectors (from a lower level of aggregation to a higher). The loss of information is reflected in the vertical axis.

As it can be observed, a kind of inverted L is represented, where the lowest values are reached in the initial and central parts, while the highest are in the opposite end. When the table is very aggregated, a new aggregation generates a sharp reduction in the amount of information; this performance coincides with the one that was seen in Graph 1.

3.5. Relative loss

The loss of information due to the aggregation depends both on the changes undergone by the matrix coefficients and on the variation in the number of sectors. In other words, even though the distribution of the flows between sectors can be more uniform as a result of aggregation, the reduction in the number of sectors results in a decrease of the final value of the measure, since a lower number of sectors will provide less information regarding the economy. In order to try to eliminate the impact that the reduction in the number of sectors has on the loss of information, we propose a relative measure of loss of information by dividing the loss that was previously defined between the number of sectors that form it, but in a "softened" way, that is, starting from the logarithm of referred number (log *n*).

As it has been indicated before, Shannon entropy is delimited at the top by log *n* and at the bottom by zero. If the loss of information is expressed in relative terms (with regard to the bound) the effect caused by the change in the number of sectors can be avoided and the loss of information will only depend on the distribution.

A relative loss of information can be defined in the following terms:

$$\Delta^{R}(H) = \frac{H_0(X,Y)}{\log r} - \frac{H_1(X,Y)}{\log s} \quad (13)$$

where n is the number of coefficients before the aggregation and s is their number after this aggregation has taken place. That is to say

$$\Delta^{R}(H) = \frac{-\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} \log p_{ij}}{\log r} +$$

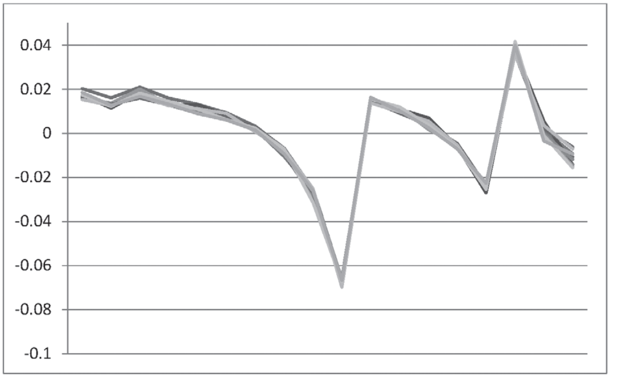
$$\frac{[\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right)\log\left(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}\right) + \sum_{j=1m+1}^{n}\left(\sum_{i=1}^{m} p_{i,m+1}\right)\log\left(\sum_{i=1}^{m} p_{i,m+1}\right) + \sum_{i=m+1}^{n}\left(\sum_{j=1}^{m} p_{i,j}\right)\log\left(\sum_{j=1}^{m} p_{i,j}\right) + \sum_{i=m+1}^{n}\sum_{j=m+1}^{n} p_{ij}\log\left(p_{ij}\right)]}{\log s}$$

log n>log s, while $H_0(X,Y)>H_1(X,Y)$, therefore the sign of the relative loss is inconclusive, in other words

$$\frac{H_0(X,Y)}{\log r} < \frac{H_1(X,Y)}{\log s} \quad (14)$$

depending on the situation.

When (log s) $H_0(X,Y)$<(log n) $H_1(X,Y)$ the amount of initial information, which is higher, is compensated by the reduction in the number of sectors, whereas if it is the case that (log s) $H_0(X,Y)$>(log n) $H_1(X,Y)$ this compensation does not take place.



Graph 4: Relative loss of information. Intermediate distributions

In this graph, it can be observed that for certain aggregations the relative loss of information is negative, which means that the information before the aggregation is less than the information contained in the table after having performed this operation. This happens in those situations where the aggregation leads to a result in which the distribution approximates more to the uniform one.

Also, it can be noticed that in very aggregated tables a slightly negative loss takes place, due to the great homogeneity derived from the resulting small number of sectors.

4. Amount and loss of information in the input-output table for Asturias

As mentioned earlier, we are going to apply the concepts which were previously explained to the Input-Output Table for Asturias corresponding to the year 2010, since this one was the latest to be published at the time of writing this study.

Different aggregations have been used in order to observe its impact on the amount of associated information. In particular, we have used as a starting point the aggregation proposed by SADEI[8] in 66 sectors; af-

[8] The classification proposed by SADEI is a modification of the CNAE 2009.

terwards, we have used the one that appears in EUROSTAT with regard to the input-output tables, in 60 sectors[9]. Moreover, we have taken into consideration the aggregation performed in the WIOD database in 32 sectors for each country, as referred database is often used by input-output analysts. In this application, we have used the classification in 21 sectors proposed by the CNAE (National Classification of Economic Activities), although as a consequence of homogeneity problems, the resulting aggregated Input-Output Table for Asturias is formed by 20 sectors in the end. Another classification that has been taken into account is the International Standard Industrial Classification of All Economic Activities (ISIC), which considers 16 sectors. Finally, aggregations in 10 sectors of EUROSTAT and in 4 of SADEI have been used; both aggregations have a very small number of sectors, but thanks to this, we have at our disposal a wide range of possibilities to analyse. All the classifications that have been used here can be found in the annex.

Table 1: Amount of information associated with the Input-Output Table for Asturias corresponding to the year 2010

SOURCE	Sectors	H(X,Y)	Upper bound	Percentage of bound
SADEI	66	5,8537	8,3793	69,859
EUROSTAT	60	5,6938	8,1551	69,819
WIOD	32	5,1830	6,9315	74,775
CNAE	21	3,9775	5,9915	66,386
CIIU	16	3,4658	5,5452	62,501
EUROSTAT	10	2,9605	4,6052	64,286
SADEI	4	1,8819	2,7726	67,874

From the previous table we can deduce that the highest amount of information is presented by the most disaggregated tables, as had been indicated before. However, if we observe the last column of the table, we notice that not always is the number of sectors directly related to the amount of information contained in it. Observe, for instance, the classification proposed in the WIOD database: even though it is in 32

[9] Given the differences with the one proposed by SADEI, the final aggregation which has been used had to be in 59 sectors.

sectors, in relative terms, it is better than the initial one proposed by SADEI or that of EUROSTAT in 59. This is due to the fact that teferred aggregation leads to structures which are more uniformly distributed, where the groupings of sectors have a more homogeneous performance.

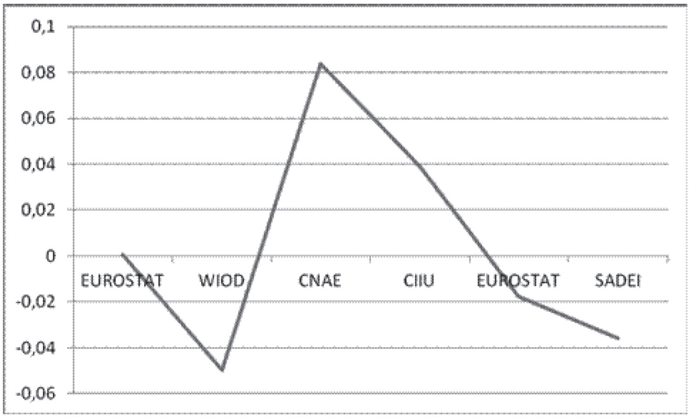
Below, we proceed to quantify the relative loss derived from the performed aggregations.

Table 2: Relative loss due to the aggregation

	H66	H59	H32	H20	H16	H10
H59	0,040					
H32	-4,915	-4,956				
H20	3,473	3,433	8,389			
H16	7,358	7,318	12,273	3,885		
H10	5,574	5,533	10,489	2,100	-1,785	
H4	1,985	1,945	6,900	-1,488	-5,373	-3,589

The previous table must be read in the following way: the initial entropies are situated in the first row, while in the first column the final aggregations are shown. In other words, $H^R(66)$-$H^R(59)$= 0.040, that is to say, the relative loss when aggregating from the 66 sectors of the initial table to the 59 is of 0,04, and so on. Consequently, if the initial table is aggregated in 32 sectors (according to the classification found in the WIOD database), a negative relative loss takes place, that is to say, a relative gain of information. Therefore, it can be noticed that the aggregations in 32 sectors (used in the WIOD database) and in 4 (SADEI) are the ones that, in relative terms, would produce a gain of information at the time of aggregating, since they lead to more uniform classifications of sectors. Nevertheless, it is also true that using a classification in 4 sectors in an analysis is not particularly helpful.

The graph below represents the relative loss when considering an initial table in 66 sectors and aggregating it according to the different classifications previously indicated. As it can be observed, the aggregations found in the WIOD database (32) and in SADEI (4) generate a relative gain of information, despite having a smaller number of sectors. In the case of the groupings of SADEI, we consider that it is due to the homogeneity derived from the small number of sectors.



Graph 5: Relative loss of information due to different aggregations

5. Aggregation and linkages

Other effects of aggregation are the ones referred to the input-output structural analysis. One of the main strengths of Input-Output analysis is its ability to evaluate the significance of the different productive sectors within the economic structure of a country (region). It is relevant to consider whether the level of segregation can influence the value of these multipliers and the results of structural analysis performed.

There are a great number of studies that deal with how the aggregation affects the results of a structural analysis. We can highlight, among them, those of Hatanaka (1952), Ara (1959), Tilanus and Theil (1965), Morimoto (1970), Blin and Cohen (1977), Lauritzen (1989) or Russo (2001). In this part of the article, we are going to study the impact that the aggregation has on the linkages.

The extraction method has its origins in the studies carried out by Strassert (1968), who presents it as a structural evaluation alternative with regard to the classic methods. This author proposes to quantify the effect that would occur in an economy if a certain sector was hypothetically extracted from it. In order to do this, the productive sector under consideration is eliminated from the matrix **A** of technical coefficients.

We start with Leontief demand model shown in the equation (2) and it is assumed that one sector is extracted from the economy. In this case, this equation can be rewritten as

$$\bar{\mathbf{x}}(k) = \left[\mathbf{I} - \bar{\mathbf{A}}(k) \right]^{-1} \bar{\mathbf{y}}(k) \quad (15)$$

where $\bar{A}$ (k) is a matrix of order $(n\text{-}1)(n\text{-}1)$, as the row and the column of the *k*th sector have been eliminated, $\bar{\mathbf{x}}(k)$ represents a vector of total output and $\bar{\mathbf{y}}(k)$, a demand vector, both of dimension n-1.

Therefore, and given the values that $\mathbf{y}(k)$ and $\bar{\mathbf{y}}(k)$ reach, $x_i(k)$ will be smaller than x_i, that is, $\bar{x}_i(k)<x_i \ \forall$ i=1,2...k-1, k+1,...n. Where $\bar{\mathbf{x}}(k)$ is obtained as if the *k*th sector did not exist in the economy and, consequently, this sector does not generate relationships with other productive sectors. Then, the sum of the differences between the elements of $\mathbf{x}$ and $\bar{\mathbf{x}}(k)$ can be considered as the measure of the interrelationships of this sector extracted from the rest. In this way, the following equation is presented:

$$L(k) = \sum_{i=1, i \neq k}^{n} \left[x_i - \bar{x}_i(k) \right] \quad (16)$$

where $L(k)$ is a global measure of the importance of the *k*th sector.

This initial idea that Strassert proposed has undergone extensions, some of which are those presented by Cella (1984), Sonis, et al (1995) and Dietzenbacher and Van der Linden (1997). In the present study, we are going to apply the extension proposed by Sonis et al. (1995).

Let us consider now an initial situation, previous to the aggregation, represented on the basis of a matrix **A** of technical coefficients.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (17)$$

where $a_{ij} = \dfrac{x_{ij}}{x_j}$

That is to say, the proportion of purchases that the sector j makes from the sector i, with regard to the totality of the purchases of sector j.

The Pure Linkage Method is based on the idea of the need to isolate one sector (or group of sectors) to determine its importance. In order to achieve this objective, it proposes the following decomposition of the matrix **A** of technical coefficients:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A_{11}} & \mathbf{A_{1r}} \\ \mathbf{A_{r1}} & \mathbf{A_{rr}} \end{bmatrix} = \begin{bmatrix} \mathbf{A_{11}} & \mathbf{A_{1r}} \\ \mathbf{A_{r1}} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A_{rr}} \end{bmatrix} = \mathbf{A_1} + \mathbf{A_r} \quad (18)$$

The sector (or group of sectors) 1 is the one which is going to be isolated from the rest of the economy. The Leontief inverse matrix (**L**) can be decomposed in the following way: $\mathbf{L=P_2P_1}$ or as $\mathbf{L=P_1P_3}$. Where $\mathbf{P_1=[I\text{-}A_r]^{-1}}$, $\mathbf{P_2=[I\text{-}P_1A_1]^{-1}}$ and $\mathbf{P_3=[I\text{-}A_1P_1]^{-1}}$.

The equation $\mathbf{L=P_2P_1}$ isolates the interaction in the rest of the economy of the sector 1. $\mathbf{P_2}$ shows the impacts derived from the demand of the sector 1 in the economy as a whole ($\mathbf{P_1A_1}$).

The Leontief inverse matrix can be expressed on the basis of the previous expressions in the following manner:

$$\mathbf{L} = \begin{bmatrix} \tilde{\Delta}_1 & \tilde{\Delta}_1\mathbf{A_{1r}} \\ \Delta_r\mathbf{A_{r1}}\tilde{\Delta}_1 & \mathbf{I} + \Delta_r\mathbf{A_{r1}}\tilde{\Delta}_1\mathbf{A_{1r}} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \Delta_r \end{bmatrix} = \mathbf{P_2P_1} \quad (19)$$

where $\tilde{\Delta}_1 = [\mathbf{I}\text{-}\mathbf{A_{11}}\text{-}\mathbf{A_{1r}}\Delta_r\mathbf{A_{r1}}]^{-1}$, $\Delta_r=[\mathbf{I}\text{-}\mathbf{A_{rr}}]^{-1}$ and $\Delta_1=[\mathbf{I}\text{-}\mathbf{A_{11}}]^{-1}$. (20)

On the other hand,

$$\mathbf{P_2} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \Delta_r\mathbf{A_{r1}} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \tilde{\Delta}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{A_{1r}} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad (21)$$

where $\mathbf{P_2=[I\text{-}B_1]^{-1}}$ and $\mathbf{B_1} = \mathbf{P_1A_{11}} = \begin{bmatrix} \mathbf{A_{11}} & \mathbf{A_{1r}} \\ \Delta_r\mathbf{A_{r1}} & \mathbf{0} \end{bmatrix}$. (22)

On the basis of this last equation, we can define the Pure Backward Linkage (PBL) as PBL=$\mathbf{i_{11}}\Delta_r\mathbf{A_{r1}x_{11}}$, where q_{11} is the total output of sector 1, that is to say

$$\mathbf{q_{11}} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_m \end{bmatrix} \quad (23)$$

Our aim is to compare the variation that the linkages experience when the aggregation is performed. To do so, we are going to consider that the group 1 is formed by those sectors that will be aggregated afterwards. Before the aggregation takes place, the expression of **A** is

$$\mathbf{A} = \begin{bmatrix} \mathbf{A_{11}} & \mathbf{A_{1r}} \\ \mathbf{A_{r1}} & \mathbf{A_{rr}} \end{bmatrix} \quad (24)$$

$\mathbf{A_{11}}$ represents a matrix of technical coefficients of m rows and m columns (which will be aggregated in a second stage). $\mathbf{A_{r1}}$ is formed by n-m rows and m columns and it presents the amounts of purchases of the group 1 from the rest of the sectors. $\mathbf{A_{1r}}$ is composed of m rows and n-m columns; it represents the sales of the group 1 to the rest of the economy. Finally, $\mathbf{A_{rr}}$ consists of n-m rows and columns and it shows the flows of the rest of the sectors.

If we apply the expressions indicated before, the Pure Linkage associated with the group of sectors (1) will be $PBL_1 = \mathbf{i_{11}} \Delta_{\mathbf{r}} \mathbf{A_{r1}} \mathbf{q_{11}}$.

The matrices A_{21} and Δ_2 are formed by the following elements

$$\mathbf{A_{r1}} = \begin{bmatrix} a_{m+1,1} & a_{m+1,2} & \dots & a_{m+1,m} \\ a_{m+2,1} & a_{m+2,2} & \dots & a_{m+2,m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,m} \end{bmatrix} \text{ and } \mathbf{\Delta}_{\mathrm{r}} = \begin{bmatrix} \beta_{m+1,m+1} & \dots & \dots & \beta_{m+1,n} \\ \beta_{m+2,m+1} & \dots & \dots & \beta_{m+2,n} \\ \dots & \dots & \dots & \dots \\ \beta_{n,m+1} & \dots & \dots & \beta_{n,n} \end{bmatrix} \quad (25)$$

Thus,

$PBL_1 = \beta_{m+1,m+1} \sum_{j'=1}^{m} a_{m+1,j'} x_{j'} + \dots + \beta_{m+1,\,n} \sum_{j'=1}^{m} a_{n,j'} x_{j'} + \dots \beta_{n,1} \sum_{j'=1}^{m} a_{n,j'} x_{j'} + \beta_{n,n} \sum_{j'=1}^{m} a_{n,j'} x_{j'} x_k$. Let us suppose that the aggregation is performed by joining the first m sectors. We will then obtain

$$\mathbf{A}^* = \begin{bmatrix} a^*_{11} & a^*_{12} & \dots & a^*_{1s} \\ a^*_{m+11} & a^*_{m+1,m+1} & \dots & a^*_{2s} \\ \dots & \dots & \dots & \dots \\ a^*_{s1} & a^*_{s2} & \dots & a^*_{s,s} \end{bmatrix} = \quad (26)$$

$$\begin{bmatrix} \dfrac{\sum_{i=1}^{m}\sum_{j=1}^{m} X_{ij}}{\sum_{j=1}^{m} x_j} & \dfrac{\sum_{i=1}^{m} X_{i,im+1}}{x_{m+1}} & \dots & \dfrac{\sum_{i=1}^{m} X_{in}}{x_n} \\ \dfrac{\sum_{j=1}^{m} X_{m+1,j}}{\sum_{j=1}^{m} x_j} & \dfrac{X_{m+1,m+1}}{x_{m+1}} & \dots & \dfrac{X_{m+1,n}}{x_n} \\ \dots & \dots & \dots & \dots \\ \dfrac{\sum_{j=1}^{m} X_{nj}}{\sum_{j=1}^{m} x_j} & \dfrac{X_{n,m+1}}{x_{m+1}} & \dots & \dfrac{X_{nn}}{x_n} \end{bmatrix}$$

If we partition the matrix, we obtain

$$\begin{bmatrix} \mathbf{A}^*_{11} & \mathbf{A}^*_{1r} \\ \mathbf{A}^*_{r1} & \mathbf{A}^*_{rr} \end{bmatrix} \quad (27)$$

Now $\mathbf{A}^*_{11}$ represents a matrix with just one element, due to the fact that it is formed by the m sectors aggregated in rows and in columns in a single sector. $\mathbf{A}^*_{1r}$ represents the matrix of technical coefficients referred to those sectors that are aggregated in rows; its dimension is 1 row and n-m columns. $\mathbf{A}^*_{r1}$ contains the technical coefficients associated with the sectors that are aggregated in columns with a dimension of n-m rows and 1 column. Last but not least, $\mathbf{A}^*_{rr}$ includes the technical coefficients derived from the non-aggregated sectors with n-m rows and columns; it coincides with $\mathbf{A}_{rr}$.

On the basis of both partitioned matrices **A** and **A*** we are going to try to compare the effects of aggregation in the calculation of the multipliers.

Now, let us calculate the value of the Pure Backward Linkage, considering that the first m sectors are aggregated. In this case, we would have a $PBL^*_1=\mathbf{i}^*_{11}\Delta^*_r\mathbf{A}^*_{r1}\mathbf{q}^*_{11}$, where the terms with an asterisk refer to the aggregated magnitudes.

$$\mathbf{A}^*_{r1} = \begin{bmatrix} a^*_{m+1,1} \\ a^*_{m+2,1} \\ \cdots \\ a^*_{n,1} \end{bmatrix} \qquad \mathbf{\Delta}^*_r = \begin{bmatrix} \beta_{m+1,m+1} & \cdots & \cdots & \beta_{m+1,n} \\ \beta_{m+2,m+1} & \cdots & \cdots & \beta_{m+2,n} \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{n,m+1} & \cdots & \cdots & \beta_{n,n} \end{bmatrix} \text{ and } q^*_{11} = x_1 + x_2 + \ldots + x_m$$

Therefore

$$PBL^*_1 = \left(\sum_{j=m+1}^{n} \sum_{i'=m+1}^{n} \beta_{m+1,j} a^*{}_{i'k} + \ldots + \sum_{j=m+1}^{n} \sum_{i'=m+1}^{n} \beta_{m+2,j} a^*{}_{i'k} + \ldots + \sum_{j=m+1}^{n} \sum_{i'=m+1}^{n} \beta_{n,j} a^*{}_{i'k} \right) q^*_{11} \quad (28)$$

with $q^*_{11} = \sum_{k=1}^{m} q_k$

$$\text{where } a^*_{m+1,1} = \frac{\sum_{j=1}^{m} X_{m+1,j}}{\sum_{j=1}^{m} x_j}, \ \ldots a^*_{n,1} = \frac{\sum_{j=1}^{m} X_{nj}}{\sum_{j=1}^{m} x_j}. \quad (29)$$

On the other hand,

$$a^{*}_{m+1,1} = \frac{\sum_{j=1}^{m} X_{m+1,j}}{\sum_{j=1}^{m} x_j} = \frac{X_{m+1,1} + X_{m+1,2} + ... + X_{m+1,m}}{\sum_{j=1}^{m} x_j} = \frac{\frac{X_{m+1,1}x_1}{x_1}}{\sum_{j=1}^{m} x_j} + \frac{\frac{X_{m+1,2}x_1}{x_2}}{\sum_{j=1}^{m} x_j} + ... + \frac{\frac{X_{m+1,m}x_m}{x_m}}{\sum_{j=1}^{m} x_j} =$$

$$= \frac{a_{m+1,1}x_1}{\sum_{j=1}^{m} x_j} + \frac{a_{m+1,2}x_2}{\sum_{j=1}^{m} x_j} + ... + \frac{a_{m+1,m}x_m}{\sum_{j=1}^{m} x_j} = a_{m+1,1}\omega_1 + a_{m+1,2}\omega_2 + ... + a_{m+1,m}\omega_m$$

Analogously

$$\mathrm{a}^{*}_{\mathrm{n},1} = \frac{\sum_{j=1}^{m} X_{nj}}{\sum_{j=1}^{m} x_j} = a_{n,1}\omega_1 + a_{n,2}\omega_2 + ... + a_{n,m}\omega_m \quad (30)$$

The rest of the addends are determined in the same way.

With

$$\omega_j = \frac{x_j}{\sum_{j=1}^{m} x_j} \forall j = 1,2...m \quad (31)$$

Consequently

$$\mathrm{a}^{*}_{\mathrm{m+1},1} = a_{m+1,1}\omega_1 + a_{m+1,2}\omega_2 + ... + a_{m+1,m}\omega_m = a_{m+1,1}\frac{x_1}{q^{*}_{11}} + a_{m+1,2}\frac{x_2}{q^{*}_{11}} + ... + a_{m+1,m}\frac{x_m}{q^{*}_{11}}$$

……

$$\mathrm{a}^{*}_{\mathrm{n},1} = a_{n,1}\frac{x_1}{q^{*}_{11}} + a_{n,2}\frac{x_2}{q^{*}_{11}} + ... + a_{n,m}\frac{x_m}{q^{*}_{11}} \quad (32)$$

Developing

$$PBL^{*}_1 = \left(\beta_{m+1,m+1} \sum_{j'=1}^{m} a_{m+1,j'} \frac{x_{j'}}{q^{*}_{11}} + ... + \beta_{m+1,n} \sum_{j'=1}^{m} a_{n,j'} \frac{x_{j'}}{q^{*}_{11}} + ... + \beta_{n,n} \sum_{j'=1}^{m} a_{n,j'} \frac{x_{j'}}{q^{*}_{11}} \right) q^{*}_{11} =$$

$$\beta_{m+1,m+1} \sum_{j'=1}^{m} a_{m+1,j'} x_{j'} + ... + \beta_{m+1,n} \sum_{j'=1}^{m} a_{n,j'} x_{j'} + ... \beta_{n,1} \sum_{j'=1}^{m} a_{n,j'} x_{j'} + \beta_{n,n} \sum_{j'=1}^{m} a_{n,j'} x_{j'}$$

That is to say, the linkages coincide if we consider in conjunction the sectors that have not been aggregated and the ones that have.

Finally, and for purposes of illustration, we are going to aggregate the agricultural, forestry and fishing[10] sectors and calculate the Pure Linkage before and after having performed this operation. To start with, we will consider the symmetric matrix for Asturias corresponding to the year 2010, formed by 66 sectors, where Agriculture, livestock breeding and hunting (1), Silviculture and forest exploitation (2) and Fishing and aquaculture (3) are three separate sectors. The PBL for the block formed by those three sectors will be determined. In this case, the matrix $\mathbf{A_{r1}}$ will be composed of 63 rows and 3 columns, $\mathbf{A_{11}}$ will have 3 rows and columns, $\mathbf{A_{1r}}$ will be formed by 3 rows and 63 columns and $\mathbf{A_{rr}}$ will consist of 63 rows and columns. The value of the Pure Linkage is calculated and, if we express it in relative terms with regard to the total output of the economy, we obtain a result of 0.55%.

After that, the first three sectors of the table are aggregated and we will then obtain a matrix of 64 sectors. Now, the matrix $\mathbf{A^*_{11}}$ will be formed by a single sector, $\mathbf{A^*_{r1}}$ will be composed of 63 rows and 1 column, $\mathbf{A^*_{rr}}$ will have 63 rows and columns and $\mathbf{A^*_{1r}}$ will consist of one row and 63 columns. Once again, the Pure Linkage for the aggregated sector is determined and the result is the same 0.55%, that is to say, the value of the linkage does not change with aggregation.

6. Conclusions

The aggregation of sectors in an input-output table may be necessary when different sources of information are combined or when it is required to obtain a symmetric matrix from supply and use tables.

A table contains a certain amount of information that can be quantified through measures of statistical information, among which we can find Shannon entropy. As a table is aggregated, its amount of information de-

[10] These three sectors represent the first one of the input-output table aggregated in 4 sectors (SADEI).

creases. Referred reduction is stronger when the tables are already very aggregated than when their level of disaggregation is higher.

A concept that contains this phenomenon is that of the loss of information, which presents a shape of an inverted L if intermediate distributions (neither degenerate nor uniform) are considered. In other words, the greatest loss takes place in very aggregated tables.

The distribution of the coefficients intervenes in the loss of information and so does the number of sectors. If we define the relative loss as the loss of information divided by its upper bound, we are eliminating the importance of the number of sectors. When we determine the loss in this case, we appreciate that there are situations in which an aggregation leads to a relative "gain" of information. These situations take place when the aggregated table has a more uniform behaviour than the table before the aggregation.

We have applied the previous concepts to the table for Asturias corresponding to the year 2010, which is initially aggregated in 66 sectors of activity. On the basis of this table, different aggregations have been applied: that of EUROSTAT referred to input-output tables, that of the WIOD database, the CNAE in 20 sectors, the ISIC, the aggregation of EUROSTAT in 10 sectors and that of SADEI in 4. Even though the amount of information decreases with the aggregation, if we take the relative loss into consideration, this loss is negative (relative gain) in the case of the aggregations proposed by the WIOD and by SADEI in 4 sectors.

Another analysed aspect involved studying what effects the aggregation has on the linkages. For this purpose, we have considered the Pure Linkage proposed by Sonis et al (1995). It has been theoretically proven that the value of the linkage does not vary when the sectors of activity are aggregated.

To conclude, we have calculated the PBL on the basis of the table for Asturias in 66 sectors, considering a block formed by the sectors of Agriculture, Silviculture and Fishing. We have obtained the same value as if it was calculated on the basis of the aggregated table, where the block would now be formed by the sum of the three indicated sectors.

References

ARA, K. “The aggregation problem in input-output analysis”. *Econometrica: Journal of the Econometric Society*, 27, 257-262, 1959.

BLIN, J. M., & COHEN, C.. “Technological similarity and aggregation in input-output systems: a cluster-analytic approach·. *The Review of Economics and Statistics*, 59, 82-91., 1977.

CABRER, B., CONTRERAS, D., & MIRAVETE, E. “Aggregation in Input–Output Tables: How to Select the Best Cluster Linkage”. *Economic Systems Research*, *3*(2), 99-110, 1991.

CELLA, G: “The Input-Output Measurement of Interindustry Linkages”, *Oxford Bulletin of Economics and Statistics*, 46(1), 73-84, 1984.

DIETZENBACHER, E. & VAN DER LINDEN, J: “Sectoral and Spatial Linkages in the EC Production Structure”, *Journal of Regional Science*, 37(2), 235-257, 1997.

DOBLADO J. M. *Teoría de la información: errores de agregación en las tablas input-output*. Universidad Nacional de Educación a Distancia, UNED, 1988.

EUROSTAT: *NACE, Rev. 2*, 2007.

HATANAKA, M. “Note on consolidation within a Leontief system”. *Econometrica*, 301-303, 1952.

INE: *CNAE 2009,* 2007.

LEONTIEF, W. “An alternative to aggregation in input-output analysis and national accounts”. *The Review of Economics and Statistics*, 49, 412-419, 1967.

LEONTIEF, W. (Ed.). *Input-output economics*. Oxford University Press, 1986.

MORIMOTO, Y. “On aggregation problems in input-output analysis”. *The Review of Economic Studies*, 27(1), 119-126, 1970.

PARDO, L: *Teoría de la Información Estadística*. Ed. Hespérides, 1998.

RUSSO, G. *Input–output analysis and aggregation bias: An empirical assessment*. 84(13). Working paper Universita degli Studi di Trieste, 2001.

SADEI: *Tablas Input-Output de Asturias (2010),* 2013.

SONIS, M. GUILHOTO, G., HEWINGS, G. & MARTINS, E: “Linkages, Key Sectors, and Structural Change: Some New Perspectives”, *The Developing Economics*, 33(3), 233-270, 1995.

THEIL, H. “Linear aggregation in input-output analysis”. *Econometrica*, 25, 111-122, 1957.

TILANUS, C & THEIL, H. “The information approach to the evaluation of input-output forecasts”. *Econometrica: Journal of the Econometric Society*, 847-862, 1965.

TIMMER, M. (ed), *The World Input-Output Database (WIOD): Contents, Sources and Methods*, WIOD Working Paper 10, 2012.

UNITED NATIONS: *CIIU, Rev. 3.1*, 2002.

Assessing the distribution and use of income and changes in income with Social Accounting Matrices

Susana Santos[1]

Abstract

Several different areas of the socio-economic activity of countries, and in particular the distribution of income, can be studied with the use of Social Accounting Matrices (SAMs).

The features underlying a SAM allow for the reading and interpretation of the socio-economic activity of countries, leading to the production of an empirical work that can highlight specific aspects of the reality under study and offers the chance to perform experiments with changes in those aspects. Thus, SAMs can also be used to support policy decision processes.

With the aid of some methodological principles based on the works of R. Stone, G. Pyatt and J. Round, a SAM-based approach to the study of income distribution is carried out, seeking to provide both an empirical and a theoretical description of the socio-economic activity of a country, respectively through a numerical and an algebraic version of a SAM. The algebraic version can also be referred to as a 'SAM-based model'.

This study uses the nomenclatures of the latest version of the System of National Accounts (2008 SNA).

By including production and institutions accounts in a matrix format, the structural features of a country's production and income distribution can be worked upon together, making it possible to capture specific networks of linkages and the corresponding multiplier effects, in subsequent modelling exercises.

[1] ISEG - Lisbon School of Economics and Management, Universidade de Lisboa
UECE (Research Unit on Complexity and Economics) and DE (Department of Economics)
E-mail: ssantos@iseg.ulisboa.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_14

A basic SAM is presented, with rows and columns representing accounts. The production accounts are represented by activities (or industries), products (or goods and services), and factors of production. The (domestic) institutions are represented by the current, capital and financial accounts. The rest of the world account represents the "external" part of the (domestic) economy.

Bearing in mind the importance of ensuring the consistency of the whole system, possible disaggregations and extensions to that basic structure are analysed.

Aggregates and balancing items that can be identified and calculated outside this matrix format are also presented.

An assessment of a country's distribution and use of income is presented, by means of the identification of some of the underlying structural features of SAMs.

Macroeconomic effects of changes in income are then identified by way of an experiment with a change in the taxes on income and wealth, paid by the households to the government, using a SAM-based approach.

The exposition is accompanied by an example which is applied to Portugal.

Keywords: Social Accounting Matrix; National Accounts; Income Distribution; Macroeconomic Modelling.

1. Introduction

The Social Accounting Matrix (SAM) is a tool that has specific features that can be used to study the socio-economic activity of countries. Such features allow for the reading and interpretation of the reality under study, leading to the production of an empirical work, which not only highlight specific aspects of that activity, but also offers the chance to perform experiments with different interventions with regard to its functioning.

Section 2 outlines how the SAMs can be used to measure and model the socio-economic activity of countries, and also how they can be used as an alternative support for studies being carried out in several areas, as well as for supporting those taking part in the policy decision process. The main features of the SAM-based approach are outlined in Subsection 2.1, by adopting a methodological framework based on Richard

Stone's research and that of his followers, according to which the SAM can describe the activity of countries, either empirically or theoretically, depending respectively on whether it is presented in a numerical or an algebraic version. In Subsection 2.2, a proposal is presented for the development of a basic SAM, together with an explanation of possible alternative taxonomies. In Subsection 2.3, the accounting multipliers are presented as a possible algebraic version of a SAM, to be used in a SAM-based approach.

In Sections 3 and 4, an assessment is made of the distribution and use of income, and also of macroeconomic effects of changes in income, respectively. In the former case, some of the structural features underlying the SAMs are identified, and, in the latter, an experiment based on a change in the taxes on income and wealth, paid by the households to the government, is performed using a SAM-based approach.

Finally, Section 5 summarises and presents some concluding remarks on how SAMs can be useful working tools for assessing the distribution and use of income, as well as, for the analysis of possible macroeconomic effects of changes in a country's income.

Throughout the research, the methodologies are accompanied by an application for Portugal.

2. Measuring and modelling the socio-economic activity of countries with SAMs

2.1. The SAM and the SAM-based approach

The Social Accounting Matrix (SAM) is a square matrix which measures and models the socio-economic activity of a country, depending respectively on whether it is in a numerical or in an algebraic version. In this way, specific aspects of that activity can be studied through a SAM-based approach, thus benefitting from a greater analytical content which is provided by the matrix format, and allowing for the capture of networks of linkages that have not been captured otherwise.

A numerical version of a SAM can be seen as an empirical snapshot of the socio-economic activity of a country, in which each cell has a specific value, with the sums of rows being equal to the sums of the columns.

A theoretical snapshot of the socio-economic activity of a country can be given by an algebraic version of a SAM, in which each cell is filled with one or more algebraic expressions which, together with those of all the other cells, form a SAM-based model, the calibration of which involves the replication of a numerical version.

Therefore, the SAM is a working tool which, on one hand, enables the reading of the reality under study and, on the other hand, offers a chance to simulate different interventions with regard to its functioning.

With a methodological framework based on the works of R. Stone, G. Pyatt and J. Round the proposal is to work SAMs with production and institutional accounts, to represent the transactions within the domestic economy, and one or more rest of the world accounts, to represent the transactions between the economy and abroad.

The level of detail of the versions of SAMs used will depend on the purposes of the study in which it can be used as a working tool, as well as the available data.

The National Accounts are the main reference source of data on the socio-economic activity of a country. In the attempt to capture more and better representative data of this activity, either the National Accounts data or the corresponding underlying system (the System of National Accounts - SNA), have been subject to successive improvements. Thus, National Accounts – Data and System (namely its nomenclatures) were considered in the specification of the contents of the SAM presented here, and were adopted as the base source of information.

In the SAM accounts representing the transactions within the economy there are production and institutional accounts. The former is divided into activity sectors (or industries), products and factors of production, which are worked together with the accounts of the institutional sectors of the latter. Thus, the representative transactions of the (primary) distribution of income are worked upon together with the redistribution and use of income, as well as capital and financial transactions. In this way, the network of linkages captured by the matrix format will allow for a better quantification of the multiplier effects in subsequent modelling.

In research using a SAM-based approach, the challenge is to find out what is the best use of the SAM, both in the numerical and the algebraic versions. Such research may support studies carried out in several areas, as well as the studies of those taking part in the policy decision process. This involves the best definition of the taxonomy of the SAM accounts and of the corresponding network of linkages.

In particular, when the focus of a study is on the role played by households (within the institutional sectors) in the socio-economic activity of a country during a pre-defined period (usually a year), all the data that can be studied by using the SAM can be further complemented with a Socio-Demographic Matrix (SDM).

Representing the households all the population of a country, the flows in its production (expressed in currency units), its consumption and also its income can be worked upon within a SAM framework. In turn, when the entire population of a country is studied within a SDM framework, its movements (expressed in numbers of human beings) can be represented in the form of a stock-flow matrix, which offers many possibilities for disaggregation. A SDM records the movements of the relationship between the opening and closing stock of a population for a particular period and also the flows of people during that period (births, deaths, immigration, emigration and changes between groups). Depending on the purpose of the study, active or passive sequences can be adopted to identify these groups. Usually the first step that is taken prior to the definition of these sequences is to identify the different age groups.

This paper reinforces the construction of numerical versions of SAMs from the SNA, and further study of possible algebraic versions of SAMs and SDMs is recommended in: Stone (1986, 1981 and 1973) and Santos (2014).

2.2. Constructing numerical versions of SAMs from the SNA

The latest versions of the SNA have devoted a number of paragraphs to discussing the question of SAMs. The 2008 version mentions SAMs in Section D of Chapter 28, which is entitled "Input-output and other matrix-based analysis", in which a matrix is presented of the all the accounts in the SNA (ISWGNA, 2009, pp.519-522). This matrix is not, however, to be confused with the SAM that is presented below, although both practically cover all the flows recorded in the SNA accounts.

In turn, the European System of National and Regional Accounts in the European Union of 2010 (the adaptation for Europe of the 2008 version of the SNA - 2010 ESA) makes a reference to the SAM, stating that, amongst other features, it can be thought of as an expanded system of labour (satellite) accounts (EUPC, 2013, p.523).

The SAM that is presented below is a result of the research undertaken by the author within the above-mentioned methodological framework. Much effort was made to reconcile that framework with that which is defined by (the successive versions of) the SNA. Seeking to include as much as possible of the flows observed by the National Accounts, a version of the SAM is proposed which accompanied the application made for Portugal in 2011 – which was the base year of the adaptation to Europe of the above mentioned latest version of the SNA.

Accordingly, a square matrix is worked upon, in which the sum of the rows is equal to the corresponding sum of the columns. In keeping with what is conventionally accepted, and after some adjustments have been made to adapt this to the SNA, the following items are represented in the entries made in the rows: resources, incomes, receipts or changes in liabilities, and net worth; whilst the entries made in the columns are: uses, outlays, expenditures or changes in assets. Each flow is therefore recorded only once, in a cell of its own.

The starting point for the construction of a numerical SAM should be its design, i.e. the classification or taxonomy of its accounts. That taxonomy, together with the levels of adopted disaggregation, will both depend on the purposes for which the SAM is to be used, as well as on the way that the available information is organised. On the other hand, there needs to be a concern that the most significant features of economy-wide transactions will be represented, and that the SAM can also be used as the base for a model. A basic structure is proposed below, which adopts the National Accounts as the base source of information, takes into account the underlying SNA, and which also highlights the consistency of the whole system. The flexibility of that basic structure is shown, together with the possibilities that it presents for characterising problems and for achieving specific purposes. These can be seen as ways for going beyond the SNA, thus facing the dangers associated with the adoption of international standards and systems of classification and their failure to recognise important issues and realities.

2.2.1. The SAM's basic structure and its consistency with the whole system

Adopting the working method recommended by Stone, the basic structure for the SAM presented here is a summary set of the National Accounts and the controlling totals for the other levels of disaggregation, all of which will be analysed later. Thus, in keeping with the conventions and nomenclatures defined by the SNA, besides a rest of the world account, the proposed SAM also includes both production and institutional accounts.

Table 1 shows that basic structure, which represents the nominal transactions ("t") with which two indexes are associated. The location of these transactions within the matrix framework is described by those indexes, the first of which represents the row account, whilst the second represents the column account. Each cell of this matrix will be converted into a sub-matrix, with the number of rows and columns corresponding to the level of disaggregation of the row and column accounts. This same table also identifies blocks which are sub-matrices, or sets of sub-matrices with common characteristics. The specification of these blocks is made below and involves an identification of the flows of the National Accounts, which will continue to be the same, even if disaggregation is carried out to some degree – thereby preserving the consistency of the whole system.

Table 1: The Basic SAM by blocks

	p	*a*	f	dic	dik	dif	rw	Total
p – products	TTM ($t_{p,p}$)	IC ($t_{p,a}$)	0	FC ($t_{p,dic}$)	GCF ($t_{p,dik}$)	0	EX ($t_{p,rw}$)	AD ($t_{p\cdot}$)
a – a ctivities	P ($t_{a,p}$)	0	0	0	0	0	0	VPT ($t_{a\cdot}$)
f – factors	0	CFP_ GAV ($t_{f,a}$)	0	0	0	0	CFP ($t_{f,rw}$)	AFIR ($t_{f\cdot}$)
dic –(domestic) institutions' current account	NTP ($t_{dic,p}$)	NTA ($t_{dic,a}$)	CFP_ GNI ($t_{dic,f}$)	CT ($t_{dic,dic}$)	0	0	CT ($t_{dic,rw}$)	AI ($t_{dic\cdot}$)
dik –(domestic) institutions' capital account	0	0	0	S ($t_{dik,dic}$)	KT ($t_{dik,dik}$)	NLB ($t_{dik,dif}$)	KT ($t_{dik,rw}$)	INVF ($t_{dik\cdot}$)
dif – (domestic) institutions' financial account	0	0	0	0	0	FT ($t_{dif,dif}$)	FT ($t_{dif,rw}$)	TFTR ($t_{dif\cdot}$)
rw – rest of the world	IM& NTP ($t_{rw,p}$)	NTA ($t_{rw,a}$)	CFP ($t_{rw,f}$)	CT ($t_{rw,dic}$)	KT ($t_{rw,dik}$)	FT ($t_{rw,dif}$)	0	TVRWP ($t_{rw\cdot}$)
Total	AS ($t_{\cdot p}$)	VCT ($t_{\cdot a}$)	AFIP ($t_{\cdot f}$)	AIP ($t_{\cdot dic}$)	AINV ($t_{\cdot dik}$)	TFTP ($t_{\cdot dif}$)	TVRWR ($t_{\cdot rw}$)	

Source: Own construction.

Note: The first three accounts (p = products (or goods and services), *a* = activities (or industries) and f = factors (of production)) are the production accounts of the economy and the next three accounts (dic = current; dik = capital; dif = financial) are the accounts of the (domestic) institutions. The last account (rw = rest of the world) represents the "external" part of the (domestic) economy.

A description of the blocks now follows (included in the texts are letters, followed by numbers between brackets, which are the flow codes of the National Accounts, in accordance with the 2008 SNA):

a) Production – P (cell: $t_{a,p}$; basic prices) – represents the output of goods and services (P1).

b) Domestic Trade is represented by the value of domestically transacted products, which can be either be domestically produced or imported.

B1) Intermediate Consumption – IC (cell: $t_{p,a}$; purchasers' prices) – consists of the value of the goods and services consumed as inputs of the process of production, excluding those fixed assets whose consumption is recorded as consumption of fixed capital (P2).

b2) Final Consumption – FC (cell: $t_{p,dic}$; purchasers' prices) – consists of the expenditure incurred by resident institutional units on goods or services which are used for the direct satisfaction of individual needs or wants, or the collective needs of members of the community (P3).

B3) Gross Capital Formation – GCF (cell: $t_{p,dik}$; purchasers' prices) – includes gross fixed capital formation, changes in inventories, and acquisitions less disposals of valuables (P5).

c) External Trade includes the transactions in goods and services from non-residents to residents, also known as imports (P7), or IM (cell: $t_{rw,p}$; purchasers' prices), and the transactions in goods and services from residents to non-residents, also known as exports (P6), or EX (cell: $t_{p,rw}$; purchasers' prices).

d) Trade and Transport Margins – TTM (cell: $t_{p,p}$) – amounts to zero and, when it is disaggregated and takes the form of a submatrix, it allocates the output of the trade and transport services used in the domestic trade to the supplied products.

e) Net taxes on production and imports

e1) Net Taxes on Production – NTA (cells: $t_{dic,a}$; $t_{rw,a}$) – represents the (other) taxes on production (D29) minus the (other) subsidies to production (D39)

e2) Net Taxes on Products – NTP (cells: $t_{dic,p}$; $t_{rw,p}$) – represents the taxes on products (D21) minus the subsidies on products (D31).

f) Compensation of Factors of Production – CFP (cells: $t_{f,a}$; $t_{dic,f}$; $t_{f,rw}$; $t_{rw,f}$) – consists of the income of the institutional sectors originating from the compensation of employees (D1) and the compensation of employers and own-account (or self-employed) workers, as well as the compensation of capital, including property income (D4; B2g-B3g). Functional distribution is represented by the Gross Added Value – GAV (cell: $t_{f,a}$), whereas institutional distribution is represented by the Gross National Income – GNI (cell: $t_{dic,f}$).

g) Current Transactions – CT (cells: $t_{dic,dic}$; $t_{dic,rw}$; $t_{rw,dic}$) – includes: current taxes on income, wealth, etc. (D5); net social contributions (D61); social benefits other than social transfers in kind (D62); other current transfers (D7), and; the adjustment made for the change in pension entitlements (D8).

h) Capital Transactions – KT (cells: $t_{dik,dik}$; $t_{dik,rw;}$ $t_{rw,dik}$) – includes: capital taxes (D91); investment grants (D92); and other capital transfers (D99).

i) Financial Transactions – FT (cells: $t_{dif,dif}$; $t_{dif,rw}$; $t_{rw,dif}$) – includes: monetary gold and special drawing rights (F1); currency and deposits (F2); debt securities (F3); loans (F4); equity and investment fund shares or units (F5); insurance, pension and standardised guarantee schemes (F6); financial derivatives and employee stock options (F7); and other accounts receivable/payable (F8).

j) Gross Saving – S (cell: $t_{dik,dic}$) – measures the portion of aggregate income that is not used for final consumption expenditure and current transfers to domestic institutions or to the rest of the world (B8g).

k) Net Lending/Borrowing – NLB (cell: $t_{dik,dif}$)

The net lending or borrowing of the total economy is the sum of the net lending or borrowing of the institutional sectors. It represents, respectively, the net resources that the total economy makes available to the rest of the world or the net resources that it receives from the rest of the world in order to finance the corresponding needs of investment funds (B9).

In this case, these amounts are recorded in the row(s) of the capital account, i.e. as changes in liabilities and net worth, and in the column(s) of the financial account, i.e. as changes in assets. This is why the mathematical signs of this item have been changed in relation to the SNA.

The construction of a SAM is easier when it is performed by blocks.

The totals of each account represent the corresponding sums of the cells in rows and in columns, with the following descriptions:

I. Products account (p): Aggregate Demand – AD, the row sum (cell: $t_{p.}$); Aggregate Supply – AS, the column sum (cell: $t_{.p}$).

II. Activities account (*a*): Production Value – VPT, the row sum (cell: $t_{a.}$); Total Costs – VCT, the column sum (cell: $t_{.a}$).

III. Factors of Production (f): Aggregate Factors Income (Received) – AFIR, the row sum (cell: $t_{f.}$); Aggregate Factors Income (Paid) – AFIP, the column sum (cell: $t_{.f}$).

IV. Current account of (domestic) institutions (dic): Aggregate Income (received) – AI, the row sum (cell: $t_{dic.}$); Aggregate Income (Paid) – AIP, the column sum (cell: $t._{dic}$).

V. Capital account of (domestic) institutions (dik): Investment Funds – INVF, the row sum (cell: $t_{dik.}$); Aggregate Investment – AINV, the column sum (cell: $t._{dik}$).

VI. Financial account of (domestic) institutions (dif): Total Financial Transactions (Received) – TFTR, the row sum (cell: $t_{dif.}$); Total Financial Transactions (Paid) – TFTP, the column sum (cell: $t._{dif}$).

VII. Rest of the world account (rw): Value of Transactions to the Rest of the World (Paid) – TVRWP, the row sum (cell: $t_{rw.}$); Value of Transactions from the Rest of the World (Received) – TVRWR, the column sum (cell: $t._{rw}$).

From the above descriptions, two further advantages of the matrix presentation of the National Accounts can be highlighted: firstly, each transaction represents a single entry and can be characterised by its position and secondly; each account is represented by a row and a column, whose balance is ensured by the equality of their sums.

The Integrated Economic Accounts Table is a summary of all the detail observed by the SNA, including the full sequence of accounts for (domestic) institutional sectors, the rest of the world and the total economy. Based on this table, and taking into consideration the previous descriptions, it was possible to construct the basic SAM that is presented in Table 2, which represents the highest possible level of aggregation of the activity of Portugal as observed in the National Accounts of 2011 (at current prices).

Taking into consideration the description given of Table 1 above, and based on the reading of the products account for Portugal in 2011, which is represented in Table 2 – row and column p, the main components of the aggregate demand and supply of goods and services transacted in the market and recorded by the Portuguese National Accounts can be identified in millions of Euros. Thus, reading in row, the total aggregate demand of 416,695 million Euros was composed of 172,577 million Euros of intermediate consumption, 150,944 million Euros of final consumption, 32,764 million Euros of gross capital formation, and 60,410 million Euros of exports. Reading in column, the total aggregate supply of 416,695 million Euros (equal to the total aggregate demand) was composed of 326,819 million Euros of output of goods and services, 22,087 million Euros of net taxes on products received by the Portuguese Government, - 163 million Euros of net taxes on products sent to European Union institutions, and 67,952 million Euros of imports. A similar reading can be made for each of the other columns and rows of Table 2, for Portugal.

Therefore, as mentioned above, the basic SAM that has just been described can be considered to be the most aggregate summary set of the National Accounts, which represents a first level of the intended hierarchical method, with all the controlling totals for the next level of that hierarchy. Based on this, the consistency of the whole (supposedly) observed system can be ensured.

Table 2: Basic SAM of Portugal in 2011(Unit: 10^6 euros)

	p	*a*	f	dic	dik	dif	rw	Total
p – products	0	172,577		150,944	32,764		60,410	416,695
a – activities	326,819							326,819
f – factors		154,848					8,048	162,895
dic – (domestic) institutions' current account	22,087	1,156	149,275	86,044			5,910	264,472
dik – (domestic) institutions' capital account				22,900	2,565	7,297	2,632	35,394
dif – (domestic) institutions' financial account						13,836	-7,397	6,439
rw – rest of the world	67,789[(a)]	- 1,761	13,620	4,584	65	- 14,694		69,603
Total	416,695	326,819	162,895	264,472	35,394	6,439	69,603	

Sources: Statistics Portugal *(INE);* Portuguese Central Bank *(Banco de Portugal).*
[(a)] 67,952 (imports) less 163 (net taxes on products sent to the institutions of the European Union)

2.2.2. *Disaggregations and extensions*

As the National Accounts cover all the details covered by the basic structure presented above, some other levels of the above mentioned hierarchical method can be identified within the National Accounts, which provide other controlling totals for greater levels of disaggregation. As will be seen below, this disaggregation can be made at the levels of the production, institutional and rest of the world accounts, either within the scope of the SNA conventions, or not. The opposite applies to the case of aggregation. Therefore, it is possible to carry out a simpler aggregation and disaggregation of the accounts without losing the consistency of the system, in accordance with the specific requirements of the exercise carried out and the information available.

It is also worth mentioning that, although the quarterly National Accounts are not as complete as the annual ones, it would, nevertheless, be possible to carry out some further disaggregations from these accounts in terms of time. Furthermore, disaggregations can be made in terms of space, as regional accounts are also considered. In this case, it is possible to work with regions and countries, either individually, or as a group. It would even be possible to consider data in world terms, if the SNA could be adopted worldwide.

Extensions are also possible, either from the National Accounts, or from other sources of information, with a convenient adjustment to, or connection with, the whole system, in order to maintain consistency.

The 2008 SNA dedicates Chapter 29 to "Satellite accounts and other extensions", where the main idea is to serve specific analytical purposes, in a way that is consistent with the central framework, although not fully integrated into it. In this respect, the author would like to support the idea of Keuning and Ruijter (1988) of a "complete data set" which "could be tentatively labelled: a System of Socio-economic Accounts".

i) Production accounts

In the basic structure proposed above, production accounts are the accounts of products, activities and factors of production. These accounts correspond respectively to the SNA accounts of goods and services, production and the primary distribution of income. Thus, within these accounts, and depending on the available level of disaggregation, one can see how the available products are used, with some details being provided about the process of production and about the way in which the income resulting from that process is distributed among institutions and activities, as well as the ownership of assets.

The SNA uses the Central Product Classification (CPC) Version 2 (completed in December 2008) to classify products, which are organised into 10 sections, with the possibility of going as far as the fifth level of disaggregation within each of these.

In turn, the International Standard Industrial Classification of All Economic Activities (ISIC) Revision 4 (officially released in August 2008) is used to classify activities, which are organised into 23 sections, and their disaggregation is perfectly consonant with that of the product classification. The Supply and Use Table provides this information, usually at an intermediate level of disaggregation.

As described above, in the characterisation of the block representing the compensation of the factors of production, the disaggregation that can be made from the above-mentioned tables of the National Accounts is between labour (or the compensation of employees) and that which has been referred to as 'compensation of other factors', which includes the compensation of employers and own-account (or self-employed) workers and also the compensation of capital, namely property income. Such information can only be derived from the Integrated Economic Accounts Table if the products and activities accounts are not disaggregated. This can also be derived from the Supply and Use Table, if the same accounts are disaggregated.

As applied to Portugal, Table 3 presents the possible disaggregation of the factors of production accounts, based on the Integrated Economic Accounts Table, in which the products and activities accounts are not disaggregated.

ii) Institutions and rest of the world accounts

In the basic structure proposed above, domestic institutions are divided into current, capital and financial accounts. These accounts correspond, respectively, to the following SNA accounts: secondary distribution of income, redistribution of income in kind and use of income; capital, and; financial accounts. Within these accounts, depending on the level of disaggregation available, the current accounts show how national income is transformed into disposable income through the receipt and payment of current transfers, and how the latter is distributed between final consumption and saving. In turn, the capital account records the transactions linked to acquisitions of non-financial assets and capital transfers involving the redistribution of wealth, whereas the financial account records the transactions in financial assets and liabilities between institutional units, and between these and the rest of the world.

All the linkages between the domestic economy and the rest of the world (i.e. all the transactions between resident and non-resident units) are recorded in both the SAM and the SNA, through the rest of the world account.

Chapter 4 of the 2008 SNA specifies the institutional sectors, including the rest of the world, as well as their possible disaggregation, which in some cases can be taken as far as the third level, although normally it cannot be taken beyond the first level. In the case of the rest of the world, such disaggregation will certainly depend on the country, or group of countries, that adopt and adapt this system[2].

At the first level of disaggregation, the accounts of both institutions and the rest of the world are part of the Integrated Economic Accounts Table. Higher levels of disaggregation, whenever these are possible, are usually published in the separate accounts of institutions. Even at the first level of disaggregation, any research carried out of the institutional sectors with transactions involving more than one row or column of the SAM, also requires the so-called "from whom to whom matrices". These matrices make it possible to fill in the cells of the sub-matrices of transactions taking place both within domestic institutions, and also between institutions and the rest of the world, which are recorded in the above--described blocks of current, capital, and financial transactions.

---

[2] In the case of the 2010 ESA (Chapter 23), a second level of disaggregation is made, with a distinction being drawn between Member States and institutions and bodies of the European Union, and non-member countries and international organisations that are not resident in the European Union. For the former, disaggregation can be made as far as the fourth level.

The disaggregation of specific institutional sectors makes it possible to analyse the most diverse aspects of the corresponding roles in the activity of countries. Santos (2007) is an example of a study of the role of government in general and its sub-sectors – central government, local government and social security funds.

The detailed study of the specific accounts of domestic institutions and their corresponding transactions also makes it possible to analyse specific aspects of the same activity, namely: the distribution and redistribution of income, using the current account; the redistribution of wealth, using the capital account; investment and its financing, and also the implicit levels of financing requirements and the availability of institutional sectors and the whole economy, using both the capital and financial accounts.

In turn, the rest of the world account can provide many possibilities for studying the international economic relations of the domestic economy.

Table 3 presents the possible disaggregation of the institutions' current and capital accounts, and it is derived from the Integrated Economic Accounts Table and the "from whom to whom matrices" for our application to Portugal. From that table, we can see a little more about the reality under study – Portugal in 2011. In the case of aggregate factors income – the so-called primary distribution of income, which amounts to 162,895 million Euros, it can now be seen that 82,006 million Euros represent the compensation of employees – row and column 3, and that 80,890 million Euros account for the compensation of other factors of production, namely employers and own-account workers, land and capital (including property income) – row and column 4.

Row 3 shows that the part of the gross-added-value generated in the domestic economy by residents and non-residents that is attributed to employee compensation, amounts to 81,617 million Euros, to which are added 388 million Euros generated abroad by residents. In turn, row 4 shows that the part of the gross-added-value generated in the domestic economy by residents and non-residents that is attributed to the above-mentioned other factors of production, amounts to 73,231 million Euros, to which are added 7,659 million Euros generated abroad by residents.

Rows 3 and 4 enable the study of aspects related to the functional distribution of income.

Column 3 shows that the part of the gross national income generated in the domestic economy and abroad by residents which is attributed to compensation of employees and is received by households, amounts to 81,554 million Euros, to which are added 452 million Euros which are generated in the domestic economy by non-residents and sent abroad. In column 4, the part of the gross national income generated in the domestic economy and abroad by residents which is attributed to the compensation of other factors and is received by the various institutional sectors, amounts to 67,721 million Euros, to which are added 13,169 million Euros generated in the domestic economy by non-residents and sent abroad. From the cells of columns 3 and 4, it is possible to study aspects related to the institutional distribution of income.

Just as the matrix form of the production accounts may be easily worked on from the supply and use tables, it would also be possible to work on the matrix form of the institutional accounts if some kind of "from whom to whom tables" were made official. This would be a crucial factor for implementing the SAM-based approach, in which SAMs with production and institutional accounts, that conveniently capture the circular flow of income and the underlying network of linkages, would form the basis for macroeconomic models which would be capable of reproducing the multiplier processes that are implicit in the socio-economic activity of countries.

Table 3 SAM of Portugal in 2011, with disaggregated factors of production and the (domestic) institutions' current and capita accounts (Unit: 10^6 Euros)

Income (receipts) \ Outlay (expenditure)				PRODUCTION					INSTITUTIONS					
						FACTORS			CURRENT ACCOUNT					
				PRO-DUCTS	ACTIVI-TIES	Labour (emplo-yees)	Other	Total	House-holds	Non-financial corporations	Financial corpora-tions	Govern-ment	NPISH	Total
				1	2	3	4		5	6	7	8	9	
PRODUCTION		PRODUCTS	1	0	172,577	0	0	0	112,611	0	0	34,983	3,351	150,944
		ACTIVITIES	2	326,819	0	0	0	0	0	0	0	0	0	0
	FACTORS	Labour (employees)	3	0	81,617	0	0	0	0	0	0	0	0	0
		Other	4	0	73,231	0	0	0	0	0	0	0	0	0
		Total		0	154,848	0	0	0	0	0	0	0	0	0
INSTITUTIONS	CURRENT ACCOUNT	Households	5	0	0	81,554	39,878	121,432	1,465	1,549	3,617	30,464	322	37,417
		Non-financial corporations	6	0	0	0	18,116	18,116	1,549	0	496	122	0	2,166
		Financial corporations	7	0	0	0	9,391	9,391	3,435	694	969	152	28	5,277
		Government	8	22,087	1,156	0	- 394	- 394	32,235	5,091	978	8	26	38,339
		Non Profit Inst. Serv. Households (NPISH)	9	0	0	0	730	730	680	152	56	1,953	2	2,844
		Total		22,087	1,156	81,554	67,721	149,275	39,364	7,485	6,117	32,699	378	86,044
	CAPITAL ACCOUNT	Households	10	0	0	0	0	0	9,458	0	0	0	0	9,458
		Non-financial corporations	11	0	0	0	0	0	0	12,547	0	0	0	12,547
		Financial corporations	12	0	0	0	0	0	0	0	8,251	0	0	8,251
		Government	13	0	0	0	0	0	0	0	0	- 7,254	0	- 7,254
		Non Profit Inst. Serv. Households (NPISH)	14	0	0	0	0	0	0	0	0	0	- 102	- 102
		Total		0	0	0	0	0	9,458	12,547	8,251	- 7,254	- 102	22,900
	FINANCIAL ACCOUNT		15	0	0	0	0	0	0	0	0	0	0	0
REST OF THE WORLD			16	67,789	- 1,761	452	13,169	13,620	1,114	300	867	2,262	42	4,584
TOTAL				416,695	326,819	82,006	80,890	162,895	162,547	20,333	15,234	62,689	3,669	264,472

Sources: Statistics Portugal *(INE);* Portuguese Central Bank *(Banco de Portugal).*

Table 3: (continued). SAM of Portugal in 2011, with disaggregated factors of production and the (domestic) institutions' current and capital accounts (Unit: 10^6 Euros)

Income (receipts) \ Outlay (expenditure)				INSTITUTIONS – CAPITAL ACCOUNT – Households	INSTITUTIONS – CAPITAL ACCOUNT – Non-financial corporations	INSTITUTIONS – CAPITAL ACCOUNT – Financial corporations	INSTITUTIONS – CAPITAL ACCOUNT – Gover-nment	INSTITUTIONS – CAPITAL ACCOUNT – NPISH	INSTITUTIONS – CAPITAL ACCOUNT – Total	INSTITUTIONS – FINANCIAL ACCOUNT	REST OF THE WORLD	TOTAL
				10	11	12	13	14		15	16	
PRODUCTION	PRODUCTS		1	6,696	18,233	904	6,120	811	32,764	0	60,410	416,695
PRODUCTION	ACTIVITIES		2	0	0	0	0	0	0	0	0	326,819
PRODUCTION	FACTORS	Labour (employees)	3	0	0	0	0	0	0	0	388	82,006
PRODUCTION	FACTORS	Other	4	0	0	0	0	0	0	0	7,659	80,890
PRODUCTION	FACTORS	Total		0	0	0	0	0	0	0	8,048	162,895
INSTITUTIONS	CURRENT ACCOUNT	Households	5	0	0	0	0	0	0	0	3,697	162,547
INSTITUTIONS	CURRENT ACCOUNT	Non-financial corporations	6	0	0	0	0	0	0	0	50	20,333
INSTITUTIONS	CURRENT ACCOUNT	Financial corporations	7	0	0	0	0	0	0	0	566	15,234
INSTITUTIONS	CURRENT ACCOUNT	Government	8	0	0	0	0	0	0	0	1,501	62,689
INSTITUTIONS	CURRENT ACCOUNT	Non Profit Inst. Serv. Households (NPISH)	9	0	0	0	0	0	0	0	95	3,669
INSTITUTIONS	CURRENT ACCOUNT	Total		0	0	0	0	0	0	0	5,910	264,472
INSTITUTIONS	CAPITAL ACCOUNT	Households	10	0	0	490	63	0	553	- 5,277	121	4,855
INSTITUTIONS	CAPITAL ACCOUNT	Non-financial corporations	11	0	0	8	491	0	499	6,239	697	19,983
INSTITUTIONS	CAPITAL ACCOUNT	Financial corporations	12	0	0	462	606	0	1,068	- 7,246	7	2,079
INSTITUTIONS	CAPITAL ACCOUNT	Government	13	4	162	9	0	4	178	12,967	1,761	7,652
INSTITUTIONS	CAPITAL ACCOUNT	Non Profit Inst. Serv. Households (NPISH)	14	0	0	26	241	0	267	614	46	825
INSTITUTIONS	CAPITAL ACCOUNT	Total		4	162	995	1,401	4	2,565	7,297	2,632	35,394
INSTITUTIONS	FINANCIAL ACCOUNT		15	0	0	0	0	0	0	13,836	-7,397	6,439
REST OF THE WORLD			16	- 1,845	1,588	181	131	9	65	-14,694		69,603
TOTAL				4,855	19,983	2,079	7,652	825	35,394	6,439	69,603	

Sources: Statistics Portugal *(INE)*; Portuguese Central Bank *(Banco de Portugal)*.

2.2.3. *Aggregates and balancing items*

As can be seen above, practically all the transactions of the National Accounts are covered by the SAM, and thus macroeconomic aggregates and balancing items can all be identified by it (see the description of the cells or blocks in Table 1, as well as Tables 2 and 3 for the application for Portugal).

Gross Domestic Product at market prices (GDP_{pm}), which is usually considered to be the main macroeconomic aggregate, can be calculated using the following three known approaches:

– Production approach: $GDP_{pm} = P - IC + NTP = t_{a,p} - t_{p,a} + (t_{dic,p} +$ (part of) $t_{rw,p})$;

– Expenditure approach: $GDP_{pm} = FC + GCF + Ex - IM = t_{p,dic} + t_{p,dik} + t_{p,rw} -$ (part of) $t_{rw,p}$;

– Income approach: $GDP_{pm} = GAV + NTP + NTA = t_{f,a} + (t_{dic,p} +$ (part of) $t_{rw,p}) + (t_{dic,a} + t_{rw,a})$.

The Portuguese GDP_{pm} in 2011 was 176,167 million Euros, which can be calculated from these three approaches as follows:

– Production approach: GDP_{pm} = 326,819 – 172,577 + (22,087 + (–163));

– Expenditure approach: GDP_{pm} = 150,944 + 32,764 + 60,410 – 67,952;

– Income approach: GDP_{pm} = 154,848 + (22,087 + (–163)) + (1,156 + (–1,761)).

Domestic Product can be converted into National Product, by adding the compensation of factors received from the rest of the world, and by deducting the compensation of factors and the net indirect taxes (on both products and production) sent to the rest of the world, when these exist. Thus, from the cells described for the basic SAM, GDP_{pm} can be converted into Gross National Product at market prices (GNP_{pm}), or Gross National Income (GNI_{pm}), as follows: $GDP_{pm} + t_{f,rw} - t_{rw,f} - t_{rw,a} -$ (part of) $t_{rw,p}$. On the other hand, as the SAM directly provides Gross National Income, this can also be calculated by simply adding the net indirect taxes (on both products and production) received by domestic institutions to the compensation of factors received by domestic institutions: $t_{dic,f} + t_{dic,p} + t_{dic,a}$. The corresponding amount for Portugal in 2011 is 172,518 million Euros, for which the underlying calculations are as follows: GNP_{pm}= 176,167 + 13,620 – 8,048 – (– 1,761) – (–163); GNI_{pm}= 149,275 + 22,087 + 1,156.

Disposable Income (Domestic or National) is very important and is calculated by adding the net current transactions received by domestic institutions to GNI_{pm}: GNI_{pm} + ((received)$t_{dic,dic}$+ $t_{dic,rw}$) – ((paid)$t_{dic,dic}$+ $t_{rw,dic}$). For Portugal this is: 172,518 + (86,044 + 5,910) – (86,044 + 4,584) = 173,844 million Euros.

In turn, gross aggregates can be converted into net aggregates (and balancing items) by deducting the consumption of fixed capital. This lies outside the basic SAM, but is part of the Integrated Economic Accounts.

Gross Saving (S) and Net Lending or Borrowing (NLB) are provided directly by the SAM, through $t_{dik,dic}$ and $t_{dik,dif}$, respectively, which, in the case of Portugal in 2011, are 22,900 and 7,297 million Euros respectively. As explained above, the latter amount represents Net Borrowing.

The main items for income and expenditure of the institutional sectors and of the rest of the world can be calculated from the respective rows and columns of the SAM. In the case of the institutional sectors: the total balancing item is the net lending/borrowing (NLB) of the respective institutional sector, with an opposite mathematical sign to that registered in the SAM; the current balancing item is the respective gross saving (S), and; the capital balancing item is the difference between the first and the second (NLB and S).

2.3. An algebraic version of a SAM: the accounting multipliers

The base methodology that is to be followed uses accounting multipliers, in keeping with the research of Pyatt and Roe (1977) and Pyatt and Round (1985).

Exogenous and endogenous accounts exist (see Table 4) and consequently the transactions in each cell of the SAM will be considered exogenous or endogenous, according to the corresponding row and column accounts.

From Table 4, it can be written that

$y_n = n + x$ (1)

$y_x = l + r$ (2)

The amount that the endogenous accounts receive is equal to the amount that they spend (row totals equal column totals). In other words, in aggregate terms, total injections from the exogenous into the endogenous accounts (i.e. the column sum of "x") are equal to total leakages from the endogenous into the exogenous accounts, i.e. considering i' to be the unitary vector (row), the column sum of "1" is:

$x * i' = l * i'$. (3)

Table 4: The SAM in endogenous and exogenous accounts

	Endogenous		Exogenous		Total
		Σ		Σ	
Endogenous	N	n	X	x	y_n
Exogenous	L	l	R	r	y_x
Total	y_n'		y_x'		

Source: Pyatt and Round (1985).

Note: As referred to above, rows represent resources, incomes, receipts or changes in liabilities and net worth. Columns represent uses, outlays, expenditures or changes in assets.

Key:

N = matrix of transactions between endogenous accounts; n = vector of the (corresponding) row sums.

X = matrix of transactions between exogenous and endogenous accounts (injections from first into second); x = vector of the (corresponding) row sums.

L = matrix of transactions between endogenous and exogenous accounts (leakages from first into second); l = vector of the (corresponding) row sums.

R = matrix of transactions between exogenous accounts; r = vector of the (corresponding) row sums.

y_n = vector (column) of the receipts of the endogenous accounts ($\hat{y}_n$: diagonal; $\hat{y}_n^{-1}$: inverse); y_n' = vector (row) of the expenditures of the same accounts.

y_x = vector (column) of the receipts of the exogenous accounts; y_x' = vector (row) of the expenditures of the same accounts.

In the structure of Table 4, if the entries in the N matrix are divided by the corresponding total expenditures, a corresponding matrix (squared) can be defined of the average expenditure propensities of the endogenous accounts within the endogenous accounts or of the use of resources within those accounts. Calling this matrix A_n, it can be written that

$A_n = N{*}\hat{y}_n^{-1}$ (4)

$N = A_n{*}\ \hat{y}_n$ (5)

Considering equation (1), $y_n = A_n{*}y_n + x$ (6)

Therefore, $y_n = (I-A_n)^{-1}{*}\ x = M_a * x$. (7)

We thus have the equation which calculates the total receipts of the endogenous accounts (y_n), by multiplying the injections "x" by the matrix of the accounting multipliers:

$M_a = (I-A_n)^{-1}$. (8)

On the other hand, if the entries in the L matrix are divided by the corresponding total expenditures, a corresponding matrix (non-squared) can be defined as the average expenditure propensities of the endogenous accounts into the exogenous accounts or as the use of resources from the endogenous accounts into the exogenous accounts. Calling this matrix A_l, it can be written that

$A_l = L{*}\hat{y}_n^{-1}$ (9)

$L = A_l{*}\ \hat{y}_n$ (10)

Considering equation (2), $y_x = A_l{*}y_n + r$ (11)

Thus, $l = A_l * y_n = A_l * (I-A_n)^{-1}{*}\ x = A_l * M_a * x$. (12)

Therefore, by using the accounting multipliers, the impact of changes in receipts is analysed at the moment when they occur, assuming that the structure of expenditure in the economy does not change. It is also assumed that excess capacity in the economy exists, as does production technology and resource endowment.

3. Identifying the structural features of a country's distribution and use of income with SAMs.

Structural indicators for the functional and institutional distribution of generated income can be calculate from a SAM, as can indicators for the distribution and use of disposable income, which allows for an assessment of the distribution and use of a country's income.

In order to prepare the experiment performed in Section 4, the same methodology will be adopted, using Portugal in 2011 as an example, but now the year 2010 will also be added. From 2010 to 2011 there was an overall absolute decrease in all the aggregates mentioned below, except for savings.

The functional distribution of generated income is given by the distribution of gross added value – GAV – among the factors of production, through the structure of the sub-matrix in cell $t_{f,a}$ of the basic SAM (Table 1), with its level of detail depending on the disaggregation of the activities (column account) and also of the factors of production (row account). Table 5 shows the results for the application for Portugal, whereby, although with a decreased difference, more than half of the domestic income (income generated in Portugal by residents and non-residents alike) is compensation of employees, in 2010 (53.3%) and 2011 (52.7%).

Table 5: Portuguese functional distribution of income generated in 2010 and 2011 (in percentage terms)

	2010	2011
Factors of Production (generated income = gross added value or gross domestic product)		
Labour (employees)	53.3	52.7
Other (employers and own-account workers; capital)	46.7	47.3
Total	100.0	100.0

Sources: Statistics Portugal *(INE);* Portuguese Central Bank *(Banco de Portugal)*.
Note: Gross domestic product (at current prices) decreased 2.1% from 2010 to 2011.

In turn, the institutional distribution of generated income is given by the distribution of gross national income – GNI, through the structure of the submatrix in cell $t_{dic,f}$ of the basic SAM (Table 1). In this case, the level of detail will depend on the disaggregation of the factors of production (column account), and of the current account of the domestic institutions (row account). Table 6 shows the results of our application, with households being the most significant, but with a decreased share of the national income (income generated in Portugal by residents and non-residents alike), from 83.6% in 2010, to 81.3% in 2011. On the contrary, non-financial and especially financial corporations have reinforced positions, which is the opposite to that of general government.

Table 6: Portuguese institutional distribution of income generated in 2010 and 2011 (in percentage terms)

	Factors of Production				
	Labour (employees)	Other (employers and own-account workers; capital)		Total	
	2010 = 2011	2010	2011	2010	2011
Institutions (generated income = gross national income)					
Households	100.0	62.6	58.9	83.6	81.3
Non-financial corporations		26.7	26.8	11.7	12.1
Financial corporations		7.5	13.9	3.3	6.3
General government		2.2	- 0.6	0.9	-0.3
Non-profit institutions serving households		1.1	1.1	0.5	0.5
Total	100.0	100.0	100.0	100.0	100.0

Sources: Statistics Portugal *(INE);* Portuguese Central Bank *(Banco de Portugal)*.
Note: Total gross national income (at current prices) decreased 0.7% from 2010 to 2011 (3.9% in the case of the households).

As described in Subsection 2.2.3, the disposable income of the institutional sectors can be calculated in the same way for the whole economy. Thus, each institutional sector obtains its disposable income by excluding from the corresponding gross national income, current transfers paid to other institutional sectors and to the rest of the world, and by including current transfers received from other institutional sectors and the rest of the world. In the case of the government, the net taxes on production and imports are also included. Disposable income is then used in final consumption and savings, except in the case of non-financial and financial corporations, which do not have any final consumption. As can be seen in Table 7, households have more than half of the disposable income of Portugal, although this share has decreased from 72.9% in 2010, to 70.2% in 2011. The same trend occurs in the case of non-financial corporations, whose disposable income, overall, is 7.3% in 2010, and 7.2% in 2011. The remaining institutional sectors have increased shares, with the emphasis being on financial corporations, whose share of the total disposable income are 2.2% in 2010, and 4.7% in 2011.

Furthermore, the share of total disposable income spent in final consumption decreased from 89% in 2010, to 86.8% in 2011. By sacrificing savings, households increased the share of their lower absolute disposable income spent on final consumption (from 90.3% in 2010, to 92.3% in 2011). With negative savings, general government and the non-profit institutions serving households exceed their corresponding disposable incomes with final consumption, although with a decreased share.

Table 7: Portuguese distribution and use of disposable income among institutions in 2010 and 2011 (in percentage terms)

	Distribution of Disposable Income		Use of Disposable Income			
			Final Consumption Expenditure		Saving	
	2010	2011	2010	2011	2010	2011
Households	72.9	70.2	90.3	92.3	9.7	7.7
Non-financial corporations	7.3	7.2	---	---	100.0	100.0
Financial corporations	2.2	4.7	---	---	100.0	100.0
General government	15.9	16.0	133.9	126.2	-33.9	- 26.2
Non-profit institutions serving households	1.6	1.9	113.9	103.1	-13.9	- 3.1
Total	100.0	100.0	89.0	86.8	11.0	13.2

Sources: Statistics Portugal *(INE);* Portuguese Central Bank *(Banco de Portugal)*.
Note: Total gross disposable income (at current prices) decreased 0.5% from 2010 to 2011 (4.2% in the case of the households).

From the above analysis it can be concluded that, between 2010 and 2011, Portugal experienced a general absolute decrease in domestic, national and disposable income. A change in the structure of the distribution and use of income can also be observed. A smaller share of income is allocated to households, whilst financial corporations nearly doubled their share in the case of national income, and in the case of disposable income, this more than doubled. The share of savings of total disposable income increased, to the detriment of the final consumption of the general government and non-profit institutions serving households, as households spent a larger share of their (smaller) disposable income in final consumption.

4. Experimenting a change in a country's income with a SAM-based approach.

Let us now study the changes described in the previous section, using the methodology described in Section 2. Taking an overall view of the content of that section, a general conclusion can be drawn - although possible, in the view of the author, this task is neither easy, nor quick. In fact, so many flows of income are involved and there are so many intervening factors, that a thorough explanation would exceed the limits of time and space of this chapter. Instead, we identify some of the many potentialities of SAMs.

Considering the importance of the households' institutional sector, which represent the total population in a country, as already mentioned, and using our application to Portugal, the following example will show a possible SAM-based approach, considering the taxes on income and wealth, paid by the households to the government.

In 2011, the Portuguese general government received 10,554.04 million Euros of taxes on income and wealth from households, which was 932.295 million Euros more than in 2010. From 2010 to 2011, income generated by the Portuguese households decreased 4.2%, and the percentage of that income that was paid to the Portuguese government in taxes on income and wealth increase from 7.6% in 2010, to 8.7% in 2011. Table 8 shows the relative importance of those amounts with regards the aggregate income of those institutional sectors, that is, in the row/column sum of the current accounts of each.

Our experiment will calculate the macroeconomic effects of the above mentioned increase of 932.295 million Euros in the taxes on income and wealth paid by the households to the government.

Table 8. Relative importance of taxes on income and wealth paid by the households to the general government in the corresponding aggregate incomes.

	2010	2011
Households	5,7%	6,5%
Government	15,4%	16,8%

Source: Statistics Portugal (INE).

Thus, from a SAM of Portugal in 2010, which is the equivalent to Table 3, and considering the methodology described in Subsection 2.3 and the purpose of working with a flow from the households to the government, the (current and capital) accounts of the households were set as exogenous, as also were the financial and the rest of the world accounts. The accounting multipliers were next calculated and run, after having increased the cell (8,5) of current transfers from the households to the government, in the amount that corresponds to the increase registered for taxes on income and wealth paid by the households to the government from 2010 to 2011 (932.295 million Euros). From the replied SAM ($2010^{after\Delta}$), the (new) macroeconomic aggregates were calculated, as explained is Subsection 2.2.3. Table 9 shows the percentage changes in these aggregates and final consumption and gross savings: between 2011 and 2010; before (2010), and; after the mentioned change ($2010^{after\Delta}$), both for the total economy and for the two institutional sectors involved in the experiment - households and the government.

Table 9: Percentage changes in macroeconomic aggregates, final consumption and gross saving in Portugal from 2010 to 2011.

	Total Economy		Households		Government	
	2011-10	$2011\text{-}10^{after\Delta}$	2011-10	$2011\text{-}10^{after\Delta}$	2011-10	$2011\text{-}10^{after\Delta}$
GDP	- 2.1	2.3	---	---	---	---
GNI	- 0.7	0.9	- 3.9	4.3	- 4.5	4.9
GDI	- 0.5	0.7	- 4.2	4.2	- 0.4	2.0
FC	- 3.0	3.5	- 2.1	2.2	- 6.1	8.2
S	19.3	- 16.7	- 23.3	30.3	- 23.1	32.1

Source: Own construction.

Key:

GDP = Gross Domestic Product Gross Domestic Product at market prices.

GNI = Gross National Income.

GDI = Gross Disposable Income.

FC = Final Consumption.

S = Gross Saving.

Thus, the columns "2011-10" show the effects of the total changes that occurred from 2010 to 2011, and in columns "2011-10$^{\text{after}\Delta}$" show the effects of the change in the taxes on income and wealth paid by the households to the government, *ceteris paribus* (that is, if everything else remained constant). Comparing these two effects, for both the total economy and for households or the government, a negative mathematical sign can generally be associated with the former, and a positive mathematical sign can generally be associated with the latter. Therefore, as the change in the taxes on income and wealth paid by the households to the government is part of the total changes, some contribution to contradict the general trend can be attributed to the experimented change. It can also be said that at a first glance, a change which worsens the financial conditions of households, also has impacts at other levels, and on the economy as a whole, which, in the end, improves the general situation. We reached a similar conclusion to that of a similar experiment that was performed for the years 1995 and 2005 (Santos, 2010). Much more could have been studied and, certainly concluded, if we had had access to information by groups of households.

Whilst recognising the limitations of the adopted algebraic version (another could have been adopted in this same experiment), the features of the presented SAM and a SAM-based approach allowed the above described experiment and the corresponding results highlighted the emphasis of specific aspects, from which an assessment of the macroeconomic effects of a change in the income of households and government was made.

This research is just an illustration of how much could be derived from the described methodology, and even from the results of its use.

5. Concluding remarks

By using the form of a square matrix, with rows and columns representing production, institutions and rest of the world accounts, a Social Accounting Matrix (SAM) can represent all the transactions within a domestic economy and also those between a domestic economy and the rest of the world.

By adopting the National Accounts - Data and System - as the base source of information, that "economy" is identified within a country, whose socio-economic activity is studied, using the SAM as a working tool. By capturing the circular flow of income and the underlying network of linkages, the SAM enables the structural features of both production and income distribution to be analysed together, providing the basis for macroeconomic models which are capable of reproducing the multiplier processes, implicit in the socio-economic activity of countries. Thus, in a SAM-based approach, this activity can be measured and modelled, depending, respectively, on whether the SAM is in a numerical or in an algebraic version.

Apart from presenting the specificities of this methodology, an empirical study was carried out, using an application for Portugal. Having identified distribution of income as being one of the various areas of the socio-economic activity of a country, which can be researched using the above-mentioned tool, some aspects of the reality under study (Portugal in 2010 and 2011) were highlighted, and an experiment was carried out which simulated an intervention in its functioning.

Numerical versions of SAMs of Portugal in 2010 and 2011 were constructed using the National Accounts, which permitted the calculation of structural indicators, from which the functional and institutional distribution of generated income and the distribution and use of disposable income could be identified. In the case of the functional distribution of (domestic) income, although this represented the major share, it was possible to identify a decrease in the importance of labour (employees) with regard to capital, and employers and own-account workers. In the case of the institutional distribution of (national) income, as well as that of disposable income, households saw their greater shares reduced in favour of other institutions, especially financial corporations. In turn, the largest share of disposable income used in final consumption was reduced, to the detriment of savings.

From numerical versions of SAMs of Portugal in 2010 and 2011, it was also possible to identify a general (nominal) decrease in the underlying macroeconomic (gross) aggregates of: domestic product, national income, disposable income.

The accounting multipliers, which were calculated from the SAM of Portugal in 2010, allowed an experiment to be carried out which simulated a change in taxes on income and wealth paid by the households to the government, which was equal to the actual increase registered for that amount from 2010 to 2011. Having calculated a new numerical version of the SAM, the macroeconomic effects of that change then permitted the identification of a general increase of the above-mentioned macroeconomic aggregates. Thus, as the change used in the above-mentioned experiment was a part of the total change that occurred from 2010 to 2011, some of the contribution that contradicted the latter can be attributed to the former. Another possible conclusion that can be drawn from the results of this experiment is that, by reproducing the multiplier processes, our working tool showed that a change which, at a first glance, worsens the financial situation of households, also has an impact at other levels, and on the economy as a whole, which, ultimately, improves the overall situation.

Much more research can be carried out using SAM-based approach methodology, especially at the level of SAM-based models, or of SAM algebraic versions, and such research will certainly continue to be carried out.

Acknowledgements: The author gratefully acknowledges the financial support from national funds provided by FCT (Fundação para a Ciência e a Tecnologia) Portugal for the research and writing of this chapter, which forms part of the Strategic Project: UID/ECO/00436/2013.

References

EUPC – European Union Parliament and Council, *European System of National and Regional Accounts in the European Union,* EU Regulation No. 549/2013, Official Journal of the European Union, L174, Volume 56, 26 June 2013.

ISWGNA – Inter-Secretariat Working Group on National Accounts (United Nations, European Commission, International Monetary Fund, Organisation for Economic Cooperation and Development and World Bank), *System of National Accounts 2008*, Series F, No. 2, Rev. 5, United Nations, New York, 2009.

KEUNING S. and RUIJTER W., "Guidelines to the construction of a Social Accounting Matrix", *Review of Income and Wealth*, 34, 1988, pp.71-100.

PYATT, G. and ROE, A. *Social Accounting for Development Planning with special reference to Sri Lanka,* Cambridge: Cambridge University Press, 1977.

PYATT, G. and ROUND, J., "Accounting and Fixed Price Multipliers in a Social Accounting Matrix Framework", in G. Pyatt, and J. Round, (coord.), *Social Accounting Matrices. A Basis for Planning*, A World Bank Symposium, World Bank, pp. 52-69, 1985.

SANTOS, S., "Approach to the socio- economic activity of countries with a Social Accounting Matrices supported by Socio-Demographic Matrices. An application to Portugal",. Working Paper Series – SSRN (Social Science Research Network) abstract=2530735, 2014.

SANTOS, S., "A quantitative approach to the effects of social policy measures. An application to Portugal, using Social Accounting Matrices". EERI (Economics and Econometrics Research Institute) RP (Research Papers) 2010/33; MPRA (Munich Personal RePEc Archive) Paper No. 23676, 2010.

SANTOS, S., "Modelling Economic Circuit Flows in a Social Accounting Matrix Framework. An Application to Portugal". *Applied Economics*, 39, 2007, pp. 1753-1771.

STONE, R., "Demographic Input-Output: An Extension of Social Accounting", in: I. Sohn (coord), *Readings in Input-Output Analysis: Theory and Applications*, Oxford University Press, USA, 1986, pp.151-172.

STONE, R., *Aspects of Economic and Social Modelling*, Editions Droz, Genève (Suisse), 1981.

STONE, R., "A System of Social Matrices", *Review of Income and Wealth*, 19, 1973, pp.143-166.

MULTI-CRITERIA SUSTAINABILITY CLASSIFICATION OF DAIRY FARMS IN A PORTUGUESE REGION

Sandra Silva[1,2], Luís Alçada-Almeida[2,3] and Luís C. Dias[2,3]

Abstract:

This paper presents a Multi-criteria Decision Aid assessment of dairy farms sustainability in the Entre Douro e Minho region in Portugal. Sustainability is a multidimensional concept encompassing often conflicting economic, social and environmental points of view, potentially involving various stakeholders and decision makers. Seven environmental and seven socio-economic criteria were considered, aided by experts from different fields of knowledge. We used the multi-criteria classification method ELECTRE TRI to sort dairy farms into three categories, where socio-economic and environment viability were studied separately. The results were then combined to summarize the socio-economic and environmental viability according to four categories: Highly Sustainable, Moderately Sustainable, Marginally Sustainable and Not Sustainable. We used a Geographic Information System for better visualization and analysis of the spatial behavior of the classification obtained.

Keywords: Multi-criteria Decision Aiding; ELECTRE TRI; Geographic Information Systems; Agricultural Sustainability.

[1] Polytechnic Institute of Viana do Castelo

[2] INESC - Coimbra, Institute for Systems Engineering and Computers

[3] Faculty of Economics, University Coimbra
E-mails: sandrasilva@esa.ipvc.pt, alcada@fe.uc.pt, lmcdias@fe.uc.pt

http://dx.doi.org/10.14195/978-989-26-1039-9_15

1. Introduction

The concept of sustainability is defined in the Brundtland Report (UN, 1987) as "development which meets the needs of the present without compromising the ability of future generations to meet their own needs" arguing that sustainability is based on three distinct pillars: economic, environmental and social. Agricultural sustainability enhances the quality of life of the population through strengthening the economy, preserving cultural heritage, protecting natural resources, and increasing food security. The principles of sustainable development thus require that multiple dimensions are taken into account in sustainability assessment.

Multi-Criteria Decision Aiding (MCDA) methods are suitable for allowing the consideration of various dimensions associated with an assessment, involving decision makers and potentially other stakeholders in the evaluation process. This work presents an MCDA methodology for the classification of dairy farms, according to their sustainability, and presents an application to the Entre Douro e Minho (EDM) Region, in Portugal. Initially, environmental concerns and socio-economic concerns are classified independently. The results are later combined into a single classification for each farm: Highly Sustainable, Moderately Sustainable, Marginally Sustainable or Not Sustainable. To perform a classification the ELECTRE TRI MCDA method (Yu, 1992; Dias et al. 2002; Figueira et al. 2010) has been selected. A Geographic Information System (GIS) was used to present the results.

2. Agricultural Sustainability and Multi-criteria Decision Aiding

Agricultural sustainability requires monitoring the agricultural activities in order to prevent unsustainable behaviors and attitudes. It is necessary to integrate environmental, social and economic factors in the analysis, so that all actors involved can coordinate efforts in the quest for sustainability throughout all levels of the agricultural system.

The way these dimensions are considered in the evaluation of sustainability varies among different authors. Some authors assume that an agricultural system is sustainable when the "trade-offs" between the objectives considered for evaluation of its performances (economic, social and environmental) reach acceptable values for the society as a whole (Becker, 1997; Hediger, 1999; Stoorvogel et al., 2004). This approach to agricultural sustainability becomes possible, while using aggregation methods to cover the three dimensions mentioned, by aggregating the set of criteria or indicators into a multidimensional simple or composite indicator. This approach has been used by several authors as Stockle et al. (1994), Andreoli and Tellarini (2000), Rigby et al. (2001), van Calker et al. (2005), Sydorovych and Wossink (2008) and Dillon et al. (2010).

However, the quantification of sustainable agriculture through a set of indicators is difficult to interpret and often criticized (Hansen, 1996; Munda, 2005; Gómez-Limón and Riesco, 2009), especially for:

1. the subjectivity of these methods (the choice of functional forms for the aggregation and weighting of the individual indicators);
2. the compensability usually aggregates the different dimensions of sustainability attributes (additive aggregation approaches) in spite of their theoretical incommensurability;
3. not facilitating the process of establishing cause-effect relationships between diverse system properties and therefore not providing mechanisms for diagnosing causes of unsustainability, or for evaluating effects of proposed interventions.

Taking into account the characteristics that are desired in this kind of evaluation, we consider that not all aggregation methods would be appropriate. For example, the requirement of non-compensation seems to be essential to avoid that a bad performance in one criterion would be cancelled by a good performance in another criterion. Thus, it was considered more appropriate to use non-compensatory methods based on the outranking concept. The outranking approach seemed more appropriate because, although it is less familiar to decision makers (compared to a weighted sum) and it requires more configuration parameters, it presents the following advantages:

- it is not compensatory and therefore compatible with the requirements demanded by the concept of sustainability;
- it introduces the relation of incomparability, useful to account for situations in which the decision maker and/or the analyst are not able to compare two actions, which can happen in the sustainability analysis;
- it enables the usage of incomplete value information and, for example, judgments on ordinal measurement scale. In the classification of sustainability, the type of information that exists may be quite diverse as it comes from various fields (economic, social and environmental) with different scales and measures.

The sustainability classification of dairy farms has been framed as a sorting problem (Roy, 1985), because there is no need to compare the farms among themselves (furthermore, the high number of farms - 1705 - is not adequate for that purpose), but there is need to compare the farms to predefined targets. ELECTRE TRI, which is a widely used MCDA Outranking Method, has been chosen because, as a multi-criteria sorting method, it classifies alternatives (dairy farms) according to preference-ordered categories.

The use of multi-criteria methods to assess and determine the sustainability of agriculture is still scarce. Raju el al. (2000) used ELECTRE TRI and other MCDA techniques for sustainable water resources planning. Antunes et al. (2011) present the development of a participatory multi-criteria analysis process using AHP (Analytic Hierarchy Process) and SMCE (Social Multi-Criteria Evaluation) to evaluate irrigation management alternatives. van Calker et al. (2006) applieed MAUT (Multi-Attribute Utility Theory) to assess economic, social and ecological sustainability of Dutch dairy farms. Dantsis et al. (2010) applied MAVT (Multi-Attribute Value Theory) to evaluate and compare the level of sustainability of farms in two Greek regions. Gómez-Limón and Riesco (2009) and Santos (2011) applied AHP to the evaluation of agricultural sustainability. Lebacq et al. (2013) review typologies of sustainability indicators that have been developed at the farm level and present a case study to help users to select a set of indicators.

3. Outranking Method - ELECTRE TRI

The outranking approach is based on pairwise comparisons between potential actions or alternatives using a binary, so-called, outranking relation (Roy, 1996): one action outranks another action if the former is considered "not worse than" ("at least as good as") the latter. ELECTRE methods comprise the construction of one or several outranking relation(s) comparing pairs of actions in a comprehensive way, followed by an exploitation procedure (Figueira et al., 2005).

The ELECTRE TRI method (Yu, 1992) provides a complete sorting of the alternatives of a set of alternatices A into two or more ordered categories C_i,i=1,...,k, where C_1 is the worst (least desirable) category and C_k the best one. In order to define the categories, ELECTRE TRI uses some reference alternatives (reference profiles) b_i, i=1,...,k−1, which can be real or hypothetical alternatives. Each reference profile b_i is simultaneously the upper bound of category C_i and the lower bound of category C_{i+1}. The assignment of each alternative $a \in A$ to a category is done by comparing it to the reference profiles. Among different ELECTRE TRI variants, we use the pessimistic (pseudo-conjunctive) procedure (Figueira et al. 2010). Each alternative is assigned to the highest category for which that alternative outranks its lower bound in the reference profile: a belongs to C_1 if it does not outrank b_1, it belongs to C_2 if it outranks b_1 but it does not outrank b_2, etc.

ELECTRE TRI requires several parameters such as preference (p_j), indifference (q_j) and veto (v_j) thresholds, and weight-importance coefficients (w_j). Weight-importance expresses the relative importance of the criteria; indifference is the largest difference in performance, for a factor, that may be considered insignificant; preference is the smallest difference in performance constituting a clear advantage and the veto threshold indicates a difference in performance (discordance) so large that it vetoes an outranking, even if all other criteria agreed to it.

Each outranking relation is established in four steps: criteria concordance indexes and global concordance indexes calculation; discordance index calculation; credibility degree calculation; and establishing the outranking relation through the cutting level ($\lambda \in [0.5, 1]$), that is defined as the lowest degree of credibility required to assign an alternative to a specific category. For further details on the algorithm and concepts of this method, see: Yu (1992), Roy and Bouyssou (1993), Roy (1996), Mousseau and Dias (2004), and Figueira et al. (2005, 2010).

4. Case Study

4.1. Characterization of Dairy Farms

The study area is corresponds to the Primary Dairy Basin in the EDM Region located in the Northwest Region of Portugal and it consists of 10 counties: Viana do Castelo, Barcelos, Esposende, Póvoa de Varzim, Vila Nova de Famalicão, Vila do Conde, Santo Tirso, Trofa, Matosinhos and Maia.

The 1705 dairy farms analyzed in this study are mainly located across the region of Barcelos (35% of total), Vila do Conde (21%), Póvoa de Varzim (12%) and Vila Nova de Famalicão (9%). Together, these farms comprise a total of 106,958 animals. Barcelos (32% of the animals), Vila do Conde (25%) and Póvoa do Varzim (14%) are the counties with more animals.

The total area of farms in the study measures 20.331,00 hectares (ha), of which 17.318 ha are Land Area Under Agricultural Production (LAUAP). The average physical dimension of the total area of each dairy farms is 12 ha.

4.2. Criteria Selection

Many authors have analyzed the criteria (variables) that explain the processes of transformation of agricultural production structures. Although their conclusions do not coincide necessarily, they generally reinforce three factors that are unanimously considered influential in the possible paths of structural adjustment (expansion or early abandonment of agricultural activity): the political and economic factors, the producer (or farmer) and his family, and also the adaptation to environmental requirements (Santos, 2011).

Sustainable livestock systems should indeed be environmentally friendly, economically viable for farmers, and socially acceptable, notably for animal welfare. In this work, the criteria are divided into two dimensions: socio-economic and environmental. The definition and evaluation of performance criteria involved two experts from Environmental and Zoo-technic fields for the environmental dimension and one expert from Agro-Economic field for the socio-economic dimension. In defining the criteria, the experts took into account the new rules for bovine licensing, the Code of Good Agricultural Practices, Manual Fertilization of crops and Water Act, and the European legislation and recommendations.

Our study focuses on data-driven assessments carried out at farm level, from available databases. For the construction of the criteria, the database obtained from the survey developed during the implementation of the Basin Plan Land Dairy Primary Entre Douro and Minho (POBLPEDM, 2007) was considered, which is the most recent and comprehensive survey developed in the region. The survey covers some environmental, social and economic issues of dairy farms; but, not having been developed with the aim of evaluating dairy farms sustainability, it does not have all the information that would be ideal to perform a complete economic and environmental analysis. Financial and economic data as net revenue and equity to debt are not available. Still, we consider (along with the experts consulted) that, despite not include the ideal set of criteria, it would be sufficient to illustrate the methodology developed in this case study.

4.2.1. Socio-Economic Criteria

For an analysis of farmers, Ondersteijn et al. (2003) points out in the first place the management capabilities of the producer, saying that they rely on a set of personal characteristics such as their influences, motivations, skills or biographical factors, age (SE1 criterion in Table 1), time dedication to work (evaluating the situation of professional activity – SE2 criterion) and also the professional development of the producers (indicating their interest in updating and acquiring new knowledge to apply to their operation – SE3 criterion). It is the producer who makes substantive decisions such as those related to the production system investments. The characterization of farmers and the important role that their families play in the management and the activities of the project are very important.

The importance of family is related to the family tradition in this business, considering the involvement of its members and the consequent existence of successors who demonstrate interest and dedication to farming (SE4 criterion). Ochoa et al. (2007) discuss the importance of the succession of generations in agriculture in order to maintain the activity. Furthermore, the farmer´s (and their family) income earned out of their dairy farming activity indicates dependence and consequently a need for business continuity (SE5 criterion).

Since dairy farming is still a particularly demanding activity in terms of investment in infrastructures and facilities, machinery and effective force, there can be an obstacle to maintain the motivation for new investments (SE6 criterion). Finally, it is important analyse the future perspectives (shutdown, decrease, relocate, stabilize and increase) of the producer regarding growth and business continuity (SE7 criterion).

The seven socio-economic criteria established are summarized in Table 1 (for more details see Silva, 2015).

Table 1: Socio-Economic Criteria and their Objectives.

ID	***Socio-Economic Criteria***	***Objective***
SE1	Producer´s Age	Minimize
SE2	Professional Dedication (%)	Maximize
SE3	Professional Development	Maximize
SE4	Successors	Maximize
SE5	External Income	Minimize
SE6	Investment Strategies	Maximize
SE7	Future Perspectives	Maximize

4.2.2. Environmental Criteria

In the more intensive systems that dominate dairy production, the main environmental issues are nutrient contamination of soil, groundwater and air pollution. The main impact is on soil integrity which is affected by increased use of fertilizers, feed additives and the more concentrated use of waste products such as manure. Other impacts derive from the pollution of groundwater with nitrates, pesticides and contaminated surface water. The impact of dairying on the atmosphere arises from de-nitrification, the production of methane, ammonia volatilization and carbon dioxide.

Consumers and society in general have forced the introduction of new attitudes with regard to questions of dairy production in particular, such as animal welfare, environmental impact and traceability and food health and safety regulations.

Thus, in the environmental dimension, seven criteria are considered that are defined on the one hand, in accordance with legal requirements and, on the other hand, with good practice to be implemented, referring to the cross-compliance, i.e. compliance with standards such as environmental health and animal-well-being, maintaining the land in good agricultural and environmental conditions, and respecting the rules on public health and pollution. These are summarized in Table 2.

Table 2: Environmental Criteria and their Objectives.

ID	*Environmental Criteria*	*Objective*
E1	Storage Capacity of Manure	Maximize
E2	Number of Livestock units per area of manure application	Minimize
E3	Excess on nitrogen balance	Minimize
E4	Total Production of Greenhouse Gases	Minimize
E5	Percentage of Storage Structures Near Water Lines	Minimize
E6	Percentage of Individualized Collection of Rainwater	Maximize
E7	Animal Well-being	Maximize

The first criterion (E1) concerns the farm's capacity to store manure produced by its animals and is based on the total storage capacity of effluents (m^3) and the number of standard livestock units (LU), according to rules for licensing bovine farming. The E2 criterion takes into account the amounts allowed for the incorporation of organic nitrogen in the soil in terms of LU/LAUAP, according the Wastewater Management Plan. The E3 criterion determines the excessive incorporation of nitrogen excreted in relation to the forage area. The E4 criterion defines the production of greenhouse gases acoording the production of milk of each dairy farm. The E5 criterion addresses the risk of contaminating water lines and consists of the ratio between the number of storage points less than 25 meters away from the water lines and the total number of existing storages on the farm. The E6 criterion addresses the use of water and consists of the proportion of the farms' storage facilities that have individualized collection of rainwater, that increases the effectiveness of storage capacity, provides fertilizers of better quality, reuses water, and thereby reduces the water burden. The E7 criterion concerns animal well-being, assessed by considering the ratio between the covered area (stall) and the number of cows on the farm. For more details on these criteria see Silva (2015).

4.3. Application of the ELECTRE TRI Method

ELECTRE TRI is a suitable MCDA method in sorting problems, as addressed in this case study of dairy farms classification according to their environmental and socio-economic viability. In a classification of dairy farms in which social, economic and environmental aspects are at stake, it is often important to use known standards (references profiles defined in legislation) to define various categories.

ELECTRE TRI assigns a set of alternatives to a group of predefined categories, considering multiple criteria. In this work, the set of alternatives are dairy farms and the predefined categories consist of three viability levels: *Not Viable* (Category 1), *Barely Viable* (Category 2) and *Viable* (Category 3). We study the environmental and socio-economic viability separately due to the distinct nature of these two areas, and then derive a broader perspective on sustainability.

The ELECTRE TRI method performs the assignment of each alternative to a category, based on how its performances compare with the reference profiles. We define two reference profiles b_1 and b_2, where b_2 indicates the minimum performance standards that a farm should have in order to be placed in Category 3 and b_1 indicates the minimum performance standards in order to be in Category 2. The profiles (Table 3) have been defined by experts.

Table 3: References Profiles (b_1 and b_2) for the Environmental and Socio-Economic classification using the ELECTRE TRI method.

	ENVIRONMENTAL						
	E1	E2	E2	E4	E5	E6	E7
b_2	1.0	1.0	0	1000	0.25	0.5	5.0
b_1	0.8	1.3	300	4000	0.50	0.2	2.5

	SOCIO-ECONOMIC						
	SE1	SE2	SE3	SE4	SE5	SE6	SE7
b_1	50	4	3	4	2	1	4
b_1	65	2	0	2	4	0	1

Table 4: Environmental and Socio-Economic Thresholds of the ELECTRE TRI method: indifference (q_j), preference (p_j), and veto (v_j) thresholds, and weight-importance coefficients (w_j).

	ENVIRONMENTAL						
	E1	E2	E3	E4	E5	E6	E7
q_j	0.10	0.10	100	300	0.05	0.10	1
p_j	0.20	0.20	150	800	0.10	0.20	2
v_j	0.50	0.50	500	6000	0.55	0.45	3
w_j	0.20	0.20	0.20	0.05	0.15	0.05	0.15

	SOCIO-ECONOMIC						
	SE1	SE2	SE3	SE4	SE5	SE6	SE7
q_j	10	1	1	1	1	0	1
p_j	15	2	2	2	2	0.5	2
v_j	20	4	3	4	3	1	3
w_j	0.2	0.1	0.1	0.2	0.1	0.15	0.15

Taking into account the inaccuracies and uncertainties in the performances of the alternatives, thresholds of indifference (q_j) and preference (p_j) were defined for each criterion, in order to better model the preferences of the decision makers. The veto threshold (v_j) that is used in the discordance test was also determined for each criterion, as well as the weight (w_j) of each criterion that is used in the concordance test (Table 4). We calculated the partial concordance index for each of the criteria and then the comprehensive concordance indices of global agreement that indicates how much is the performance of the alternatives consistent with "a outranks b_i", i.e. aSb_i. Most of the criteria (sum of weights above the cutting level) should be in favor of this statement. Next, we calculated the discordance indexes for each criterion, representing the statement opposed to aSb_i.

Finally, the credibility index ($\sigma(a,b_i)$) was calculated, which determines how credible the statement aSb_i is according to the comprehensive concordance and discordance indexes. For the classification procedure it was also necessary to define a cutting level (λ), in this case λ = 0.6, to determine the binary relationship between the alternatives and reference profiles, such that: $aSb_i \Leftrightarrow \sigma(a,b_i) \geq \lambda$. The cutting level indicates that 60% of the criteria (considering their weights) have to agree in order to assign an alternative to a specific category.

Individual meetings were held with each expert, where the method was presented and the values of its parameters were elicited.

5. Results

A Web Multicriteria Spatial Decision Support System (Web MC-SDSS) named *ELECTRE TRI in ArcGIS* has been applied in this work (for details see Silva et al., 2014) to classify the environmental and socio-economic viability of dairy farms, obtaining the results presented in Figure 1 and the map with the viability classification.

Concerning environmental viability, we conclude that the majority (60%) of dairy farms are classified as *Environmentally Not Viable* (Category 1). On the other hand, the assessment of the Socio-economic Viability, considering the defined criteria, indicates that only a small minority (3%) is classified as *Not Viable*. The analysis of Figure 1 shows that environmental issues are those that mostly affect the viability of dairy farms in the EDM Region, for non-compliance with environmental requirements. Another important aspect is the noticeable difference between the classifications obtained in the evaluation of Environmental and Socio-economic Viability of farms. Let us note that these two dimensions are not completely independent: a farm could improve its environmental condition by making more investments, at the cost of hindering its economic condition, unless subsidies or other incentives were made available (unfortunately in the current situation of this region such investments are not likely). Thus, the classification *Environmentally Not Viable* should be understood as meaning that a farm is hardly viable due to its environmental condition.

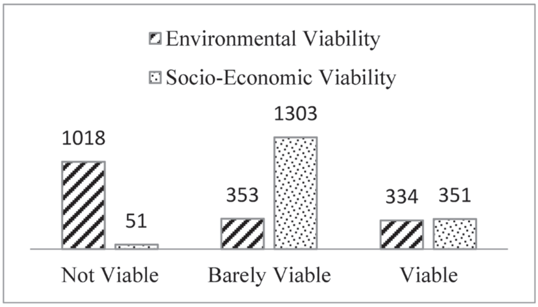


Figure 1: Number of farms classified in each category.

Table 5 compares the classifications obtained according to Environmental Viability and the Socio-economic Viability. Only 2% (30) of the dairy farms are classified as *Environmentally Not Viable* and *Socio-economically Not Viable*; 16% (271) of the farms are classified as *Barely Viable*, considered to be environmental and Socio-economic *Barely Viable* simultaneously. Only 3% (48) of the farms are classified as *Viable* according to the Environmental and the Socio-economic perspectives. The largest group, with 44% (754) of the total number dairy farms, is classified as both *Socioeconomically Barely Viable* and *Environmentally Not Viable*.

Table 5: Cross tabulation of the classification of Environmental and Socio-Economic Viability.

		Socio-Economic Viability			
		Not Viable	**Barely Viable**	**Viable**	
Environmental Viability	**Not Viable**	30	754	234	1018
	Barely Viable	13	271	69	353
	Viable	8	278	48	334
		51	1303	351	1705

After obtaining these results, the experts were again involved in a new discussion and analysis, considering that in order to conduct the evaluation of a farm as to the overall Sustainability, it would make sense to reclassify (or sub-divide) the categories obtained in the study of Environmental and Socioeconomic Viability into the following new set of categories:

Highly Sustainable - dairy farms classified as Viable on Socio-Economic and Environmental Viability, simultaneously;

Moderately Sustainable - dairy farms classified as Barely Viable in Environmental Viability and Socio-Economic Viability in Viable or vice versa;

Marginally Sustainable - dairy farms classified as Barely Viable in Environmental and Socioeconomic Viability;

Not Sustainable - all other.

This reclassification yields the results depicted in Figure 2.

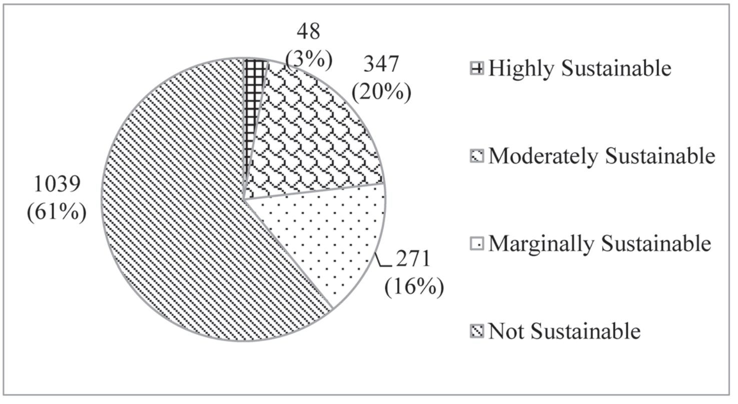


Figure 2: Number and percentage of dairy farms classified in each of four categories of Sustainability

Figure 3 presents the corresponding kernel density maps for a better visualization of the distribution of the density of each of these four categories. The Kernel Density (available in ArcGIS software) consists in an interpolation which is intended to generate a continuous surface through sample points (in this case, exploits), which is determined by number of farms classified in each category in relation to the area (km^2).

Not Sustainable	Marginally Sustainable	Moderately Sustainable	Highly Sustainable

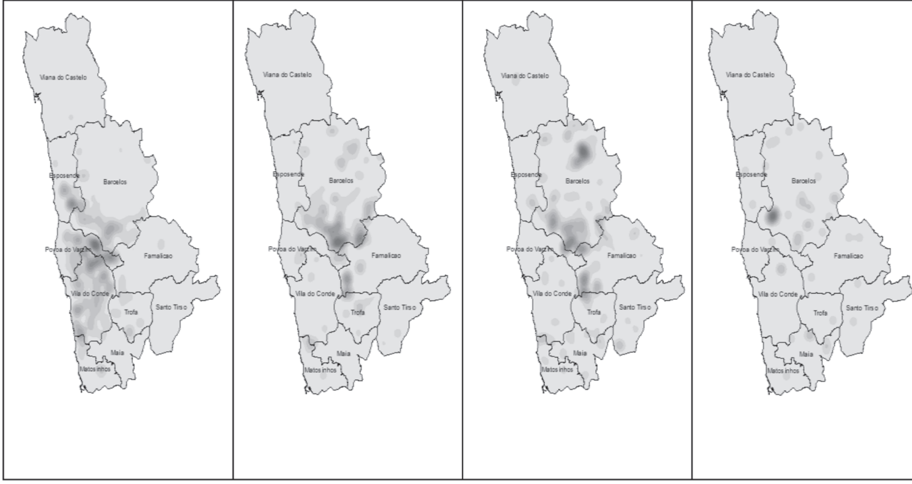


Figure 3: Density of distribution of dairy farms according to the new (aggregated) perspective of Sustainability

Analyzing briefly the spatial distribution of the sustainability classification, we can highlight the following: about 90% of the dairy farms in the county of Vila do Conde are classified as Not Sustainable, as well as 88% of the dairy farms in Póvoa de Varzim and 82% of those in Esposende – which are very high percentages. The counties of Maia, Santo Tirso and Barcelos are those with better classifications: there is a higher percentage of farms classified as Highly Sustainable and the lowest percentage of dairy farms classified as Not Sustainable (but still around 40%). The dairy farms classified as Not Sustainable, have in general a high animal density and reduced dimensions, and have difficulties in growth due to their proximity to urban areas. In turn, in dairy farms classified as High Sustainable there is a better fit between the number of animals and their dimension but especially they have improved environmental conditions.

6. Sensitivity Analysis

Setting the required input data for the ELECTRE TRI method (weights, thresholds of indifference and preference, veto, reference profiles and cutting level) is a complex and ambiguous task because of the associated uncertainties. To check how changes in the parameters influence the results, we performed a sensitivity analysis, considering different scenarios, in order to verify the robustness of the results.

Two scenarios were defined, in order to develop a sensitivity analysis in the study of Environmental and Socio-economic Viability, which we will call the Benevolent and Stringent scenarios. These scenarios are characterized by a variation of 5% in the profiles of the categories, in the preference, indifference and veto thresholds and in the cutting level (there was no change in the weights in any of the scenarios). We aim at evaluating changes in the results obtained in these two scenarios when compared with the classification of originally obtained data, which we shall call the Original scenario.

The Benevolent Scenario is characterized by changing the parameters of the ELECTRE TRI method as follows: the cutting level is decreased by 0.03 (becoming 0.57), the reference profiles of the categories increase 5% if a criterion is to be minimized and decrease 5% if a criterion is to be maximized, and thresholds of preference, indifference and veto are increased by 5%. This is a less demanding scenario than the Original one.

In turn, the Stringent Scenario considers the changes referred to in the following parameters of the ELECTRE TRI method: the cutting level is increased by 0.03 (becoming 0.63), the reference profiles of the categories decrease 5% if the criterion is to be minimized, and increase 5% if a criterion is to be maximized, and thresholds of preference, indifference and veto are decreased by 5%. This is a more demanding scenario than the original one.

The results of comparing the three scenarios are summarized in Figure 4.

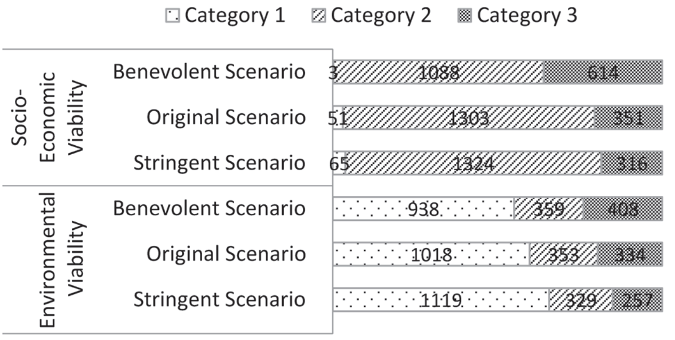


Figure 4: Comparison of the three scenarios regarding the classification of Socio-Economic and Environmental Viability (Category 1 - Not Viable, Category 2 – Barely Viable, Category 3 - Viable).

In the Stringent Scenario, the percentage change in the socio-economic assessment is lower than in the environmental assessment. The percentage change in each category is equal to or less than the variation caused in the parameters and, as such, it can therefore be said that despite the change in the final results this is not very relevant.

In the Benevolent Scenario, it is noted that the number of dairy farms changing classification in environmental analysis is less than the number changing in the socio-economic analysis. This means that, even when requiring less compliance with environmental requirements, no major changes in the classification of farms are visible. Despite the fact that there were some changes obtained in the classification, those were not very significant.

We considered also other Stringent and Benevolent scenarios characterized by a greater variation: 15% in the profiles of the categories, in the preference, indifference and veto thresholds and cutting level (becoming 0.51 in Benevolent and 0.69 in Stringent scenarios). In both new Stringent and Benevolent scenarios the changes in the classification, in all four sustainability categories, is less than percentage change, not exceeding 7%, compared with results of Original scenario. In Figure 5, it is possible compare the number of dairy farms classified in each of four categories of Sustainability for the 5 scenarios analysed: Stringent

(5% and 15%), Original and Benevolent (5% and 15%) scenarios. These different scenarios considered support the stability of the results obtained by ELECTRE TRI method.

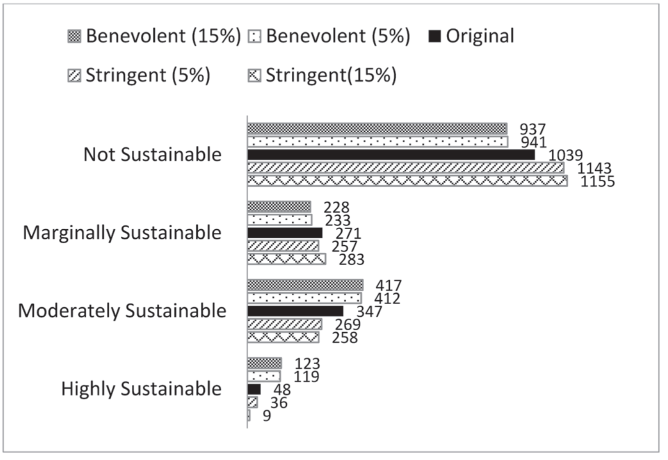


Figure 5: Number of dairy farms classified in each of four categories of Sustainability by comparing the five scenarios.

7. Discussion

Considering the results, one notes the worrying situation in which many farms in the EDM Region are found regarding their sustainability. In this region, the economic component contributes greatly to the sustainability of farms, but the breach of environmental requirements is having a greater influence on their sustainability. There are reasons, natural and historical, which tend to hinder the resolution of environmental problems, such as the ownership structure of dairy farms and their dispersion, the morphological and geological nature of the terrains and the proximity to urban areas. The solution to this environmental problem is an evolution of the structure of dairy farms to reduce the intensity of production, but also for obtaining dimension which will amortize investments in equipment and training. However, the economic framework is itself heavily influenced by the nature of the support regime (largely price support), technology development and structural change in the production sector (plus structural change in the upstream and downstream supply chain). So it is necessary to find diverse environmentally and economically sustainable solutions, for those farms that are moderately or marginally sustainable. As for the farms that are not sustainable, although solutions should be sought to enable the continuation of the dairy farming activity, a reconversion to a different activity is also an alternative to consider.

For this study, the ELECTRE TRI method seemed appropriate as it sorts alternatives (dairy farms) by preference-ordered categories. Indeed, in classification of dairy farms, where various environmental, economic and social aspects are at stake, it is often important to use known standards (reference profiles) to define various categories of sustainability. This requires capability of evaluating each farm in absolute terms, not just in comparison with other peers, as well as the need to include evaluation aspects expressed in different units, using any type of scales (including qualitative). The ELECTRE TRI refuses the possibility of total compensation between the alternative's performances on the criteria. Thus it prevents that a farm with very good performance on some criterion compensates a very poor performance on another criterion and achieves the best category despite that major weakness.

Finally, the sensitivity analysis developed verifies the robustness of the results of dairy farms sustainability assessment.

8. Conclusions

The complexity of sustainability assessment can be mitigated with the use of MCDA techniques. In this study, we found that MCDA can be useful in the evaluation of sustainability that encompasses multiple dimensions. The Outranking Method ELECTRE TRI proved to be an adequate method, preventing unwarranted compensation of poor performances, and yielding a classification of dairy farms on sustainability according to the parameters and values assigned by experts. The integration of GIS capabilities and MCDA techniques reinforces the advantages of both tools.

The main highlights of the methodology followed in this paper are the following. First, the sustainability assessment is based explicitly on multiple criteria, rather than trying to find an indicator such as carbon footprint or a financial ratio that would synthetize environmental and economic sustainability, respectively. Second, the evaluation is grounded on a classification method which is parsimonious in terms of the value judgments it requires, namely not requiring substitution rates among the criteria, nor pairwise comparisons among a large number of alternatives. Although the results are somewhat coarse (when compared to a full ranking), classification results are familiar and widely accepted by decision makers (e.g., efficiency labeling). Third, a separate assessment of sustainability dimensions is made, directing attention to different potential weaknesses of the farms. Nevertheless, these two dimensions of sustainability can be conveniently summarized into an easy to understand overall classification, namely to present graphically in maps.

In this work it was also possible to analyze the geographical distribution of the classification of dairy farms and to realize that the areas which reflect higher levels of sustainability are those that have an increasing concentration of farms, whereas areas that reflect the worst levels of sustainability are those that are subject to abandonment.

The structure, process and application of sustainability assessment can be extended to other areas besides agriculture. In this decision process, options were taken in each phase, which leave space for future studies and the development of new approaches, such as the aggregation of these or other criteria, with the involvement of more experts with the application of other(s) multi-criteria method(s), with application for more updated reinvestigation.

References

Andreoli, M., Tellarini, V., 2000. Farm sustainability evaluation: methodology and practice. *Agriculture, Ecosystems and Environment,* 77(1): 43-52.

Antunes, P., Karadzic, V., Santos R., Beça P., Osann, A., 2011. Participatory multi-criteria analysis of irrigation management alternatives: the case of the Caia irrigation district, Portugal. *International Journal of Agricultural Sustainability*, 9(2): 334-349.

Becker, B., 1997. Sustainability assessment: a review of values, concepts, and methodological approaches. Consultative Group on International Agricultural Research, The World Bank, Washington, D.C., 63 pp.

Dantsis, T., Douma, C., Giourga, C., Loumou, A., Polychronaki, E.A., 2010. A methodological approach to assess and compare the sustainability level of agricultural plant production systems. *Ecological Indicators*, 10: 256–263.

Dias, L., Mousseau, V., Figueira, J., Clímaco, J., 2002. An aggregation/disaggregation approach to obtain robust conclusions with ELECTRE TRI. *European Journal of Operational Research*, 138: 332–348.

Dillon, E. J., Hennessy T., Hynes, S., 2010. Assessing the sustainability of Irish agriculture. *International Journal of Agricultural Sustainability*, 8(3): 131-147

Figueira, J., Mousseau, V., Roy, B., 2005. Electre Methods. In: Figueira, J.; Greco, S.; Ehrgott, M. Multiple Criteria Decision Analysis: State of The Art Survey. New York: Springer, Cap. 4.

Figueira, J.R., Greco S., Roy B., Słowinski R., 2010. ELECTRE methods: Main features and recent developments. In C. Zopounidis and P. Pardalos (Eds.), Handbook of Multicriteria Analysis, Chapter 4, New York, USA: Springer.

Gómez-Limón, J. A., Riesgo, L., 2009. Alternative Approaches to the construction of a composite indicator of agricultural sustainability: An application to irrigated agriculture in the Duero basin in Spain. *Journal of Environmental Management*, 90: 3345-3362.

Hansen, J.W., 1996. Is agricultural sustainability a useful concept? *Agricultural Systems*, 50(1): 117-143.

Hediger, W., 1999. Reconciling 'weak' and 'strong' sustainability. *International Journal of Social Economics*, 26:7/8/9, 1120-1143.

Lebacq, L., Baret, P.V., Stilmant, D. 2013. Sustainability indicators for livestock farming. A review. *Agronomy for Sustainable Development*, 33:311-327.

Mousseau, V., Dias, L., 2004. Valued outranking relations in ELECTRE providing manageable disaggregation procedures. *European Journal of Operational Research*, 156: 467–482.

Munda, G., 2005. Measuring sustainability: a multi-criterion framework. *Environment, Development and Sustainability,* 7(1):117-134.

Ochoa, A.M.A, Oliva, V.C, Sáez, C.A, 2007. Explaining farm succession: the impact of farm location and off-farm employment opportunities. *Spanish Journal of Agricultural Research*, 5(2): 214-225.

Ondersteijn, C.J.M., Giesen, G.W.J., Huirne, R.B.M., 2003. Identification of farmer characteristics and farm strategies explaining changes in environmental management and environmental and economic performance of dairy farms. *Agricultural Systems*, 78: 31-55.

POBLPEDM, 2007. Plano de Ordenamento da Bacia Leiteira Primária do Entre Douro e Minho. Escola Superior Agrária do Instituto Politécnico de Viana do Castelo, Instituto para o Desenvolvimento Agrário da Região Norte e Universidade do Porto – CIBIO.

Raju, K.S., Duckstein, L., Arondel, C., 2000. Multicriterion Analysis for Sustainable Water Resources Planning: A Case Study in Spain. *Water Resources Management*, 14: 435–456.

Rigby, D., Woodhouse, P., Young, T. and Burton, M., 2001. Constructing a farm level indicator of sustainable agricultural practice. *Ecological Economics*, 39(3): 463-478.

Roy, B., 1985. Méthodologie multicritère d'aide à la decision. Economica, Paris.

Roy B., Bouyssou, D., 1993. Aide multicritère à la décision: méthodes et cas, Economica, Paris.

Roy, B., 1996. Multicriteria Methodology for Decision Aiding. Kluwer Academic Publishers.

Santos, J.C.S.M., 2011. Viabilidad de las explotaciones lecheras en un escenario de desconexión del apoyo y condicionalidad: Un estudio de caso en el noroeste de Portugal. Ph.D. Dissertation, Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros Agrónomos.

Silva, S., 2015. Sistema de Apoio à Decisão Espacial Multicritéio na localização de centrais de biogás. Tese de Doutoramento em Gestão-Ciência Aplicada à Decisão. Faculdade de Economia, Universidade de Coimbra.

Silva, S., Almeida-Alçada, L., Dias L., 2014. Development of a Web-based Multi-criteria Spatial Decision Support System for the assessment of environmental sustainability of dairy farms, *Computers and Electronics in Agriculture*, 108: 46–57.

Stockle, C.O., Papendick, R.I., Saxton, K.E., Campbell, G.S., van Evert, F.K., 1994. A framework for evaluating the sustainability of agricultural production systems. *American Journal of Alternative Agriculture*, 9(1-2): 45-50.

Stoorvogel, J.J., Antle, J.M., Crissman, C.C. and Bowen W., 2004. The tradeoff analysis model: integrated biophysical and economic modeling of agricultural production systems. *Agricultural Systems*, 80(1): 43-66.

Sydorovych, O., Wossink, A., 2008. The meaning of agricultural sustainability: Evidence from a conjoint choice survey. *Agricultural Systems*, 98(1): 10-20.

United Nations, 1987. Brundtland Report, Report of the World Commission on Environment and Development: Our Common Future, Oxford University Press.

van Calker, K., Berentsen, P., Romero, C., Giesen, G., Huirne, R., 2006. Development and application of multi-attribute sustainability function for Dutch dairy farming systems. *Ecological Economics*, 57: 640-658.

van Calker, K., Berentsen P., Giesen G., Huirne, R., 2005. Identifying and Ranking Attributes that Determine Sustainability in Dutch Dairy Farming. *Agriculture and Human Values*, 22: 53-63.

Yu, W., 1992. ELECTRE TRI: Aspects méthodologiques et manuel d'utilisation. Document du LAMSADE No. 74, Université Paris-Dauphine.

Pedro Godinho is an Assistant Professor at the Faculty of Economics of the University of Coimbra, which he joined in 1993. He holds a PhD in Management, a Master in Financial Economics and a degree in Informatics Engineering from the University of Coimbra. He is also a Researcher at GEMF – Group for Monetary and Financial Studies, Coimbra. His research interests include project appraisal, genetic algorithms, real options, game theory and project management. He is the author or co-author of more than forty refereed publications in the fields of finance and operations research, and two books. He is a reviewer for several international journals, and has a vast experience in teaching both graduate and undergraduate students.

Joana Matos Dias is an Assistant Professor at the Faculty of Economics, University of Coimbra, and a researcher at Instituto de Engenharia de Sistemas e Computadores (INESC-Coimbra). She has a BSc in Computers Engineering by the University of Coimbra, a MSc in Operations Research by the University of Lisbon, a MSc in Quantitative Finance by the University of London, and a PhD in Management Science by the University of Coimbra.
She is author or co-author of more than 50 papers in international journals, book chapters or conference proceedings. She is a reviewer of several international journals, and has a vast experience in teaching, to both graduate and undergraduate students, courses in the area of operational research, informatics, logistics, among others.

Série Investigação

•

Imprensa da Universidade de Coimbra
Coimbra University Press
2015

mais CENTRO
Programa Operacional Regional do Centro

QREN
QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL
PORTUGAL 2007.2013

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

• U C •

IMPRENSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS