NOTAS ECONÓMICAS 56

ARTICLES

Giuseppe Munda
Evaluation and Assessment of Sustainability Policies

Renato Rosa / Ramiro Parrado / Francesco Bosello
Trading REDD Credits in International Carbon Markets: Interactions among International Trade, Carbon and Agricultural Markets

R. Quentin Grafton / Safa Fanaian
Responding to the Global Challenges of 'Too Much, Too Little and Too Dirty' Water: Towards a Safer and More Just Water Future

LETTERS

Carlos Carreira / Pedro Silva
The Determinants of Trade Credit During and After the 2008 International Financial Crisis

Nuno Gonçalves / Carlos Carreira
Intangible Capital and Productivity of Portuguese Firms in the Last Decade (2010-2019)

Imprensa da Universidade de Coimbra
Coimbra University Press

Faculdade de Economia da
Universidade de Coimbra

NOTAS ECONÓMICAS 56

Informações aos autores (Submission guidelines)
https://impactum-journals.uc.pt/notaseconomicas/about/submissions

Publicação resumida e indexada por (Indexing and Abstracting)
EconLit
RePEc
REDIB
latindex

Design
Imprensa da Universidade de Coimbra

Execução gráfica
BOOKPAPER

Tiragem 200 exemplares

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

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

FCT Fundação para a Ciência e a Tecnologia

Número avulso
Portugal 7,5 €
Outros países 15 €

Condições de assinatura bienal
Portugal: 20 €
Outros países: 40 €

ISSN 0872-4733

ISSN Digital 2183-203X

DOI https://doi.org/10.14195/2183-203X_56

Depósito Legal 64932/93

Contents

Articles

Letters

Giuseppe Munda
European Commission, Joint Research Centre
(JRC), Ispra, Italy
giuseppe.munda@ec.europa.eu
orcid.org/0000-0003-3398-2971

DOI: https://doi.org/10.14195/2183-203X_56_1

Evaluation and Assessment of Sustainability Policies

Avaliação de Políticas Sustentáveis

Giuseppe Munda

Received for publication: January 10, 2022
Revision accepted for publication: March 28, 2022

ABSTRACT
Sustainability policy evaluation and assessment seeks to answer the key question, *sustainability of what and whom*? Consequently, sustainability issues are multidimensional in nature and feature a high degree of conflict, uncertainty and complexity. Social multi-criteria evaluation (SMCE) has been explicitly designed for public policies; it builds on formal modelling techniques whose main achievement is the fact that the use of different evaluation criteria translates directly into plurality of values and dimensions underpinning a policy process. SMCE aims at being inter/multi-disciplinary (with respect to the technical team), participatory (with respect to the community) and transparent. SMCE can help deal with three different types of sustainability-related policy issues: 1) epistemological uncertainty (human representation of a given policy problem necessarily reflects perceptions, values and interests of those structuring the problem); 2) complexity (the existence of different levels and scales at which a hierarchical system can be analyzed implies the unavoidable existence of non-equivalent descriptions of it both in space and time); and 3) mathematical manipulation rules of relevant information (compensability versus non-compensability, preference modelling of intensities of preference, mixed information on criterion scores, weights as trade-offs versus weights as importance coefficients, choice of a proper ranking algorithm). This paper focuses on the these three issues and provides an overview of the SMCE approaches to them.
Keywords: Complexity theory; social multi-criteria evaluation; history of economic thought: social choice; SOCRATES software.

JEL Classification: B16; B23; C44; D04; D71; D81; H4; Q01.

Acknowledgement: This article has been developed in the context of the activities of the JRC Competence Centre on Modelling and Decision Analysis. The views expressed are purely those of the writer and may not in any circumstances be regarded as stating an official position of the European Commission.

1. Complexity, incommensurability and sustainability policies

> "… there is such a long tradition in parts of economics and political philosophy of treating one allegedly homogeneous feature (such as income or utility) as the sole 'good thing' that could be effortlessly maximized (the more the merrier), that there is some nervousness in facing a problem of valuation involving heterogeneous objects,… And yet any serious problem of social judgement can hardly escape accommodating pluralities of values,… We cannot reduce all the things we have reason to value into one homogeneous magnitude." (Sen, 2009, p. 239)

Since its origins, the concept of sustainable development is often considered a policy framework for win-win strategies (e.g. Barbier, 1987), allowing the full achievement of a plurality of goals in a variety of domains; but is this possible? A legitimate question is: sustainable development *of what and whom?* (Allen et al., 2002). Norgaard (1994, p.11) writes: "*consumers want consumption sustained, workers want jobs sustained. Capitalists and socialists have their "isms", while aristocrats and technocrats have their "cracies*".

Complexity arises when something is difficult to understand and impossible to analyse by using simple frameworks. However, when dealing with sustainability policy problems, there is a natural temptation to try to reduce them to simpler, more manageable elements. Although many definitions of complexity exist, a key common characteristic of complex systems is that the information space required to represent *relevant* aspects of a complex system cannot be compressed without losing relevant information (Gell-Mann, 1994; Prigogine and Stengers, 1981; Rosen, 1985, 1991; Simon, 1962).

To make things more difficult, systems involving humans are *reflexively* complex. Reflexive systems display two peculiar characteristics: "*awareness*" and "*purpose*", both requiring an additional "jump" in describing complexity. The presence of self-consciousness and purpose (*reflexivity*) means that these systems can continuously add new relevant qualities/attributes to be considered when explaining, describing or forecasting their behaviour (i.e. human systems are learning systems); this implies that complex adaptive systems become something else over time (Funtowicz et al., 1999).

Moreover, the existence of *different levels and scales* on which a hierarchical system can be analysed implies the unavoidable existence of non-equivalent descriptions of it. Even a simple "objective" description of a geographical orientation is impossible without taking an arbitrary subjective decision on the relevant system scale. In fact, the same geographical place, for example in Europe, may be considered to be in the north, south, east or west according to the scale chosen as a reference point (the whole Europe, a single State, a region, a specific place, etc.). Therefore, the problem of *multiple identities* in complex systems cannot be interpreted solely in terms of *epistemological plurality* (non-equivalent observers), but also necessarily in terms of *ontological characteristics* of the observed system (non-equivalent observations) (Giampietro et al., 2006, 2012; Giampietro and Mayumi, 2000, 2018). There is an unavoidable political dimension in any scientific description in as much as some decision is required regarding how to frame a policy problem. Therefore, to reach a ranking of policy options, there is a prior need to decide *what is important* for different social actors as well as *what is relevant* for the representation of the real-world entity described in the model.

No mathematical model, even if legitimate in its own terms, can be sufficient for a complete analysis of the reflexive properties of a real-world problem. These reflexive properties include the human dimensions of e.g. the ecological change and the transformations of human perceptions along the way. The learning process that takes place while analyzing the issue and defining policies will itself influence perceptions and alter significantly the decisional space in which alternative strategies are chosen. At the other end, institutional and cultural representations of the same system, while also legitimate, are on their own insufficient to define what should be done in any particular case.

In general, these concerns were not considered very relevant by scientific research as long as time was considered an infinite resource. On the other hand, the new nature of the problems faced in this third millennium implies that, when dealing with problems that may have long term consequences, we are confronting issues "*where facts are uncertain, values in dispute, stakes high and decisions urgent*" (Funtowicz and Ravetz, 1991).

Scientists cannot therefore provide any useful input without interacting with the rest of society while the rest of the society cannot make any sound decision without interacting with scientists. That is, the question of "how to improve the quality of a policy process" must be put, rather quickly, on the agenda of "scientists", "policy-makers" and indeed of society as a whole. This extension of the "peer community" is essential for maintaining the quality of the process of decision-making when dealing with reflexive complex systems. In relation to this objective Funtowicz and Ravetz have developed a new epistemological framework called *"Post-Normal Science",* with which it is possible to deal better with two crucial aspects of science in the policy domain: *uncertainty* and *value conflict*. The term "post-normal" signals a divergence from the puzzle-solving exercises of normal science, in the Kuhnian sense.

In operational terms, one should admit that there is no optimal solution to the management of complex systems. If we want to avoid reductionism, it will be necessary to take incommensurable dimensions into account and to use different scientific languages describing disparate but legitimate representations of the same system. Accepting the complexity of natural and social systems is the first step in understanding how policy problems should be structured. A second step is to choose appropriate management and policy tools: *those that address rather than ignore complexity.*

Multi-criteria decision analysis is becoming more and more popular both in the private and public sectors (see e.g. Figueira et al., 2016). Arrow and Raynaud (1986) considered the so-called *"industrial outranking problem"*, where a typical business-person is the reference decision-maker, who wishes *"to make safer the equilibrium of the productions of the firm"* (Arrow and Raynaud, 1986, p. 9). Typical business criteria may be market standing, innovation level, productivity, profitability, physical and financial resources, etc. In empirical evaluations of public projects and public provided goods, multi-criteria decision analysis seems to be an adequate policy tool as well, since it allows taking into account a wide variety of evaluation criteria (e.g. environmental impact, employment, distributional equity, and so on) which can measure the effects on the social welfare.

Social *multi-criteria evaluation (SMCE) techniques* have the potential to take into account conflictual, multidimensional and uncertain properties of policy decisions (Munda, 2004, 2008). SMCE can therefore provide insights into the nature of conflicts and complexity and facilitate the process of reaching political compromises by explaining divergent values and increasing the transparency of the decision process.

SMCE proceeds on the basis of following main concepts: dimensions, objectives, criteria, weights, criterion scores, impact matrix and compromise solution. Dimension is the highest hierarchical level of analysis and indicates the scope of objectives, criteria and criterion scores. The general categories of economic, social and environmental impacts are dimensions. Objectives indicate the direction of change desired, e.g. growth has to be maximised, social exclusion has to be minimised, carbon dioxide emissions have to be reduced. A criterion is a function that associates each alternative action with a variable indicating its performance from a specific point of view. Weights are often used to represent the relative importance attached to dimensions, objectives and criteria. The idea behind this practice is very intuitive and easy, that is, to place the greatest number in the position corresponding to the most important factor.

In operational terms, the application of a SMCE framework involves the following seven main steps (Munda, 2008):

i. Description of the relevant social actors. For example, institutional analysis may be performed on historical, legislative and administrative documents to provide a map of the relevant social actors.
ii. Definition of social actors' values, desires and preferences by using focus groups or other participatory techniques such as anonymous questionnaires and personal interviews.
iii. Generation of policy options and selection of evaluation criteria is a process of co-creation resulting from a dialogue between analysts and social actors.
iv. Construction of the multi- criteria impact matrix synthesising the scores of all criteria for all alternatives, i.e. the performance of each alternative according to each criterion.
v. Construction of a social impact matrix (i.e. a matrix showing the impacts of the alternatives on the various social actors).
vi. Application of a mathematical procedure to aggregate criterion scores and obtain a final ranking of the available alternatives. The importance of mathematical approaches is their ability to allow a consistent aggregation of the diverse information.
vii. Sensitivity analyses help elucidating conflicts among alternatives and objectives and testing the robustness of the model. Expressing results in terms of sensitivities, both to uncertainties in the model as well as divergent values, reveals model biases as rank orders of alternatives potentially change (Saltelli et al., 2004, 2013).

These seven steps are not rigid. On the contrary, flexibility and adaptability to actual situations are among the main advantages of SMCE. As a tool for policy evaluation and conflict management, SMCE has demonstrated its applicability to problems in various geographical and cultural contexts. A recent and exhaustive overview of world-wide SMCE applications can be found in Etxano and Villalba-Eguiluz (2021).

In my experience, the empirical argument that SMCE deals with complex issues in an effective way is accepted in policy contexts, but often it is not a sufficient one for scholars.

Therefore, more formal arguments have to be developed; in this context analytical philosophy is very useful. The starting point is the relationship between comparability and commensurability (Chang, 1997; O'Neill, 1993). From a philosophical perspective, it is possible to distinguish between the concepts of a) *strong comparability* (there exists a single comparative term by which all different actions can be ranked), implying *strong commensurability* (a common measure of the various consequences of an action based on an interval or ratio scale of measurement, such as money or energy), or *weak commensurability* (a common measure based on an ordinal scale of measurement, such as consumer's utility); and b) *weak comparability,* which implies incommensurability (Martinez-Alier et al., 1998). Incommensurability can be further distinguished into technical and social ones (Munda 2004). *Technical incommensurability* refers to the impossibility of compressing different dimensions into a single metric consistent will all the original dimensions and *social incommensurability* refers to the existence of an irreducible value conflict among social actors, when deciding what common comparative term should be used to rank alternative options.

Two other useful concepts are *set and rod commensurability* (Munda, 2016). Commensurability, a necessary condition for strong comparability, can be implemented in two different ways:

1. By looking for a more general category (set) that can contain all the characteristics of the objects we wish to compare; these characteristics are described by using adjectives. This can be defined as "set commensurability" (e.g. apples and oranges are legitimately lumped together as fruit, along with grapes, bananas, etc.).
2. By finding one property common to all objects to be compared and measurable by using one measurement unit, obviously comparison of objects is possible according to the characteristics of this property only. This can be defined as "rod commensurability".

Of course, when possible, set commensurability is the most attractive one since apparently no information is lost in the comparison process, while rod commensurability always requires a kind of reductionism. Here the question is: *when set commensurability is possible and correct?* Geach's (1956) distinction between *attributive and predicative adjectives* can help us in answering this question. In Geach's own words: *"There are familiar examples of what I call attributive adjectives. Big and small are attributive;* x *is a big flea does not split up into* x *is a flea and* x *is big, nor* x *is a small elephant into* x *is an elephant and* x *is small; for if these analyses were legitimate, a simple argument would show that a big flea is a big animal and a small elephant is a small animal. Again, the sort of adjective that the mediaevals called alienans is attributive;* x *is a forged banknote does not split up into* x *is a banknote and* x *is forged, nor* x *is the putative father of* y *into* x *is the father of* y *and* x *is putative. On the other hand, in the phrase a red book, red is a predicative adjective in my sense, although not grammatically so, for is a red book logically splits up into is a book and is red. I can now state my first thesis about good and evil: good and bad are always attributive, not predicative, adjectives"* (Geach, 1956, p. 32).

Although Geach's arguments were developed in the context of moral philosophy, they have an extraordinary explicative power for evaluation problems too. In fact, evaluation is all about an option *a* being declared better, worse or equal than another option *b*. However, although Geach saw the clear difference between predicative and attributive adjectives, he only gave examples of them but no general definition was provided, the new concepts of absolute and relative predicative adjectives were then recently introduced (Munda, 2016). An adjective is *absolute predicative* if its meaning does not change in relation to the subsets considered. It is an intrinsic characteristic of the object considered. The characteristic of being a red-headed person does not change if we consider subsets such as police officers, politicians, scientists or basketball players. In terms of measurement theory, an absolute predicative adjective is always measured on a nominal scale i.e. individual characteristics are grouped into a set of equivalence classes.

An adjective is *relative predicative* if it does not hold its meaning once one switches to a larger or different set of objects. It describes a characteristic that is dependent on the relative comparisons among the objects considered. In terms of measurement theory, a relative predicative adjective is always measured on an ordinal scale. An adjective is *attributive* if it does not have any meaning when referred to a different set or problem framework. A good person can be a bad basketball player and a good economist can be a bad person.

At this stage, the following conclusion can be derived: *when considering adjectives,* set commensurability is correct only if the adjectives considered are *absolute predicative* ones. An adjective *Z* is absolute predicative if it passes the *ontological* check of the two following logical tests: test (1) implies statements such as *"if x_1 is red and it is a car then x_1 is a red car"* and test (2) *"if x_1 is a red car and all cars are a mean of transport then x_1 is a red mean of transport".* Adjectives that fail such tests are relative predicative or attributive adjectives, which always imply weak comparability based on incommensurability. For example, the adjective "good" clearly fails (2), statements such as *"x_1 is a good car, all cars are a mean of transport, and therefore x_1 is a good mean of transport"* or *"x_1 is a good scientist, all scientists are human beings, and therefore x_1 is a good human being"* are invalid arguments on the light of a real-world corroboration.

In summary, the point is that different metrics are also linked to different social objectives and values; in this context, the statement *"x is better than y"* implies an answer to two questions*: 1) according to what? 2) According to whom?* To use only one measurement unit for incorporating a plurality of dimensions, objectives and values, implies reductionism necessarily. If evaluative adjectives like "good" and "valuable" are attributive in standard uses, this does not however preclude the possibility of rational choices between objects, which do not fall into the range of a single comparative. *Weak comparability* based on *incommensurability* is compatible with the existence of such limited ranges; for example, regional sustainability is not evaluated as good or bad as such, but rather in relation to different descriptions or indicators. It can be at one and the same time a *"good income per capita"* and a *"bad social inclusion"*, a *"beautiful landscape"* and a *"heavy pollution"*. The use of these value terms in such contexts is attributive clearly.

In summary, we can conclude that incommensurability does not imply incomparability; on the contrary, it is in terms of weak comparability that evaluation has to take place in practice. This is exactly the basic idea of social multi-criteria evaluation.

2. Tackling the Discrete Multi-Criterion Problem in a SMCE Framework

> "Non zeli ad zelum, nec meriti ad meritum, sed solum numeri ad numerum fiat collatio" (Gregorius X (1210-1276, Papa, 1271), VI Decretalium, lib. I, tit. VI, cap. 9)

Results of a real-world policy exercise depend strongly on the way a given problem is structured during the evaluation process obviously, but mathematical models play a very important role: the one of guaranteeing consistency between assumptions used and results obtained. This implies to take into account the technical *uncertainties* properly, such as:

i. *Compensability versus non-compensability.*
ii. *A relevant preference modelling of intensities of preference.*
iii. *Mixed information on criterion scores (i.e. various measurement scales and related uncertainty).*
iv. *Weights as trade-offs versus weights as importance coefficients.*
v. *A proper ranking algorithm.*

Here, I will make an overview of the main solutions proposed inside the SMCE framework to deal with these issues and that have been implemented in a software tool called SOCRATES (SOcial multi-CRiteria AssessmenT of European policieS) (all methodological and mathematical details behind the SOCRATES software can be found in Azzini and Munda, 2020; Munda, 2004, 2009, 2012, 2022)[1].

The *discrete multi-criterion problem* can be described in the following way: *A* is a finite set of *N* feasible options (or alternatives); *M* is the number of different points of view or evaluation criteria g_m $m=1, 2, \dots, M$ considered relevant in a policy problem, where the option *a* is evaluated to be better than option *b* (both belonging to the set *A*) according to the *m-th* point of view if $g_m(a)>g_m(b)$. In synthesis, the information contained in the impact matrix useful for solving the so-called multi-criterion problem is:

i. *Intensity of preference* (when quantitative criterion scores are present).
ii. *Number* of criteria in favour of a given alternative.
iii. *Weight* attached to each single criterion.
iv. *Relationship* of each single alternative with all the other alternatives.

Combinations of this information generate different aggregation conventions, i.e. manipulation rules of the available information to arrive at a preference structure. The aggregation of several criteria implies taking a position on the fundamental issue of compensability. *Compensability* refers to the existence of trade-offs, i.e. the possibility of offsetting a disadvantage on some criteria by a sufficiently large advantage on another criterion, whereas smaller advantages would not do the same. Thus, a preference relation is non-compensatory if no trade-off occurs and is compensatory otherwise. The use of weights with intensity

[1] See also https://knowledge4policy.ec.europa.eu/modelling/topic/social-multi-criteria-evaluation-policy-options_en/socrates_en

of preference originates compensatory multi-criteria methods and gives the meaning of trade-offs to the weights. On the contrary, the use of weights with ordinal criterion scores originates non-compensatory aggregation procedures and gives the weights the meaning of importance coefficients (Bouyssou, 1986; Bouyssou and Vansnick, 1986; Keeney and Raiffa, 1976; Podinovskii, 1994; Roberts, 1979; Vansnick, 1986).

The concept of importance I am using along this paper can be classified as symmetrical importance, that is "if we have two non-equal numbers to construct a vector in R^2, then it is preferable to place the greatest number in the position corresponding to the most important criterion" (Podinovskii, 1994, p. 241).

A common practice is the pragmatic solution of no criterion weighting. However, the fact that all criteria have the same weight does not guarantee at all that dimensions have the same weight. This would be guaranteed only under the condition that all the dimensions have the same number of criteria; this of course is quite unnatural and artificial, and even dangerous. On the contrary, different criterion weights can guarantee that all the dimensions are considered equal. A reasonable practice can be to start by giving the same weight to each dimension and then splitting each weight among the criteria of any dimension proportionally. Figures 1 and 2, obtained by means of the SOCRATES software, represent these situations in a graphical way. As one can see in this case the relation dimensions/criteria is a very peculiar one. In fact, most of criteria belong to the economic dimension, while other dimensions are much less populated. This implies that the starting weighting assumption can be only equal dimension weights because otherwise (under the equal criterion weighting assumption) the economic dimension would dominate since its weights would be higher than 50% of all dimensions considered (in technical terms it would become a *dictator*).

Figure 1: Equal criterion weighting (economic dimension receives 61.54%)

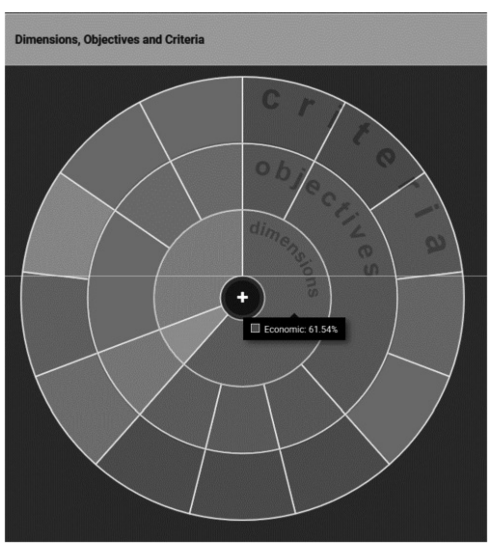
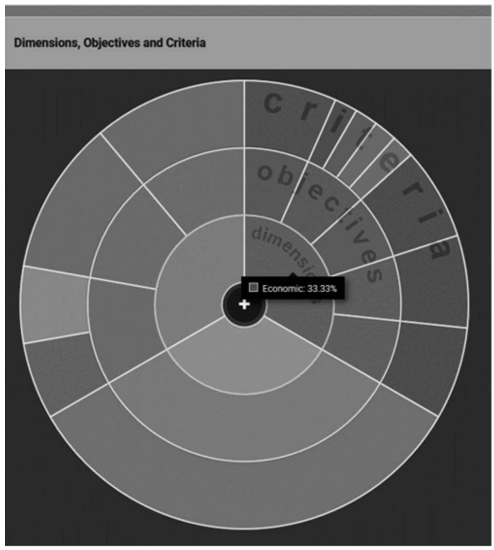


Figure 2: Equal dimension weighting (economic dimension receives 33.33%)



Of course, one could assume that some dimensions are more important than other ones, and thus their weight should be higher, but this should be justified. Finally, one should note that weights can be used in the way described here, only if they have the *meaning of importance*, which depends on the fact that they are combined with *non-compensatory aggregation mathematical rules*.

2.1. Pair-wise comparison of alternatives

The famous bald paradox in Greek philosophy (how many hairs one has to cut off to transform a person with hairs to a bald one?), later on Poincaré (1935, p. 69) and finally Luce (1956) made the point that the transitivity of indifference relation is incompatible with the existence of a sensibility threshold below which an agent either does not sense the difference between two elements, or refuses to declare a preference for one or the other. Luce was the first one to discuss this issue formally in the framework of preference modelling. Mathematical characterisations of preference modelling with thresholds can be found in Roubens and Vincke (1985).

By introducing a positive constant indifference threshold q the resulting preference model is the *threshold model*:

$$\begin{Bmatrix} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) + q \\ a_j I a_k \Leftrightarrow | \, g_m(a_j) - g_m(a_k) \, | \leq q \end{Bmatrix} \tag{1}$$

where a_j and a_k belong to the set A of alternatives and g_m to the set G of evaluation criteria.

Real life experiments show that often there is an intermediary zone inside which an agent hesitates between indifference and preference. This observation led to the so-called *double threshold model* where variable indifference and preference thresholds are introduced, that is:

$$\begin{cases} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) + p(g_m(a_k)) \\ a_j Q a_k \Leftrightarrow g_m(a_k) + p(g_m(a_k)) \geq g_m(a_j) > g_m(a_k) + q(g_m(a_k)) \\ a_j I a_k \Leftrightarrow \begin{cases} g_m(a_k) + q(g_m(a_k)) \geq g_m(a_j) \\ g_m(a_j) + q(g_m(a_j)) \geq g_m(a_k) \end{cases} \end{cases} \quad (2)$$

For any $m = 1, 2, \ldots, M$, being p a positive preference threshold. Relation Q has been called "weak preference" by Roy (1985, 1996). It translates the decision-maker's hesitation between indifference and preference and not "less strong" preference as its name might lead to believe. A criterion with both preference and indifference thresholds is called a pseudo-criterion. A *pseudo-order structure* is a double threshold model upon which the following consistency condition is imposed:

$$g_m(a_j) > g_m(a_k) \Leftrightarrow \begin{cases} g_m(a_j) + q(g_m(a_k)) > g_m(a_k) + q(g_m(a_k)) \\ g_m(a_j) + p(g_m(a_k)) > g_m(a_k) + p(g_m(a_k)) \end{cases} \quad (3)$$

A problem is that the modelling procedure based on the notion of a pseudo-criterion may present a serious lack of stability. Such undesirable discontinuities make a sensitivity analysis (or robustness analysis) necessary; however, this important analysis step is very complex to manage because of the combinatorial nature of the various sets of data. One should combine variations of 2 thresholds (indifference and preference) and k possible scores of the M criteria. A solution to this problem may come from the concept of *valued preference relations*, that is a preference relation where there is a need to assign to each ordered pair of alternatives (a_j, a_k) a value $v(a_j, a_k)$ representing the "strength" or the "degree of preference" (Fishburn, 1970, 1973a; Ozturk et al., 2005; Roubens and Vincke, 1985).

In this framework, an interesting concept is the one of a *fuzzy preference relation* (Kacprzyk and Roubens, 1988). If A is assumed to be a finite set of N alternatives, a *fuzzy preference relation* is an element of the $N \times N$ matrix $R = (r_{jk})$, i.e.

$$r_{jk} = m_R(a_j, a_k), \text{ with } j, k = 1, 2, \ldots, N \quad \text{and} \quad 0 \leq r_{jk} \leq 1. \quad (4)$$

$r_{jk} = 1$ indicates the maximum credibility degree of preference of a_j over a_k; each value of r_{jk} in the open interval $(0.5, 1)$ indicates a definite preference of a_j to a_k (a higher value means a stronger credibility); $r_{jk} = 0.5$ indicates the indifference between a_j and a_k. This definition implies that fuzzy preference relations can be used as mathematical models of intensity of preference.

Usually, fuzzy preference relations are assumed to satisfy two properties:

(a) reciprocity, i.e. $r_{jk} + r_{kj} = 1$;
(b) max-min transitivity, i.e. if a_i is preferred to a_j and a_j is preferred to a_k, then a_i should be preferred to a_k with at least the same credibility degree, that is

$$r_{ij} \geq 0.5, \quad r_{jk} \geq 0.5 \quad \Rightarrow r_{ik} \geq \min(r_{ij}, r_{jk}). \tag{5}$$

By using a fuzzy preference modelling since small variations of input data (scores and thresholds) modify in a continuous way; the consequential preference model can allow one to avoid the drawbacks of the pseudo-criterion model.

Let's now consider any criterion g_m belonging to the set G and any pair of alternatives a_j and a_k belonging to the set A. The criterion scores $g_m(a_j)$ and $g_m(a_k)$ are measured on an interval or ratio scale. Let p_m be a constant preference threshold and q_m a constant indifference threshold for the criterion g_m. Then the credibility degree m of preference (P) and indifference (I) relations between a_j and a_k can be computed as follows:

$$\begin{cases} \mu_{(a_j P a_k)} = \left[1 + c_{pm}\left(g_m(a_j) - g_m(a_k)\right)^{-2}\right]^{-1} \\ \mu_{(a_j I a_k)} = e^{-c_{qm}\left|g_m(a_j) - g_m(a_k)\right|} \\ \mu_{(a_k P a_j)} = \left[1 + c_{pm}\left(g_m(a_k) - g_m(a_j)\right)^{-2}\right]^{-1} \end{cases} \tag{6}$$

where $\mu_{(a_j I a_k)} \quad \forall \; g_m(a_j)$ and $g_m(a_k)$ and

$$\mu_{(a_j P a_k)} \quad \text{if} \quad g_m(a_j) - g_m(a_k) > 0 \tag{7}$$

$$\mu_{(a_k P a_j)} \quad \text{if} \quad g_m(a_j) - g_m(a_k) < 0. \tag{8}$$

It has to be admitted that the shape of the function representing the credibility degrees of the preference and indifference relations is arbitrary. However, some consistency requirement such as that the functions are continuous and monotonic and that $p_m > q_m$ exist, thus reducing considerably the degree of arbitrariness.

2.2. The case of mixed information on criterion scores

Ideally the information available for a policy problem should be precise, certain, exhaustive and unequivocal. But in real-world problem, it is often necessary to use information which does not have these characteristics and thus to deal with uncertainty of a stochastic and/or fuzzy nature present in the data. Let's then introduce a more realistic assumption, i.e. that the set of evaluation criteria $G = \{g_m\}$, $m = 1, 2, \ldots, M$, on the set $A = \{a_n\}$, $n = 1, 2, \ldots, N$ of potential alternatives may include either crisp (that is impacts measured on interval or ratio scales), stochastic and fuzzy criterion scores. A very useful concept for quantifying vagueness on criterion scores is the one of a fuzzy number. A *fuzzy number* is simply a fuzzy set in the real line and is completely defined by its membership function such as $\mu_x: R \rightarrow [0, 1]$. For computational purposes, in general this definition is restricted to those fuzzy numbers which are both normal and convex.

Normality: $\sup\{\mu(x)\} = 1 \quad with \quad x \in R$.

Convexity: $\mu\{\lambda x_1 + (1 - \lambda)x_2\} \geq \min\{\mu(x_1), \mu(x_2)\} \quad with \quad x \in R \quad and \quad \lambda \in \{0, 1\}$.

The requirement of convexity implies that the points of the real line with the highest membership values are clustered around a given interval (or point). This fact allows one to easily understand the semantics of a fuzzy number by looking at its distribution.

A general type of fuzzy number is the so-called *L-R fuzzy number*; it is defined as follows:

$$\mu(x) = \begin{cases} \dfrac{F_L(x-m)}{\alpha} & if \quad -\infty < x < m, \alpha \in R^+ \\ 1 & if \quad x = m \\ \dfrac{F_R(x-m)}{\delta} & if \quad m < x < +\infty, \delta \in R^+ \end{cases} \tag{9}$$

where *m, a, d,* are the "middle" value, the left-hand and the right-hand variation, respectively. $F_L(x)$ is a monotonically increasing membership function and $F_R(x)$, not necessarily symmetric to $F_L(x)$, is a monotonically decreasing function.

The treatment of mixed information on criterion scores proposed here is mainly based on the semantic distance originally developed in Munda (1995) and furtherly formalised in Munda (2012). This semantic distance allows dealing consistently with an impact matrix which may include crisp, stochastic or fuzzy measurements of the performance of an alternative with respect to an evaluation criterion. Therefore, the multi-criterion problem is considered in its more general form. The only restriction is that in the case of fuzzy information, continuous, convex membership functions allowing also a definite integration are required.

Let's start with the case of fuzzy criterion scores:
if $m_1(x)$ and $m_2(x)$ are two fuzzy numbers, one can write (see Ragade and Gupta, 1977, for a formal proof):

$$f(x) = k_1\mu_1(x) \quad \text{and} \quad g(y) = k_2\mu_2(x), \tag{10}$$

where $f(x)$ and $g(y)$ are two functions obtained by rescaling the ordinates of $m_1(x)$ and $m_2(x)$ through k_1 and k_2, such as

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{+\infty} g(y)\,dy = 1 \tag{11}$$

The distance between all points of the membership functions is computed as follows:

If *f(x)* is defined on $X = [x_L, x_U]$ and *g(y)* is defined on $Y = [y_L, y_U]$, where sets X and Y can be non-bounded from one or either sides, then

$$S_d(f(x), g(y)) = \iint_{x,y} |x - y| f(x) g(y) \, dydx \tag{12}$$

If the intersection between the 2 membership functions is empty, it is $x > y \ \forall x \in X$ and $\forall y \in Y$, it follows that a continuous function in 2 variables is defined over a rectangle. Therefore, the double integral can be calculated as iterated single integrals; the result is

$$S_d(f(x), g(y)) = |E(x) - E(y)|, \tag{13}$$

where *E(x)* and *E(y)* are the expected values of the 2 membership functions.

When the intersection between 2 fuzzy sets is not empty, their distance is greater than the difference between the respective expected values since $|x - y|$ is always greater than $(x - y)$. In this case one finds

$$S_d(f(x), g(y)) = \int_{-\infty}^{+\infty} \int_{x}^{+\infty} (y - x) f(x) g(y) \, dydx + \int_{-\infty}^{+\infty} \int_{-\infty}^{x} (x - y) f(x) g(y) \, dydx \tag{14}$$

This is the case of a double integral over a general region; since this is not vertically simple or horizontally simple, it is not possible its computation by means of iterated integration, but it is necessary to take the limit of the Riemann sum. This problem can be easily overcome by means of numerical analysis.

From a theoretical point of view, the following main conclusions can be drawn:

(a) the absolute value metric is a particular case of the semantic distance;
(b) the comparison between a fuzzy number and a crisp number is equal to the difference between the expected value of the fuzzy number and the value of the crisp number considered;
(c) stochastic information can be taken into account too.

In sum, this semantic distance allows one to deal with fuzzy numbers, probability distributions and crisp numbers with the theoretical guarantee that all these sources of information are tackled equivalently, thus solving an open problem for multi-criterion methods dealing with mixed information. Of course, this search for an equivalent treatment of available information implies a trade-off with precision. For example, if stochastic information only is available, a stochastic dominance approach is more effective (see e.g. Markowitz, 1989, Martel and Zaras, 1995), or if fuzzy numbers only have to be compared, Matarazzo and Munda (2001) present a more sophisticated approach based on area comparison. However, in the case of mixed information in a multi-criterion framework, the semantic distance illustrated here is probably the best available compromise solution between generality and precision. Moreover, the use of this semantic distance allows a homogeneous preference modelling on all the criteria, impossible otherwise. Going back to the pair-wise comparison of alternatives, let's assume $f(x) = g_m(a_j)$ *and* $g(y) = g_m(a_k)$, where g_m is any criterion belonging to the set G and a_j and a_k any pair of alternatives belonging to the set A. The criterion scores $g_m(a_j)$ and $g_m(a_k)$ are *fuzzy or stochastic* in nature. Let p_m be a preference threshold and q_m an indifference threshold for the criterion g_m. Then it is:

$$\begin{cases} \mu_{(a_jPa_k)} = \left[1 + c_{pm}\left(\iint_{x,y}(x-y)\,f(x)\,g(y)\,dydx\right)^{-2}\right]^{-1} \\ \mu_{(a_jIa_k)} = e^{-c_{qm}\left(\iint_{x,y}|x-y|\,f(x)g(y)dydx\right)} \\ \mu_{(a_kPa_j)} = \left[1 + c_{pm}\left(\iint_{y,x}(y-x)\,f(x)\,g(y)\,dydx\right)^{-2}\right]^{-1} \end{cases} \tag{15}$$

where $\mu_{(a_jIa_k)}$ $\forall$ x, y and

$$\mu_{(a_jPa_k)} \qquad if \quad \iint_{x,y}(x-y)\,f(x)\,g(y)\,dydx > 0 \tag{16}$$

$$\mu_{(a_kPa_j)} \qquad if \quad \iint_{x,y}(x-y)\,f(x)\,g(y)\,dydx < 0. \tag{17}$$

One should note that the comparison between the criterion scores of each pair of actions is carried out by means of the semantic distance. Since the absolute value metric is a particular case of this distance, fuzzy, stochastic and crisp criterion scores are dealt with equivalently.

2.3. Introducing weights as importance coefficients

At this point, a very sensitive step has still to be tackled i.e. the exploitation of the inter-criteria information in the form of weights. Let's then assume the existence of a set of criterion weights $W = \{w_m\}$, $m = 1, 2,..., M$, with $\Sigma_{m=1}^{M} w_m = 1$ derived as importance coefficients. The problem here is the theoretical guarantee that weights are really treated as importance coefficients and not as trade-offs. The point is that no connection must be done between criterion weights and the corresponding criterion intensity of preference. Our objectives are then:

(a) to find a way to combine weights with credibility degrees without a direct interpretation of the latter as intensity of preference;
(b) to divide each criterion weight in 2 parts proportionally to the credibility degrees of the indifference and preference fuzzy relations. In doing so, the requirement that $\Sigma_{m=1}^{M} w_m = 1$ should not be lost.

Let's define μ_p as the fuzzy preference relation between a pair of alternatives and μ_I as the fuzzy indifference relation between the same pair of alternatives. Let's put

$\mu_{min} = \min(\mu_p, \mu_I)$ and $\mu_{max} = \max(\mu_p, \mu_I)$. Clearly, it is $\mu_p = \mu_{min}$ on the left of the intersection point between the indifference and the preference fuzzy relations and vice-versa on the right. A criterion weight w_m is divided proportionally to μ_p and μ_I, according to equation (18):

$$\begin{cases} W_{m1} = W_m \dfrac{\mu_{\min}}{\mu_{\max} + \mu_{\min}} \\ W_{m2} = W_m \dfrac{\mu_{\max}}{\mu_{\max} + \mu_{\min}} \end{cases} \tag{18}$$

Equation (18) presents the following properties:

$$W_{m1} + W_{m2} = W_m \tag{19}$$

$$\textit{if } \mu_{\min} = 0 \Rightarrow W_{m2} = W_m \tag{20}$$

$$\textit{if } \mu_{\min} = \mu_{\max} = 0 \Rightarrow W_m = 0 \tag{21}$$

$$\textit{if } \mu_{\min} = \mu_{\max} \Rightarrow W_{m1} = W_{m2} = \frac{1}{2} W_m \tag{22}$$

As a consequence, equation (18) fits our objective that the addition of all weights should be kept equal to one perfectly. Moreover, in equation (18) no direct use of the concept of intensity of preference is done; as a result, we can be sure that criterion weights are used consistently with their nature of importance coefficients. Finally, if a criterion score is *ordinal in nature*, it can be considered a particular case where $\mu_{\min} = 0$. Again, *the treatment of crisp, fuzzy, stochastic and ordinal criterion scores is perfectly equivalent.* Moreover, when indifference and preference thresholds are not used, the corresponding criteria can be dealt with as ordinal criteria, where

$$\begin{cases} a_j P a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \\ a_j I a_k \Leftrightarrow g_m(a_j) > g_m(a_k) \end{cases} \tag{23}$$

Now a $N \times N$ matrix E can be built, where any generic element e_{jk}, with $j \neq k$, is the result of the pair-wise comparison, according to all the M criteria, between alternatives j and k. Such a global pair-wise comparison is obtained by means of equation (24):

$$e_{jk} = \Sigma_{m=1}^{M} \left(w_m(P_{jk}) + \frac{1}{2} w_m(I_{jk}) \right) \tag{24}$$

where $w_m(P_{jk})$ and $w_m(I_{jk})$ are derived from μp and μI through equation (18). It is

$$e_{jk} + e_{kj} = 1. \tag{25}$$

Property (25) is very important since it allows us to consider matrix E as a *voting matrix* i.e., a matrix where instead of using criteria, alternatives are compared by means of voters' preferences (with the principle one agent, one vote). This analogy between the multi-criterion problem and the social choice one, as noted by Arrow and Raynaud (1986), is very useful for tackling the step of ranking the N alternatives in a consistent axiomatic framework.

2.4. Ranking algorithm

Vansnick (1990) showed that the two main approaches in multi-criteria decision analysis i.e., the compensatory and non-compensatory ones can be directly derived from the seminal work of Borda (1784) and Condorcet (1785). Indeed, looking at social choice literature, one can realize that various ranking procedures used in multi-criterion methods have their origins in social choice. Just to give a few examples, the weakness-strength approach, typical of outranking methods (Roy, 1985, 1996), has a clear derivation from two Condorcet consistent rules, i.e. the Copeland (1951) and Simpson (1969) rules; Arrow-Raynaud propose a sequential procedure (building on Köhler, 1978) which is also based on some principles of the Condorcet rule; the so-called frequency matrix approach (Hinloopen et al., 1983; Matarazzo, 1988) comes directly from Borda algorithm, or the permutation method (Paelinck, 1978), has a strict connection with an original Condorcet approach developed to tackle cycles, and so on.

Given that there is a consensus in the literature that the Condorcet' theory of voting is non-compensatory (Vansnick, 1986) and useful for generating a ranking of the available alternatives while Borda's one is more useful for isolating one alternative, considered the best (Moulin, 1988; Truchon, 1995; Young, 1988, 1995), here clearly it is advisable to follow the Condorcet tradition[2] (since in a SMCE framework non-compensability and a complete ranking of alternatives are considered desirable properties (Munda, 2004)).

A basic problem inherent in the Condorcet's approach is the presence of cycles, i.e. cases where *aPb, bPc* and *cPa* may be found. This problem has been studied by various scientists (e.g., Fishburn, 1973; Kemeny, 1959; Moulin, 1985; Truchon, 1995; Young and Levenglick, 1978, Vidu, 2002; Weber, 2002). Now the question is: Is it possible to tackle the cycle issue in a general way? The answer to this question is yes, and it is generally known in social choice as the Kemeny method. However, in reality other scientists, including Condorcet himself, have contributed to the development of this ranking method. The historical reconstruction of this method and a deeper methodological analysis can be found in Munda (2008, Chapter 6). Here, I just synthesise some main points.

[2] Arrow and Raynaud (1986) also arrive at the conclusion that a Condorcet aggregation algorithm has in general to be preferred in a multi-criterion framework. They show that whenever the majority rule can be operationalized, it should be applied. However, the majority rule often produces undesirable intransitivities, thus "*more limited ambitions are compulsory. The next highest ambition for an aggregation algorithm is to be Condorcet*" (Arrow and Raynaud, 1986, p. 77).

Condorcet himself was aware of the problem of cycles in his approach; he built examples to explain it and he was even close to find a consistent rule able to rank any number of alternatives when cycles are present. Attempts of clarifying, fully understanding and axiomatizing Condorcet's approach for solving cycles have been mainly done by Kemeny (1959) who made the first intelligible description of the Condorcet approach, and by Young and Levenglick (1978) who made its clearest exposition and complete axiomatization. For this reason, I call this approach the Condorcet-Kemeny-Young-Levenglick ranking procedure, or the C-K-Y-L ranking procedure. Its main methodological foundation is the maximum likelihood concept. The maximum likelihood principle selects as a final ranking the one with the maximum pair-wise support. This selected ranking is also the one which involves the minimum number of pair-wise inversions. The selected ranking is also a median ranking for those composing the profile (in multi-criteria terminology it is the "compromise ranking" among the various conflicting points of view), for this reason the corresponding ranking procedure is often known as the Kemeny median order.

A problem of the C-K-Y-L ranking procedure is that it does not respect the axiom of independence of irrelevant alternatives (Arrow, 1963). However, two considerations have to be made on this subject.

1. A Condorcet consistent rule always presents smaller probabilities of the occurrence of a rank reversal in comparison with any Borda consistent rule (Moulin, 1988; Young, 1995). This is a strong argument in favor of a Condorcet consistent rule.
2. Young (1988, p. 1241) claims that the C-K-Y-L ranking procedure is the *"only plausible ranking procedure that is locally stable"*. Where *local stability* means that the ranking of alternatives does not change if only an interval of the full ranking is considered. It is interesting to note that this property was also studied by Jacquet-Lagrèze (1969), one of the first researchers in multi-criteria analysis, who called it the median procedure.

Other properties of the C-K-Y-L ranking procedure are the following (Young and Levenglick, 1978).

i. *Neutrality:* it does not depend on the name of any alternative, all alternatives are equally treated.
ii. *Unanimity* (sometimes called *Pareto Optimality*): if all criteria prefer alternative *a* to alternative *b* than *b* should not be chosen.
iii. *Monotonicity:* if alternative *a* is chosen in any pair-wise comparison and only the criterion scores of *a* are improved, then *a* should be still the winning alternative. Monotonicity is an essential property in a SMCE framework since dominated alternatives are not advised to be deleted from the analysis.
iv. *Reinforcement:* if the set *A* of alternatives is ranked by 2 subsets G_1 and G_2 of the criteria set *G*, such that the ranking is the same for both G_1 and G_2, then $G_1 \cup G_2 = G$ should still supply the same ranking. This general consistency requirement is very important in a multi-criterion framework where to test robustness

of results, one may wish to apply the criteria belonging to each single dimension first and then pool them in the general model.

It has to be noted that the C-K-Y-L ranking procedure is the only Condorcet consistent rule which holds the reinforcement property and as noted by Arrow and Raynaud, reinforcement *"… has definite ethical content and is therefore relevant to welfare economics and political science."* (Arrow and Raynaud, 1986, p. 96). Given that Arrow and Raynaud deal with the *"industrial outranking problem"* relevant for business people they do think that in this framework, reinforcement is less important that independence of irrelevant alternatives. On the contrary, in the framework of public policy, dealt with here, reinforcement becomes a decisive argument in favour of the C-K-Y-L ranking procedure

Although as one can see, the theoretical characterization of the C-K-Y-L ranking procedure is not that easy, the algorithm *per se* indeed is very simple. The maximum likelihood ranking of alternatives, in a multi-criterion framework, is the ranking supported by the maximum number of criteria for each pair-wise comparison, summed over all pairs of alternatives. More formally, the C-K-Y-L ranking procedure can be adapted to a multi-criterion framework as follows.

All the $N(N-1)$ pair-wise comparisons compose the matrix E, where let's remember that $e_{jk} + e_{kj} = 1$, with $j \neq k$. Let's call R the set of all the $N!$ possible complete rankings, of alternatives, $R=\{r_s\}$, $s=1, 2,..., N!$. For each r_s, let's compute the corresponding score φ_s as the sum of all e_{jk} over all the $\binom{N}{2}$ pairs jk of alternatives, such that a_j is preferred to a_k in the ranking r_s. More formally, let us denote $a_j >^s a_k$ the fact that a_j is preferred to a_k in the ranking r_s, then

$$\varphi_s = \sum_{j=1}^{N} \sum_{k=1}^{N} e_{jk}^{s}, \tag{26}$$

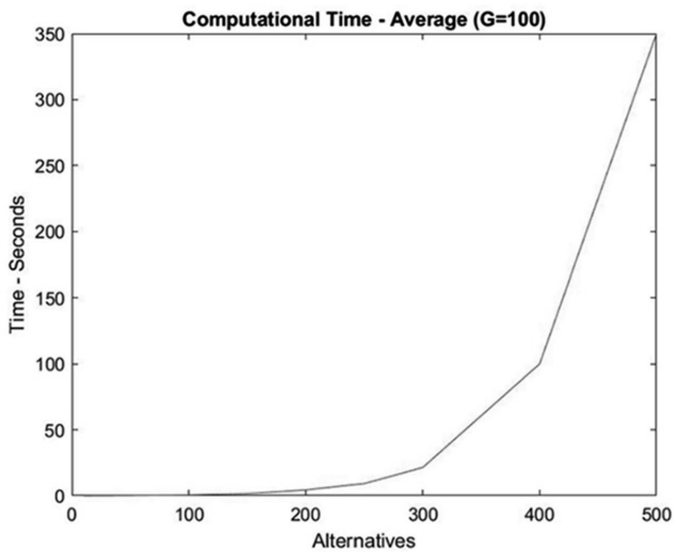
where $j \neq k$ and $e_{jk}^{s} = \begin{cases} e_{jk} & \text{if } a_j >^s a_k \\ 0 & \text{otherwise} \end{cases}$.

The final ranking is r_t, $t \in \{1, 2,..., N!\}$, such that

$$\varphi_t = \max \varphi_s, \ s = 1, 2, \ldots, N! \tag{27}$$

The computational problem is a clear drawback of this approach. One should note that the number of permutations can easily become unmanageable; for example, when 10 alternatives are present, it is 10! = 3,628,800. A numerical algorithm solving this computational drawback in an efficient way has been developed recently (Azzini and Munda, 2020). In 1000 simulations, keeping constant the number of criteria (G=100), the average computational time is about 1 second till 100 alternatives and it reaches a maximum of 350 seconds for 500 alternatives (see Figure 3).

Figure 3: Computational time needed by increasing the number of alternatives



2.5. Introducing the minority principle: a borda approach

At this point, we have to refer to the normative tradition in political philosophy, which has also an influence in modern social choice (Moulin, 1981) and public policy (Mueller, 1978). The basic idea is that any coalition controlling more than 50% of votes may be converted in an actual dictator. As a consequence, the *"remedy to the tyranny of the majority is the minority principle, requiring that all coalitions, however small, should be given some fraction of the decision power. One measure of this power is the ability to veto certain subsets of outcomes..."* (Moulin, 1988, p. 272). The introduction of a veto power in a multi-criterion framework can be further justified in the light of the so-called *"prudence"* axiom (Arrow and Raynaud, 1986, p. 95), whose main idea is that it is not prudent to accept alternatives whose degree of conflictuality is too high (and thus the decision taken might be very vulnerable).[3] The point is then how can we implement this idea of veto power in a multi-criterion framework?

Historically, the first attempt was done by Roy (1985, 1996) in the so-called ELECTRE methods. Basically, Roy proposed that for any pair of alternatives one should look at the majority principle expressed as a concordance index and to the minority one in the form of the discordance index. The discordance index is calculated according to the intensity of

[3] It has to be noted that to mitigate the vulnerability of the C-K-Y-L ranking procedure is very important since this is one of the main criticism against this method.

preference any single criterion has against the concordance coalition. This means that on each single criterion a veto threshold needs to be defined.

In my opinion, the implementation of the veto power in a SMCE framework needs three desirable properties:

1. To be independent of arbitrary ad hoc thresholds.
2. To consider the global opposition against the final ranking and not against a pair of alternatives, or any specific possible ranking.
3. No specific intensity of preference should be considered (if one combines a weight with a veto threshold on each single criterion, the resulting concept of criterion importance depends on the intensity of preference too, this means that probably weights cannot anymore considered as importance coefficients).

It is interesting to note that the approach fitting these requirements can again be found in classical social choice and in particular, this time in the Borda's approach. The Borda rule is normally used to find a Borda winner, where the winner is the alternative which receives the highest score in favour (an alternative receives *N-1* points if ranks first and so on till *0* score if it ranks last on a given criterion). In the same way, a *Borda loser* can be defined as the alternative which receives the highest score against (where *N-1* points are assigned to the last alternative in the ranking and so on till *0* points are given to the alternative which ranks first).

Formally the procedure I am proposing can be described as follows by taking inspiration from the concept of frequency matrices (Hinloopen et al., 1983, Matarazzo, 1988). Let's call F the matrix where any element f_{ij} means that a given criterion g_m scores alternative a_j in the *i-th* ordinal position. Now it is possible to define the $N \times N$ matrix $\mathbf{\Phi}$ where any element φ_{ij} represents the summation of the weights of criteria which score alternative j at the *i-th* position; that is

$$\varphi_{ij} = \sum_{m \in G_i} w_m \qquad (28)$$

$$\textit{where } G_i = \{g_m : g_m(a_j) = f_{ij}\} \quad \textit{with} \quad G_i \subset G \qquad (29)$$

I = 1, 2,…, N and j = 1, 2,…, N.

It is obviously:

$$\sum_{i=1}^{N} \varphi_{ij} = 1 \quad \forall\, a_j \in A \qquad \text{and} \qquad (30)$$

$$\sum_{j=1}^{N} \varphi_{ij} = 1 \quad \textit{with } j = 1, 2, ..., N \qquad (31)$$

Now for any alternative a_j let's apply the Borda rule in search of the Borda looser, that is

$$B(a_j) = \sum_{i=1}^{N} (\varphi_{ij} \times b_i), \;\; b_i = N-1, N-2, ..., 0 \quad \textit{with } i = N, N-1, ..., 1. \qquad (32)$$

The vetoed alternative $\bar{a}_j$ is the Borda looser, i.e. the a_j for which $B(a_j)$ = max.

One should note that by means of this procedure, weights are never combined with intensities of preference and no *ad hoc* parameter is needed. Consistently with the Borda approach only one alternative, considered the one with the highest opposition, is selected as alternative to be vetoed. It has to be remembered that the Borda procedure respects all the properties of the C-K-Y-L one, except local stability. This is the main reason why Borda consistent rules are more adequate for the selection of one alternative only and not for the generation of rankings.

Finally, a question to be answered is: do Borda and Condorcet rules normally lead to different solutions? One can in fact think that the divergence of solutions is a very special case and thus the value added of introducing the Borda looser is very limited. This question has been answered by Fishburn (1973b) and Moulin (1988), who proved that Condorcet consistent rules and Borda voting rules are deeply different in nature and consequently it is useful to combine them in a complementary way.

3. Conclusion

> "We live in a world of contradiction and paradox, a fact of which perhaps the most fundamental illustration is this: that the existence of a problem of knowledge depends on the future being different from the past, while the possibility of the solution of the problem depends on the future being like the past." (Knight, 1921, p. 313)

This article has illustrated how SMCE can help in dealing with three different types of sustainability related policy issues: 1) epistemological uncertainty 2) complexity 3) mathematical manipulation rules of relevant information. In summary, we can conclude that:

In sustainability policies evaluation and assessment, key questions to be answered are sustainability of *what and whom*? Consequently, sustainability problems are multidimensional in nature and characterised by a high degree of conflict, uncertainty and complexity.

Complexity arises when something is difficult to understand and impossible to analyse by using simple frameworks.

To reach a ranking of policy options, there is a prior need to decide *what is important* for different social actors as well as *what is relevant* for the representation of the real-world entity described in the model.

In operational terms, social *multi-criteria evaluation (SMCE) techniques* have the potential to take into account conflictual, multidimensional and uncertain properties of policy decisions. SMCE can therefore provide insights into the nature of conflicts and complexity and facilitate the process of reaching political compromises by explaining divergent values and increasing the transparency of the decision process.

From a theoretical perspective, we can conclude that commensurability, a necessary condition for strong comparability, can be implemented by means of "set commensurability" and "rod commensurability"; both of them are not of a general applicability. Different metrics are linked to different objectives and values. To use only one measurement unit for incorporating a plurality of objectives and values, implies reductionism necessarily, therefore weak comparability grounded on incommensurability, can be implemented by using social multi-criteria evaluation.

No mathematical model, even if legitimate in its own terms, can be sufficient for a complete analysis of the reflexive properties of a real-world problem. Results of a real-world policy exercise depend strongly on the way a given problem is structured during the evaluation process obviously, but mathematical models play a very important role: the one of guaranteeing consistency between assumptions used and results obtained. This implies to take into account the technical *uncertainties properly*; consequently, this article has presented a mathematical aggregation convention useful for the solution of the so-called discrete multi-criterion problem in a SMCE context. This mathematical aggregation procedure is a "reasonable" approach based on theoretical and empirical grounds, all of them made explicit and thus easy to evaluate in relation with a particular use.

Throughout the whole pair-wise comparison step, it is guaranteed that ordinal, crisp, stochastic and fuzzy criterion scores are tackled equivalently. To deal with the lack of stability of the pseudo-order structure, valued preference relations modelled by means of fuzzy preference relations are introduced. Weights are never combined with intensities of preference, as a consequence the theoretical guarantee they are importance coefficients exists. The pair-wise comparisons can be synthesised in an outranking matrix, which can be interpreted as a voting matrix. The information contained in the voting matrix is exploited to rank all alternatives in a complete pre-order by using a Condorcet consistent rule. Consistently with the normative tradition in political philosophy and following the prudence axiom, the minority principle is introduced by means of a veto power, grounded on the original Borda approach implemented by using a frequency matrix.

References

Arrow, K.J. (1963) *Social Choice and Individual Values*, 2d edition. New York, Wiley.

Arrow, K.J.; Raynaud H. (1986) *Social Choice and Multicriterion Decision Making*. Cambridge, M.I.T. Press.

Azzini, I; and Munda, G. (2020) A new approach for identifying the Kemeny Median Ranking. *European Journal of Operational Research, 281*, 388-401.

Barbier, E.B. (1987) The concept of sustainable economic development. *Environmental Conservation*, *14*(2), 1-110.

Borda, J.C. de (1784) Mémoire sur les élections au scrutiny. In Histoire de l' Académie Royale des Sciences, Paris.

D. (1986) Some remarks on the notion of compensation in MCDM. *European Journal of Operational Research*, *26*, 150-160.

Bouyssou, D.; Vansnick, J.C. (1986) Noncompensatory and generalized noncompensatory preference structures. *Theory and Decision*, *21*, 251-266. Chang, R. (ed.) (1997) *Incommensurability, Incomparability, and Practical Reason*. Cambridge, Harvard University Press.

Condorcet, M. de (1785) Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Probabilité des Voix. Paris, De l' Imprimerie Royale.

Copeland, A. H. (1951) *A reasonable social welfare function*, University of Michigan, mimeo.

Etxano, I.; Villalba-Eguiluz, U. (2021) Twenty-five years of social multi-criteria evaluation (SMCE) in the search for sustainability: Analysis of case studies. *Ecological Economics, 188*, 107-131.

Figueira, J.; Greco, S.; Ehrgott, M. (eds.) (2016) *Multiple-Criteria Decision Analysis. State of the Art Surveys*. New York, Springer International Series in Operations Research and Management Science.

Fishburn, P.C. (1970) Utility theory with inexact preferences and degrees of preference. *Synthese*, *21*, 204-222.

Fishburn, P.C. (1973a) Binary choice probabilities: on the varieties of stochastic transitivity. *Journal of Mathematical psychology*, *10*, 327-352.

Fishburn, P.C. (1973b) *The Theory of Social Choice*. Princeton, University Press.

Funtowicz, S.; Martinez-Alier, J.; Munda, G.; Ravetz, J. (1999) *Information tools for environmental policy under conditions of complexity (Report* Environmental Issues Series, No. 9), Luxembourg, Office for Official Publications of the European Communities.

Funtowicz, S.O.; Ravetz, J. R. (1991) A new scientific methodology for global environmental issues. In R. Costanza (ed.) *Ecological Economics*, New York, Columbia, pp. 137-152.

Geach, P. T. (1956) Good and Evil. In Philippa (ed.) *Theories of Ethics*. Oxford, University Press, pp. 64-73.

Gell-Mann, M. (1994) The Quark and the Jaguar: Adventures in the Simple and the Complex. New York, W.H. Freeman.

Giampietro, M.; Mayumi K.; Munda G. (2006) Integrated assessment and energy analysis: quality assurance in multi-criteria analyses of sustainability. *Energy, 31*(1), 59-86.

Giampietro, M.; Mayumi, K. (2018) Unraveling the Complexity of the Jevons Paradox: The Link Between Innovation, Efficiency, and Sustainability. *Frontiers in Energy Research*, *6*. https://doi.org/10.3389/fenrg.2018.00026

Giampietro, M.; Mayumi, K. (2000) Multiple-scale integrated assessment of societal metabolism: integrating biophysical and economic representations across scales. *Population and Environment*, *22* (2), 155-210.

Giampietro, M.; Mayumi, K. (2000) Multiple-scale integrated assessment of societal metabolism: Introducing the approach. *Population and Environment*, *22*(2), 109-154.

Giampietro, M.; Mayumi, K.; Sorman, A.H. (2012) *The Metabolic Pattern of Societies: Where Economists Fall Short.* Routledge.

Hinloopen, E.; Nijkamp, P.; Rietveld, P. (1983) Qualitative discrete multiple criteria choice models in regional planning. *Regional Science and Urban Economics*, 77-102.

Jacquet-Lagrèze, E. (1969) L'agrégation des opinions individuelles. Informatiques et Sciences Humaines, 4.

Kacprzyk, J.; Roubens, M. (Eds.). (1988) *Non-Conventional Preference Relations in Decision Making*. Heidelberg, Springer-Verlag. https://doi.org/10.1007/978-3-642-51711-2

Keeney, R.; Raiffa, H. (1976) Decision with Multiple Objectives: Preferences and Value Trade-offs. New York, Wiley.

Kemeny, J. G. (1959) Mathematics without Numbers. *Daedalus*, *88*(4), 577-591.

Knight, F. (1921) Risk, Uncertainty and Profit. Cambridge, Che Riverside Press.

Köhler, G. (1978) *Choix multicritère et analyse algébrique des données ordinales*, Ph.D. Thesis, Université Scientifique et Médicale de Grenoble, France.

Luce, R.D. (1956) Semiorders and a theory of utility discrimination. *Econometrica*, *24*, 178-191.Markowitz; H.M. (1989). Mean-variance Analysis in Portfolio Choice and Capital Markets. Oxford, Basil-Blackwell.

Martel, J.M.; Zaras, K. (1995) Stochastic dominance in multicriteria analysis under risk. *Theory and Decision*, *35*(1), 31-49.

Martinez-Alier, J.; Munda, G.; O'Neill, J. (1998) Weak comparability of values as a foundation for ecological economics. *Ecological Economics*, *26*, 277-286.

Matarazzo, B. (1988) Preference ranking global frequencies in multicriteria analysis (PRAGMA). *European Journal of Operational Research*, *36*, 36-49.

Matarazzo, B.; Munda, G. (2001) New approaches for the comparison of L-R fuzzy numbers: A theoretical and operational analysis. *Fuzzy Sets and Systems*, *118* (3), 407-418.

Moulin, H. (1981) The proportional veto principle. *Review of Economic Studies*, *48*, 407-416.

Moulin, H. (1985) From social welfare orderings to acyclic aggregation of preferences, *Mathematical Social Sciences*, *9*, 1-17.

Moulin, H. (1988) *Axioms of Co-operative Decision Making. Econometric Society Monographs*, Cambridge, Cambridge University Press.

Mueller, D. (1978) Voting by veto. *Journal of Public Economics*, 10, 57-75.

Munda, G (2009) A conflict analysis approach for illuminating sustainability distributional issues. *European Journal of Operational Research*, *194*, 307-322

Munda, G (2022) Qualitative reasoning or quantitative aggregation rules for impact assessment of policy options? A multiple criteria framework. *Quality & Quantity. International Journal of Methodology,* 56, 3259-3277.

Munda, G. (1995) Multicriteria Evaluation in a Fuzzy Environment. In *Contributions to Economics Series,* Heidelberg, Physica-Verlag, pp. 131-148.

Munda, G. (2004) Social multi-criteria evaluation (SMCE): Methodological foundations and operational consequences. *European Journal of Operational Research*, *158* (3), 662- 677.

Munda, G. (2008) *Social Multi-criteria Evaluation for a Sustainable Economy*. Heidelberg, Springer-Verlag,

Munda, G. (2012). Intensity of preference and related uncertainty in non-compensatory aggregation rules. *Theory and Decision*, *73*(4), 649-669. https://doi.org/10.1007/s11238-012-9317-4

Munda, G. (2016) Beyond welfare economics: Some methodological issues. *Journal of Economic Methodology*, *23* (2), 185-202Norgaard, R. B. (1994) *Development Betrayed.* London, Routledge.

O'Neill, J. (1993) *Ecology, Policy and Politics*. London, Routledge.

Ozturk, M.; Tsoukias, A.; Vincke, Ph. (2005) Preference modelling. In Figueira, J.; Greco, S. and Ehrgott, M. (eds.) *Multiple-criteria Decision Analysis. State of the Art Surveys*. New York, Springer International Series in Operations Research and Management Science, pp. 27-71.

Paelinck J.H.P. (1978) Qualiflex, a flexible multiple criteria method. *Economic Letters*, 3, 193-197.

Podinovskii, V.V. (1994) Criteria importance theory. *Mathematical Social Sciences*, *27*, 237- 252.

Poincaré, H. (1935) *La Valeur de la Science*. Flammarion.

Prigogine, I.; Stengers, I. (1981) *Order Out of Chaos*. New York, Bantam Books.

Ragade, R.K. and Gupta, M.M. (1977) Fuzzy set theory: Introduction. In M.M. Gupta, G.N; Saridis and B.R. Gaines, *Fuzzy Automata and Decision Processes*. Amsterdam, North-Holland, pp. 105-131

Roberts, F. S. (1979) Measurement theory with applications to decision making, utility and the social sciences, London, Addison-Wesley.

Rosen, R. (1985) Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations. New York, Pergamon Press.

Rosen, R. (1991) Life Itself: A Comprehensive Inquiry into the Nature, Origin and Fabrication of Life. New York, Columbia University Press.

Roubens; M.; Vincke, Ph. (1985) *Preference Modelling*. Heidelberg, Springer-Verlag,

Roy, B. (1996) Multicriteria Methodology for Decision Analysis. Kluwer, Dordrecht.Roy; B. (1985) Méthodologie Multicritere d' Aide à la Decision. *Economica*, Paris.

Saltelli, A.; Tarantola S.; Campolongo, F. and Ratto, M. (2004) *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. New York, John Wiley & Sons publishers,

Saltelli, A.; Van der Sluijs, J.; Guimarães Pereira Â.; Funtowiz, S.O. (2013) What do I make of your Latinorum? Sensitivity auditing of mathematical modelling. *International Journal Foresight and Innovation Policy*, *9*(2/3/4), 213-234.

Sen, Amartya. (2009) *The Idea of Justice*. The Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Simon, H. A. (1962) *The architecture of complexity*. Proc. Amer. Philos. Soc. *106*, 467-482.

Simpson, P. (1969) On defining areas of voter choice: Professor Tullock on stable voting. *Quarterly Journal of Economics*, 83, 478-490.

Truchon, M. (1995) Voting games and acyclic collective choice rules. *Mathematical Social Scien*ces, *25*, 165-179.

Vansnick, J. C. (1990) Measurement theory and decision aid. In Bana e Costa C.A. (ed.), *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin, pp. 81-100.

Vansnick, J.C. (1986) On the problem of weights in multiple criteria decision making (the non-compensatory approach). *European Journal of Operational Research*, *24*, 288-294.

Vidu, L. (2002) Majority cycles in a multi-dimensional setting. *Economic Theory*, *20*, 373-386.

Weber, J. (2002) How many voters are needed for paradoxes? *Economic Theory*, *20*, 341-355.

Young, H.P. (1988) Condorcet's theory of voting. *American Political Science Review*, *82* (4), 1231-1244.

Young, H.P. (1995) Optimal voting rules. *Journal of Economic Perspectives*, *9*, 51-64.

Young, H.P.; Levenglick, A. (1978) A consistent extension of Condorcet's election principle, *SIAM Journal of Applied Mathematics*, 35, 285-300.

Renato Rosa
Univ Coimbra, CeBER, Faculty of Economics, Coimbra, Portugal
renato.rosa@fe.uc.pt
orcid.org/0000-0002-1155-9681
Ramiro Parrado
RFF-CMCC European Institute on Economics and the Environment (EIEE)
and Centro Euro-Mediterraneo sui Cambiamenti Climatici, Università Ca' Foscari,
Venice, Italy
ramiro.parrado@cmcc.it
orcid.org/0000-0002-0951-1013
Francesco Bosello
RFF-CMCC European Institute on Economics and the Environment (EIEE),
Centro Euro-Mediterraneo sui Cambiamenti Climatici, Università Ca' Foscari, and
Department of Environmental Sciences, Informatics and Statistics University Ca' Foscari
of Venice, Italy
francesco.bosello@cmcc.it

DOI: https://doi.org/10.14195/2183-203X_56_2

Trading REDD Credits in International Carbon Markets: Interactions among International Trade, Carbon and Agricultural Markets

Transação de Créditos *REDD* em Mercados Internacionais de Carbono: Interações entre Comércio Internacional, Mercados de Carbono e Mercados Agrícolas

Renato Rosa
Ramiro Parrado
Francesco Bosello

Received for publication: March 10, 2023
Revision accepted for publication: May 31, 2023

ABSTRACT

While several economic studies have looked into the role of REDD in climate policy, the interlinks between climate policy, international trade and agricultural markets have been only marginally considered. This paper adds to that discussion by developing a policy simulation exercise in which REDD credits can be traded in an international carbon market using a recursive dynamic computable general equilibrium model. The model was extended to incorporate abatement cost curves of avoided deforestation from a partial equilibrium study, and to account for the corresponding induced effects on land and timber markets. We conclude that REDD may significantly reduce policy costs. A large number of REDD credits entering the carbon market would allow regions pertaining to the climate policy agreement to systematically emit above their targets. These results confirm that policy design may require limits to the use of REDD credits along with the creation of long-term incentives to promote a greener economy. Finally, when international competitiveness effects are taken into account, we show that the use of REDD as a means to foster developing countries' participation in climate policy may not be sufficient.

Keywords: Reduced emission from deforestation and forest degradation; climate change policy; carbon markets; agricultural markets; general equilibrium.

JEL Classification: Q4; Q5; F1.

Acknowledgments: CeBER's research is funded by national funds through FCT–Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020.

Renato Rosa
Ramiro Parrado
Francesco Bosello

Trading REDD Credits in International Carbon Markets: Interactions among International Trade, Carbon and Agricultural Markets

1. Introduction

Forests are a two-edged sword in global climate policy. On the one hand, the reduction of forested areas is one of the major contributors to increasing average global temperatures. Tropical deforestation has been recognized as the second largest driver of anthropogenic global warming (IPCC 2007, 2014), accounting for roughly 17% to 20% of greenhouse gases (GHG) released during the 1990s (Gullison et al., 2007; Strassburg et al., 2010). On the other hand, by keeping current forest stocks, increasing forest areas or changing timber management practices, forests may help stabilize or even decrease current GHG concentrations.

Given the significant role played by forests in regulating climate and their potential contribution to an optimal climate change policy, it is not surprising that they have since long been central in international climate negotiations. Historically, however, issues like permanence, uncertainty or additionality have seriously hindered the inclusion of forests-based carbon sequestration activities into climate agreements. Despite those concerns, REDD has been supported by a large number of economic studies. These can be divided into two major categories. The first uses partial equilibrium forest/land use models to derive costs of reduced emissions from avoided deforestation. By comparing three global forestry and land use models Kindermann et al. (2008) offers a good synthesis of that literature. According to those authors, regions with the lowest avoided deforestation costs could provide 2.8-4.7 of Gt of reduced CO_2 emissions during 2005–2030 at 100$ per ton of CO_2. The second branch of that literature combines/links forest/land use models with macroeconomic models, which provide a more comprehensive description of the economic system. By doing so, these studies can jointly investigate the potential of forest-based carbon sequestration with other carbon mitigation options. Sohngen and Mendelsohn (2003), Tavoni et al. (2007) and Bosetti et al. (2011) provide good examples of such studies. The sectorial disaggregation in those studies, however, tends to be rather coarse and international trade absent or marginally taken into account. These two aspects are, however, particularly important to capture when examining the policy impacts of REDD. In fact, REDD directly affects carbon prices (and therefore energy-intensive sectors) and agricultural land availability (consequently, agricultural sectors). Impacts that will not only be differently disseminated throughout the production chain but that directly affect two sectors where international trade is particularly relevant and intense.

This study addresses those issues building upon a Computable General Equilibrium (CGE) model improved to take into account land use change and timber effects resulting from REDD activities. The explicit representation of international and intersectoral trade flows make CGE models particularly apt to this task. Factors of production are mobile between sectors within a country while commodities are exchanged in international markets, responding to scarcity signals provided by changes in relative prices. Therefore, when some 'perturbation' is applied to the economies under investigation, the model provides the induced final implications on their GDP, which is considered market-driven adaptation (all adjustments at work in the economic system that could smooth or amplify the initial impact).

The CGE land use modelling approach in this study builds upon a previous methodology developed in Bosello et al. (2015). Business-as-usual deforestation rates and carbon emission reduction resulting from financing REDD activities are provided by a global forest land use model, the IIASA model cluster (Gusti et al., 2008). The original CGE model is thus modified to account for new regional carbon emissions and changes in both agricultural land availability and timber flows due to avoided deforestation. Our methodology shares therefore some common aspects with that of Hertel et al. (2009). In contrast to that study, however, we do not apply the so-called Agro-Ecological Zone (AEZ) approach (Lee et al., 2009), but develop an alternative methodology enabling us to capture the trade-offs resulting from avoided deforestation, as reduced deforestation translates both in less land available to agricultural activities and to a lower natural resource input to the timber industry.

The role of forest carbon sequestration in global mitigation of climate change has been studied in Hertel et al. (2009), Golub et al. (2009), Golub et al. (2012) and Hussein et al. (2013) using a modelling approach considering the mitigation potential from CO_2 and non-CO_2 emissions as well as a carbon sequestration incentives. These studies use a comprehensive approach considering afforestation, avoided deforestation, and forest management, which correspond to the REDD+ definition. Golub et al. (2009) extends the analysis of Hertel et al. (2009), considering land-based and industrial mitigation, and find that land based sectors could contribute up to half of near-term mitigation at modest carbon prices, with most of the abatement coming from forests. Golub et al. (2012) find that a forest carbon sequestration incentive in developing countries is effective in controlling emission leakage in agricultural sectors under a unilateral mitigation policy only in Annex I countries. Hussein et al. (2013) conduct a disaggregated CGE analysis of the impacts of forest carbon sequestration incentive on poverty in developing countries and find that the overall effect of the incentive is to raise poverty in the majority of developing countries.

Only a restricted number of studies analyses REDD using a CGE framework. At the same time, most of them focus on assessing REDD mitigation potential and its associated costs.[1] Rose et al. (2012) analyze the implications of Total Factor Productivity growth patterns on deforestation. Overmars et al. (2014) couple a CGE model (LEITAP) with an Integrated assessment model (IMAGE) to estimate the opportunity costs of protecting forested areas. In contrast, Gurgel et al. (2007) uses a CGE framework that accounts for deforestation for policy analysis, but focuses on the economic consequences of biofuel's potential production. In fact, to the best of our knowledge, only Bosello et al. 2015 explicitly addresses the role of REDD in an international climate policy. The authors offer, however, a static exercise to investigate the mitigation potential of avoided deforestation resulting from introducing REDD credits in the European Trading Scheme. In this paper, we improve that analysis by using a refined version of the CGE model in a recursive dynamic setup with yearly time steps. In particular, our goal is to develop a simulation exercise that allows for the study of the interlinks between climate policy, international trade and agricultural markets when REDD credits can be traded in an international carbon market. To that end, we setup a policy scenario where a comprehensive climate agreement is in place assuming the Copenhagen Accord pledges. While this policy scenario does not correspond to the most recent state of affairs in international climate negotiations, note that it still serves the purposes of our analysis. It is in that context that the baseline and policy scenarios used in this study should be considered.

[1] For summarized reviews of these studies, we refer to Bosello et al. (2015) regarding REDD and climate policy and to Overmars et al. (2014) for the effects avoided deforestation and mitigation potential.

By using the above-mentioned framework, we add to the literature on the role of REDD in climate policies by studying the following questions: (i) to what extent the use of REDD credits can reduce deforestation rates? (ii) will REDD credits eventually flood international carbon markets? (iii) How will the selling of REDD credits affect REDD regions? (iv) What are the effects of using REDD on economic/carbon leakage? (v) What are the likely effects of REDD on world food production and prices? To answer these questions, we design a policy scenario where an international carbon market is implemented and all countries within the Copenhagen Accord have committed themselves to their announced high pledges. While this may seem a somewhat optimistic assumption, it provides a background that better enables us to assess the consequences of using REDD credits in an international climate policy agreement as: (i) REDD is most likely to be introduced into an international agreement involving a large number of participants; (ii) avoided deforestation has been often presented as an incentive to bring developing countries into the climate policy zone; (iii) concerns on an eventual flood of REDD credits in the carbon market require ambitious mitigation goals. Finally, taking into account this last political concern, different scenarios in which the use of REDD credits is limited are also considered.

Finally, taking into account this last political concern, different scenarios in which the use of REDD credits is limited are also considered. Section two introduces the CGE model and the corresponding modifications to include REDD as a carbon abatement alternative. Section three describes the selected scenarios for policy analysis, while section four discusses the implications of introducing REDD credits exchange in an international carbon market. Section 5 discusses the study's main findings.

2. Modelling Framework

The present analysis relies on ICES (Intertemporal Computable Equilibrium System), a recursive-dynamic CGE model. It is based on the Global Trade Analysis Project (GTAP) model (Hertel, 1997) as well as the GTAP-E model (Burniaux and Truong, 2002), and has been widely used for climate change impact and policy analysis (Bosello et al., 2015; Parrado and De Cian, 2014; Bosello et. al. 2012; Bosello et. al. 2011; and Eboli et, al. 2010). For this particular analysis, ICES has been modified to assess the implications of introducing REDD credits in a carbon market. A detailed model description of the model with the corresponding modifications are described in Appendix 1 and can be found as well in Bosello et al. (2015) and Parrado and De Cian (2014). More details about the aggregation, production tree and baseline assumptions are available on the Supplementary Materials.

On what follows we reproduce a summarized description of the main modifications done by Bosello et al. (2015), highlighting additional changes made to improve the modelling of avoided deforestation and its implications on an international carbon market. The climate policy module originally designed to induce emission reductions from fossil fuel use has been extended to account for emission reductions from avoided deforestation and the trading of the corresponding carbon credits. In addition, the effects of avoided deforestation have been taken into account through three different channels.

First, following Bosetti et al. (2011), we introduce avoided deforestation marginal abatement cost curves estimated by simulations of the International Institute for Applied Systems Analysis (IIASA) model cluster (Gusti et al. 2008), prepared for the Eliasch (2008) report. These abatement curves are time specific, providing the mitigation response to different carbon prices, changing every five years and are available for the following areas: Africa, Central and South America and Southeast Asia. These regions, according to Kindermann et al. (2008), correspond to the areas where avoided deforestation may be supplied at lowest possible costs. In addition, according to the deforestation rates reported by the model cluster (Gusti et al. 2008), these areas cover more than 94% of total world deforestation activity (2000 data). Emission reductions (abatement) from REDD (REDD_CO_2) are a function of the abatement cost in terms of price per ton of CO_2 (pco_2) as in equation (1):

$$REDD\_CO_2 = f(pco_2). \tag{1}$$

This abatement is then subtracted from gross total emissions ($GROSSTCO_2$) originated by the ICES model in each region to get total emissions (TCO_2) following equation (2):

$$TCO_2 = GROSSTCO_2 - REDD\_CO_2. \tag{2}$$

In addition, we allow, for each region providing abatement from REDD, to sell REDD_CO_2 credits in the international carbon market in exchange of emission reduction efforts. The revenues associated to the selling of REDD credits add to sellers' national income and reduce that of the buyers. This implies that the initial gross quota set for each region participating to a carbon market ($GROSSQCO_2$) is corrected by the abatement accomplished thanks to REDD efforts, and therefore in the carbon market the quota (QCO_2) becomes:

$$QCO_2 = GROSSQCO_2 - REDD\_CO_2. \tag{3}$$

Secondly, changes in deforestation due to REDD activities decrease available land for agricultural, forestry and pasture uses. This reduction in available land is defined with respect to baseline land availability under "business as usual deforestation rates". Data for baseline regional land availability were estimated using the IIASA model cluster. These data consist in baseline emissions from deforestation that were converted to additional available land for agriculture and pasture using information from the Food and Agriculture Organization (UN FAO, 2006).

Then, land availability is endogenously corrected in response to (lower) deforestation under different carbon prices according to the following equation:

$$LANDAGR_{r,t} = LANDAGR_{r,t}^{BAU} - LANDAGR_{r,t}^{REDD}, \tag{4}$$

where for each region r, at time t, the amount of available agricultural land in each simulation (LANDAGR), is corrected by subtracting from the available agricultural land under business-as-usual ($LANDAGR^{BAU}$), the amount corresponding to policy induced change in land due to avoided deforestation ($LANDAGR^{REDD}$).

We refine the land effects modelling from Bosello et al. (2015) to correct the fact that not all the land cleared from deforestation ($LANDREDD^{REDD}$) becomes available for agricultural purposes. To calculate the amount of land entering large scale agriculture after deforestation ($LANDAGR^{REDD}$) we use the conversion coefficient $\alpha_r<1$ in equation (5) following UN FAO (2001), and multiply it by the total land related to REDD. According to UN FAO (2001), roughly 10% of deforestation in Africa was due to conversion to this type of land use, while for Latin America and Asia this numbers is equal to 46% and 30%, respectively:

$$LANDAGR_{r,t}^{REDD} = \alpha_{r,t} * LANDREDD_{r,t}^{REDD}. \quad (5)$$

It is important to highlight two points: i) only α_r is valid for land use effects, therefore the remaining ($1-\alpha_r$) simply represents land not used for agriculture, and ii) all abatement related to REDD efforts is considered when calculating net CO_2 emissions as well as for the exchange of REDD credits.

Thirdly, reduced deforestation decreases the volume of timber entering timber markets (TIMBSUPP). This is captured in the model following the same methodology as in equation (4):

$$TIMBSUPP_{r,t} = TIMBSUPP_{r,t}^{BAU} - TIMBSUPP_{r,t}^{REDD}. \quad (6)$$

Business as usual timber supply ($TIMBSUPP^{BAU}$) is endogenously adjusted to account for lower harvesting ($TIMBSUPP^{REDD}$) resulting from lower deforestation rates. To calculate the impact of non-harvested hectares on timber production from primary forest (cubic meters) we coupled data from FAO (UN FAO, 2006) with Brown (2000). This last provides information on harvesting from both primary forests and forest plantations.

3. Scenarios Description

In this section, we present the scenarios used in our simulation exercises. As mentioned above, the primary goal of this study is to shed new light on the interactions between international trade, carbon and agricultural markets resulting from the introduction of REDD credits in an international carbon market. The avoided deforestation marginal abatement cost curves estimated using the International Institute for Applied Systems Analysis (IIASA) model cluster (Gusti et al. 2008) are region and time specific. The cost curves, however, have been simulated for time steps of five years ending in 2020. Having that in mind, and for the sake of consistency, the model baseline year and scenarios for the world economy in this study thus refer to projections available during that period. In particular, we assume the national emission-reduction commitments following the Copenhagen Accord. More details are provided in the text below. While we acknowledge that using more recent data and emission reduction targets may be of higher interest, note that the assumptions here considered still serve our study's purposes. Finally, and to avoid misinterpretations, when presenting the results of our analysis in section 4, we refer to time periods instead of the corresponding calendar time.

Our simulations compare four different scenarios. The first one is the Reference scenario which is a no climate policy, business as usual benchmark spanning from 2001 to 2020. It is obtained perturbing the calibration year equilibrium (2001) in order to replicate the regional GDP growth paths of the A2 IPCC SRES scenario. This baseline also incorporates medium-term price evolution of major fossil fuels according to EIA (2009).

In the second scenario, under the name High Pledges, all countries commit themselves to the high pledges defined in the Copenhagen Accord (see Table 1), but REDD policies are not implemented. A fully integrated carbon market in the form of an Emission Trading Scheme (ETS) is implemented only for countries with emission reduction targets. Accordingly, China and India, whose targets are defined in carbon intensity terms, pursue independent domestic policies consisting in the introduction of a carbon tax to comply with their pledges. Both SSA and ROW regions have no commitment nor participate to the carbon market.

Table 1: Emissions reduction from High Pledges scenario for the ICES regions

Region	Target for 2020	With respect to 2001 levels
Australia	25% against 2000 levels	-33,3%
New Zealand	20% against 1990 levels	-51,9%
China	GDP carbon intensity reduction: 45% with respect to 2005	-
Japan	25% against 1990 levels	-41,2%
South Korea	30% against baseline	-22,1%
India	GDP carbon intensity reduction: 25% with respect to 2005	-
Canada	17% against 2005 levels	-24,5%
USA	3% against 1990 levels	-20,3%
EU27	30% against 1990 levels	-37,3%
Russia	25% against 1990 levels	7,7%
South Africa	34% against baseline	-31,8%
NORICE	39% against 1990 levels	-69,4%
EASIA *	(Indonesia)*	63,9%
LACA *	(Brazil and Mexico)*	7,8%
SSA	No target	-
ROW	No target	-

Note: For the regions flagged with * the target is defined imposing the emission reduction required for the individual countries inside it that have a commitment under Copenhagen: EASIA – Indonesia 26% emission reduction against baseline by 2020; LACA – Mexico 30% emission reduction against baseline by 2020; Brazil 39% emission reduction against baseline by 2020.

In the third scenario, High Pledges + REDD, mitigation policy targets are defined as above, but with the additional possibility for SSA, LACA and EASIA thereafter to enter the ETS selling REDD credits. Therefore, LACA and EASIA can potentially sell emission reduction credits coming from both reduced emissions compared to their targets and REDD activities. SSA that does not hold any pledge on emission reductions is, however, allowed to sell REDD credits on the basis of proven reduction in its "business as usual" deforestation activities. This option has been chosen as it provides the highest incentive for REED countries to engage in avoided deforestation actions and allows us to better evaluate its role in this policy context. Finally, in the fourth scenario, we simulate different restrictions to the use of REDD credits (High Pledges + Limited REDD).

A final remark regarding the policy modelling procedure. Given the dynamic nature of the model it is assumed that the desired mitigation target is gradually imposed starting from 2010 and becoming linearly more stringent until 2020 when all regions comply with their respective commitments. In what follows we refer to time using "time-periods" instead of calendar time, implying that our scenarios start at period 1, end at period 20 and climate policy is enacted at period 10.

4. Trading REDD Credits in the Carbon Market

4.1. Climate policy without redd – high pledges scenario

To better understand the implications of a climate change policy, and in particular of an international carbon market, it is first necessary to evaluate how the different regional annual emission targets compare to their business-as-usual emissions. Therefore, in this section we compare the "High Pledges" and the "Reference" scenarios. In absolute terms, the top 3 regions with higher emission reduction levels vis-à-vis to BAU are the USA, EU27 and Japan with a decrease of, respectively, 2695, 2108 and 675 Million Tons of CO_2 in period 20 (see Table A3 in the supplementary materials). This could be referred to as the absolute mitigation effort made by those countries. In relative terms however, where annual relative reduction is defined as percentage of the BAU emission scenario, this ordering changes to NORICE, New Zealand and Japan with a decrease of 69%, 60% and 46% in period 20, respectively (see Table A4 in the supplementary materials). This represents their relative mitigation effort.

The mitigation policy implemented originates a carbon price rising from 4.4$/t CO_2 in period 10 to 77$/t CO_2 in period 20 (see Table 2). The magnitudes of transactions, and the respective role different regions play in the international carbon market, tend to reflect the relative positions of their targets with respect to business-as-usual emissions. The main buyers of carbon credits in absolute terms are EU27 and Japan, while the main sellers are USA, EASIA and Russia (see Table 3). In relative terms, defined as the percentage of emissions traded credits with respect to the annual target, the main buyers are NORICE, New Zealand, Japan and EU27; while USA no longer ranks among the top 3 sellers that are now constituted by South Africa, EASIA and Russia.

Table 2: GDP and CO_2 prices with respect to BAU in period 20

		High Pledges	High Pledges + Limited access to REDD credits in the ETS market				High Pledges
		Without REDD	25%	50%	75%	100%	unlimited REDD
GDP % w.r.t B.A.U	Australia	-1,96%	-1,91%	-1,86%	-1,81%	-1,76%	-0,84%
	New Zealand	-1,39%	-1,36%	-1,32%	-1,28%	-1,24%	-0,60%
	China	0,49%	0,46%	0,43%	0,41%	0,38%	0,00%
	Japan	-0,58%	-0,57%	-0,55%	-0,54%	-0,52%	-0,21%
	South Korea	-2,70%	-2,63%	-2,56%	-2,48%	-2,41%	-1,24%
	India	0,99%	0,96%	0,93%	0,90%	0,87%	0,39%
	Canada	-1,32%	-1,28%	-1,24%	-1,20%	-1,16%	-0,47%
	USA	-1,10%	-1,07%	-1,04%	-1,01%	-0,99%	-0,48%
	EU27	-0,65%	-0,62%	-0,60%	-0,58%	-0,55%	-0,15%
	Russia	-8,99%	-8,75%	-8,51%	-8,28%	-8,05%	-3,98%
	South Africa	-6,76%	-6,61%	-6,46%	-6,31%	-6,17%	-3,27%
	NORICE	0,18%	0,17%	0,17%	0,17%	0,16%	0,20%
	EASIA	-2,11%	-2,05%	-1,99%	-1,93%	-1,88%	-0,98%
	LACA	-0,80%	-0,76%	-0,72%	-0,68%	-0,64%	0,00%
	SSA	2,12%	2,15%	2,16%	2,17%	2,17%	1,39%
	ROW	2,54%	2,48%	2,41%	2,35%	2,28%	1,38%
	Climate Policy Region	-0,87%	-0,85%	-0,82%	-0,80%	-0,77%	-0,34%
CO_2 Price I (period 20)	CO_2 ETS Price $/t	76,7	74,9	73,1	71,4	69,7	36,2
	% reduction w.r.t. policy without REDD	-//-	-2%	-5%	-7%	-9%	-53%
	China's Carbon tax $/t	7,5	7,5	7,5	7,4	7,4	6,9
	India's Carbon tax Price $/t	10,0	9,9	9,8	9,8	9,7	8,0
	REDD in period 20 as % of BAU Deforestation	-//-	3%	5%	8%	11%	75%

The resulting cost for the policy-participating countries as a whole equals a loss of 0.87% of GDP compared to baseline. Whilst in absolute terms the USA and the EU27 are the regions bearing higher policy costs; in relative terms as % of GDP, losses are higher in Russia, South Africa and South Korea (9%, 7% and 3% respectively – see Table 2). Interestingly enough, India and China observe a higher GDP growth in the policy than in the baseline scenario (1 and 0.5% respectively). Note that both regions pursue domestic polices targeting carbon intensity, which in fact allow them to increase emissions even though less than in the baseline. Consequently, they face significantly lower carbon prices than those observed inside the ET market (7.5$/t CO_2 for China and 10$/t CO_2 for India in period 20). As a result, these two regions become relatively more competitive, they produce more, especially carbon intensive commodities, (the so-called economic leakage effect), and enjoy more growth overcompensating the cost of reducing carbon intensity. This effect is even stronger in those regions without any pledge (ROW and SSA). Emissions outside the "climate policy zone" increase with a carbon leakage effect of 9% if measured against reductions from countries with emission targets, and 7% considering also the mitigation effort from China and India.[2]

Table 3: Carbon market trading in period 20

	ETS TRADE (Mtons of CO_2) *				ETS TRADE (2001 US$ million)			
	NO REDD	REDD			NO REDD	REDD		
		unlimited REDD	restriction			unlimited REDD	restriction	
			100%	50%			100%	50%
Australia	-5.87	61.73	2.92	-1.50	450	-2234	-203	110
New Zealand	12.77	18.27	13.57	13.17	-979	-661	-946	-963
Japan	329.70	485.82	353.29	341.57	-25277	-17582	-24632	-24980
South Korea	-49.83	39.75	-37.32	-43.58	3820	-1439	2602	3187
Canada	10.11	98.85	23.06	16.61	-775	-3578	-1608	-1214
USA	-488.78	646.88	-339.36	-414.41	37473	-23411	23660	30306
EU27	955.44	1437.12	1027.48	991.72	-73250	-52011	-71637	-72527
Russia	-263.39	64.09	-216.55	-240.01	20193	-2319	15099	17552
South Africa	-86.55	-20.34	-78.39	-82.52	6636	736	5466	6035
NORICE	31.02	35.65	31.79	31.41	-2378	-1290	-2217	-2297
EASIA	-341.34	-639.10	-356.15	-338.06	26169	23130	24831	24723
LACA	-103.28	-1985.40	-278.52	-196.84	7918	71853	19419	14395
SSA	0.00	-243.32	-145.81	-77.57	0	8806	10166	5673

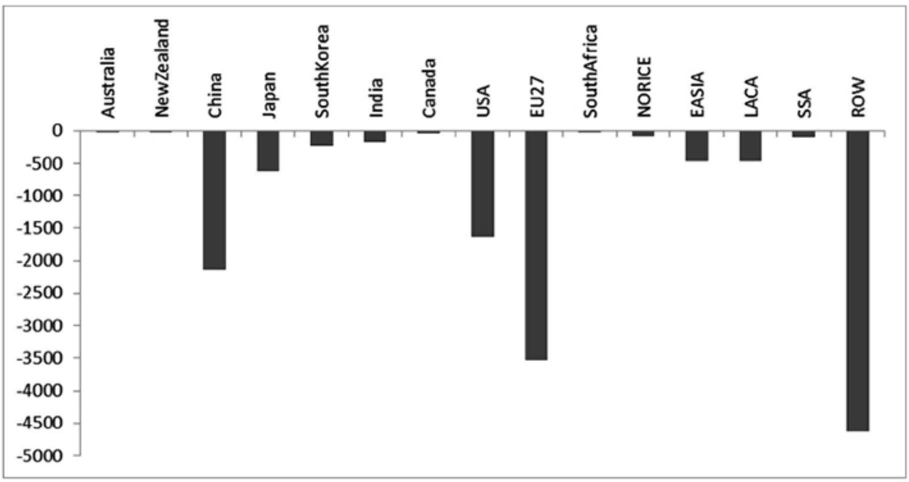
Note: Negative numbers are credit supplies.

The high policy costs occurring in Russia and South Africa deserve a more detailed analysis. The Russian sectors that most contribute to national exports, and that also rank highly in terms of total production, are all energy intensive. In particular, the main destination of these exports is EU27, that highly decreases its imports from Russia when climate policy is implemented, substituting them from regions outside the "climate policy zone". This decrease in Russian exports towards EU27 is particularly strong for Energy intensive industries (see Figure 1). In addition to this, China and ROW are, after EU27, the other two most important destinations of Russian exports. As EU27, these regions also sharply reduce their imports from Russia. In fact, together with EU27, they make up the top 3 regions with higher decrease in Russian exports. The reduction of Russian imports in these regions are due to a substitution of imports of energy intensive products for national production, a direct result of their productive systems being more competitive as they are outside the implemented carbon policy. In a nutshell, carbon leakage severely damages Russian production as, on the one hand its mains importers shift their demand towards areas not subject to carbon price and, on the other hand, regions outside the "carbon policy zone" substitute

[2] We present this difference due to the divergence of targets between China and India (defined as carbon intensity) with respect to the remaining countries inside the Copenhagen Accord (quantitative emission reduction targets)

their imports for domestic production. The importance of these industries in the Russian economy, combined with the major role played by climate policy free riding economies in Russian bilateral trade, make this country particularly vulnerable to carbon leakage under this policy design and explain the high policy costs observed for this region.

Figure 1: Russian energy intensive industries exports change wrt BAU in period 20 (2001 Million U.S. dollars)



In a similar way, the vulnerability of South Africa to carbon leakage explains its high policy costs. The main destination regions for South African exports are EU27, USA, Japan, and SSA. After the policy has been implemented these regions decrease their imports of energy intensive products from South Africa, and source them from relatively more competitive exporting regions.

4.2. Climate policy with redd – high pledges + redd

As expected, and in line with previous literature, we observe that climate policy costs can be significantly reduced by opening the carbon market to an unrestricted use of REDD credits. Those are now only 0.34% of GDP compared to baseline. The carbon price starts from 2.5$/t CO_2 in period 0 to reach around 36$/t CO_2 in period 20. The large number of REDD credits entering the carbon market allows regions participating to the climate policy agreement to systematically emit above their targets. Had the announced targets be met without REDD, total emission of the countries with pledges would have equalled 14.3 billion tonnes of CO_2 in period 20. Under the unrestricted REDD scenario, they reach 17.1 billion. This implies that the 2.8 billion tonnes of CO_2 emissions increase is compensated by avoided deforestation. REDD revenues as a share of GDP could represent up to 2% for LACA and SSA and 1% for EASIA (see Table A6 in the supplementary materials). The only seller of carbon credits that is not a REDD region is South Africa (note that South Africa was already the main seller in the market in the No REDD credits scenario). From the REDD regions group only EASIA sell carbon credits in addition to the ones resulting from REDD, and LACA sells only a fraction of the abatement attained with REDD using the rest to offset emissions within the region (see Table 4). While it is economically sounding that abatement is shifted to lower abatement costs activities, these results confirm that policy design may require the creation of long-term incentives to promote a greener economy. In effect, the option of limiting the number of REDD credits allowed in the market has been widely proposed in the policy arena. We discuss this option in the next subsection.

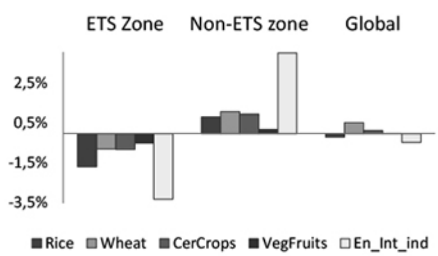
Note also that REDD reduces the costs of all countries initially loosing with the climate policy, but decreases the benefits of those gaining (i.e. SSA, China, India and ROW). Exceptions are Norway and Iceland that remain basically unaffected (see table2). This result is related with carbon leakage and international competitiveness effects. By reducing the abatement effort of countries with binding emission reductions, REDD alleviates the burden on their energy intensive industries, goods and services. These are thus more competitive in international markets as there is a less stringent environmental policy (read a lower carbon price signal). Conversely, for regions without emission reduction targets benefiting in the policy scenario due to the existence of a leakage effect, REED credits generates higher indirect costs than direct benefits, as they face regions with binding emission reduction that are now more competitive in the international market. Thus, for instance, SSA, would prefer not to sell REDD credits and loose the related financial inflows than to sell REDD credits to a group of countries whose products would consequently become cheaper and more competitive. This is a typical situation in which higher order effects, through competitiveness, prevail over first order effects (cost savings or direct money inflows). In fact, almost all agricultural sectors deteriorate their trade balances when comparing the No REDD scenario versus the REDD one (see Table A5 in the supplementary materials). Finally, REDD also helps to mitigate the increase in emissions occurring outside the climate policy zone. This is now equal to 4% if measured against reductions from countries with emission targets, compared to a 9% increase in the scenario without REDD credits.

4.2.1. *Effects on agricultural production and prices*

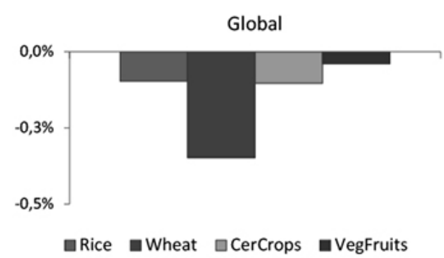
We start this section by first observing the effects on agricultural production and prices of a climate policy without REDD credits (*High Pledges scenario*). The introduction of a carbon price reduces world production of energy intensive sectors but increases the demand for agricultural products. Looking to this with more detail, one finds that agricultural production actually decreases in regions pertaining to the global carbon market (ETS zone), but that this reduction is compensated by the increase in demand occurring in the rest of the world (see Figure 2a). The aggregate increase in world production is therefore triggered by economic leakage resulting from climate policy implementation. Regions outside the climate policy zone experience higher GDP growth and consequently increase their demand, including the one for agricultural products. This increase compensates the reduction in regions pertaining to the policy area and, as a result of a higher demand, global agricultural products' prices increase (see Figure 3a). The only exception to this occurs in the rice sector, where the increase in demand of regions outside the carbon market is not enough to balance the reduction occurring in important producing regions belonging to that zone like Japan, South Korea and EASIA.

Figure 2: Change in production

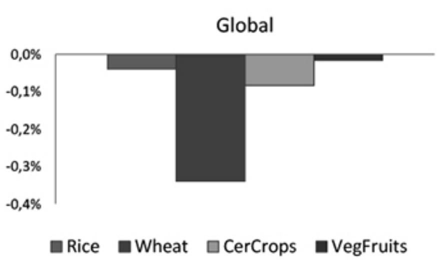
a) No REDD Policy vs BAU



b) Policy REDD with land effects vs no REDD



c) Policy REDD without land effects
vs Policy No REDD



d) Policy REDD with land effects
vs Policy REDD without land effects

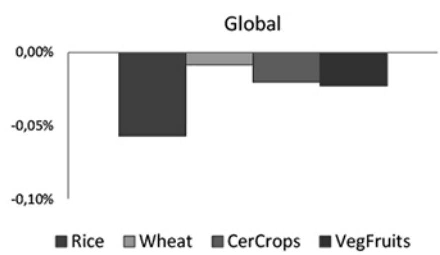
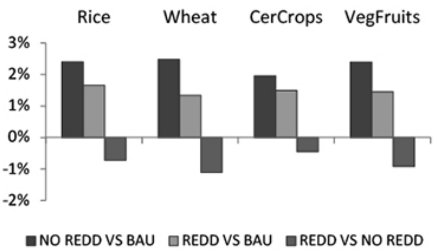
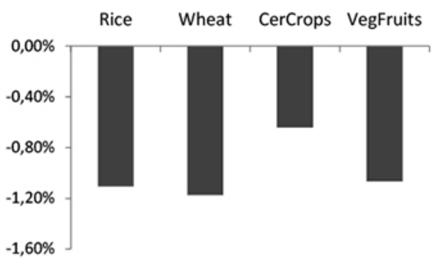


Figure 3. Change in world agricultural prices

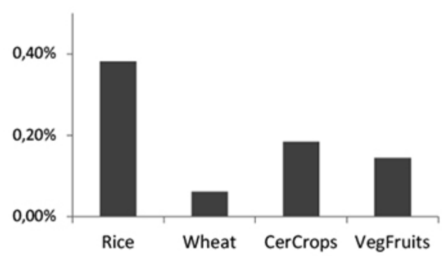
a) Policy scenarios vs BAU



b) Policy REDD without land effects
vs Policy No REDD



c) Policy REDD with land effects
vs Policy REDD without land effects



When REDD credits are allowed to enter the international carbon market one should note that two conflicting impacts occur:

(i) The first is a supply effect. REDD reduces the total amount of available agricultural land and negatively impacts agricultural supply. C*eteris paribus*, this reduces world agricultural production and increase prices vis-à-vis a no REDD scenario.

(ii) The second is a demand side effect consisting of two parts. On the one hand, as carbon price decreases, economic leakage reduces. Regions not pertaining to the carbon market therefore experience lower GDP growth rates compared to a no REDD scenario and, as a result, reduce their demand for agricultural products. On the other hand, regions within the policy zone benefit from a lower carbon price and grow at higher rates than under a No REDD scenario, thus increasing their demands for agricultural products. The demand side effect is therefore ambiguous. Whether the final effect on world aggregate demand for agricultural products will be positive or negative will depend on which of these forces dominate.

How the introduction of REDD credits affect world agricultural production and prices is, therefore, ultimately an empirical question. In case the reduction of economic leakage dominates the demand effect, i.e. world agricultural demand is lower vis-à-vis to a no REDD scenario, world agricultural production will be lower under the REDD scenario while the effect on prices in unclear Alternatively, if world demand for agricultural products increases vis-à-vis to a no REDD scenario then prices should increase while changes in production are uncertain.

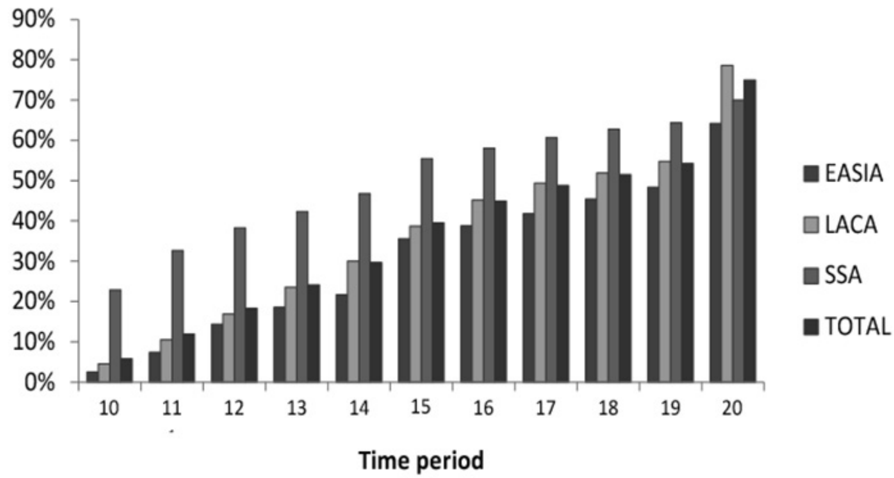
To disentangle these counteracting effects we run an additional simulation using the original ICES model, i.e., when the modifications described in section 2 are not active. By doing so, REDD credits will be available in the carbon market but will not affect agricultural land availability, i.e., we control for the supply effect identified in (i).

Results from this simulation reveal that world agricultural production decreases (see Figure 2c), reducing also agricultural products' prices (see Figure 3b). The demand side effect identified in (ii) is therefore negative. The reduction in agricultural production triggered by reduced economic leakage is stronger than the increase in demand occurring in the regions that benefit from a lower carbon price. This is a point that is noteworthy to highlight: the reduced leakage effect resulting from the introduction of REDD credits reduces world agricultural production vis-à-vis to a policy scenario without REDD credits, even if agricultural land availability remains unchanged.

It remains to answer, however, if lower agricultural land availability may still induce higher agricultural products' prices. Running now a simulation allowing for REDD credits, and using the modified model including land effects, we observe that global agricultural production further decreases while prices only marginally increase with respect to the simulation using the original ICES model simulation (see Figures 2d and 3c). This slight increase being so small that world agricultural product prices are still inferior to the climate policy scenario without REDD credits (see Figure 3a).

Concluding, the lower global demand of agricultural products induced by lower carbon prices outweighs the reduced land availability triggered by REDD credits. As a result, and while this may at first seem counter-intuitive, the introduction of REDD credits reduces both world agricultural production and prices vis-à-vis to a policy scenario without REDD (see Figures 2b, and 3a). As expected, the only exception to this general conclusion occurs on REDD regions (EASIA, LACA and SSA), where the land effect prevails and agricultural production decreases but prices increase due to higher scarcity of agricultural land. Globally, however, we have once again a fine example where indirect effects (carbon price reduction) of a climate policy prevail over direct ones (reduction in agricultural land).

Figure 4: Avoided deforestation (% BAU deforestation rates)



Finally, note that despite the sharp carbon price reduction, the use of REDD credits is still enough to trigger substantial amounts of avoided deforestation. In fact, even if in period 10 only 6% of business as usual deforestation is avoided, this number rapidly increases to 75% in period 20. While in absolute terms LACA is the region with higher avoided deforestation levels, SSA is the region with the highest avoided deforestation rate defined as a percentage of BAU deforestation (see Figure 4).

4.3. Introducing limits to the use of redd credits

In the unrestricted scenario a fairly large number of REDD credits enter the market and the carbon price drops to 36$ per tons of CO_2 in period 20. In order to prevent such a flooding into the carbon market, it has been often proposed the imposition of restrictions to the use of this type of credits. Such a policy envisages regulating carbon prices' decrease keeping thus incentives sufficiently high to foster research and development of renewable and more efficient technologies but also as an incentive for early participation of REDD countries in global climate policy.

In the present analysis, REDD restrictions are defined in terms of emission reduction efforts. According to our business-as-usual scenario, in the year immediately before climate policy implementation the aggregate emissions of regions participating in the international carbon market amount to 18676 million tonnes of CO_2. By period 20 this number has to decrease to 14305. A restriction of 100% therefore implies that during the time policy horizon the total amount of REDD credits allowed to enter the market cannot be superior to the required reduction, i.e. 4371 million tonnes of CO_2. Accordingly, for a restriction of 25% this last figure is equal to 1092 million tonnes of CO_2. With this in mind we have considered 4 restriction levels, 25%, 50%, 75% and 100%. We observe that, for the restrictions here considered, the carbon price decrease is significantly reduced. For a restriction equal to 25%, carbon price drops only by 2% in period 20 while this number equals 9% if the level of restriction is 100% (see Table 2). As expected, including REDD credits restrictions still generates policy costs savings, but to a much lower extent if compared to the unrestricted scenario. For the 25% restriction scenario, policy costs equal 0.85% in terms of GDP, while for the 100% restriction scenario GDP is reduced by 0.77% in period 20. Finally, a heavy restriction in the use of REDD credits also undermines the use of such a policy as a way to significantly reduce deforestation rates. For the 100% restriction scenario, avoided deforestation amounts to 6% of period 10 business as usual deforestation, rising only to 11% in period 20.

In light of the discussion made in the previous subsection, we conclude that EASIA and LACA are increasingly worse off as the restriction to REDD credits is more stringent (see table2). Accordingly, while such a policy aiming to control for carbon price decreases may create a more favourable economic environment to the development of cleaner technologies, conversely it may prevent countries with higher deforestation rates from entering into a global climate policy agreement. On the other hand, however, SSA who does not have a binding emission reduction target is actually unambiguously better off under the restriction scenarios here considered. In section 4.2 we have observed that the introduction of REDD vis-à-vis a policy scenario without REDD credits, actually damaged this region as it reduced carbon leakage. In contrast, however, when the use of REDD credits are limited, carbon prices only slightly reduce allowing this region to still reap the benefits resulting from the leakage effect. In addition, REDD revenues are still high for this region under restriction scenarios, as for SSA the reduction in quantities of sold REDD credits is almost compensated by the carbon price increase. By period 20, this last effect is actually so strong that SSA REDD revenues are higher under the 100% restriction scenario than under the unrestricted one.

5. Final Remarks

By using a modified global CGE model to take into account avoided deforestation induced effects this paper sheds new light on the use of REDD credits in an international carbon market. In addition to confirm previous results on the major role that such credits may play in climate change policy, we also reveal that changes occurring in international markets, namely in energy intensive sectors, are crucial in the design of optimal REDD policies. Those changes may be so important that they can actually dominate direct effects resulting from avoided deforestation activities. Such an analysis is out of the scope of typical partial equilibrium models or macroeconomic models that do not explicitly take into account international trade. Two examples are noteworthy to highlight. First, indirect effects occurring on international carbon markets may prevail over direct impacts. This is for instance the case regarding impacts on agricultural markets. By reducing carbon leakage, the use of REDD credits reduces agricultural products demand by regions outside the climate policy zone vis-à-vis to a policy without such credits. This demand reduction effect turns out to be stronger than the direct impact of reduced agricultural land availability triggered by avoided deforestation activities. As a consequence, the use of REDD credits alleviates pressure on world agricultural prices. Second, the use of REDD as a means to foster developing countries participation into climate policy may not be sufficient. Financial flows accruing from REDD revenues may not be enough to compensate for a reduction in free riding benefits. REDD credits provide a sounding instrument to reduce the increase of emissions occurring outside the climate policy zone, significantly reduces climate policy costs, may provide an effective instrument to reduce deforestation rates but other instruments are likely to be necessary to make REDD countries positively engage in international negotiations.

References

Bosello, F.; Eboli, F.; Parrado, R.; Nunes, P.; Ding, H.; Rosa, R. (2011) The economic assessment of changes in ecosystem services: An application of the CGE methodology. *Economia Agraria y Recursos Naturales*, 11(1), 161-190.

Bosello, F.; Nicholls, R.J.; Richards J., Roson, R.; Tol, R. (2012) Economic impacts of climate change in Europe: Sea-level rise. *Climatic Change*, 112, 63-81.

Bosello, F.; Parrado, R.; Rosa, R.;Eboli, F. (2015) REDD in the carbon market: A general equilibrium analysis. *Environmental Modeling & Assessment*, 20, 103-115.

Bosetti, V.; Lubowski, R.; Golub, A.; Markandya, A. (2011) Linking reduced deforestation and a global carbon market: Implications for clean energy technology and policy flexibility. *Environment and Development Economics*, 16, 479-505.

Brown, C. (2000) The global outlook for future wood supply from forest plantations. *FAO (working paper N°: GFPOS/WP/03)*.

Burniaux, J.M.; Truong, T.P. (2002) GTAP-E: An Energy-Environmental Version of the GTAP Model. *GTAP Technical Paper n.16*.

Eboli, F.; Parrado, R.; Roson, R. (2010) Climate-change feedback on economic growth: Explorations with a dynamic general equilibrium model. *Environment and Development Economics*, 15(5), 515-533.

EIA – Energy Information Administration. (2009) *International Energy Outlook 2009*. Report n° DOE/EIA-0484(2009).

Eliasch, J. (2008*) Climate Change: Financing Global Forests*. Norwich, UK: The Stationery Office.

Golub, A.; T.W. Hertel, H-L. L.; Rose, S.; Sohngen, B. (2009) The opportunity cost of land use and the global potential for greenhouse gas mitigation in agriculture and forestry. *Resource and Energy Economics*, 31(4), 299-319.

Golub, A.; Henderson, B., Hertel, T. Gerber, P.; Rose, S.; Sohngen, B. (2012) Global Climate Policy Impacts on Livestock, Land Use, Livelihoods and Food Security. Proceedings of the National Academy of Sciences.

Gullison, R.E.; Frumhoff, P.C.; Canadell, J.G.; Field, C.B., Nepstad, D.C.; K. Hayhoe, R.; Avissar, L.M.; Curran, P; Friedlingstein, C.D. J.; Nobre, C. (2007) Tropical forests and climate policy. *Science*, 316(5827), 985-986.

Gurgel A.; Reilly, J.; Paltsev, S. (2007) Potential land use implications of a global biofuels industry. *Journal of Agricultural & Food Industrial Organization*, 5, Article 9, Special issue: Explorations in Biofuels Economics, Policy, and History.

Gusti, M.; Havlik, P.; Oberstenier, M. (2008) Technical description of the IIASA model cluster. International Institute for Applied System Analysis (IIASA).

Hertel, T.W. (1997). *Global Trade Analysis: Modeling and applications*. Cambridge: Cambridge University Press.

Hertel, T.; Lee, H-L; Rose, S.; Sohngen, B. (2009) Modeling land-use related greenhouse gas sources and sinks and their Mitigation potential. In Hertel, T.W.; Rose, S.; Tol, R.S.J. (Eds.) *Economic Analysis of Land Use in Global Climate Change Policy*. Routledge Explorations in Environmental Economics.

Hussein, Z., Hertel, T.; Golub, A. (2013). Climate change mitigation policies and poverty in developing countries. *Environmental Research Letters*, 8 035009 (10pp). DOI 10.1088/1748-9326/8/3/035009

IPCC. (2007) Climate Change 2007: Synthesis Report. *Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, Pachauri, R. K and Reisinger, A. (eds.)]. IPCC, Geneva, Switzerland.

IPCC. (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Barros, V.R.; Field, C.B.; Dokken, D.J.; Mastrandrea, M.D.; Mach, K.J.; Bilir, T.E. Chatterjee, M.; Ebi; K.L. Estrada, Y.O.; Genova, R.C.; Girma, B.; Kissel, E.S.; Levy, A.N.; MacCracken, S.; Mastrandrea, P.R.;White, L.L. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 688.

Kindermann, G.; Obersteiner, M.; Sohngen, B.; Sathaye, J.; Andrasko, K.;Rametsteiner, E.; Schlamadinger, B.; Wunder, S.; Beach, R. (2008) *Global cost estimates of reducing carbon emissions through avoided deforestation*. Proceedings of the National Academy of Sciences. 105(30), 10302–10307.

Lee, H.L.; Hertel, T.; Rose, S.; Avetisyan, M. (2009) An integrated global land use database for CGE analysis of climate policy options. In Hertel T. W.; Rose, S. K., Tol, R. S. J., (Eds.), *Economic Analysis of Land Use in Global Climate Change*Policy. Routledge Explorations in Environmental Economics.

Overmars, K. P.; Stehfest, E.; Tabeau, A.; van Meijl, H.; Mendoza Beltrán, A.; Kram, T. (2014) Estimating the opportunity costs of reducing carbon dioxide emissions via avoided deforestation, using integrated assessment modelling. *Land Use Policy* 41, 45-60.

Parrado, R.; De Cian, E. (2014) Technology spillovers embodied in international trade: Intertemporal, regional and sectoral effects in a global CGE framework. *Energy Economics*, 41, 76-89.

Rose, S.K.; Golub, A.; Sohngen, B. (2012) Total factor and relative agricultural productivity and deforestation. *American Journal of Agricultural Economics*, 95(2), 426-434.

Sohngen, B. ; Mendelsohn, R. (2003) An optimal control model of forest carbon sequestration. *American Journal of Agricultural Economics*, 85(2), 448-457.

Strassburg, B.B.N.; Kelly, A.; Balmford, A.; Davies, R.G., Gibbs, H.K.; Lovett, A.; Miles, L.; Orme, C.D.L.; Price,J.; Turner, J. R.K.; Rodrigue, A.S.L. (2010) Global congruence of carbon storage and biodiversity in terrestrial ecosystems. *Conservation Letters*, 3(2), 98-105.

Tavoni, M.; Sohngen, B.; Bosetti, V. (2007) Forestry and the carbon market response to stabilize climate. *Energy Policy*, 35(11), 5346-5353.

UN FAO – Food and Agricultural Organization (2001) *Forest Resources Assessment 2000: Main Report*. Chapter 46. Rome.

UN FAO – Food and Agricultural Organization. (2006) *Global Forest Resources Assessment 2005: Progress towards sustainable forest management*. FAO Forestry paper 147. Rome, Italy.

APPENDIX

Baseline assumptions and database aggregation

The regional and sectoral detail of the model used for this study are represented in the following:

Table A1: Regional and sectoral disaggregation of the ICES model

Region	Sectors	
Australia	Rice	Non-Market Services
New Zealand	Wheat	
China	Other Cereal	
Japan	Vegetable Fruits	
South Korea	Animals	
India	Forestry	
Canada	Fishing	
USA	Coal	
EU27	Oil	
Russia	Gas	
South Africa	Oil Products	
NORICE	Electricity	
EASIA	Water	
LACA	Energy Intensive industries	
SSA	Other industries	
ROW	Market Services	

Note: NORICE denotes Norway and Iceland; EASIA denotes East Asia; LACA denotes Latin America and the Caribbean; SSA denotes Sub-Saharan Africa; and ROW denotes Rest of the World.

ICES solves recursively a sequence of static equilibria linked by endogenous investment determining the growth of capital stock from 2001 to 2050. For the baseline or Business as Usual scenario we relied in exogenous drivers for population, energy efficiency as well as fossil fuel prices projections. Assumptions on the evolution of population were taken from UNPD (2008), energy efficiency from Bosetti et. al., (2006), while major fossil fuel prices are based on EIA (2007) and EIA (2009). Regarding GDP, growth rates for the selected regions are reported in Table A2, and we used as reference the IPCC A2 scenario. Labour stock grows at the same pace as population while capital is cumulated following the recursive dynamics of the model. Finally, we changed labour productivity in order to replicate the target GDP growth rates.

Table A2: Selected growth rates for BAU scenario (% 2001-2020)

Region	GDP growth	Population	CO_2 emissions
Australia	52.4	20.9	17.5
NewZealand	59.2	18.3	20.8
China	222.7	11.1	133.3
Japan	35.5	-2.2	8.4
SouthKorea	46.8	0.0	11.3
India	142.9	30.3	62.7
Canada	54.1	18.0	12.2
USA	57.9	19.0	25.2
EU27	42.7	3.0	9.3
Russia	95.9	-9.8	46.2
South Africa	37.4	11.8	0.3
NORICE	30.4	12.7	-1.7
EASIA	177.1	24.4	79.9
LACA	92.5	24.3	32.4
SSA	122.9	58.1	85.6
ROW	120.9	31.8	60.8
World	67.72	23.5	45.75

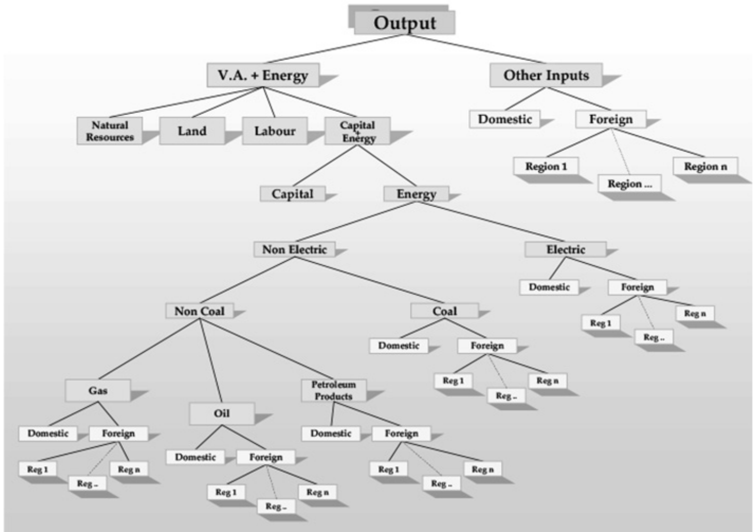
ICES technical description

ICES (Inter-temporal Computable Equilibrium System) is a top-down recursive-dynamic, multi-sector and multi-region CGE model developed mainly with the aim of analyzing climate change impacts and policies. In contrast to integrated assessment models, climate change damages are not endogenous to the model. However, ICES can be used to simulate the economy-wide impacts of climate change imposed as exogenous shocks to inputs of the model (Bosello et al., 2006, 2007, 2008; Bosello and Zhang, 2006; Eboli et al., 2010).

Supply side structure of firms

Each industry is modeled as a cost-minimizing representative firm taking prices as given. Output prices are given by average production costs. The production functions are specified via a series of nested CES functions. Domestic and foreign inputs are not perfect substitutes, according to the so-called "Armington" assumption. The production tree is reported in Figure A1.

Figure A1: Nested tree structure for industrial production processes of the ICES model



Regional subscripts have been omitted for convenience in the following equations. For a complete detail of all the remaining equations, interested readers may refer to Hertel (1997).

Final output of sector j (Y_j) is a function of a technological index (A_j), aggregate value added-energy composite (VAE_j), other intermediate inputs (M_j), and α_j are distribution parameters. The elasticity of substitution for the top nest (σ_M) has been set equal to 0, therefore, representing a Leontieff specification:

$$Y_j = A_j \left[\alpha_{VAE,j} VAE_j^{\frac{\sigma_{M-1}}{\sigma_M}} + \alpha_{M,j} M_j^{\frac{\sigma_{M-1}}{\sigma_M}} \right]^{\frac{\sigma_M}{\sigma_{M-1}}} \quad \text{(A1)}$$

Aggregate value added-energy output, VAE_j, is produced with Z_i primary factors (i = land, labor, natural resources, and a capital-energy composite, KE), with an elasticity of substitution σ_{VAE} and a distribution parameter, δ_{ij}:

$$VAE_j = \left[\sum_i \delta_{i,j} Z_{i,j}^{\frac{\sigma_{VAE}-1}{\sigma_{VAE}}} \right]^{\frac{\sigma_{VAE}}{\sigma_{VAE}-1}} \quad \text{(A2)}$$

The capital-energy composite (*KE*) is produced by combining capital (*K*) and energy (*E*) as illustrated by equation A3:

$$KE_j = \left[\alpha_{k,j} K_j^{\frac{\sigma_{KE}-1}{\sigma_{KE}}} + \alpha_{e,j} E_j^{\frac{\sigma_{KE}-1}{\sigma_{KE}}}\right]^{\frac{\sigma_{KE}}{\sigma_{KE}-1}} \quad \text{(A3)}$$

The Energy (*E*) nest compounds Electricity (*EL*) with Non-Electric energy (*NEL*) and an elasticity of substitution ($\sigma_{ELY}=1$):

$$E_j = \left[\alpha_{EL,j} EL_j^{\frac{\sigma_{ELY}-1}{\sigma_{ELY}}} + \alpha_{NEL,j} NEL_j^{\frac{\sigma_{ELY}-1}{\sigma_{ELY}}}\right]^{\frac{\sigma_{ELY}}{\sigma_{ELY}-1}} \quad \text{(A4)}$$

Non-electric energy (*NEL*) is composed of Coal and Non-Coal energy, assuming an elasticity of substitution of $\sigma_{COAL}=0.5$:

$$NEL_j = \left[\alpha_{COAL,j} COAL_j^{\frac{\sigma_{COAL}-1}{\sigma_{COAL}}} + \alpha_{NCOAL,j} NCOAL_j^{\frac{\sigma_{COAL}-1}{\sigma_{COAL}}}\right]^{\frac{\sigma_{COAL}}{\sigma_{COAL}-1}} \quad \text{(A5)}$$

Liquid fossil fuels (*F*) are combined in a composite (*NCOAL*) also following a CES production function with the elasticity of substitution ($\sigma_{FF}=1$):

$$NCOAL_j = \left[\Sigma_i \, \beta_{i,j} F_{i,j}^{\frac{\sigma_{FF}-1}{\sigma_{FF}}}\right]^{\frac{\sigma_{FF}}{\sigma_{FF}-1}} \qquad i = \text{oil, gas, oil products.} \quad \text{(A6)}$$

The “Armington” assumption makes domestic (*DOM*) and foreign (*IMP*) commodities imperfect substitutes in accounting for product heterogeneity:

$$Mi = \left[\alpha_{dom,i} DOM_i^{\frac{\sigma_{dom}-1}{\sigma_{dom}}} + \alpha_{imp,i} IMP_i^{\frac{\sigma_{dom}-1}{\sigma_{dom}}}\right]^{\frac{\sigma_{dom}}{\sigma_{dom}-1}} \quad \text{(A7)}$$

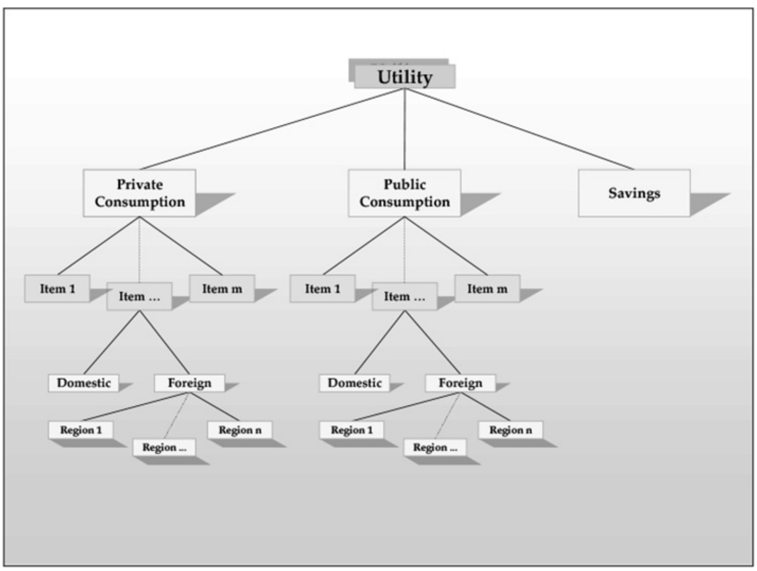
Imported commodities are a composite of commodity *i* from all source regions (*s*):

$$IMPi = \left[\Sigma_s \, O_{i,s} Y_{i,s}^{\frac{\sigma_{imp}-1}{\sigma_{imp}}}\right]^{\frac{\sigma_{imp}}{\sigma_{imp}-1}} \quad \text{(A8)}$$

Households' demand side structure

A representative consumer in each region receives income, defined as the service value of national primary factors (natural resources, land, labour, capital, see Figure A2). Capital and labour are perfectly mobile domestically but immobile internationally. Land and natural resources, on the other hand, are industry-specific. Income is used to finance three classes of expenditure: aggregate household consumption, public consumption and savings. The expenditure shares are generally fixed, which amounts to saying that the top-level utility function has a Cobb-Douglas specification.

Figure A2: Nested tree structure for final demand of the ICES model



The top-level demand system is described by a Cobb-Douglas utility function where the aggregate utility involves the per-capita utility from private and government consumption, and real savings:

$$U = CU_P^{w_p} U_G^{wG} U_S^{wS}, \quad \text{(A9)}$$

where U is the per-capita aggregate utility while U_P, U_G, and U_S are, respectively, the per-capita utility from private and government consumption, and real savings; whilst ω_i represent their distributional parameters. Public consumption is split in a series of alternative consumption items, again according to a Cobb-Douglas specification. However, almost all expenditure is actually concentrated in one specific industry: Non-market Services. Private consumption is analogously split in a series of alternative composite Armington aggregates.

However, the functional specification used at this level is the Constant Difference in Elasticities (CDE) form: a non-homothetic function, which is used to account for possible differences in income elasticities for the various consumption goods. The CDE demand system is characterized by an indirect utility function of the form:

$$1 = \Sigma_i B_i U^{Y_i R_i} \left(\frac{P_i}{X} \right)^{Y_i} \tag{A10}$$

with P_i, being price of commodity i, X the household expenditure, while B_i, y_i, and R_i are positive parameters.

Investment is internationally mobile: savings from all regions are pooled and then investment is allocated so as to achieve equality of expected rates of return to capital. In this way, savings and investments are equalized at the world, but not at the regional level. Because of accounting identities, any financial imbalance mirrors a trade deficit or surplus in each region.

Recursive dynamics: Capital and debt accumulation

The ICES model generates a sequence of static equilibria under myopic expectations linked by capital and international debt accumulation. Growth is driven by changes in primary resources (capital, labor, land and natural resources). Dynamics are endogenous for capital and exogenous for other primary factors. Capital accumulation is the outcome of the interaction of: i) investment allocation between regions and ii) debt accumulation. Savings are pooled by a world bank and allocated as regional investments according to:

$$\frac{I_r}{Y_r} = \phi_r \, exp\left[\rho_r \left(r_r - r_w\right)\right], \tag{A11}$$

where I_r is regional annual investment, Y_r is regional income, r_i is regional and world returns on capital. ϕ_r is a given parameter that represents the average propensity to save and ρ_r is a flexibility parameter related to investment supply sensitivity to return differentials. The rationale of equation (A12), follows the ABARE GTEM model (Pant, 2002). Capital stock accumulates over time in a standard relationship with a constant depreciation:

$$K_r^{t+1} = I_r^t + (1 - \delta) K_r^t. \tag{A12}$$

There is no equalization of regional investments and savings from equation (A12), so any excess of savings over investments equals the regional trade balance (*TB*). The stock of debt evolves by considering the trade balance as follows:

$$D_r^{t+1} = TB_r^t + D_r^t. \tag{A13}$$

Finally, foreign debt is serviced every period on the basis of the world interest rate r_w.

CO_2 emissions

The GTAP-E model uses average emission coefficients for each fossil fuel (Coal, Oil, Gas and Oil products) which are constant across sectors and regions of the world economy (Truong and Lee, 2003). We applied the same average emission coefficients in ICES to compute the corresponding emissions to the combustion or use of fossil fuels, but not their transformation as in the case of oil being refined and processed to obtain oil products. This means that the database we used provides information about emissions released to the atmosphere when a fossil fuel is burnt during the production process of a commodity or final consumption by households.

Table A3: Absolute mitigation effort s with respect to BAU in period 20 (Mtons of CO2)

Region	Pledge	High Pledges	High Pledges + Limited access to REDD credits in the ETS market				High Pledges
		Without REDD	25%	50%	75%	100%	unlimited REDD
Australia	-200	-206	-204	-202	-199	-197	-138
NewZealand	-26	-13	-13	-13	-12	-12	-7
China	-	-1910	-1911	-1913	-1915	-1917	-1943
Japan	-675	-345	-339	-333	-328	-322	-189
SouthKorea	-183	-233	-229	-226	-223	-220	-143
India	-	-130	-131	-131	-132	-132	-140
Canada	-216	-206	-202	-199	-196	-193	-117
USA	-2695	-3184	-3147	-3110	-3072	-3035	-2048
EU27	-2108	-1153	-1135	-1117	-1099	-1081	-671
Russia	-584	-848	-836	-824	-813	-801	-520
SouthAfrica	-121	-208	-206	-204	-202	-200	-142
NORICE	-40	-9	-9	-9	-9	-9	-5
EASIA	-135	-477	-470	-464	-457	-450	-292
LACA	-341	-444	-435	-426	-417	-409	-216
SSA	-	31	32	33	34	35	20
ROW	-	635	622	610	597	584	330
World	-	-8698	-8613	-8528	-8443	-8357	-6221

Table A4: Relative mitigation effort s with respect to BAU in period 20 (in percentage)

Region	Pledge	High Pledges	High Pledges + Limited access to REDD credits in the ETS market				High Pledges
		Without REDD	25%	50%	75%	100%	unlimited REDD
Australia	-43%	-45%	-44%	-44%	-43%	-43%	-30%
NewZealand	-60%	-30%	-30%	-29%	-29%	-28%	-17%
China	-	-22%	-22%	-22%	-22%	-22%	-22%
Japan	-46%	-23%	-23%	-23%	-22%	-22%	-13%
SouthKorea	-30%	-38%	-38%	-37%	-37%	-36%	-23%
India	-	-8%	-8%	-8%	-8%	-8%	-8%
Canada	-33%	-31%	-31%	-30%	-30%	-29%	-18%
USA	-36%	-43%	-42%	-42%	-41%	-41%	-28%
EU27	-43%	-23%	-23%	-23%	-22%	-22%	-14%
Russia	-26%	-38%	-38%	-37%	-37%	-36%	-23%
SouthAfrica	-32%	-55%	-54%	-54%	-53%	-53%	-37%
NORICE	-69%	-16%	-16%	-15%	-15%	-15%	-8%
EASIA	-9%	-31%	-31%	-30%	-30%	-29%	-19%
LACA	-19%	-24%	-24%	-23%	-23%	-22%	-12%
SSA	-	11%	11%	11%	12%	12%	7%
ROW	-	15%	14%	14%	14%	13%	8%
World	-	-24%	-23%	-23%	-23%	-23%	-17%

Table A5: Changes in setoral trade balances in period 20 (REDD vs NoREDD) (2001 US$ million)

	Australia	New Zealand	China	Japan	South Korea	India	Canada	USA	EU27	Russia	South Africa	NORICE	EASIA	LACA	SSA	ROW
Rice	-0.1	0.0	0.3	8.2	-0.2	0.3	0.0	-0.7	0.8	-0.3	0.0	0.0	-10.0	0.6	-1.4	1.7
Wheat	-8.5	0.0	4.7	1.2	-0.3	8.7	4.7	-3.3	-3.2	-2.2	1.7	0.3	0.8	-13.4	-14.2	24.5
Other Cereal	4.6	0.8	56.2	15.3	-7.6	39.1	14.4	28.0	176.6	-11.5	18.6	0.9	-53.1	4.9	-354.9	98.5
Vegetable Fruits	12.0	2.4	167.7	15.7	-3.3	57.7	20.8	5.8	82.7	-35.5	22.5	1.2	-86.5	-119.6	-208.0	98.3
Animals	-21.9	-4.6	85.4	3.6	-0.1	12.8	-2.2	-17.9	2.5	-1.7	4.3	0.9	-18.9	-12.5	-36.1	14.6
Forestry	0.9	3.6	94.8	7.4	-2.7	23.2	4.1	15.5	58.7	-9.2	1.6	0.5	29.4	-1.4	-209.7	18.9
Fishing	-0.7	-0.3	3.7	18.8	4.3	1.2	0.1	4.2	20.9	-3.7	1.4	6.8	-3.5	-4.0	-35.4	-2.0
Coal	64.2	0.9	65.1	-108.4	-47.9	-7.9	-3.4	40.5	-124.7	-8.9	30.8	0.2	10.3	-14.3	-2.0	49.1
Oil	-71.4	**-19**.8	-141.9	-1357.5	-753.2	-212.0	-26.4	-3181.4	-3000.5	853.2	-154.4	674.8	-612.1	442.8	928.0	6498.5
Gas	29.5	0.0	-8.5	-273.2	-77.4	0.0	296.3	-276.8	-785.5	172.4	0.0	122.0	122.5	-20.4	32.2	610.3
Oil Products	1.5	-13.3	221.8	25.8	46.0	-47.5	45.2	-327.2	168.6	287.6	160.4	17.5	-234.0	-34.7	-308.4	-39.0
Electricity	0.0	0.0	-36.6	0.0	0.0	-1.3	-91.0	165.7	749.4	332.4	118.5	-38.4	28.1	58.5	-169.1	-1116.2
Water	-0.4	0.0	0.2	0.3	0.9	0.1	-0.2	4.4	-0.1	-0.5	0.2	-0.2	3.7	-1.3	-3.6	-3.4
Energy Intensive industries	278.4	27.5	-1529.4	244.4	470.6	-201.5	335.1	1720.1	492.5	1379.0	317.7	-87.3	904.7	-636.9	-1197.0	-2372.6
Other industries	28.0	**-41**.2	1268.4	1601.0	567.2	467.3	-264.5	10373.9	-2.5	-254.2	118.7	-359.5	-828.1	-7600.7	-4770.3	-101.1
Market Services	12.0	-2.4	390.9	375.8	92.2	106.6	27.8	1957.4	28.3	-201.1	37.8	-251.2	471.4	-1314.5	-1597.2	-326.8
Non-Market Services	-1.0	-0.8	52.9	27.7	15.0	13.1	-10.5	520.1	9.7	-59.5	3.0	-22.5	31.8	-194.2	-322.9	-61.8

Table A6: REDD revenues as a share of GDP for the High pledges and unlimited REDD scenario (in percentage)

Region	Period										
	10	11	12	13	14	15	16	17	18	19	20
EASIA	0.0%	0.0%	0.1%	0.2%	0.2%	0.4%	0.5%	0.6%	0.7%	0.9%	1.0%
LACA	0.0%	0.1%	0.1%	0.3%	0.4%	0.6%	0.8%	1.0%	1.2%	1.5%	2.0%
Sub Saharan Africa	0.1%	0.2%	0.4%	0.6%	0.8%	0.9%	1.1%	1.4%	1.6%	1.9%	1.9%

References

Bosello, F.; Zhang, J. (2006) Gli effetti del cambiamento climatico in agricoltura. *Questione Agraria*, 1-2006, 97-124.

Bosello, F.; Roson, R., Tol, R.S.J. (2006) Economy wide estimates of the implications of climate change: human health. *Ecological Economics*, 58, 579-591.

Bosello, F.; Roson, R.; Tol, R.S.J. (2007) Economy wide estimates of the implications of climate change: sea level rise. *Environmental and Resource Economics*, 37, 549-571.

Bosello, F.; Roson, R.; Tol, R.S.J. (2008) Economy wide estimates of the implications of climate change: human health: A rejoinder. *Ecological Economics*, 66, 14-15.

Bosetti, V., Carraro, C.; Galeotti, M.; Massetti, E.; Tavoni, M. (2006) WITCH: A World Induced Technical Change Hybrid Model. [Special Issue. Hybrid Modeling of Energy-Environment Policies: Reconciling Bottom-up and Top-down]. *Energy Journal*, 13-38.

Eboli, F., Parrado, R.; Roson, R. (2010) Climate-change feedback on economic growth: explorations with a dynamic general equilibrium model. *Environment and Development Economics*, *15*(5), 515-533.

EIA-Energy Information Administration (2007) Annual Energy Outlook 2007, with projections to 2030. *Report DOE/EIA-0383(2007)*. Retrieved from www.eia.doe.gov/oiaf/aeo/

EIA-Energy Information Administration. (2009) International Energy Outlook 2009. *Report no. DOE/EIA-0484*. Retrieved from http://www.eia.doe.gov/oiaf/ieo/index.html

Hertel, T.W. (1997) *Global Trade Analysis: Modeling and Applications*. Cambridge: Cambridge University Press.

Pant, H. (2002) Global Trade and Environment Model (GTEM): A Computable General Equilibrium Model of the Global Economy and Environment, Canberra: Australian Bureau of Agricultural and Resource Economics.

Truong, P. T.; Huey-Lin, L. (2003) GTAP-E Model and the 'new' CO_2 Emissions Data in the GTAP/EPA Integrated Data Base – Some Comparative Results, *GTAP Resource #1296*.

UNPD (2008) World population prospects. *The 2008 revision*. Retrieved form : http://esa.un.org/unpp/index.asp

R. Quentin Grafton
Crawford School of Public Policy,
The Australian National University, Australia
quentin.grafton@anu.edu.au
orcid.org/0000-0002-0048-9083
Safa Fanaian
Crawford School of Public Policy,
The Australian National University, Australia
Safa.Fanaian@anu.edu.au

DOI: https://doi.org/10.14195/2183-203X_56_3

Responding to the Global Challenges of 'Too Much, Too Little and Too Dirty' Water: Towards a Safer and More Just Water Future

Respondendo aos Desafios Globais de ''Too Much, Too Little and Too Dirty Water': Rumo a um Futuro Mais Seguro e Justo

R. Quentin Grafton
Safa Fanaian

Received for publication: May 31, 2023
Revision accepted for publication: June 30, 2023

ABSTRACT
The world water crisis is manifest through 'Too Much, Too Little and Too Dirty' water at multiple scales from the local to the global. Understanding the key drivers and consequences of this water crisis, and who bears the biggest costs, is necessary to develop appropriate responses, at scale and over time. Using four framings: one, water stocks and limits; two, water rights and responsibilities; three, water values and prices; and four, green and grey water infrastructure, we review the challenges and possible responses. Using a water justice lens, we highlight the transitional and transformational pathways towards a safer and more just water future.
Keywords: World water crisis; justice; floods; droughts; WASH; infrastructure.

JEL Classification: Q25; Q57; Q58.

Acknowledgement: This research was funded, in part, by the Australian Research Council grant FL190100164: 'Water Justice: Indigenous Water Valuation and Resilient Decision-making', the 'Hilda John Endowment' of the Australian National University and a subsidy from the Dutch Enterprise Agency. The authors also gratefully acknowledge insightful discussions with Joyeeta Gupta, Aromar Revi, Jagdish Krishnaswamy, Hilmer Bosh, Neha Sami, and Julia Schaef.

> *"We shall overcome because the arc of the moral universe is long, but it bends toward justice."*
>
> Dr. Rev. Martin Luther King Jnr, 31 March 1968

1. The World Water Crisis

In this review, we describe the world water crisis in its multiple dimensions and the consequences that manifest themselves as 'Too Much, Too Little, and Too Dirty' water (Chen, 2018; Fanaian, 2022). Too Much water is primarily associated with flooding events that expose at least 20 percent of humanity to flood risks (Tellman et al., 2021). In coastal areas, Too Much water from storm surges exacerbates saline intrusion associated with sea-level rise (Mohammed and Scholz, 2018). Too Little water is primarily about hydrological droughts that arise from both meteorological and human actions, such as excessive water withdrawals (Agha Kouchak, 2021). Too Little water also includes the limited water access of billions of people due to exclusion from formal piped water systems and/or from the high economic costs of access to safe water supplies (Rusca and Cleaver, 2022). Too Dirty water is about water pollution; most visible with inadequate Water Sanitation and Hygiene (WASH) for many vulnerable communities in the Global South[1] (Dados and Connell, 2012). All three dimensions will worsen with climate change (Flörke et al., 2018; IPCC, 2022; Pokhrel et al., 2021; Satoh, 2022).

Collectively, floods and droughts increase mortality and morbidity, contribute to declines in ecosystem services, create food price spikes, displace people, damage infrastructure, reduce economic activity and contribute to conflicts (The World Bank, 2016). Too much water is not just a result of excess precipitation but is caused by land-use planning that unnecessarily exposes people to flood risks, inadequate or improper infrastructure that transfers, and may magnify, downstream and coastal flooding risks, and the degradation of green infrastructure (e.g., wetlands loss, deforestation, etc.) that would otherwise mitigate flood events (WMO, 2021). Too Little water arises from hydrological droughts, defined as low water availability that can arise from multiple factors including reduced precipitation and excessive water withdrawals (Grafton et al., 2022a; Mukherjee et al., 2018). Hydrological droughts can be particularly devastating, especially if they are multi-year phenomena, and have multiple, and sometimes persistent, negative health and economic impacts, especially on poor and vulnerable communities (Damania et al., 2018). An historical review of global droughts indicates that the severity of hydrological droughts that impose costs on agriculture and ecosystem services is increasing (Vincente-Serrano et al. 2022). Too Dirty water means that globally some 2 billion people are forced to drink unsafe water which has a disproportionate negative impact on both children and women (WHO 2019, 2021; WHO and UNICEF, 2022). Failing to deliver safe water and sanitation causes premature deaths, globally, of about one

[1] "The term Global South functions as more than a metaphor for underdevelopment. It references an entire history of colonialism, neo-imperialism, and differential economic and social change through which large inequalities in living standards, life expectancy, and access to resources are maintained" (Dados and Connell, 2012, p. 13).

million people per year and widespread morbidities associated from water borne diseases and parasites (e.g., cholera, dysentery, schistosomiasis, etc.).

In section 2, we describe the world water crisis in its three critical dimensions (Too Much, Too Little and Too Dirty water) and present some consequences at both a global and regional level. Our regional focus includes seven countries (Australia, China, France, India, Nigeria, South Africa, and the United States of America) across five continents in relation to Too Much and Too Little water. In section 3, and with a water justice lens, we present four framings to better understand the world water crisis: one, water flows and limits; two, water rights and responsibilities; three, water values and prices; and four, green and grey water infrastructure. In section 4, we highlight possible transitional and transformational pathways to mitigate the world water crisis. We offer our conclusions in section 5.

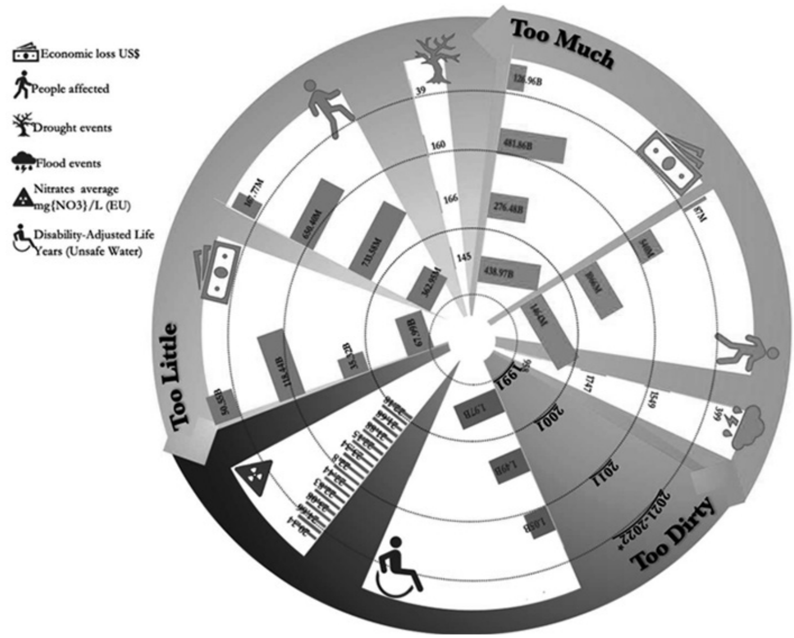
2. **TOO MUCH, TOO LITTLE AND TOO DIRTY WATER**

2.1. Too much water

Figure 1 shows an increasing number of flooding events over time; this is consistent with a growing intensity of rainfall events associated with climate change (IPCC, 2022). For example, for the two years 2021 and 2022 there were more than a quarter of the flood events of the previous decade. The number of floods, however, is not necessarily commensurate with the intensity of the flood events, as measured by economic costs. Globally, between 2001 and 2010, there were some large-scale 1,700 flood events that generated total damages of US$276 Billion (adjusted for inflation). By comparison, between 2011 and 2020 there were some 1,500 flood events with reported damages of US$481 billion.

Despite an increase in flood events over time, human adaptation (Jongman, 2018; Islam et al. 2018) in the form of flood warning systems, flood protection infrastructure, flood risk land-use planning, and nature-based solutions has resulted in a global decline in the reported global number of people impacted by floods in both high and low-income per capita countries (Figure 1). In India and China, the reported number of those affected by flooding in 2010-2020 was less than a third of what it was in 1991-2000 (Figure 2). Africa, however, is not experiencing a downward trend in the numbers affected by floods. Further, in some locations, the consequence of flooding events appears to be increasing. For example, in Australia, flooding events in 2021-22 alone affected more people than in the previous two decades (2000-2020).

Figure 1: The world water crisis: Too Much, Too Little and Too Dirty Water



Source: Authors; for data sources and detailed notes see Appendix.

2.2. Too little water

Due to a changing climate, including increases in atmospheric evaporative demand, extended hydrological droughts are intensifying in the 21st Century (Haile et al., 2020; Vincente-Serrano et al., 2022). Some countries, such as China (Figure 2), have reduced the economic costs and number of people affected by hydrological droughts. In the case of China, adaptation to hydrological droughts has included huge infrastructure investments, especially in large inter-basin water transfers (Sun et al., 2021).

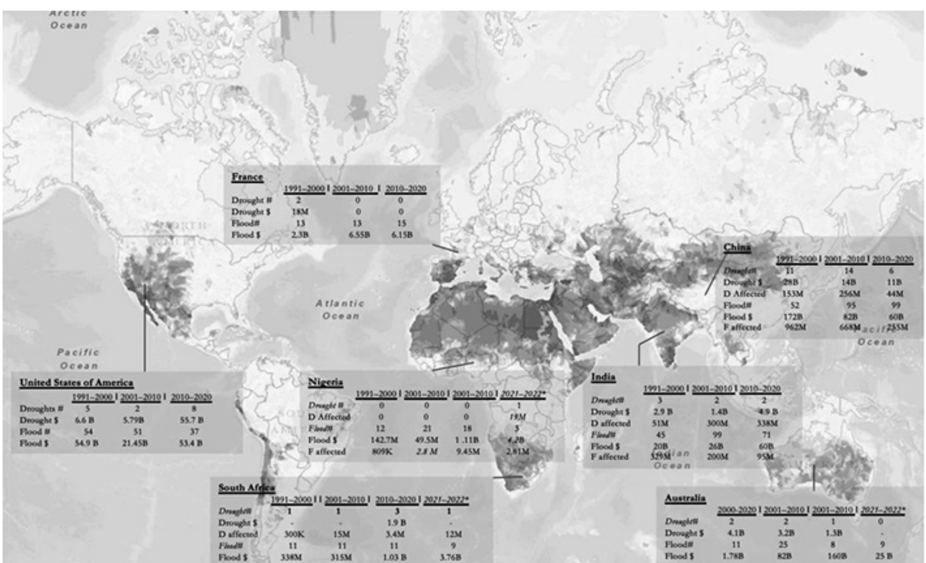
By contrast, the number of people affected by hydrological droughts in India has got worse, not better, increasing by more than two-thirds between the decades 1991-2000 and 2010-2020. In part, this is a result of population growth with some countries, such as Pakistan, experiencing large (80 percent) declines in water availability per capita (World Bank, 2023). Climate change may also mean that countries which have not historically been subject to extended hydrological droughts, such as Nigeria (Shiru, 2020), are particularly vulnerable because of limited experience in adapting to less water.

2.3. Too dirty water

Increasing pollution in rivers, lakes, wetlands, and groundwater has multiple and negative consequences on human and ecosystems health. A global study on the burden of disease (IHME, 2020) shows that unsafe water sources led to as many as 1.7 million deaths in 2017 and caused disabilities (Disability-adjusted life years) for more than 87 million. By comparison, in 2019 the global annual water-related mortality due to unsafe water source was three times larger than the world's deaths due to homicide (IHME, 2020).

The sources of water pollution are diverse and include domestic waste and pollution from agriculture and industry (Mekonnen and Hoekstra, 2018). While there is an extensive network to capture data on water availability, much less water quality data are available (only 37 countries report a broad range of water quality measures to the United Nations) with especially sparse reporting from the Global South (Damania et al., 2019; Grafton et al., 2023a). For one key measure, the reported median level of nitrates in groundwater, the trend is getting worse, not better, in the European Union. Importantly, without regular and widespread water quality reporting, and not just for drinking water, it will be impossible to identify the direct sources of water pollution and/or to measure the progress of mitigating actions.

Figure 2: Global map of water scarcity with Too Much and Too Little Water for Seven Countries (Australia, China, France, India, Nigeria, South Africa, and the United States of America)



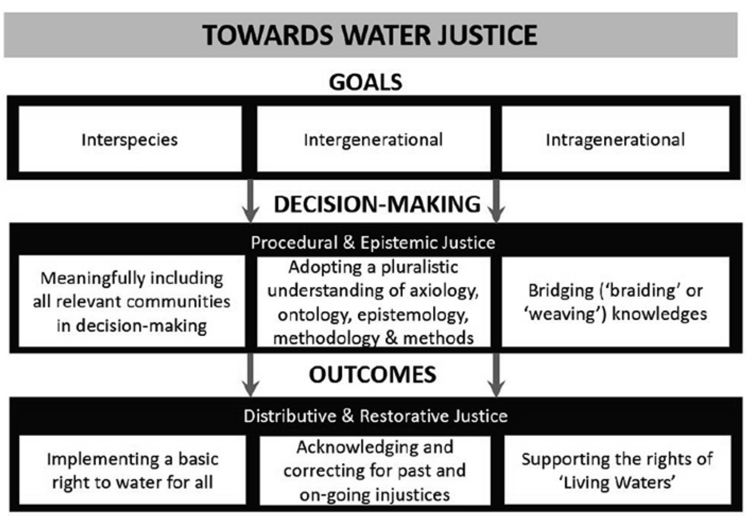
Source: The Authors; for data sources and detailed notes see Appendix.

3. Four Framings of the World Water Crisis

There are multiple ways to describe the world water crisis that include perspectives on; environmental (Gupta et al., 2023; Gupta and Lebel, 2010) and water justice (Grafton et al., 2022; Savelli et al., 2023; Zwarteveen and Boelens, 2014), WASH (WHO and UNICEF, 2022), ecosystem sustainability (Green et al., 2015; Pastor et al., 2022; Vörösmarty et al., 2010), water withdrawals (Rodell et al., 2018; Scanlon et al., 2023; Yao et al., 2023), water scarcity (Dalstein and Naqvi, 2021; Distefano and Kelly, 2017; Kummu et al., 2010; Mekonnen and Hoekstra, 2016), water insecurity (Garrick and Hahn, 2021; Grafton, 2017), food and water insecurity (Rosegrant et al., 2009), water governance (Fanaian and Fanaian, 2023; Grafton et al., 2013; OECD, 2018), planetary tipping points (Lenton and Williams, 2013) and boundaries (Wang-Erlandson et al., 2022), among others (Grafton et al., 2023b).

Connecting all these perspectives on the world water crisis is water justice (Figure 3). At a minimum, water justice requires: one, everyone's basic water needs are met; two, procedural justice such that all those materially affected by water decisions have a respected 'voice' at the table; three, substantive justice such that actions are taken to correct for past and continuing water injustice (Grafton et al., 2022b; Gupta et al., 2023; Syme et al., 1999); four, epistemic justice such that decision-makers value and respect all knowledges and experiences (Mehltretter et al., 2023); and, five, justice for 'living waters' that goes beyond an exclusive anthropogenic and/or utilitarian view of water (Bates et al., 2023; Mcgregor et al., 2020). These five underpinnings of water justice are consistent with the three I's of Earth System Justice; Interspecies, Intergenerational and Intragenerational equity (Gupta et al., 2023).

Figure 3: Towards water justice



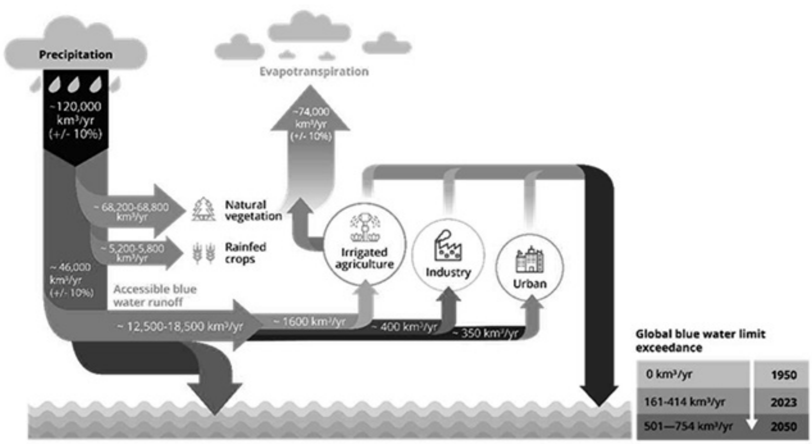
Source: Authors, adapted from or inspired by Bates et al. (2023); Gupta et al. (2023, Figure 2); McGregor, (2018); and Mehltretter et al. (2023).

Including water justice as a connecting theme, we present four framings to better understand the causes, consequences and possible actions required to respond to the world water crisis.

3.1. Water flows and limits

Fresh water availability and accessibility, including both surface and groundwater, have had an enormous impact on human social, cultural, and economic development. People have, over millennia (Hosseiny et al., 2021), developed successful strategies to mitigate against water scarcity and water variability (Hall et al., 2014), such as building or enhancing water storages, water transfers and, more recently, desalination. Nevertheless, local and regional social, and economic development progress remains closely tied to both the quantity and quality of freshwater available for household use, and the production of food and fibre, especially for irrigated agriculture.

Figure 4: The water cycle, global water consumption by sector and blue water consumption exceedance

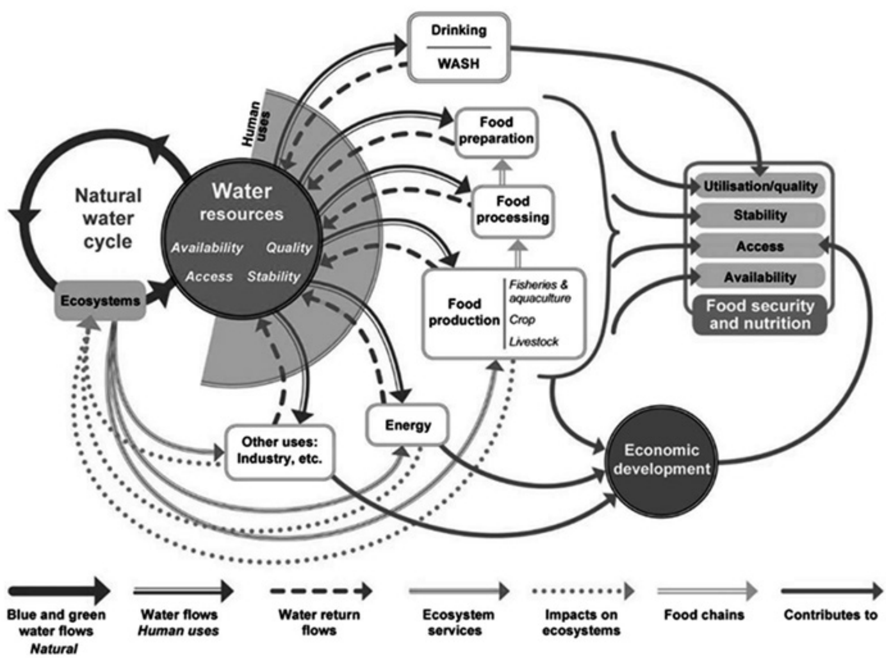


Source: Grafton, Krishnaswamy and Revi (2023).

Much of the global terrestrial freshwater flow is consumed via evapotranspiration from natural vegetation (more than 50% of annual precipitation); rain-fed crops consume about 5% of the total terrestrial precipitation. Accessible runoff, the water available in accessible streams and rivers, represents about 10-15% of the total land precipitation. Of this total runoff, irrigated agriculture accounts for over 80% of human water consumption, via evapotranspiration (see Figure 3), and produces about 30-40% of the world's food (Rosegrant et al., 2009). About one half of the agricultural production from irrigation is associated with unsustainable water consumption (Rosa et al., 2019). Further, about a quarter of the world's food is traded which means that unsustainable water consumption in irrigated agriculture poses systemic risks for global food security (D'Ordorico et al., 2014).

Systemic risks between water and food (Figure 4), and between food and energy because of the intensity of fossil use in intensive agriculture (Rosa et al., 2021), are increasing. This is because: one, the global food trade has increased by more than one half since the mid-1980s, two, the calories from food trade per volume of water withdrawn has declined (D'Ordorico et al., 2014), and three, projected declines in food availability to 2050 and 2100 under multiple climate change scenarios from increased water stress and heat stress (Kompas et al., 2023). Multiple and important connections exist between water resources, water use and water consumption, and food security (Figure 5). Water insecurity in terms of gaps in availability, access, stability, and quality is a key contributor to food insecurity via constraints on food production and of inadequate WASH services.

Figure 5: Food and water interconnections



Source: HLPE (2015, Figure 1).

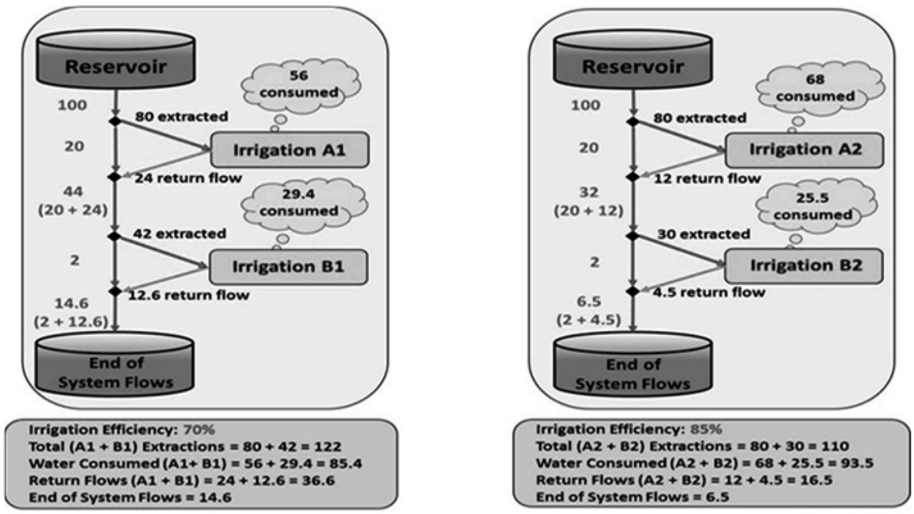
Human blue water withdrawals and consumption (evapotranspiration) account, respectively, for about 35 percent (Postel et al., 1996) and 20 percent (see Figure 4) of the accessible annual water run-off. A key challenge with the world water crisis is that, at a global level, is the annual rate of blue water (water in rivers, lakes, groundwater, and human-made water storages) consumption exceeds the sustainable limit. This exceedance is expected to double by 2050 under business as usual (Grafton, Krishnaswamy and Revi, 2023).

The proximate cause of the exceedance of blue water consumption limits are twofold. First, groundwater depletion, especially in arid and semi-arid locations, that arises from rates of water withdrawal that exceed aquifer recharge (Wada et al., 2012; Yao et al., 2023). Second, excessive surface water withdrawals that have reduced stream flows in streams and rivers below minimum environmental flows (Richter et al., 2012), and this projected to get worse to 2050 (Zhang et al., 2023), degrade ecosystem services (Poff and Zimmerman, 2010). The rates of blue water exceedance are, typically, the greatest in arid and semi-arid areas that have high population densities such as Northern India and Northern China. Overall, the world's current blue water consumption exceeds the sustainable level of blue water consumption, is increasing and, with business as usual, could be twice as large by 2050 (Figure 4).

A common response to blue water exceedance and increasing water scarcity has been to subsidise and/or promote increase in water-use efficiency (HLPW, 2018). In the case of irrigation, water-use efficiency is defined as the ratio of the water consumed in beneficial plant growth to the total water withdrawals measured at either the field, farm, or catchment scale (Figure 6) and is known as irrigation efficiency. While increasing irrigation efficiency benefits irrigators by increasing the returns from any additional volume of water that is withdrawn, this typically reduces the blue water that would otherwise have returned to groundwater and streams and rivers, known as return flows (Willardson et al., 1994). The paradox of irrigation efficiency is that increasing water-use efficiency will, typically, not increase water availability for other purposes, such as for environmental flows (Grafton et al., 2018), and frequently reduces return flows and end-of-system flows (Figure 6), both of which can generate large economic benefits and support water justice (Owens et al., 2022).

Instead of subsidising increases in irrigation efficiency, water accounting complemented by water consumption caps, are much more likely to control anthropogenic blue water consumption (Grafton et al., 2023a). To ensure global food sufficiency from an increased global population and lower growth (or no growth) in yields due to climate change (Grafton et al., 2017; Kompas et al., 2023), there is also a need to substitute unsustainable water withdrawals in irrigated agriculture with green water (soil moisture available from plant growth) for rain-fed agricultural food production (Rosa et al., 2020).

Figure 6: Irrigation efficiency, return flows and end-of-system flows



Source: Perry et al. (2023, Figure 2).

3.2. Water rights and responsibilities

Rights to access, use, consume water and then to dispose of wastewater determine the 'who gets what' of water. Safe drinking water and sanitation are considered a basic human right consistent with Resolution 64/292 of the UN General Assembly. Delivering this right requires much larger than current investments in grey (human built) or green (nature-based) infrastructure and the delivery of affordable (Al-Ghuraiz and Enshassi, 2005) basic water services to the poor (Tortajada and Biswas, 2017).

Both a lack of safe access, especially in rural and urban areas of the Global South, and affordability, explain why some 2 billion people lack access to safely managed drinking water services (WHO, 2021) and some 3.6 billion lack access to improved sanitation services (WHO and UNICEF, 2022). Beyond a right to basic water services and investment in water services, countries need regulatory frameworks to allocate and to reallocate water among sectors (e.g., industry, agriculture, household), and across individual water users to ensure just outcomes. Without proper consideration of 'winners and losers' from water infrastructure investments, such as for large dams, water injustices can, and have, been exacerbated (Blake and Barney, 2021; Duflo and Pande, 2007). Attention must also be given to the scale, distribution and diversity of infrastructure, and their ownership, management, and control (Schwartz et al. 2018; Fanaian and Fanaian, 2023).

Typically, water justice is not prioritised when reallocating water across time and place. Importantly, the responsibility to deliver the basic human right to water and water justice is not only a moral obligation but is closely connected to sustainability of ecosystem services that affects both the rich and poor (Gupta et al., 2023; Rammelt et al., 2023). This means that for those with well-defined water rights and services, there is a responsibility to act to ensure that those who do not have their basic water needs met will, ultimately, achieve this basic human right.

The provision of rights to water must pay special attention to those who have been dispossessed of their rights, including Indigenous peoples' rights (Jackson, 2018) recognised in the United Nations Declaration of the Rights Indigenous Peoples (UN General Assembly 2007), known as UNDRIP. Exercising Indigenous rights as part of UNDRIP should encompass the EAUX principles (Mehltretter et al., 2023) of: Equity (honoring Indigenous Peoples' sovereignty), Access (recognising and affirming Indigenous rights), Usability (benefits Indigenous peoples), and eXchange (on-going flow of information among diverse groups for mutual understanding).

Water rights are increasingly being traded and water markets are expanding in several countries, ostensibly to overcome water insecurity (Wheeler, 2021). Without careful design and regulatory oversight, however, water markets will not deliver efficiency, equity, or sustainability (Grafton, Horne and Wheeler, 2022). That is, there must be, at a minimum, water accounting (Vardon et al., 2023) about 'who gets what and when' and rules about 'how' water is used and consumed to mitigate the external costs imposed on others from any given water use. Where there are water rights and water markets, there must also be: one, responsibilities in relation to fairness in the initial allocation of water rights and, two, complementary regulations and market rules to ensure water withdrawals and consumption are sustainable and do not impose unacceptable costs on those without water rights and the environment.

3.3. Water values and prices

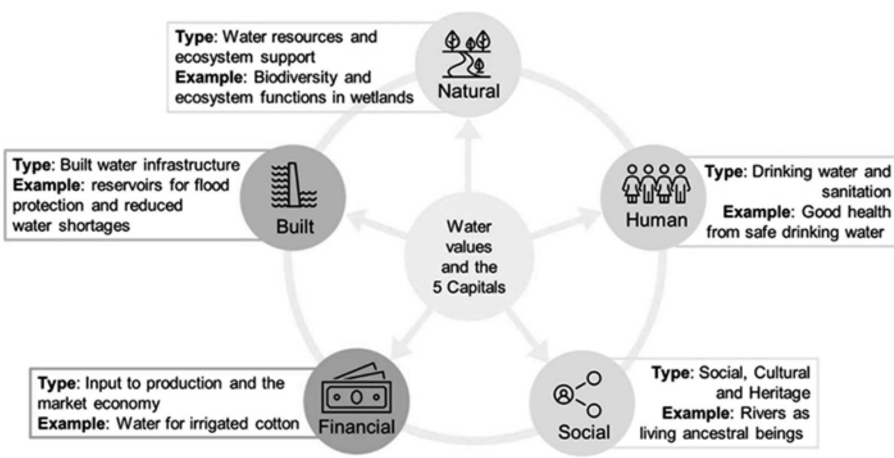
The value of water is the benefit (direct and indirect) to users from access, use and/or consumption of a given volume of water at a particular place and time. By contrast, the price of water is the amount paid (typically in monetary units) by a user (individual, household, community, business, etc.) for a given volume of water of perceived quality at a particular place and time (Grafton et al., 2023c). Thus, while price and value are related (e.g., the higher the value of water the higher price that a user is willing to pay for water) they are not the same.

Water prices that adjust to changes in water availability can provide incentives to conserve water when there is less available (Grafton et al., 2011). The challenge in incentivising water conservation from higher water prices is that, typically, the poor already suffer from inadequate access to safe drinking water and sanitation and frequently pay the highest volumetric price for water (Kariuki and Schwartz, 2005). This is because many poor are not connected to safely managed water supply systems and are forced to rely on private water vendors or collect water themselves. Thus, many of the poor in the Global South are not beneficiaries of water subsidies that primarily go to those with access to centralised water distribution systems. Consequently, when deciding on financial allocations to deliver WASH goals on the basis of water justice, subsidies need to be based on need (Whittington et al., 2015) rather than be determined by those with preferred access to existing water infrastructure (Andrés et al., 2021).

Water pricing also includes pricing 'bads' either directly through pollution charges and fines for non-compliance or indirectly through regulations (Olmstead, 2010). Whatever the pricing approach, active price intervention should include incentives for polluters to reduce pollution, investments to reduce discharges and/or to treat discharges, and appropriate monitoring compliance and enforcement. If fines for pollution are established without complementary public policy interventions and diligent monitoring, there will continue to be large and negative impacts on both people and the environment from poor water quality (Damania et al., 2019).

A comprehensive review of water values (United Nations, 2021) connects water to the major types of human and nature capital. These five capitals include: (1) built infrastructure (e.g., dams); (2) natural infrastructure (e.g., wetlands); (3) human (e.g., public health); (4) cultural (e.g., sacred rivers); and (5) financial (e.g., market benefits from industrial water use). As shown in Figure 7, supporting water values is not simply about investing in built or grey infrastructure. Instead, it requires a comprehensive response to the world water crisis that embraces the values included in human, nature, and cultural capital.

Figure 7: Water values and capital stocks



Source: Grafton et al. (2023c, Figure 2).

A key challenge is that many of the values prioritised in water decision-making are market values, such as the value of water as an input into a production process. This exclusive market and financial focus mean that many uses, including in-situ (e.g., stream flows) uses of freshwater, that may have high non-market values (e.g., wetland's ecosystem services), are frequently treated as having a zero value because they are neither monetised nor easily measured (Manero et al., 2021).

3.4. Green and grey water infrastructure

The natural environment provides, at no charge, huge and multiple benefits (Costanza et al., 1997) for biodiversity, climate change mitigation and adaptation and many non-market values. Green water infrastructure supports groundwater recharge, reduces storm runoff, and promotes higher water quality, among other benefits. These benefits are very large; conserving nature for water is estimated to be worth some USD 3 trillion by 2050 in terms of avoided replacement costs for human-made water infrastructure (Vörösmarty et al., 2021). In the case of New York City, for example, conserving its water source catchments resulted in avoided grey infrastructure capital costs, associated with water filtration plants, of at least USD 6 billion (Chichilnisky and Heal, 1998).

The grey infrastructure investments needed to achieve SDG 6 Targets are very large, in the order of USD1.5 trillion annually. Many of these grey infrastructure investments need to be spent in the Global South (United Nations, 2021) on WASH, flood control and hydropower, among other needs (Figure 8). To some extent, grey infrastructure can be substituted by conserving key aspects of nature such as wetlands and forests. Depending on the context, green infrastructure investments can effectively respond to Too Much (e.g., mangroves protect from storm surges), Too Little (e.g., wetlands can provide natural water storages and increase availability in periods of low inflows) and Too Dirty water (e.g., protected watersheds provide better quality water).

Figure 8: Grey and green infrastructure

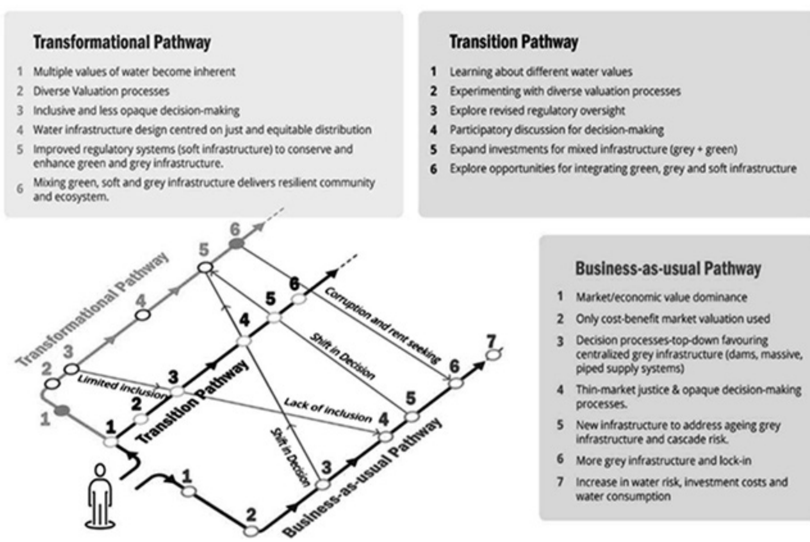
SERVICE	GREY INFRASTRUCTURE COMPONENTS	EXAMPLES OF GREEN INFRASTRUCTURE COMPONENTS AND THEIR FUNCTION
Water supply and sanitation	Reservoirs, treatment plants, pipe network	**Watersheds:** Improve source water quality and thereby reduce treatment requirements
		Wetlands: Filter wastewater effluent and thereby reduce wastewater treatment requirements
Hydropower	Reservoirs and power plants	**Watersheds:** Reduce sediment inflows and extend life of reservoirs and power plants
Coastal flood protection	Embankments, groynes, sluice gates	**Mangrove forests:** Decrease wave energy and storm surges and thereby reduce embankment requirements
Urban flood management	Storm drains, pumps, outfalls	**Urban flood retention areas:** Store stormwater and thereby reduce drain and pump requirements
River flood management	Embankments, sluice gates, pump stations	**River floodplains:** Store flood waters and thereby reduce embankment requirements
Agriculture irrigation and drainage	Barrages/dams, irrigation and drainage canals	**Agricultural soils:** Increase soil water storage capacity and reduce irrigation requirements

Source: Browder et al. (2019, p. 5).

4. Transitional and Transformational Pathways

Multiple actions are required to respond to Too Much, Too Little and Too Dirty water from a local to global scale. The specific actions, especially their prioritisation and sequencing, must be context specific and adapted to local circumstances. Here, we highlight just four, among the many actions, needed to effectively respond to the world water crisis. These actions include: one, valuing water (United Nations 2021), including non-market water values of all peoples, and including these values in decision-making; two, effectively responding to unequal power relationships (Molle et al., 2009; Tetrault and McCuligh, 2018; Wade, 1982) that contribute to rent-seeking behaviour and regulatory capture (Grafton and Williams, 2020) and prevent water being reallocated for sustainability and justice; three, improved water governance in the form of planning and regulation that delivers transformative change, includes water pricing, water accounting, water consumption limits, land-use planning, etc. (OECD, 2010), and avoids attributing much or all the blame for water scarcity on climate change (Grafton et al., 2022a; Muller, 2018); and four, much greater finance for both grey and green infrastructure which prioritises the basic human right to water for all (Tortajada and Biswas, 2017) and the sustainability of key ecosystem services (Green et al., 2015; Vörösmarty et al., 2010).

Figure 9: Pathways towards a safer and more just water future



Source: Grafton et al. (2023b, Figure 3.8).

Sustainable pathways represent a shift from 'business as usual' decision-making that has contributed to the world water crisis and inhibited meaningful transformations (Figure 9). Transitional and transformational pathways towards a safer and more just water future require appropriate and measurable goals that encompass secure food systems, ecosystem health, public health, sustainable cities, innovation, among others (Grafton et al., 2023b).

Transformational pathways require 'positive tipping points' whereby relatively small interventions and actions eventually have large impacts (Lenton et al., 2022). Positive tipping points require enabling conditions that connect socio-economic-ecological systems to create change from the local to the global. We highlight just two key elements to enable improved water governance: one, participatory decision-making that meaningfully includes all affected stakeholders, and draws from and builds upon broad-based inclusive knowledges (Mehltretter et al., 2023); and, two, the inclusion of risk and system-based thinking (Sterman, 2002) into decision-making at all levels, especially the evaluation and mitigation of systemic risks in the food, energy, environment and water nexus (Katic and Grafton, 2023).

5. Conclusions

The world faces critical choices about how to respond to three, global and inter-related crises of biodiversity loss, climate change, and the water crisis. Despite progress towards Sustainable Development Goals (SDGs) by 2030, the world is not on track to deliver on the SDG water targets or to achieve a safer and more just water future.

In terms of the 'glass half full', over the past few decades considerable progress has been made on delivering improved WASH services, including in the Global South. Communities, and some national governments, have substantially reduced the number of their citizens who are subject to severe flooding events and hydrological droughts. In the Global North, and some countries in the Global South, improvements have been made in some measures of water quality. These successes, however, are not universal and have required large infrastructure investments complemented by substantial improvements in how water is governed and how water is (re)allocated.

In terms of the 'glass half empty', billions of people remain without access to the basic human right to water. Much of the economic growth of the past few decades has been at the expense of natural capital that provides key environmental services and on which many poor are reliant for their survival. Despite an increasing recognition of systemic risks, there has been little practical action to mitigate the risks of water insecurity for food security. Nor is mitigation of greenhouse gas emissions currently sufficient to avoid what will likely be catastrophic climate change in the decades to come and that will be manifest through Too Much, Too Little and Too Dirty water.

A much greater and more co-ordinated set of actions, and at all scales, is needed to mitigate the world water crisis. Multiple and context-specific responses are required that include, but are not limited to: one, prioritising and investing in delivering the basic right to water for all; two, financing investments and establishing planning, regulations and incentives to reduce the impacts of flooding and hydrological droughts; three, monitoring and reducing water pollution via vigilant regulation and the pricing of 'bads'; four, water accounting, regulations and incentives to cap blue water consumption where it is unsustainable; and five, pro-active conservation of natural capital (e.g., wetlands), human and social capital that are critical to a sustainable and just water future.

Establishing transitional and transformational pathways for water is a huge global challenge but is not insurmountable. Both local successes and failures can be adapted noting that almost all transformations begin small before they 'take off'. Importantly, actions by those who benefit the most from the status quo must have an effective response or change will be slowed or stopped. To effect the change needed, a convincing narrative of how transformational change can be implemented beyond grey infrastructure, and a greater awareness of the risks to ecosystems and food security of business as usual, are urgently required.

References

AghaKouchak, A.;Mirchi, A, Madani, K.; Di Baldassarre, G.; Nazemi, A.; Alborzi, A.; Anjileli, H.; Azarderakhsh, M.; Chiang, F.; Hassanzadeh, E.; Huning, L.S.; Mallakpour, I; Martinez, A.; Mazdiyasni, O.; Moftakhari, H.; Norouzi, H.; Sadegh, M.; Sadeqi, D.; Van Loon, A.F.; Wanders, N. (2021) Anthropogenic Drought: Definition, Challenges, and Opportunities. *Reviews of Geophysics*, *59,* e2019RG000683.

Al-Ghuraiz, Y.; Enshassi, A. (2005) Ability and willingness to pay for water supply service in the Gaza Strip. *Building and Environment*, *40*, 1093-1102.

Andrés, L.A.; Saltiel, G.; Misra, S.; Joseph, G.; Lombana Cordoba, C.; Thibert, M.; Fenwick, C. (2021) *Troubled Tariffs: Revisiting Water Pricing for Affordable and Sustainable Water Services Issue*. World Bank.

Barbier, E.B. (2022) The economics of of managing water crises. *Philosophical Transactions of the Royal Society A*, 380/2238.

Bates, W.B.; Chu, L.; Claire, H.; Colloff, M.J.; Cotton, R.; Davies, R.; Larsen, L.; Loughrey, G.; Manero, A.; Marshall, V.; Martin, S.; Nguyen, N-M.; Nikolakis, W.; Poelina, A., Schulz, D.; Taylor, K.S.; Williams, J.; Wyrwoll, P.; Grafton, R.Q. (2023) A Tale of Two Rivers – Baaka and Martuwarra, Australia: Shared Voices and Art towards Water Justice. *The Anthropocene Review* (in press).

Blake, D.J.H.; Barney, K. (2021) Impounded rivers, compounded injustice: contesting the social impacts of hydraulic development in Laos. *International Journal of Water Resources Development,* 38(1), 1-22.

Browder, G.; Ozment, S.; Rehberger Bescos, I.; Gartner, T.; Lange, G-M. (2019) *Integrating Green and Gray: Creating Next Generation Infrastructure*. Washington, DC: World Bank and World Resources Institute. http://hdl.handle.net/10986/31430

Chen, G. (2018) Tackling a crisis of too much, too little, too polluted. The Water Blog of the World Bank. Accessed 23 June 203 at: https://blogs.worldbank.org/water/tackling-crisis-too-much-too-little-too-polluted

Chichilnisky, G.; Heal, G. (1998) Economic returns from the biosphere. *Nature, 391,* 629-630.

Costanza, R.; d'Arge, R.; de Groot, R.; Farber, S.; Grasso, M.; Hannon, B.; Limburg, K.; Naeem, S.; O'Neill, R.V.; Paruelo, J.; Raskin, R.G.; Sutton, P.; van den Belt, M. (1997) The value of the world's ecosystem services and natural capital. *Nature, 387*, 253-260.

Dados, N.; Connell, R. (2012) The Global South. *Contexts* 11(1), 12-13.

Dalstein, F.; Naqvi A. (2022) 21st Century water withdrawal decoupling: A pathway to a more water-wise world? *Water Resources and Economics,* 38, 100197.

Damania, R.; Desbureaux, S.; Hyland, M.; Islam, A.; Moore, S.; Rodella, A-S; Russ, J.; Zaveri, E. (2018) *Uncharted Waters The New Economics of Water Scarcity and Variability*. Washington, DC: World Bank. doi:10.1596/978-1-4648-1179-1.

Damania, R.; Desbureaux, S.; Rodella, A-S.; Russ, J.; Zaveri, E. (2019) *Quality Unknown: The Invisible Water Crisis*. Washington, DC: World Bank.

Distefano, T.; Kelly, S. (2017) Are we in deep water? Water scarcity and its limits to economic growth? *Ecological Economics, 132,* 130-137.

D'Odorico P.; Carr J.A.; Laio F.; Ridolfi, L.; Vandoni, S. (2014) Feeding humanity through global food trade. *Earth's Future, 2*, 458-469.

Duflo, E.; Pande, R. (2007) Dams. *Quarterly Journal of Economics,* 122(2), 601-646.

Fanaian, S. (2022) Too much, too little, too dirty: The evolution of water risks and governance in Guwahati, India. Doctor of Philosophy Thesis, University of Oxford.

Fanaian, S.; Fanaian, F. (2023) A tug of war between centralization and decentralization:the co-evolution of urban governance and water risks in Guwahati, India. *Environmental Research Communications*, 5(6), 065012. https://doi.org/10.1088/2515-7620/acdc86

Flörke, M.; Schneider, C.; McDonald, R.I. (2018) Water competition between cities and agriculture driven by climate change and urban growth. *Nature Sustainability*, 1, 51-58.

Garrick, D.E.; Hahn, R.W. (2021) An Economic Perspective on Water Security. *Review of Environmental Economics and Policy* 15(1), 45-66.

Gleeson, T.; Befus, K.; Jasechko, S.; Luijendijk, E.; Bayani Cardenas, M. (2016) The global volume and distribution of modern groundwater. *Nature Geoscience, 9*, 161-167.

Grafton, R.Q. (2017) Responding to the 'Wicked Problem' of water insecurity. *Water Resources Management,* 31(10), 3023-3041.

Grafton, R.Q.; Biswas, A.K.; Bosch, H.; Fanaian, S.; Gupta J.; Revi, A.; Sami, N.; Tortajada, C. (2023a) Goals, progress and priorities from Mar del Plata in 1977 to New York in 2023. *Nature Water,* 1, 230-240.

Grafton, R.Q.; Chu, L.; Kingsford, R.T., Bina, G.; Williams, J. (2022a) Resilience to Hydrological Droughts in the Northern Murray-Darling Basin. *Philosophical Transactions of the Royal Society A, 380*, 20210296.

Grafton, R.Q.; Fanaian, S.; Sacco G.; Liberman L. (2022b) Bending towards water justice: Pathways for truth, reconciliation, inclusion and transformative actions. *International Journal of Water Resources Development,* 38, 1-10.

Grafton, R.Q.; Horne, J.; Wheeler, S.A. (2022) Rethinking Water Markets. In *The Oxford Research Encyclopaedia of Environmental Science*. https://doi.org/10.1093/acrefore/9780199389414.013.800

Grafton, R.Q., Krishnaswamy, J.; Revi, A. (2023) *The Water Cycle and the Economy*. (Technical Note), Global Commission on the Economics of Water. https://www.waterjusticehub.org/the-water-cycle-and-the-economy/

Grafton, R.Q; Gupta, J.; Revi, A.; Mazzucato, M.; Okonjo-Iewala, N.; Rockström, J.; Shanmugaratnam, T.; Aki-Sawyerr, Y.; Bárcena Ibarra, A.; Cantrell, L.T.; Espinosa, M.F.; Ghosh A.; Ishii, N.; Jintiach, J.C.; Qui B.; Ramphele, M.; Urrego, M.R.; Serageldin, I.; Damania, R.; Dominique, K.; Esty, D.; Ovink, H.W.J.; Rao-Monari, U.; Selassie, A.; Andersen, L.S.; Beejadur, Y.A.; Bosch, H.; von Burgsdorff, L.K.; Fanaian, S.; Krishnaswamy, J.; Lim, J.; Portal, M.; Sami, N.; Schaef, J.; Bazaz, A.; Beleyur, P.; Fahrlander, S.; Ghoge, K.; Ragavan, K.V.S.; Vijendra, M.; Wankhade, K.; Zaqout, M.; Dupont, A.; Lefaive, X.; Réalé I. (2023b) The What, Why and How of the World Water Crisis: Global Commission on the Economics of Water Phase I Review and Findings. DOI 10.25911/GC7J-QM22. https://openresearch-repository.anu.edu.au/handle/1885/285201?mode=simple

Grafton, R.Q.; Manero, A.; Chu, L.; Wyrwoll, P. (2023c) The Price and Value of Water: An Economic Review. *Cambridge Prisms: Water,* 1, 1-39. doi:10.1017/wat.2023.2.

Grafton, R.Q.; Perry, C.J.; Williams, J. (2020). Confronting the Paradox of irrigation Efficiency. *Atlas of Science* https://atlasofscience.org/confronting-the-paradox-of-irrigation-efficiency/

Grafton, R.Q.; Pittock, J.; Davis, R.; Williams, J.; Fu, G.; Warburton, M.; Udall, B.; McKenzie, R.; Yu, X.; Che, N.; Connell, D.; Jiang, Q.; Kompas, T.; Lynch, A.; Norris, R.; Possingham, H.; Quiggin, J. (2013) Global insights into water resources, climate change and governance. *Nature Climate Change,* 3, 315-321.

Grafton, R.Q.; Ward, M.B.; To, H.; Kompas, T. (2011) Determinants of residential water consumption: Evidence and analysis from a 10-country household survey. *Water Resources Research,* 47, W08537.

Grafton, R.Q.; Williams, J.; Jiang Q. (2017) Possible pathways and tensions in the food and water nexus. *Earth's Future,* 5, 449-462.

Grafton, R.Q.; Perry, C.J.; Williams, J. (2020). Confronting the Paradox of irrigation Efficiency. *Atlas of Science* https://atlasofscience.org/confronting-the-paradox-of-irrigation-efficiency/

Grafton, R.Q.; Williams, J. (2020) Rent-seeking behaviour and regulatory capture in the Murray-Darling Basin, Australia. *International Journal of Water Resources Development*, 36/2–3, 484–504.

Grafton, R.Q.; Williams, J.; Perry, C.J.; Molle, F.; Ringler, C.; Steduto, P.; Udall, B.; Wheeler, S.A.; Wang, Y.; Garrick, D.; Allen R. (2018) The Paradox of Irrigation Efficiency. *Science,* 361(6404), 748-750.

Green, P.A.; Vörösmarty, C.J.; Harrison, I.; Farrell, T.; Sáenz, L.; Fekete, B.M. (2015) Freshwater ecosystem services supporting humans: Pivoting from water crisis to water solutions. *Global Environmental Change*, 34, 108-118.

Gupta, J.; Lebel, L. (2010) Access and allocation in earth system governance: Water and climate change compared. *International Environmental Agreements: Politics, Law and Economics,* 10(4), 377–395.

Gupta, J.; Liverman, D.; Prodani, K.; Aldunce, P.; Bai, X.; Broadgate, W.; Ciobanu, D.; Gifford, L.; Gordon, C.; Hurlbert, M.; Inoue, C.Y.A.; Jacobson, L.; Kanie, N.; Lade, S.J.; Lenton, T.M.; Obura, D.; Okereke, C.; Otto, I.M; Pereira, L.; Rockström, J.; Scholtens, J.; Rocha, J.; Stewart-Koster, B.; Tàbara, J.D.; Rammelt, C.; Verburg, P.H. (2023) Earth system justice needed to identify and live within Earth system boundaries. *Nature Sustainability,* 6, 630-638.

Haile, G.G.; Tang, Q.; Hosseini-Moghari, S.-M.; Liu, X.; Gebremicael, T.G.; Leng, G.; Kebede, A.; Xu, X.; Yun, X. (2020). Projected impacts of climate change on drought patterns over East Africa. *Earth's Future,* 8, e2020EF001502.

Hall, J.W.; Grey, D.; Garrick, D.; Fung, F.; Brown, C.; Dadson, S.J.; Sadoff, C.W. (2014). Coping with the curse of freshwater variability. *Science,* 346, 429-430.

High Level Panel of Experts (HLPE) (2015) *Water for Food Security and Nutrition. A Report by the High Level Panel of Experts on Food Security and Nutrition of the Committee on World Food Security.* Food and Agriculture Organization of the United Nations, Rome. https://www.unscn.org/uploads/web/news/Water-for-food-security-and-nutrition.pdf

High Level Panel on Water (HLPW) (2018) *Making every drop count: An agenda for water action*. High Level Panel on Water Outcome Document, United Nations, World Bank Group, Washington, DC.

Hosseiny, S.H.; Bozorg-Haddad, O.; Bocchiola, D. (2021) 9 Water, culture, civilization, and history. In *Economical, Political, and Social Issues in Water Resource* (Ed. Omid Bozorg-Haddad), AmsterdaM, Elsevier, pp. 189-261.

Institute for Health Metrics and Evaluation (IHME) (2020) Global Burden of Disease Collaborative Network. Global Burden of Disease Study (GBD 2019) Results. Seattle, United States: Institute for Health Metrics and Evaluation. Accessed 23 June at: https://vizhub.healthdata.org/gbd-results/.

Intergovernmental Panel on Climate Change (IPCC) (2022) *Climate change 2022: Impacts, adaptation and vulnerability. Summary for policy makers.* IPCC WGII sixth assessment report. Geneva: Intergovernmental Panel on Climate Change.

Islam, M.R., Ingham, V.; Hicks, J.; Kelly, E. (2018) *International Journal of Disaster Risk Reduction,* 28, 531-538.

Jackson, S. (2018) Water and Indigenous rights: Mechanisms and pathways of recognition, representation, and redistribution. *Wiley Interdisciplinary Reviews: Water,* 5(6), e1314.

Jongman, B. (2018) Effective adaptation to rising flood risk. *Nature Communications,* 9, 1986.

Kariuki M.; Schwartz, J. (2005). *Small-scale Private Service Providers of Water Supply and Electricity: A review of Incidence, Structure, Pricing and Operating Characteristics*. Policy Research Working Paper: No.3727, World Bank: Washington DC.

Katic, P.; Grafton, R.Q. (2023) Towards Operationalising the Water-Energy-Food Nexus (Technical Report). Global Commission on the Economics of Water, Paris. https://www.waterjusticehub.org/towards-operationalising-the-water-energy-food-nexus/

King Jr., M.L. (1968). *Remaining Awake Through a Great Revolution*. Speech given at the US national Cathedral, 31 March 1968. https://www.si.edu/spotlight/mlk

Kompas, T.; Che, T.N.; Grafton, R.Q. (2023) The Impact of Water and Heat Stress from Global Warming on Agricultural Productivity and Food Security (Technical Report). Global Commission on the Economics of Water, Paris. https://www.waterjusticehub.org/the-impact-of-water-and-heat-stress-from-global-warming-on-agricultural-productivity-and-food-security/

Kummu, M.; Ward, P.J.; de Moel,H.; Varis, O. (2010) Is physical water scarcity a new phenomenon? Global assessment of water shortage over the last two millennia. *Environmental Research Letters,* 5(3), 034006.

Lenton, T.M.; Benson, S.; Smith, T.; Ewer, T.; Lanel, V.; Petykowski, E.; Powell, T.W.R.; Abrams, J.F.; Blomsma, F.; Sharpe, S. (2022) Operationalising positive tipping points towards global sustainability. *Global Sustainability,* 5, E1 DOI: https://doi.org/10.1017/sus.2021.30

Lenton, T.M.; Williams, H.T.P. (2013) On the Origin of Planetary Scale Tipping Points. *Trends in Ecology & Evolution, 28(*7), 380-382.

Manero, A.; Taylor, K.S.; Nikolakis, W.; Adamowicz, V.; Marshall, V.; Spencer-Cotton, A.; Nguyen, M-N.; Grafton, R.Q. (2022). A systematic literature review of non-market valuation of Indigenous peoples' values: Current knowledge, best-practice and framing questions for future research. *Ecosystem Services,* 54, 101417.

McGregor D.; Whitaker S.; Sritharan M. (2020) Indigenous environmental justice and sustainability. *Current Opinion in Environmental Sustainability,* 43, 35-40.

McGregor, S.L.T. (2018) Philosophical Underpinnings of the Transdisciplinary Research Methodology. *Transdisciplinary Journal of Engineering & Science,* 9, 182-198.

Mehltretter S.; Longboat S.; Luby, B.; Bradford, A. (2023). Indigenous and Western Knowledge: Bringing Diverse Understandings of Water Together in Practice (Technical Report). Global Commission on the Economics of Water, Paris. https://www.waterjusticehub.org/indigenous-and-western-knowledge-bringing-diverse-understandings-of-water-together-in-practice/

Mekonnen, M.; Hoekstra, A.Y. (2016) Four billion people facing severe water scarcity. *Science Advances,* 2, e1500323.

Mekonnen, M.M.; Hoekstra, A.Y. (2018) Global anthropogenic phosphorus loads to freshwater and associated grey water footprints and water pollution levels: A high-resolution global study. *Water Resources Research*, 54(1), 345-358.

Mohammed, R.; Scholz, M. (2018) Critical review of salinity intrusion in rivers and estuaries. *Journal of Water and Climate Change*, 9(1), 1-16.

Molle, F.; Mollinga, P.P.; Wester, P. (2009) Hydraulic bureaucracies and the hydraulic mission: Flows of water, flows of power. *Water Alternatives*, 2/3, 328-349.

Mukherjee, S.; Mishra, A.; & Trenberth, K.E. (2018) Climate change and drought: A perspective on drought indices. *Current Climate Change Reports, 4*, 145-163.

Muller, M. (2018) Cape Town's drought: don't blame climate change. *Nature,* 559, 174-176.

OECD (2010) Managing Water for All: An OECD Perspective on Pricing and Financing Key Messages for Policy Makers. OECD: Paris.

OECD (2018) *The OECD Water Governance Indicator Framework*, Organisation for Economic Co-operation and Development, Paris, www.oecd.org/cfe/regionaldevelopment/oecd-water-governance-indicator-framework.htm.

Olmstead, S.M. (2010) The Economics of Water Quality. *Review of Environmental Economics and Policy,* 4(1), 44-62.

Our World in Data (2023) Clean Water https://ourworldindata.org/water-access#unsafe-water-is-a-leading-risk-factor-for-death

Owens, K.; Carmody, E.; Grafton, R.Q.; O'Donnell, E.; Wheeler, S.A.; Gooden, L.; Allen, R.; Lyster, R.; Steduto, P.; Jiang, Q.; Kingsford, R.; Quiggin, J. (2022) Delivering global water security: Embedding water justice as a response to increased irrigation efficiency *WIRE*s *Water,* 22; e1608.

Pastor A.V.; Biemans H.; Franssen W.; Gerten D.; Hoff H.; Ludwig F.; Kabata P. (2022) Understanding the transgression of global and regional freshwater planetary boundaries. *Philosophical Transactions of the Royal Society A,* 380(2238), 20210294.

Poff, N.L.; Zimmerman, J.K.H. (2010) Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshwater Biology,* 55, 194-295.

Pokhrel, Y.; Felfelani, F.; Satoh, Y.; Boulange, J.; Burek, P.; Gädeke, A.; Gerten, D.; Gosling, S.N.; Grillakis, M.; Gudmundsson, L.; Hanasaki, N.; Kim, H.; Koutroulis, A.; Liu, J.; Papadimitriou, L.; Schewe, J.; Schmied, H.M.; Stacke, T.; Telteu, C-E.; Thiery, W.; Veldkamp, T.; Zhao, F.; Wada, Y. (2021) Global terrestrial water storage and drought severity under climate change. *Nature Climate Change,* 11, 226–233.

Postel, S.L.; Daily, G.C.; Ehrlich, P.R. (1996) Human Appropriation of Renewable Fresh Water. *Science,* 271, 785-788.

Rammelt, C.F.; Gupta, J.; Liverman, D.; Scholtens, J.; Ciobanu, D.; Abrams, J.F.; Bai, X.; Gifford, L.; Gordon, C.; Hurlbert, M.; Inoue, C.Y.A.; Jacobson, L.; Lade, S.J.; Lenton, T.M.; Armstrong McKay, D.I.; Nakicenovic, N.; Okereke, C.; Otto, I.M.; Pereira, L.M.; Prodani, K.; Rockström, J.; Stewart-Koster, B.; Verburg, P.H.; Zimm, C. (2023) Impacts of meeting minimum access on critical earth systems amidst the Great Inequality.*Nature Sustainability,* 6, 212-221.

Richter, B.D.; Davis, M.M.; Apse, C.; Konrad, C. (2012) A presumptive Standard for Environmental Flow Protection. *River Research and Applications*, 28(8), 1312-1321.

Rodell, M.; Famiglietti, J.S.; Wiese, D.N.; Reager, J.T.; Beaudoing, H.K.; Landerer F.W.; Lo, M.-H. (2018) Emerging trends in global freshwater availability.*Nature,* 557, 651–659. https://doi.org/10.1038/s41586-018-0123-1.

Rosa, L.; Rulli, M.C.; Ali, S.; Chiarelli, D.D.; Dell'Angelo, J.; Mueller, N.D.; Scheidel, A.; Siciliano, G.; D'Odorico, P. (2021) Energy implications of the 21st century agrarian transition. *Nature Communications* 12, 2319.

Rosa, L; Chiarelli, D.D.; Tu, T.; Rulli, M.C.; D'Odorico, P. (2019) Global unsustainable virtual water flows in agricultural trade. *Environmental Research Letters,* 14(11), 114001.

Rosa, L.; Chiarelli, D.D. ; Rulli, M.C.; Dell'Angelo, J.; D'Odorico, P. (2020) Global agricultural economic water scarcity. *Science Advances,* 6**,** eaaz6031.

Rosegrant, M.W.; Ringler, C.; Zhu, T. (2009) Water for Agriculture: Maintaining Food Security under Growing Scarcity. *Annual Review of Environment and Resources,* 34, 205-222.

Rusca, M.; Cleaver, F. (2022) Unpacking everyday urbanism: Practices and the making of (un)even urban waterscapes. *WIREs Water,* 9(2). https://doi.org/10.1002/wat2.1581.

Satoh, Y.; Yoshimura, K.; Pokhrel, Y.; Kim, H.; Shiogama, H.; Yokohata, T.; Hanasaki, N.; Wada, Y.; Burek, P.; Byers, E.; Schmied, H.M.; Gerten, D.; Ostberg, S.; Gosling, S.N.; Boulange, J.E.S.; Oki, T. (2022) The timing of unprecedented hydrological drought under climate change. *Nature Communications,* 13, 3287-3287.

Scanlon, B.R.; Fakhreddine, S.; Rateb, A.; de Graaf, I.; Famiglietti, J.; Gleeson, T.; Grafton, R.Q.; Jobbagy, E.; Kebede, S.; Kolusu, S.R.; Konikow, L.F.; Long, D.; Mekonnen, M.; Müller Schmied, H.; Mukherjee, A.; MacDonald, A.; Reedy, R.C.; Shamsudduha, M.; Simmons, C.T.; Sun, A.; Taylor, R.G., Villholth, K.G.; Vörösmarty, C.J.;Zheng, C. (2023) Global water resources and the role of groundwater in a resilient water future. *Nature Reviews Earth & Environment*, 4(2), 87-101.

Schwartz, K.; Gupta, J.; Tutusaus, M. (2018) Editorial - Inclusive development and urban water services. *Habitat International*, *73*, 96–100. https://doi.org/10.1016/j.habitatint.2018.02.006

Shiru, M.S.; Shahid, S.; Dewan, A.; Chung, E.S.; Alias, N.; Ahmed, K.; Hassan, Q.K. (2020) Projection of meteorological droughts in Nigeria during growing seasons under climate change scenarios. *Scientific Reports*, 10(1), 1-18.

Sterman, J.D. (2002) *System Dynamics*: *Systems Thinking and Modelling for a Complex World*. Working Paper Series ESD-WP-2003-01.13. Massachusetts Institute of Technology Engineering Systems Division.

Sultana, F. (2022) Critical climate justice. *The Geographical Journal*, 188, 118–124.

Sun, S.; Zhou, X.; Liu, H.; Jiang, Y.; Zhou, H.; Zhang, C.; Fu, G. (2021) Unraveling the effect of inter-basin water transfer on reducing water scarcity and its inequality in China. *Water Research,* 194*,* 116931.

Syme, G.J.; Nancarrow, B.E.; McCreddin, J.A. (1999) Defining the components of fairness in the allocation of water to environmental and human uses. *Journal of Environmental Management*, 57/1, 51–70.

Tellman, B.; Sullivan, J.A..; Kuhn, C.; Kettner, A.J.; Doyle, C.S.; Brakenridge, G.R.; Erickson, T.A.; Slayback, D.A. (2021) Satellite imaging reveals increased proportion of population exposed to floods. *Nature,*596*,* 80-86.

Tetreault, D.; McCulligh, C. (2018) Water grabbing via institutionalised corruption in Zacatecas, Mexico *Water Alternatives*, 11/3, 572.

Tortajada, C.; Biswas, A. (2017) Water as a Human Right. *International Journal of Water Resources Development* 33(4), 509-511.

United Nations (2021) *The United Nations World Water Development Report 2021: Valuing Water Issue*. https://www.unwater.org/publications/un-world-water-development-report-2021/

United Nations General Assembly (2007) *United Nations Declaration on the Rights of Indigenous Peoples*, United Nations.

Vardon, M.J.; Le, T.H.L; Martinez-Lagunes, R.; Pule, O.B.; Schenau, S.; May, S.; Grafton, R.Q. (2023) Water Accounts and Water Accounting (Technical Report). Global Commission on the Economics of Water, Paris. https://www.waterjusticehub.org/water-accounts-and-water-accounting/

Vicente-Serrano, S.M.; Peña-Angulo, D.; Beguería, S.; Domínguez-Castro, F.; Tomás-Burguera, M.; Noguera, I.; Gimeno-Sotelo, L.; El Kenawy, A. (2022) Global drought trends and future projections. *Philosophical Transactions of the Royal Society A*, 380, 20210285.

Vörösmarty, C.J.; McIntyre, P.B.; Gessner, M.O.; Dudgeon, D.; Prusevich, A.; Green, P.; Glidden, S.; Bunn, S.E.; Sullivan, C.A.; Liermann, C.R.; Davies, P.M. (2010) Global threats to human water security and river biodiversity. *Nature,* 467, 555-561.

Vörösmarty, C.J.; Stewart-Koster B.; Green, P.A.; Boone, E.L.; Flörke, M.; Fischer, G; Wiberg, D.A.; Bunn, S.E.; Bhaduri, A.; McIntyre, P.B.; Sadoff, C.; Liu, H.; Stifel, D. (2021) A green-gray path to global water security and sustainable infrastructure, *Global Environmental Change*, 70, 102344.

Wada, Y.; van Beek L.P.H.; Bierkens, M.F.P. (2012) Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resources Research*, 48(6).

Wade, R. (1982) The system of administrative and political corruption: Canal irrigation in South India. *The Journal of Development Studies*, 18(3), 287-328.

Wang-Erlandsson, L.; Tobian, A.; van der Ent, R.J.; Fetzer, I.; te Wierik, S.; Porkka, M.; Staal, A.; Jaramillo, F.; Dahlmann, H.; Singh, C.; Greve, P.; Gerten, D.; Keys, P.W.; Gleeson, T.; Cornell, S.E.; Steffen, W.; Bai, X.; Rockström, J. (2022) A planetary boundary for green water. *Nature Reviews Earth & Environment*, 3(6), 380-392.

Wheeler, S.A. (2021) *Water Markets: A Global Assessment*, Edward Elgar Publishing, Cheltenham.

Whittington, D.; Nauges, C.; Fuente, D.; Wu, X. (2015) A diagnostic tool for estimating the incidence of subsidies delivered by water utilities in low- and medium-income countries, with illustrative simulations. *Utilities Policy, 34*, 70-81.

WHO (2019) *Safer Water, Better Health*, World Health Organization, Geneva, https://apps.who.int/iris/handle/10665/329905.

WHO (2021) *Progress on Household Drinking Water, Sanitation and Hygiene 2000-2020: Five Years into the SDGs* (Joint Monitoring Programme), World Health Organization, Geneva.

WHO and UNICEF (2021). *Progress on Household Drinking Water, Sanitation and Hygiene 2000-2020: Five Years into the SDGs*. Geneva, World Health Organization (WHO) and the United Nations Children's Fund (UNICEF).

WHO and UNICEF (2022) *State of the World's Drinking Water: An Urgent Call To Action To Accelerate Progress On Ensuring Safe Drinking Water For All*, WHO, UNICEF, World Bank, www.who.int/publications-detail-redirect/9789240060807.

WMO (2021) *WMO Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes*, *World Meteorological Organisation*. World Meteorological Organization (WMO). Available at: https://library.wmo.int/doc_num.php?explnum_id=10989.

Willardson, L.S.; Allen, R.G.; Frederiksen, H.D. (1994) Elimination of irrigation efficiencies, 13th Technical Conference USCID, Denver, CO, October 19-22, 1994.

World Bank (2016) *High and Dry: Climate Change, Water, and the Economy*. Working Paper, World Bank, Washington, DC, https://doi.org/10.1596/K8517.

World Bank (2023) 6 Clean Water and Sanitation in *Atlas of Sustainable Development Goals 2023*. Accessed 23 June 2023 at: https://sdga2022.github.io/sdga2022/goal-6-clean-water-and-sanitation

Yao, F.; Livneh, B.; Rajagopalan, B.; Wang, J.; Crétaux, J-F; Wada, Y.; Berge-Nguyen, M. (2023) Satellites reveal widespread decline in global lake water storage. *Science,* 380, 743-749.

Zhang, Y.; Zheng, H.; Zhang, X.; Leung, L.R.; Liu, C.; Zheng, C.; Guo, Y.; Chiew, F.H.S.; Post, D.; Kong, D.; Beck, H.E.; Li, C.; Blöschlet, G. (2023) Future global streamflow declines are probably more severe than previously estimated. *Nature Water* 1, 261-271.

Zwarteveen, M.Z.; Boelens, R. (2014) Defining, researching and struggling for water justice: Some conceptual building blocks for research and action. *Water International*, 39/2, 143-158.

Appendix

Data sources for Figure 1:

Too much, too little, too dirty- Floods (flood events, economic losses and affected people, droughts (flood events, economic losses and affected people) (Source: EM-DAT: The OFDA/CRED International Disaster Database) and water quality (Nitrates in water in Europe (Source: EIONET Central Data Repository http://discomap.eea.europa.eu/data/wisesoe/deriveddata/T_WISE4_AggregatedDataByWaterBody/0.html), disability due to unsafe water sanitation and hygiene (Source. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Results))

Data sources for Section 2: EM-DAT, CRED / UCLouvain, Brussels, Belgium – *www.emdat.be*

The OFDA/CRED International Disaster Database (EM-DAT, CRED / UCLouvain, Brussels, Belgium).

The database is compiled from various sources including UN, governmental and non-governmental agencies, insurance companies, research institutes and press agencies (see Table 2). As there can be conflicting information and figures, CRED has established a method of ranking these sources according to their ability to provide trustworthy and complete data. In most cases, a disaster will only be entered into EM-DAT if at least two sources report the disaster's occurrence in terms of deaths and/or affected persons.

Disasters: Flood (events, Total Affected, Damages, Adjusted (US$)); Droughts (events, Total Affected, Damages, Adjusted (US$)).

Countries: Australia, China, India, Nigeria, South Africa, France, United States of America.

Timeframe: 1991-2022

Website: www.emdat.be

Version: 2023-06-13

Definitions:

Disaster Events: Count of number of times flood and droughts were listed in the database. "A disaster meeting the EM-DAT criteria and which is recorded in EM-DAT. A disaster event can affect one country or several [see «Country-level disaster»]. In the case of the latter, the disaster event will result in several country-level disasters being entered into the database. A disaster event will always have a unique DISNO identifier."

Disaster criteria: EM-DAT includes all disasters from 1900 until the present, conforming to at least one of the following criteria:

– 10 or more people dead
– 100 or more people affected
– The declaration of a state of emergency
– A call for international assistance

Damages, Adjusted (US$): "A value of all damages and economic losses directly or indirectly related to the disaster. The information may include the breakdown figures by sectors: Social, Infrastructure, Production, Environment and other (when available). Adjusted value indicates that Consumer Price Index was used to convert the damages (which are given at the time the disaster occurred) to the current US$ value."

Total affected: "The total affected is the sum of injured, affected and homeless. Injured: People suffering from physical injuries, trauma, or an illness requiring immediate medical assistance as a direct result of a disaster. The number of injured people is entered when the term "injured" is written in the source. The injured are always part of the "total affected". Any related word like "hospitalized" is considered as injured. If there is no precise number is given, such as "hundreds of injured", 200 injured will be entered (although it is probably underestimated). Affected people: People requiring immediate assistance during an emergency situation. The indicator affected is often reported and is widely used by different actors to convey the extent, impact, or severity of a disaster in non-spatial terms. The ambiguity in the definitions and the different criteria and methods of estimation produce vastly different numbers, which are rarely comparable. Homeless: Number of people whose house is destroyed or heavily damaged and therefore need shelter after an event."

Floods: "A general term for the overflow of water from a stream channel onto normally dry land in the floodplain (riverine flooding), higher-than- normal levels along the coast and in lakes or reservoirs (coastal flooding) as well as ponding of water at or near the point where the rain fell (flash floods)."

Drought: An extended period of unusually low precipitation that produces a shortage of water for people, animals, and plants. Drought is different from most other hazards in that it develops slowly, sometimes even over years, and its onset is generally difficult to detect. Drought is not solely a physical phenomenon because its impacts can be exacerbated by human activities and water supply demands. Drought is therefore often defined both conceptually and operationally. Operational definitions of drought, meaning the degree of precipitation reduction that constitutes a drought, vary by locality, climate, and environmental sector.

Data used in Section 2.3:

Nitrates- figure shows the trends in nitrate in European groundwater in mg No3/l. The timeframe is from 2000-2022. Data from Europe (1258), Albania (7), Austria (41), Belgium (34), Cyprus (14), Czechia (22), Denmark* (38), Estonia (36), Finland** (70), France** (241), Germany (122), Iceland (1), Ireland** (17), Italy (25), Latvia (16), Lithuania (22), North Macedonia (18), Poland (16), Romania (89), Serbia (34), Slovakia (8), Slovenia (8), Spain** (250), Sweden* (113), Switzerland (16).

Data from: European Environmental Agency https://www.eea.europa.eu/data-and-maps/daviz/nitrate-in-groundwater-and-rivers-1#tab-chart_2

Disability-adjusted life year: due to unsafe water: GBD Results tool: Use the following to cite data included in this download: Global Burden of Disease Collaborative Network.

Global Burden of Disease Study 2019 (GBD 2019) Results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2020. Available from https://vizhub.healthdata.org/gbd-results/.

DALY is an abbreviation for disability-adjusted life year. It is a universal metric that allows researchers and policymakers to compare very different populations and health conditions across time. DALYs equal the sum of years of life lost (YLLs) and years lived with disability (YLDs). One DALY equals one lost year of healthy life. DALYs allow for the estimation of the total number of years lost due to specific causes and risk factors at the country, regional, and global levels.

Data sources for Figure 2:
Through a base global map of water scarcity, regional insights into too much and too little (Source: Base map- WWF Risk Filter Suite: riskfilter.org; regional data- EM-DAT: The OFDA/CRED International Disaster Database); and WWF Risk Filter Suite: riskfilter.org).

Water scarcity definition:
"Water scarcity refers to the physical abundance or lack of freshwater resources, which can significantly impact business such as production/supply chain disruption, higher operating costs, and growth constraints. Water scarcity is human-driven and can be aggravated by natural conditions (e.g., aridity, drought periods), and it is generally calculated as a function of the volume of water use/demand relative to the volume of water available in a given area.
The Water Risk Filter risk category water scarcity is a comprehensive and robust metric as it integrates a total of 7 best available and peer-reviewed datasets covering different aspects of scarcity as well as different modelling approaches: aridity index, water depletion, baseline water stress, blue water scarcity, available water remaining, drought frequency probability, and projected change in drought occurrence." (WWF 2021, p.9)
Citation: WWF 2021 *WWF Water Risk Filter Methodology Documentation, January 2023* Online: https://cdn.kettufy.io/prod-fra-1.kettufy.io/documents/riskfilter.org/WaterRiskFilter_Methodology.pdf

Regional data information:
Source: EM-DAT, CRED / UCLouvain, Brussels, Belgium
Database: EM-DAT: The OFDA/CRED International Disaster Database
Disasters: Flood (events, Total Affected, Damages, Adjusted (US$)); Droughts (events, Total Affected, Damages, Adjusted (US$)).
Countries: Australia, China, India, Nigeria, South Africa, France, United States of America.
Timeframe: 1991-2022
Website: www.emdat.be
Version: 2023-06-13
Definitions of indicators as listed above.

Carlos Carreira
Univ Coimbra, CeBER, Faculty of Economics
ccarreir@fe.uc.pt
orcid.org/0000-0002-4786-5605
Pedro Silva
Univ Coimbra, CeBER, Faculty of Economics
pedro.ja.silva@outlook.pt

DOI: https://doi.org/10.14195/2183-203X_56_4

The Determinants of Trade Credit During and After the 2008 International Financial Crisis

Os Determinantes do Crédito Comercial Durante e Após a Crise Financeira Internacional de 2008

Carlos Carreira
Pedro Silva

Received for publication: April 15, 2023
Revision accepted for publication: June 25, 2023

ABSTRACT
This work analyses the determinants of trade credit granted and received and the effect of the 2008 financial crisis on it. Using a sample of 96,417 Portuguese SMEs from the non-financial sector for the period 2010–2019, we found that trade credit plays an important role in firms' financing policies. Firms with better access to the credit market act as financial intermediaries and grant financing to firms that have difficulty accessing credit. Moreover, the use of trade credit seems to be a substitute for bank financing. We also found that firms use trade credit as a marketing tool to increase their sales. Finally, we found a slump in credit granted to customers after the 2008 financial crisis, which seems to mimic the contraction in aggregate bank credit.
Keywords: Trade Credit; SMEs; Financing; Commercial policy.

JEL Classification: G32; L29; G01.

Acknowledgment: We would like to thank Laura Hirche Almeida for her contribution to the literature review section as a scholarship of the “Summer with Science 2022” program, funded by the Foundation for Science and Technology (FCT). This work was funded by the FCT, R&D unit CeBER (UIDB/05037/2020).

1. Introduction[1]

Trade credit enables firms to separate the payment cycle from the delivery schedule. When looking for sources of financing, it is important to consider other funding options besides bank credit, especially trade credit. Indeed, the volume of trade credit is higher than short-term loans received from banks in all developed economies (Cuñat and García-Appendini, 2012).

Trade credit varies substantially across firms and industries. The main objective of this work is to analyze the determinants of trade credit, using a sample of Portuguese small and medium-sized enterprises (SMEs) from the non-financial sector for the period 2010–2019. This, in turn, enables us to test several theories of trade credit provided by the literature. In addition, we also test whether firms are more likely to use or apply for trade credit in response to an aggregate contraction in bank credit.

Since Meltzer's (1960) research, several studies have investigated the determinants of trade credit. This work adds to this literature by analyzing the determinants of trade credit in the context of Portugal. In particular, the study provides some important insights for corporate governance in a way that firms can decide on their trade credit behavior while considering the firm-specific characteristics. By this we mean that firms with better access to the credit market could act as financial intermediaries and grant financing to firms that have difficulty accessing credit. Further, trade credit received could be a substitute for bank financing. Trade credit may play an important role not only in financing policy, but also as a marketing tool to increase sales. Finally, policy makers should consider financing by firms as a key issue when aggregate bank credit shrinks, as firms may also reduce the credit extended to their customers.

2. Theories of Trade Credit

Trade credit can be defined as a source of short-term financing provided by suppliers to their customers. It allows the customer to purchase goods or services on credit and defer payment to a later date.

The balance sheets of most firms contain accounts receivable (i.e. financing granted to customers) as well as accounts payable (i.e. financing received from suppliers). For example, Garcia-Teruel and Martinez-Solano (2010) report that the accounts receivable (accounts payable) over assets ranged from 39.3% in Spain (28.5% in France) to 19.2% in Finland (13.2% in Finland) in the period 1996–2002. In the case of Portugal, Giannetti (2003) found that the amount of accounts payable represented up to three times other types of short-term debt in a sample of large firms from 1993 to 1997.

---

[1] A previous version of this work was presented by Pedro Jorge de Almeida e Silva as a Master's Thesis, under the title "Crédito Comercial nas PMEs Portuguesas: Análise das determinantes e dinâmicas de financiamento na indústria transformadora", under the supervision of Prof. Carlos Carreira, at the University of Coimbra, Faculty of Economics.

Why Does Trade Credit Exist? Several theories can be found in the literature to explain the existence and use of trade credit, based specially on financial, operational and commercial motives (Petersen and Rajan, 1997).

2.1. Financial motives

Close relationships between suppliers and customers can mitigate the information asymmetry between creditors and debtors. Suppliers have a comparative advantage over financial institutions regarding information acquisition, contract enforcement, and the liquidation process (Petersen and Rajan, 1997; Delannay and Weill, 2004; Huyghebaert, 2006). This advantage allows firms with easier access to credit markets to serve as financial intermediaries for firms with limited credit access.

Therefore, the level of trade credit will depend on the creditworthiness of the firm, and the availability and cost of financial resources from banks. We expect that firms with easier access to bank loans will grant more trade credit, while those with fewer financial options will resort more to trade credit from their suppliers. However, since suppliers have an information advantage, trade credit extension may be considered by banks as a signal of the quality of a borrower, and therefore induce banks to grant credit (Biais and Gollier, 1997). Consequently, trade credit and bank credit can be complementary and not substitutes, as stated above.

2.2. Operational motives

Trade credit enables firms to separate the payment cycle from the delivery schedule, reducing the transaction costs, especially in seasonal markets or with highly uncertain demand. Instead of overinvesting in fixed assets or holding excess inventory, firms can use trade credit to smooth demand (Ferris, 1981; Emery, 1987). Therefore, firms have operating motives to use trade credit—to stimulate sales in times of low demand. We thus expect that firms may use more trade credit when their sales growth is low.

However, sales growth is also a factor that affects the demand for finance in general, and for trade credit in particular. Therefore, we can also expect that firms with greater increases in sales will use more trade credit in order to finance their new investments (Garcia-Teruel and Martinez-Solano, 2010).

2.3. Commercial motives

There are also commercial motives for granting trade credit. Trade credit can be used by firms as a form of price discrimination. Prolonging the period of credit or increasing the discount for prompt payment effectively equates to a price reduction (Brennan et al., 1988). Firms operating with high contribution margins have a strong incentive to induce additional sales without cutting the price by extending trade credit instead (Petersen and Rajan, 1997). Consequently, we expect firms with higher profit margins to grant more trade credit.

Another commercial motive for trade credit is its use by suppliers to offer an implicit guarantee of quality. Indeed, suppliers can transmit information about the quality of their products by agreeing to credit terms that allow their customers a period of evaluation (Smith, 1987; Lee and Stowe, 1993). Small firms, which are typically younger and lack a solid reputation in the market, tend to grant more trade credit than large firms, which benefit from a well-established reputation (Long et al., 1993). Therefore, we expect that firms with high product quality will offer more trade credit to their customers in order to allow them to evaluate product quality.

3. Data and Methodology

3.1. The dataset

The dataset used in this study was originally compiled by Carreira et al. (2022), who extracted the raw data from the Integrated Business Accounts System (SCIE, Portuguese acronym), administered by the Portuguese Statistical Office (INE). Our sample specifically covers the whole population of small and medium-sized firms operating in Portugal from 2010 to 2019, except for the financial sector, and education, health and cultural services. The SMEs are defined according to the requirements established by the European Commission recommendation 2003/361.[2]

After this preliminary filtering, the information obtained was refined. Observations with unreasonable values (*e.g.*, non-positive turnover or total assets) were discarded. In addition, we truncated 1% of the extreme ratios (percentiles 1 and 99) presented by the variables defined in the next section. Our final sample comprises an unbalanced panel of 96,417 firms making up 488,694 year-firm observations.

3.2. Model specification and description of variables

Firms using trade credit as both suppliers and customers (Petersen and Rajan, 1997). We will examine these two sides of trade credit by first looking at firms as lenders (suppliers) and then as borrowers (customers). As proxies for how much a firm lends to its customers and borrows from its suppliers, we use accounts receivable (normalized by sales) and accounts payable (normalized by assets), respectively (Petersen and Rajan, 1997; Niskanen and Niskanen, 2006; Garcia-Teruel and Martinez-Solano, 2010).

[2] Specifically, the firms in the sample met the following conditions: (1) under 250 employees; (2) an annual turnover of up to €50 million; (3) total assets of up to €43 million; (4) not classified as a microenterprise (i.e., under 10 employees and an annual turnover or total assets of up to €2 million).

We considered the following models to investigate trade credit determinants:

$$RECEIV_{it} = \alpha_0 + \alpha_1 X_{it} + \alpha_2 Y_{it} + \varphi_i + \lambda_t + \mu_i + \epsilon_{it}, \quad (1)$$
$$PAYAB_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Z_{it} + \varphi_i + \lambda_t + \mu_i + \epsilon_{it}, \quad (2)$$

where $RECEIV_{it}$ represents the trade credit granted by firm i at time t to its customers, calculated as the ratio of accounts receivable to sales; and $PAYAB_{it}$ the finance received from its suppliers, defined as the ratio of accounts payable to total assets.

Regarding the explanatory variables that may simultaneously impact accounts receivable and accounts payable, X_{it}, we first use size and age as proxies for the firm's creditworthiness and, therefore, its ability to access alternative sources of finance. Size (LSIZE) and age (LAGE) are calculated as the logarithm of book value of assets and (1+age), respectively, where age is the number of years since the firm was incorporated. Typically, both larger and older firms have better creditworthiness and consequently easier access to finance than smaller and younger firms (Carreira and Silva, 2010). Consequently, these firms are likely to grant more credit to their customers and to use less credit from their suppliers as they have other sources of finance to fall back on (Schwartz, 1974; Petersen and Rajan, 1997). Conversely, it can be argued that larger and older firms generally have better trade reputations and are therefore less likely to be forced to offer credit to their customers to guarantee their products (Long et al., 1993). In addition, larger firms may have greater relative bargaining power in trade relations between suppliers and customers, allowing them to impose stricter payment terms on their customers. Larger and older firms are also offered trade credit more often due to their better creditworthiness. Although size and age may affect trade credit in different directions, we generally expect a positive relationship between accounts receivable and both variables, and a negative relationship in the case of accounts payable.

To control for the firm's ability to generate internal resources, we use the cash flow, defined as the ratio of net profits plus depreciation to sales/assets for accounts receivable/payable (CFLOW1 and CFLOW2, respectively). Firms with greater internal cash tend to grant more credit to their customers and need less external financing and therefore have lower accounts payable (Garcia-Teruel and Martinez-Solano, 2010). Thus, the relationship between internal financing and accounts receivable is expected to be positive, but negative in regard to accounts payable.

To control for the cost of external finance, we use the cost of financial debt (FCOST), defined as the ratio of finance costs to external financing excluding trade creditors. High financial costs lead firms to have less incentive to grant financing to their customers and more incentives to demand financing from their suppliers. So, we expect the cost of financial debt to be related negatively with accounts receivable and positively with accounts payable.

To capture the effect of possible shocks in sales on trade credit, we use the annual sales growth rate (GROWTH). A firm willing to grow may choose a strategy of granting more trade credit. Thus, growth should be positively related to accounts receivable. However, a firm whose sales are declining may also react by offering more trade credit to enhance their sales (Petersen and Rajan, 1997). To address these contradictory effects, we divide the variable into PGROWTH, which equals GROWTH when sales growth is positive and zero otherwise, and NGROWTH, which equals the absolute value of GROWTH when growth is negative and zero otherwise (Petersen and Rajan, 1997; Niskanen and Niskanen, 2006; Garcia-Teruel and Martinez-Solano, 2010). Despite the contradiction in the above arguments about the effect associated with declining sales, we expect a positive relationship between accounts receivable and both growth variables.

Growing firms have more investment opportunities, so they will have an increased demand for funds and consequently for trade credit. However, the effect depends on the substitution or complementarity between bank credit and trade credit, as sales growth is also a positive signal for the health of firms. Nevertheless, we expect a positive relationship between accounts payable and positive growth, and a negative relationship in the case of negative growth.

The second set of covariates, Y_{it}, comprises three specific independent variables to trade credit received. First, to measure the ability of a firm to access external financing, we include the short-term finance (STLEV), calculated as the ratio of current liabilities to sales. This variable can be interpreted in much the same way as the creditworthiness, that is, we expect firms that are able to obtain more short-term resources to also be able to grant more credit to their customers.

Trade credit can also be used to transmit information about the quality of products. Buyers consider trade credit extension as a signal of trust, as the credit period gives them time to assess product quality before paying (Long et al., 1993). To capture this effect, we use the ratio of sales to assets, deducting accounts receivable (TURN) (Garcia-Teruel and Martinez-Solano, 2010). We expect a negative relationship, as firms with lower sales turnover produce higher quality goods. However, we should also note that larger firms, which generally have a better reputation, may not need to signal the quality of their products by granting more commercial credit (Long et al., 1993; Garcia-Teruel and Martinez-Solano, 2010).

Firms with larger operating margins have a greater incentive to increase sales by extending trade credit. The high profit margin offsets the financial cost of offering extra trade credit. Thus, firms can use trade credit as a price discrimination mechanism, which can also help firms keep a long-term relationship with customers (Petersen and Rajan, 1997; Niskanen and Niskanen, 2006). The variable operating margins is calculated as the ratio of gross profit to sales (MARGIN).

Finally, the third set of covariates, Z_{it}, includes three specific independent variables for accounts payable. First, to assess the substitution or complementarity between bank loans and accounts payable, we include the short-term financial debt (STFIND), measured as the ratio of short-term financial debt to assets (Burkart and Ellingsen, 2004). To test whether there is a substitution effect between long-term debt and debt provided by the suppliers, we also include the variable long-term debt (LTDEBT), defined as the ratio of long-term debt to assets. If bank loans and accounts payable were substitutable, we expect to observe a negative link between this variable and accounts payable ratio.

Firms tend to match the maturity of their liabilities and the liquidity of their assets (Morris, 1976). To control this effect, we introduce the ratio of current assets to total assets (CURRAS) as one of the independent variables. We expect firms that have invested more in current assets to use more short-term finance in general, and more trade credit, in particular.

We also include industry dummies ($_i$) at the two-digit NACE level to control for the well-known impact of industry structures. Indeed, empirical evidence shows that trade credit terms vary widely across industries but have only limited variation within industries

(Petersen and Rajan, 1997; Ng et al., 1999). Furthermore, we include time dummies (λ_t) to capture the influence of macroeconomic factors (e.g. credit rationing) that may affect the decision to grant or use trade credit (Casey and O'Toole, 2014). Finally, parameter $_{i}$ is the unobservable individual effect to control for the unique characteristics of each firm, and $_{it}$ is the random disturbance.

Tables A1 and A2 in the Appendix report the descriptive statistics and the correlation matrix of covariates, respectively.

4. **Results**

4.1. The use of trade credit by portuguese smes

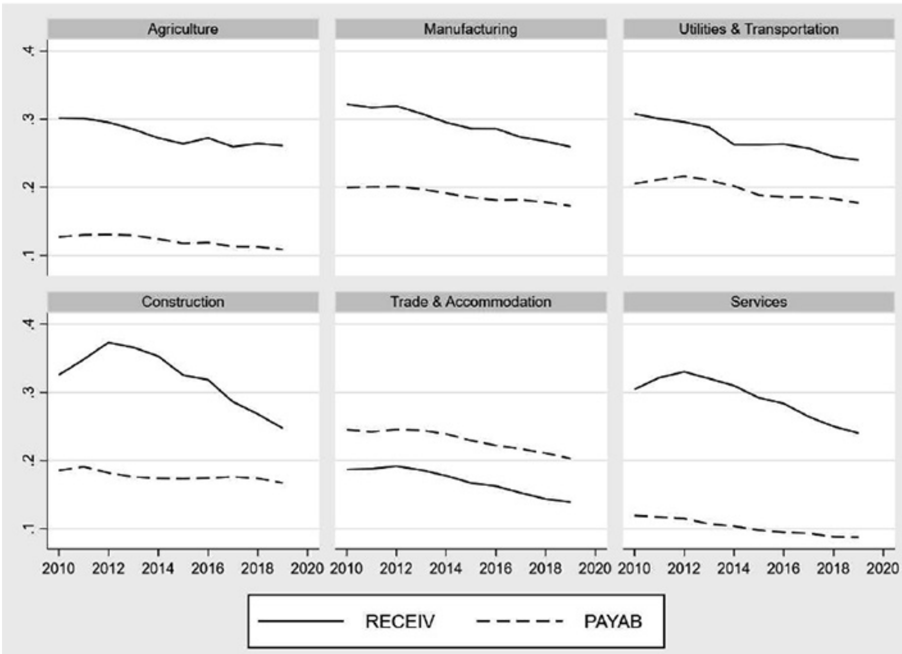
As can be seen in Table A1 in the Appendix, trade credit is an important source of external financing. The accounts receivable represent, on average, about a quarter of sales (25.2%), which is broadly in line with the 26.5% reported by Garcia-Teruel and Martinez-Solano (2010) for Spain in the period 1996–2002. It is noteworthy that, according to these authors, trade credit granted to customers is much higher in the Euro-Mediterranean countries than in the Nordic countries (Greece has the highest rate, at 33.1%, while the lowest rate is in Finland, at 9.2%).

Regarding the accounts payable, we observe that they represent about 18.3% of assets. This value is higher than the mean of the other forms of financial debt, including short-term financial debt (about 8.3%) and long-term debt (15.8%), which reveals the importance of supplier financing for firms (Table A1 in the Appendix). However, finance received from suppliers in Portugal is relatively lower than the average figures documented by Garcia-Teruel and Martinez-Solano (2010) for other Euro-Mediterranean countries (ranging from 24.9% in Spain to 28.5% in France). In the Scandinavian countries, by contrast, the level of accounts payable is somewhat lower (13.2% for Finland and 16.4% for Sweden).

The trade credit granted and received varies considerably across industries. Figure 1 shows the evolution of the trade credit over the 10 years of the sample (2010–2019) by industry. *Construction* has the highest average level of accounts receivable (at 32.1% of sales), followed by *Manufacturing* and *Services* (with 29.3% and 29.2%, respectively), while the lowest figure is seen in *Trade and Accommodation* (at 16.9%). In contrast, *Trade and Accommodation* presents the highest average accounts payable (at 23.0% of assets), while the lowest values are in *Services* and *Agriculture* (at 10.2% and 12.1%, respectively). Thus, firms in *Trade and Accommodation* make most use of financing from suppliers, while they grant the least credit to their customers. These results are broadly consistent with those of Garcia-Teruel and Martinez-Solano (2010).

Both measures of trade credit follow the same downward trend over the sample period. Specifically, on average, accounts receivable (payable) have declined by 7.0 (2.9) percentage points between 2012 and 2019.

Figure 1: Trade credit by industry, 2010-2019



Notes: Annual means. RECEIV is the ratio of accounts receivable (*i.e.* finance granted to customers) to sales; PAYAB is the ratio of accounts payable (*i.e.* finance received from suppliers) to total assets.

4.2. Determinants of trade credit

Table 1 reports the results of Models (1) and (2), accounts receivable and accounts payable, respectively. The two estimations have been performed using the fixed effect model because, first, the Breusch-Pagan test identifies the existence of individual effects (*i.e.* rejects the null hypothesis that the preferred model is pooled OLS) and, second, the Hausman test rejects random effects in favor of the fixed effects model. Moreover, in both cases, the *F*-test rejects the null hypothesis of joint insignificance of the coefficients at the 1% level.[3]

[3] Note that Petersen and Rajan (1997), Delannay and Weill (2004) and Niskanen and Niskanen (2006), *e.g.*, obtain an R^2 of similar magnitude to ours.

Table 1: Determinants of trade credit

Variables	Model (1)	Model (2)
	Accounts receivable	Accounts payable
lnSIZE	0.063*** (0.001)	-0.017*** (0.001)
lnAGE	0.026*** (0.002)	-0.030*** (0.001)
CFLOW1	-0.024*** (0.001)	
CFLOW2		-0.172*** (0.001)
FCOST	-0.248*** (0.010)	0.231*** (0.006)
PGROWTH	-0.003*** (0.000)	0.002*** (0.000)
NGROWTH (absolute value)	0.176*** (0.002)	-0.035*** (0.001)
STLEV	0.003*** (0.000)	
TURN	0.006*** (0.000)	
MARGIN	0.020*** (0.001)	
STFIND		-0.087*** (0.002)
LTDEBT		-0.084*** (0.001)
CURRAS		0.066*** (0.001)
Constant	-0.730*** (0.015)	0.501*** (0.010)
Industry dummy	YES	YES
Year dummy	YES	YES
No. of observations	423,467	423,467
No. of firms	80,921	80,921
R^2 (overall)	0.087	0.120
F statistic	467.98***	382.19***

Notes: Fixed-effects regressions of Models (1) and (2), respectively, accounts receivable and accounts payable. RECEIV – ratio of accounts receivable to sales; PAYAB – ratio of accounts payable to total assets; lnSIZE – log of assets; lnAGE – log of firm age; CFLOW1 and CFLOW2 – ratio of net profits plus depreciation to sales and assets, respectively; FCOST – ratio of finance costs to financial debt; GROWTH – annual sales growth rate; STLEV – ratio of current liabilities to sales; STFIND – ratio of short-term financial debt to assets; LTDEBT – ratio of long-term debt to assets; TURN – turnover of sales over assets; MARGIN – ratio of gross profit to sales; CURRAS – ratio of current assets to total assets. Coefficients of industry (two-digit level NACE-Rev.2 classification) and time (2010–2019) dummies not reported (in both cases, the *F*-test rejects the null hypothesis that the dummy coefficients are jointly equal to zero at the 1% level). Firm-cluster robust standard errors are given in parentheses. ***, ** and * statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Firm size and age seem to be a determinant factor of the trade credit. Indeed, as expected, we found a positive (negative) and significant relationship between accounts receivable (payable) and both size and age. A firm with €3.30 million in assets (the 75th percentile) grants 12.4% more of its sales in trade credit and uses 3.4% less credit from its suppliers than a firm with €0.46 million in assets (the 25th percentile). Increasing the firm's age from eight to 25 years old (25th and 75th percentiles) increases (decreases) the ratio of accounts receivable (payable) by 2.7 (3.1) percentage points. These results suggest that larger and older firms are more likely to benefit from access to bank credit due to their higher creditworthiness and are more willing to provide trade credit to their customers. While this last finding supports the financial motive for trade credit, it does not support the commercial motive, according to which firms use their greater bargaining power in trade negotiations.

Surprisingly, the capacity to generate internal funds (CFLOW) is negatively correlated with accounts receivable. We expected that firms with more internal cash would be able to extend more credit to their customers. Given that, we examined in more detail the effect of this variable on a firm's decision to finance its customers by splitting the internal resources generated into positive and negative cash flows (the latter in absolute values). Now we find significantly positive coefficients (PCFLOW=0.018 and NCFLOW=0.034; regression not reported) indicating different motives for trade credit between profitable and unprofitable firms. Indeed, as expected, the most profitable firms tend to extend more credit to their customers. But the unprofitable firms also tend to extend more credit, which can be explained by the fact that distressed firms use the extension of credit to attempt to maintain their sales. These firms are also in a worse bargaining position and are then unable to obtain fast payment from their customers (Petersen and Rajan, 1997).

To explore this further, we also divided losses into those when the firm has positive sales growth and losses when the firm has negative sales growth. The positive coefficient for the former is larger than for the latter. So firms that grow fast (and incur losses) seem to extend more credit ("buy" sales), but distressed firms (negative sales growth and negative income) also offer more trade credit, which seems to support the argument that debtors are less willing to repay distressed firms. Finally, when we include in the regression the square of CFLOW, both terms are statistically significant, depicting a U-shaped relationship across the entire range of data (i.e. the most profitable and the most unprofitable firms tend to grant more credit).[4]

Since a firm's ability to extend credit depends on its ability to raise funds, not only internally but also externally, we also control for the availability and cost of external financing. The coefficients of short-term financing (STLEV) and cost of financing (FCOST) are positive and negative, respectively, indicating that firms with greater access to short-term financing and cheaper external financing provide more financing to their customers.

In the case of accounts payable, the results confirm a substitution effect between supplier-provided credit and other sources of financing (internal and external). First, there is an inverse relationship between credit received from suppliers and resources generated internally (CFLOW). That is, as the ability to generate internal funds increases, firms tend to reduce their use of trade credit received. Second, short-term financial debt (STFIND) is significantly negative. Thus, the firms reduce the weight of accounts payable when they have access to other short-term financial resources. Third, the coefficient of long-term debt (LTDEBT) is also negative, again supporting the substitution hypothesis. Finally, as also expected, the relationship between accounts payable and financing costs (FCOST) is positive, which means that firms that incur higher costs in their external financing tend to resort to more financing from their suppliers.

Firms which have had positive sales growth (PGROWTH) report slightly fewer receivables (the coefficient is significantly negative, but economically small). This suggests that these firms are less dependent on their customers and consequently can influence the commercial

[4] These results are available from the authors upon request.

negotiations in their favor by reducing delays of payment. In contrast, firms whose sales have declined increase the proportion of financed sales (note that NGROWTH is defined in absolute value). Thus, firms try to limit the decline of their sales by offering more favorable terms of payment. In particular, for each euro less in sales, they grant more credit to their customers by about 17.6 cents. This result suggests that firms are mainly using trade credit as a marketing tool to improve their sales figures.

The level of accounts payable is positively affected by positive sales growth. Thus, firms with growth opportunities, which consequently have a higher demand for funds to invest, rely on the support of their suppliers to finance this growth. On the other hand, firms whose sales fall have lower accounts payable. This means that suppliers act as typical financial intermediaries and try to limit their risk as they tend to reduce the amount of credit granted to customers in trouble.

Regarding the effect of the firm's gross profit margin (MARGIN) on trade credit granted, the positive sign of the estimated coefficient seems to confirm the price discrimination theory. Indeed, firms that charge high prices (hence high margins) seem to use trade credit as a strategic tool to increase sales.

The results obtained do not allow us to support the quality-signaling hypothesis of Long et al. (1993). In fact, contrary to our expectations, we found a positive relationship between accounts receivable and TURN. This suggests that firms mainly sell products, the quality of which does not need to be transmitted, by extending more trade credit, which is consistent with the previously found result of the reputation of larger and older firms.

Firms attempt to match the maturities of assets and liabilities. The relationship found between the accounts payable and the weight of current assets (CURRAS) is positive, meaning that firms that invest more in current (short-term) assets tend to use more current debt such as trade credit.

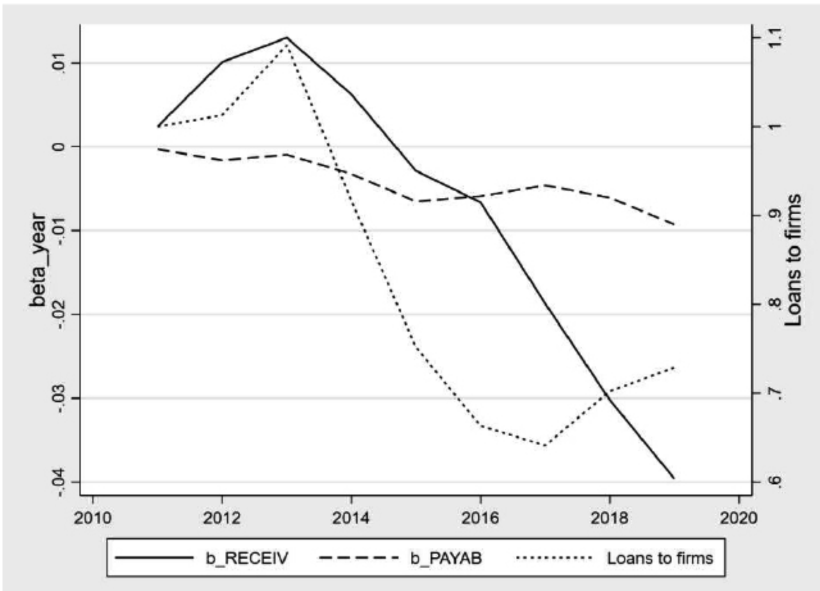
As can be seen in Tables A3 and A4 in the Appendix, the patterns of determinants of accounts receivable and accounts payable are robust to the industry disaggregation, with minor differences. Indeed, the motives behind a firm's decision to offer trade credit seem to be similar regardless of the firm's industry.

4.3. Bank lending constraints and trade credit

After the 2008 global financial crisis and the 2010–2014 Eurozone debt crisis, bank lending has been in sharp decline—by 41% between 2013 and 2017 (Figure 2).[5] Year dummies allow us to test whether firms experiencing bank lending constraints are more likely to use and grant trade credit.

[5] The severe recession following the 2008 global financial crisis left numerous European banks with non-performing loans (NPLs). In reaction, the European Banking Authority (EBA) and the European Central Bank (ECB) increased the banks' capital requirement in 2011 and deployed a series of actions to strengthen the prudential supervision of credit institutions in the Eurozone, namely the creation of a Single Supervisory Mechanism of banks in 2013 and the adoption of EBA definition of NPLs for the assessment of bank health in 2015 (Blattner et al., 2023).

Figure 2. Estimated coefficients of year dummies and bank loans to firms



Notes: b_RECEIV and b_PAYAB report the estimated coefficients of year dummies of accounts receivable and accounts payable estimations, respectively. Loans to firms is the index (2011=1) of the amount of bank loans to firms (source: Banco de Portugal).

We would expect that when bank credit is rationed, firms are more likely to use other sources of finance, including trade credit (Danielson and Scott, 2004; Casey and O'Toole, 2014). As can be seen in Figure 2, surprisingly, there is a slight negative impact on the financing received from suppliers over the decade. In particular, all else being constant, the accounts payable ratio decreases by 0.9 percentage points from 2010 to 2019. However, it should be noted that the amount of accounts payable depend not only on the demand of the firm but also on the supply of trade credit to the firm (Petersen and Rajan, 1997), which has decreased significantly over the decade (Figure 1).

We also find that there is a positive effect of the year on accounts receivable during the crisis, followed by an increasingly negative effect after the crisis period—all else remaining constant, the accounts receivable ratio falls by 5.3 percentage points from 2013 to 2019. Therefore, as bank credit shrinks, firms appear to extend less trade credit to their customers, which is consistent with the redistribution view of trade credit provision, whereby bank credit is redistributed from financially stronger firms to weaker firms via trade credit (Love et al., 2007).

5. **Conclusion**

This paper provides an empirical examination of the determinants of the trade credit policies of Portuguese firms. Using panel data of small and medium-sized firms for the period 2010–2019, we find that trade credit is an important source of external resources, financing about 18.3% of assets vis-à-vis to 15.8% and 8.3% from long-term and short-term financial debt, respectively. Portuguese SMEs also invest about a quarter of their sales in accounts receivable.

Our results seem to support the theory of financial motives for the use of trade credit. Indeed, we find that firms that have easier access to external financing at lower costs grant more trade credit to their customers (financial intermediation), while they rely less on trade credit from their suppliers (substitution effect). Firms that generate more (positive) internal resources also extend more credit to their customers and receive less credit from their suppliers. Moreover, firms with growing sales tend to rely on the support of their suppliers to finance new investments, namely in inventories, while firms whose sales decline receive less financial support from suppliers, which seems to provide further arguments for the theory that suppliers act as typical financial intermediaries.

The data do not seem to support the hypothesis of the use of trade credit as a way of transmitting information about the quality of the firm's products. However, in the case of the price discrimination theory, trade credit seems to be an appropriate marketing tool. Moreover, firms facing a decline in sales and negative internal financing respond by increasing the credit granted to customers in an attempt to stem falling sales.

Our results also show that the trade credit decisions made by firms are generally influenced by the same factors, regardless of the industry in which they operate. Finally, we find an increase in credit extended to customers to at the peak of 2008 financial crisis, followed by a subsequent collapse of this source of financing right after the crisis, which appears to mimic the contraction in bank credit in the Portuguese economy.

References

Biais, B.; Gollier, C. (1997) Trade credit and credit rationing. *Review of Financial Studies,* 10(4), 903-937.

Blattner, L.; Farinha, L.; Rebelo, F. (2023) When losses turn into loans: The cost of weak banks. *American Economic Review,* 113(6), 1600-1641.

Brennan, M.; Maksimovic, V.; Zechner, J. (1988) Vendor financing. *Journal of Finance,* 43, 1127-1141.

Burkart, M.; Ellingsen, T. (2004) In-kind finance: a theory of trade credit. *American Economic Review,* 94(3), 569-590.

Carreira, C.; Silva, F. (2010) No deep pockets: Some stylized empirical results on firms' financial constraints. *Journal of Economic Surveys,* 24(4), 731-753.

Carreira, C.; Teixeira, P.; Nieto-Carrillo, E. (2022) Recovery and exit of zombie firms in Portugal. *Small Business Economics,* 59(2), 491-519.

Casey, E.; O'Toole, C.M. (2014) Bank lending constraints, trade credit and alternative financing during the financial crisis: Evidence from European SMEs. *Journal of Corporate Finance,* 27, 173-193.

Cuñat, V.; García-Appendini, E. (2012) Trade Credit and its Role in Entrepreneurial Finance. In D. Cumming (ed.), *The Oxford Handbook of Entrepreneurial Finance.* Oxford, Oxford University Press, 526-557.

Danielson, M. G.; Scott, J.A. (2004) Bank Loan Availability and Trade Credit Demand. *Financial Review,* 39, 579-600.

Delannay, A.-F.; Weill, L. (2004) The determinants of trade credit in transition countries. *Economics of Planning,* 37,173-193.

Emery, G.W. (1987) An optimal financial response to variable demand. *Journal of Financial and Quantitative Analysis*, 22, 209-225.

Ferris, J.S. (1981). A transactions theory of trade credit use. *Quarterly Journal of Economics,* 96(2), 243-270.

Garcia-Teruel, P.J.; Martinez-Solano, P. (2010) Determinants of trade credit: A comparative study of European SMEs. *International Small Business Journal,* 28(3), 215-233.

Giannetti, M. (2003) Do better institutions mitigate agency problems? Evidence from corporate finance choices. *Journal of Financial and Quantitative Analysis,* 38(1), 185-212.

Huyghebaert, N. (2006) On the determinants and dynamics of trade credit use: Empirical evidence from business start-ups. *Journal of Business Finance and Accounting,* 33(1 & 2), 305-328.

Lee, Y.W.; Stowe, J.D. (1993) Product risk, asymmetric information, and trade credit. *Journal of Financial and Quantitative Analysis*, 28(2), 285-300.

Long, M.S.; Malitz, I.B.; Ravid, S.A. (1993) Trade credit, quality guarantees, and product marketability. *Financial Management*, 22(4), 117-127.

Love, I.; Preve, L.A.; Sarria-Allende, V. (2007) Trade credit and bank credit: Evidence from recent financial crises. *Journal of Financial Economics,* 83, 453-469.

Meltzer, A.H. (1960) Mercantile credit, monetary policy, and size of firms. *Review of Economics and Statistics,* 42, 429-437.

Morris, J.R. (1976) On the Corporate Debt Maturity Strategies. *Journal of Finance,* 31, 29-37.

Ng, C.K.; Smith, J.K.; Smith, R.L. (1999) Evidence on the Determinants of Credit Terms Used in Interfirm Trade. *Journal of Finance,* 54(3), 1109-1129.

Niskanen, J.; Niskanen, M. (2006) The determinants of corporate trade credit policies in a bank-dominated financial environment: The case of Finnish small firms. *European Financial Management,* 12(1), 81-102.

Petersen, M.A.; Rajan, R.G. (1997) Trade credit: Theories and evidence. *Review of Financial Studies,* 10(3), 661-691.

Schwartz, R. (1974) An economic model of trade credit. *Journal of Finance and Quantitative Analysis,* 9(4), 643-657.

Smith, J.K. (1987) Trade credit and informational asymmetry. *Journal of Finance,* 42(4), 863-872.

Appendix

Table A1: Descriptive statistics

Variables	N	Mean	Std. Dev.	Min.	Max.
RECEIV	488,694	0.252	0.263	0	1
PAYAB	488,694	0.183	0.195	0	1
lnSIZE	488,694	14.020	1.457	0	17.577
lnAGE	488,694	2.646	0.862	0	5.323
CFLOW1	488,694	0.033	0.622	-4.023	3.169
CFLOW2	488,694	0.045	0.166	-0.898	0.475
FCOST	488,694	0.024	0.032	0	0.192
GROWTH	423,467	0.132	0.726	-1.000	5.431
STLEV	488,694	2.278	10.190	0	87.910
STFIND	488,694	0.083	0.150	0	1
LTDEBT	488,694	0.158	0.221	0	1
TURN	488,694	2.729	3.980	0.000	27.290
MARGIN	488,694	0.097	0.615	-2.999	4.063
CURRAS	488,694	0.671	0.281	0	1

Notes: Pooled yearly values, 2010–2019. RECEIV – ratio of accounts receivable to sales; PAYAB – ratio of accounts payable to total assets; lnSIZE – log of assets; lnAGE – log of firm age; CFLOW1 and CFLOW2 – ratio of net profits plus depreciation to sales and assets, respectively; FCOST – ratio of finance costs to financial debt; GROWTH – annual sales growth rate; STLEV – ratio of current liabilities to sales; STFIND – ratio of short-term financial debt to assets; LTDEBT – ratio of long-term debt to assets; TURN – turnover of sales over assets; MARGIN – ratio of gross profit to sales; CURRAS – ratio of current assets to total assets.

Table A2: Correlation across covariates

Variable	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
[1] RECEIV	1												
[2] PAYAB	0.040*	1											
[3] lnSIZE	0.208*	-0.110*	1										
[4] lnAGE	0.119*	-0.109*	0.318*	1									
[5] CFLOW1	-0.051*	-0.081*	0.059*	0.027*	1								
[6] CFLOW2	-0.07*	-0.209*	0.010*	0.023*	0.331*	1							
[7] FCOST	0.030*	0.121*	0.060*	0.068*	-0.018*	-0.019*	1						
[8] GROWTH	-0.104*	-0.037*	0.053*	-0.154*	0.048*	0.065*	-0.027*	1					
[9] STLEV	0.167*	-0.089*	0.179*	-0.026*	-0.206*	-0.081*	-0.065*	-0.034*	1				
[10] TURN	-0.071*	0.287*	-0.370*	-0.205*	-0.019*	-0.039*	0.013*	0.029*	-0.129*	1			
[11] MARGIN	-0.007*	-0.139*	0.142*	0.017*	0.890*	0.308*	-0.001	0.042*	0.009*	-0.074*	1		
[12] STFIND	0.042*	0.010*	0.137*	0.054*	-0.068*	-0.116*	0.123*	-0.017*	0.113*	-0.047*	-0.047*	1	
[13] LTDEBT	-0.037*	-0.136*	0.096*	-0.029*	-0.091*	-0.149*	0.034*	0.044*	0.018*	-0.151*	-0.020*	-0.120*	1
[14] CURRAS	0.163*	0.254*	-0.198*	-0.032*	-0.105*	-0.008*	0.024*	-0.005*	-0.087*	0.315*	-0.158*	-0.042*	-0.270*

Notes: See notes to Table A1. * Statistical significance at the 0.05 level.

Table A3: Determinants of accounts receivable by industry

Variables	Agriculture (1)	Manufacturing (2)	Utilities & Transportation (3)	Construction (4)	Trade & Accommodation (5)	Other Services (6)
lnSIZE	0.060*** (0.006)	0.089*** (0.001)	0.091*** (0.003)	0.071*** (0.002)	0.047*** (0.001)	0.056*** (0.002)
lnAGE	0.044*** (0.013)	0.020*** (0.003)	0.002 (0.007)	0.040*** (0.007)	0.013*** (0.002)	0.014** (0.006)
CFLOW1	-0.034*** (0.006)	-0.050*** (0.005)	0.019*** (0.007)	-0.014*** (0.003)	-0.055*** (0.003)	-0.019*** (0.002)
FCOST	-0.189** (0.081)	-0.282*** (0.017)	-0.201*** (0.038)	-0.404*** (0.031)	-0.169*** (0.011)	-0.190*** (0.034)
PGROWTH	-0.006*** (0.002)	-0.006*** (0.001)	-0.001 (0.002)	-0.001 (0.001)	-0.002*** (0.001)	-0.006*** (0.001)
NGROWTH	0.146*** (0.011)	0.206*** (0.003)	0.243*** (0.008)	0.191*** (0.004)	0.163*** (0.002)	0.136*** (0.004)
STLEV	0.001*** (0.000)	0.004*** (0.000)	0.005*** (0.000)	0.003*** (0.000)	0.004***(0.000)	0.003*** (0.000)
TURN	0.009*** (0.001)	0.010*** (0.000)	0.006*** (0.000)	0.006*** (0.000)	0.003*** (0.000)	0.005*** (0.000)
MARGIN	0.037*** (0.006)	0.034*** (0.005)	-0.009 (0.007)	0.015*** (0.003)	0.048*** (0.003)	0.018*** (0.002)
Constant	-0.710*** (0.087)	-1.021*** (0.019)	-1.064*** (0.043)	-0.817*** (0.034)	-0.545*** (0.014)	-0.601*** (0.033)
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	14,443	114,234	23,033	58,348	143,068	70,341
No. of firms	2,783	19,083	4,080	13,731	27,058	15,435
R^2 (overall)	0.083	0.109	0.122	0.146	0.210	0.081
F statistic	35.26***	749.11***	146.09***	423.90***	769.98***	267.59***

Notes: Fixed-effects regression of Model (1). See notes to Table 1. Coefficients of time (2010–2019) dummies not reported. Firm-cluster robust standard errors are given in parentheses. ***, ** and * statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table A4: Determinants of accounts payable by industry

Variables	Agriculture (1)	Manufacturing (2)	Utilities & Transportation (3)	Construction (4)	Trade & Accommodation (5)	Other Services (6)
lnSIZE	0.004* (0.002)	-0.029*** (0.001)	-0.010*** (0.002)	0.001 (0.001)	-0.024*** (0.001)	-0.010*** (0.001)
lnAGE	-0.020*** (0.005)	-0.025*** (0.002)	-0.017*** (0.006)	-0.024*** (0.004)	-0.047*** (0.002)	-0.004 (0.003)
CFLOW2	-0.106*** (0.007)	-0.174*** (0.003)	-0.211*** (0.007)	-0.201*** (0.004)	-0.191*** (0.003)	-0.120*** (0.003)
FCOST	0.170*** (0.032)	0.223*** (0.012)	0.305*** (0.029)	0.162*** (0.016)	0.291*** (0.012)	0.135*** (0.015)
PGROWTH	-0.000 (0.001)	0.009*** (0.001)	0.007*** (0.002)	0.002*** (0.001)	0.003*** (0.001)	0.001** (0.000)
NGROWTH	-0.017*** (0.004)	-0.048*** (0.002)	-0.046*** (0.006)	-0.021*** (0.002)	-0.074*** (0.003)	-0.013*** (0.002)
STFIND	-0.056*** (0.007)	-0.074*** (0.003)	-0.116*** (0.007)	-0.069*** (0.005)	-0.124*** (0.003)	-0.039*** (0.003)
LTDEBT	-0.060*** (0.006)	-0.085*** (0.003)	-0.130*** (0.006)	-0.079*** (0.004)	-0.103*** (0.003)	-0.040*** (0.003)
CURRAS	0.046*** (0.006)	0.059*** (0.003)	0.094*** (0.006)	0.064*** (0.004)	0.076*** (0.003)	0.048*** (0.003)
Constant	0.118*** (0.033)	0.645*** (0.013)	0.373*** (0.032)	0.200*** (0.018)	0.679*** (0.013)	0.256*** (0.014)
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	14,443	114,234	23,033	58,348	143,068	70,341
No. of firms	2,783	19,083	4,080	13,731	27,058	15,435
R^2 (overall)	0.106	0.103	0.213	0.110	0.137	0.116
F statistic	25.49	609.38	124.57	197.89	741.31	154.83

Notes: Fixed-effects regression of Model (2). See notes to Table 1. Coefficients of time (2010–2019) dummies not reported. Firm-cluster robust standard errors are given in parentheses. ***, ** and * statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Nuno Gonçalves
Univ Coimbra, Faculty of Economics
npocinhogoncalves@outlook.pt
Carlos Carreira
Univ Coimbra, CeBER, Faculty of Economics
ccarreir@fe.uc.pt
orcid.org/0000-0002-4786-5605

DOI: https://doi.org/10.14195/2183-203X_56_5

Intangible Capital and Productivity of Portuguese Firms in the Last Decade (2010-2019)

Capital Intangível e Produtividade das Empresas Portuguesas na Última Década (2010-2019)

Nuno Gonçalves
Carlos Carreira

Received for publication: April 15, 2023
Revision accepted for publication: June 25, 2023

ABSTRACT
This article analyzes the effects of intangible capital on the productivity of Portuguese firms in the last decade. Intangible assets can increase the productivity of labor and productive factors. Although no consensus has been reached on standard principles and uniform methods for measuring intangible assets, the attempts of various investigators, such as those proposed in this research, pave the way for the development of a framework. To achieve this objective, a Cobb-Douglas production function was estimated at the firm level, where intangible capital is assumed as a productive factor. To perform a sectoral analysis, the model was estimated by activity sector. We also estimated the evolution of the contribution of intangible capital in two distinct periods 2010-14 (recession) and 2015-19 (recovery). The results obtained were to some extent expected, confirming the evidence of the positive effect of intangible assets on productivity. The intangible effect is greater in the sectors of Manufacturing and Construction, and inside the Manufacturing sector, the Textile industry is where the effect is larger. For the Trade and Business Services sector the effect is negative or null. Although intangible capital has a strong influence at the aggregate level, it has gradually lost its relevance. This result is understandable, given the low and decreasing levels of intangible investment and the continuous decrease in intangible capital during the decade.
Keywords: intangible capital; intangible investment; productivity; Cobb-Douglas production function; Portugal.

JEL Classification: C33; D24; L60; L80; O34.

Acknowledgment: We would like to thank Inês Milheiras for her contribution to the literature review section as a scholarship of the "Summer with Science 2022" program, funded by the Foundation for Science and Technology (FCT). Funding for this work was provided by the FCT, R&D unit CeBER (UIDB/05037/2020).

Nuno Gonçalves
Carlos Carreira

Intangible Capital and Productivity of Portuguese Firms in the Last Decade (2010-2019)

1. Introduction

In today's economies, intangible assets play a central role in improving the competitiveness and growth of firms and, consequently, tangible assets are no longer as determinant as they once were. Thus, intangible investment became an essential prerequisite for technological progress (Yallwe and Buscemi, 2014).

By nature, an intangible asset is a non-physical asset, examples being licenses, designs, patents, copyrights, software, marketing, branding, organizational or human capital, as well as R&D. Measuring such assets, however, is a challenge, even today. Because there is a lack of physical substance, firms typically do not properly report intangible assets on their balance sheets, so that quantifying their impact on productivity is difficult or inaccurate. Intangible investments are financially constrained, especially R&D investment, seeing that the results are unpredictable and drawn out. The funding of projects is also accompanied by adverse selection, moral hazard, and information asymmetry, and due to their nature, such assets are rarely used as collateral (Silva and Carreira, 2012).

This study analyzes the importance of intangible assets, especially in relation to the way they may have affected the productivity of Portuguese firms in the last decade. This work is important for entrepreneurs to better understand the potential benefits of intangible assets for output growth and to identify differences between intangible assets and performance in different sectors.

2. Related Literature

The role as a principal driver of economic growth was assigned by Solow (1957) to productivity, along with capital and labor force accumulation. Later, neoclassical growth models tried to explain how productivity grows endogenously, including R&D investment and other intangibles as a main source (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992). The importance of intangible assets as a fundamental component for productivity growth is underlined in many macroeconomic studies (Corrado et al., 2009; Corrado et al., 2013; Corrado et al., 2016), as well as in many microeconomic studies (Marrocu et al., 2012; Niebel et al., 2017; Piekkola, 2018; Criscuolo et al., 2021).

Corrado et al. (2009) estimated the effect of intangible assets at the macro level and their importance for economic growth, finding that intangible investment nurtured labor productivity by 0.84 percentage points (p.p.) in the United States of America. In European countries the effect was smaller but still significant: in the UK by 0.58 p.p., in Germany by 0.53 p.p., in Italy by 0.34 p.p., and in Spain by 0.19 p.p. From the mid-1990s to the period of the 2008 financial crisis, the USA's stronger labor productivity growth over that of the European countries is explained by Europe's low levels of investment in intangible assets (Corrado et al., 2013). Despite the low levels of intangible investment in these countries, the elasticities of intangible investment productivity for 10 countries (Austria, Czech Republic, Denmark, Spain, Finland, France, Germany, Italy, Netherlands, and UK) were greater than their respective share factor (Niebel et al., 2017). The work of Niebel et al., (2017) in conjunction with the study conducted by Corrado et al. (2016)

pointed to the fact that the elasticities found at the aggregate level are greater than those found by sector and differ markedly between the manufacturing industries and service sectors. Naturally, there is a heterogeneous effect of intangibles across firms. On average, there is a gap of one-third in labor productivity between firms in the top 10% of more productive firms and those in the 40-60 percentile. This could be explained by the frontier firms' greater use of a highly skilled workforce that is more creative and innovative, while the remaining firms engage in more routine work and employ a less-qualified labor force (Criscuolo et al., 2021).

Marrocu et al., (2012) investigated the impact of intangible capital on the productivity level of firms in a 6-country European panel (France, Italy, Netherlands, Spain, Sweden, and the United Kingdom), between 2002 and 2006, relying on the companies' balance sheets. They estimated a Cobb-Douglas production function, finding a highly significant effect of intangible capital on productivity. The intangible capital effect on productivity is still less than the physical capital effect, roughly half of the latter; nevertheless, the impact intangible capital has on a firm's performance is still relevant. During the post-crisis period of 2008 to 2013, for EU-28, Piekkola (2018) did not find a strong effect of intangible capital on labor productivity at the sectoral level. More noteworthy was the conclusion that intangible capital negatively affected the labor productivity growth during the period.

Using the Community Innovation Survey database from 2006 to 2018, Roth et al., (2022) estimated a production function for German firms. For the first time, intangible investment equaled the tangible investment in Germany. The positive impact of intangibles on the firm-level productivity is mainly driven by non-R&D intangibles, such as software and databases, training, advertising, and marketing. The study highlighted the fact that the impact of non-R&D intangibles on firm-level productivity was stronger in the services sector than in that of production, but on other hand, R&D is a strong driver of productivity, specifically in high-tech industries. Also, using a panel of data for the German industry during the period between 2006-2010, Crass and Peters (2014) drew several conclusions about the relation between intangible assets and productivity. They found that R&D, Brand, and Human Capital had significant positive effects on productivity. The most interesting findings were: (i) short-run productivity is increased with training expenditure, which is stronger than an increase in R&D or marketing expenditure; (ii) however, a firm's stock of patents granted, and trademarks slightly increases long-run productivity; (iii) lastly, it was found that the patent stock and skilled labor force, like patent stock and marketing, are complements. Companies belonging to the high-tech industry exhibit a certain degree of complementarity between different types of intangible assets and stable knowledge accumulation, which has a greater effect on technical efficiency (Turovets, 2021).

A few studies were conducted on the Portuguese reality, specifically exploring the regional spillover effects (Carreira & Lopes, 2018). Nunes and Almeida (2009) allude to a quadratic relationship between intangible assets and growth in Portuguese SMEs. The level of intangible assets is only a catalyst factor for growth in Portuguese firms at high levels of intangible assets, being limited for low levels of intangible assets. Other studies point out that intangible assets, in conjunction with net income, goodwill and other intangible assets, are highly important to the value of stock prices. Intellectual property and R&D investment, however, are not value-relevant factors for shareholders (Oliveira et al., 2010). The profitability of Portuguese SMEs neither increased nor diminished with an increased investment on intangible assets between 2001 and 2009 (de Carvalho et al., 2013).

3. Model Specification, Estimation Technique and Data

3.1. Model specification

We estimate the following Cobb-Douglas production function of firm i at time t:

$$Y_{i,t} = A_{i,t} M_{i,t}^{\alpha} K_{i,t}^{\beta} L_{i,t}^{\gamma} I_{i,t}^{\delta} \quad (1)$$

where $Y_{i,t}$ represents the gross output, $A_{i,t}$ denotes the total factor productivity (TFP), $M_{i,t}$ the inputs or intermediate consumptions, $K_{i,t}$ the physical capital, $L_{i,t}$ the labor and $I_{i,t}$ intangible capital. We do not impose any restriction on the elasticity parameters (i.e., we do not consider $\alpha + \beta + \gamma + \delta = 1$). When we log-normalize equation (1), we get the following equation:

$$Y_{i,t} = a_{i,t} + \alpha m_{i,y} + \beta k_{i,t} + \gamma l_{i,t} + \delta i_{i,t} \quad (2)$$

We consider the productivity term $a_{i,t}$ to be composed of a common factor z, and by an unobservable productivity term $p_{i,t}$ known by the company. We add a time dummy d_t designed to capture the macroeconomic effects, which vary over time but not across firms; productivity is composed of a vector of control variables $x_{i,t}$, and by an error term $\epsilon_{i,t}$. This gives us the following final equation:

$$y_{i,t} = z + p_{i,t} + d_t + \alpha m_{i,y} + \beta k_{i,t} + \gamma l_{i,t} + \delta i_{i,t} + \theta x_{i,t} + \epsilon_{i,t}. \quad (3)$$

3.2. Estimation technique

The estimation of equation (3) is likely to suffer from endogeneity. The endogeneity problem arises from the fact that consumption demand functions are determined by the firm's knowledge of its own productivity level. When choosing inputs, firms try to identify the last impact of inputs on productivity, thus adjusting inputs for each new production. So, inputs will be correlated with productivity and hence the error term in the productivity equation. Another issue to consider when deciding which estimator to use is to consider that firms do not adjust their investment policy every year, which translates into several zeros in the investment. To circumvent this problem, Levinsohn and Petrin (hereafter LP) (2003) proposed a two-stage semi-parametric method, using intermediate consumptions as a proxy for productivity; it is less costly to adjust intermediate inputs to the productivity shocks than to redefine investment policy. This is the main approach used in many studies (Marrocu et al., 2014; Crass & Peters, 2014; Roth et al., 2022). Other methodologies, like that of

Olley & Pakes (1992), are commonly used, although Eberhardt and Helmers (2010) alert to the fact that these estimators are conceptually quite similar for estimating Cobb-Douglas production functions, but the choice of the estimation method might influence the empirical results. The STATA command *prodest* was used for LP estimations (Rovigatti & Mollisi, 2018).

3.3. Data

The database used in this study was prepared by the researchers of the ENtRY project (funded by FCT-Fundação para a Ciência e Tecnologia, PTDC/EGEECO/31117/2017), and was extracted from the Sistema de Contas Integradas das Empresas (SCIE), administered by the Instituto Nacional de Estatística (INE). The final sample consists of an unbalanced panel of 511 687 active companies operating in Portugal, taken from the manufacturing and service industries, excluding public services, the financial sector, and social services, for the period 2010-2019. The *tobacco industry*, *manufacturing of petroleum products*, *pharmaceuticals*, and *other transport equipment*, as well as *air transport* and *water collection, treatment and distribution* were later excluded, due to the small number of observations in these categories of economic activity.

Gross output is measured as the value of sales of goods and services, less the value of purchases of goods for resale, so it is adjusted for the change in stocks of final goods and other operating income. This variable was deflated by the two-digit industry-level producer price index obtained from INE. Labor is the 12-month average of employment. Intermediate consumption includes the cost of materials and services purchased and was deflated by the GDP deflator index. The stock of tangible and intangible capital was obtained by applying the perpetual inventory method, considering the respective values of the annual investment. For the first year of a firm's time series, the book value of tangible and intangible assets was deflated by the GFCF deflator and the GDP deflator, respectively, to derive the capital stock. For subsequent years, investments are added, and depreciation rates are subtracted yearly (10% for tangible capital and 33.33% for intangible capital). As Crass and Peters (2014) have shown that productivity estimation results based on intangible capital stocks and intangible investment expenditures are almost identical[1], we use the firms' investment expenditures in training, R&D, software, industrial property, and other intangible assets, as their intangible investments. For control variables, we use dummies for exporting firms, business cycle variation (change in real GDP) to capture the effects of the Great Depression, the age of the firm, the dimension class of the firm and industry dummies (see Table A.1 in annex). All monetary variables are measured in constant euros from 2016. For the dimension class of firms, the number of employees was considered, a firm being classified as micro if it has fewer than 10 employees, as small if it has more than 10 employees, as medium if it has more than 50 employees and as large if it has more than 250 employees. In line with those parameters, we have 435 774 (85.16%) micro firms, 65 066 (12.72%) small firms, 9 444 (1.85%) medium-sized, and 1 403 (0.27%) large firms.

[1] This suggests that the amount of investment for a specific intangible is a very good proxy for the firm's capital stock of this intangible.

Table 1 presents the descriptive statistics of the main variables presented previously. On average, the output of Portuguese firms is 660 thousand euros, with almost 9 workers per firm. On average, Portuguese firms spend 423 thousand euros on intermediate consumptions. The mean value for tangible capital is 294 thousand euros, which is higher than intangible capital, with 37.5 thousand. For investments, the mean is higher for tangible assets, 45.9 thousand euros, compared to intangible investments of 6.7 thousand. On average, the firms report lower expenses in training, and more on other types of intangible capital. In our sample, and, as was evidenced by Kaus et al., (2020) for the German reality, in Portugal many firms invest nothing or very little, but a few invest large amounts in intangibles, so the variable of investment in intangible capital is highly right skewed.

Table 1: Descriptive statistics of the main variables

Variables	Mean	Standard-Deviation	Minimun	Maximum (in th)	Skewness	Kurtosis
Output	660 000	10 300 000	1	3 784 000	121.850	25 116
Labour	8.84	98	1	26.857	133.326	26 619
Intermediate consumptions	423 000	7 340 000	1	3 211 000	132.032	33 644
Tangible capital	294 000	8 070 000	1	3 263 000	248.267	87 714
Intangible capital	37 560	3 900 000	1	3 285 000	406.245	255 5148
Investment in Tangible Capital	45 900	1 370 000	0	707 008	201.816	60 405
Investment in Intangible Capital	6 697	595 000	0	432 300	299.544	471
Training	378.87	11 028	0	7 800	223.364	064
R&D	676.53	80 103	0	67 317	458.994	621
Software	1 272.26	156 000	0	96 076	344.473	431
Industrial Property	1 386.65	242 000	0	161 000	383.054	185 062
Other Intangible Capital	2 640.97	434 000	0	432 300	510.289	397 795

Notes: The number of observations is 2 795 705; the values correspond to raw state.

4. **Data Analysis, Results, and Discussion**

4.1. Preliminary analysis

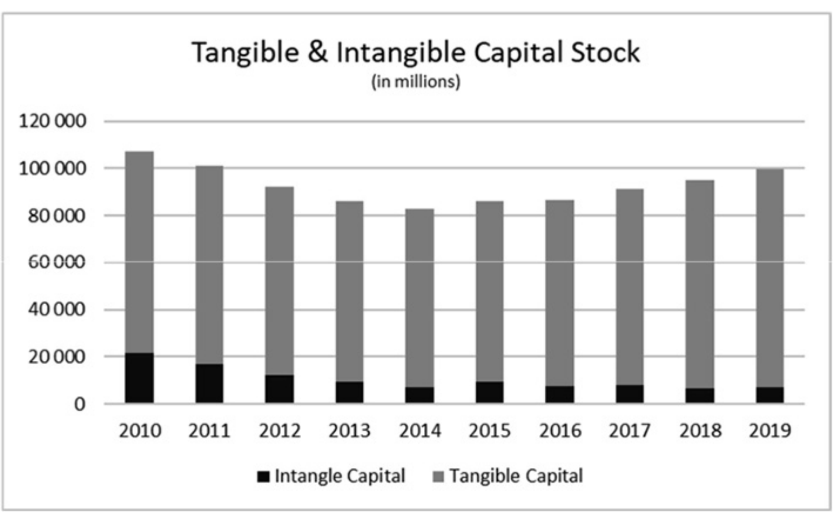
This section presents a first analysis of the data carried out with the aim of presenting some stylized facts about intangible investment at the firm level. We start by analyzing the number of firms that invested in intangibles during the period between 2010 and 2019 (see Table 2). Only 91 972 firms, corresponding to 17.97% of the sample, invested in intangibles during the last decade; 16.45% of firms invested in employee training, which is the most common type of investment, and in second place was software investment, with 14.50% of the firms involved. The type of investment least utilized was in industrial property, with only 5.65% of the firms employing it, and there were similar results in R&D investment (5.96% of firms) and other types of intangible investments (6.74%). The values presented in Table 2 tell us that a minority of Portuguese firms invested in intangible assets during the last years.

Table 2: Number of firms that invested in intangibles, or in a certain type, between 2010 and 2019

The firm invested?	Intangible Investment	Training	R&D	Software	Industrial Property	Other Intangible
Yes	91 972 (17.97%)	84 159 (16.45%)	30 508 (5.96%)	74 186 (14.50%)	28 899 (5.65%)	34 512 (6.74%)
No	419 715 (83.02%)	427 528 (83.55%)	487 719 (94.04%)	437 501 (85.50%)	482 788 (94.35%)	477 175 (93.26%)

Next, we analyze the tangible and intangible capital stock for Portuguese companies by year (in Graph 1). We find that tangible and intangible stock between 2010-2014 was in a downward trend, and in the following years there was a recovery, but not enough to reach the levels of 2010 and 2011. The levels of intangible capital stock decreased during the period, as the intangible stock in 2019 was less than half the value in 2010. Here we have a clear downward trend, with the tangible capital stock increasing in the second half of the period.

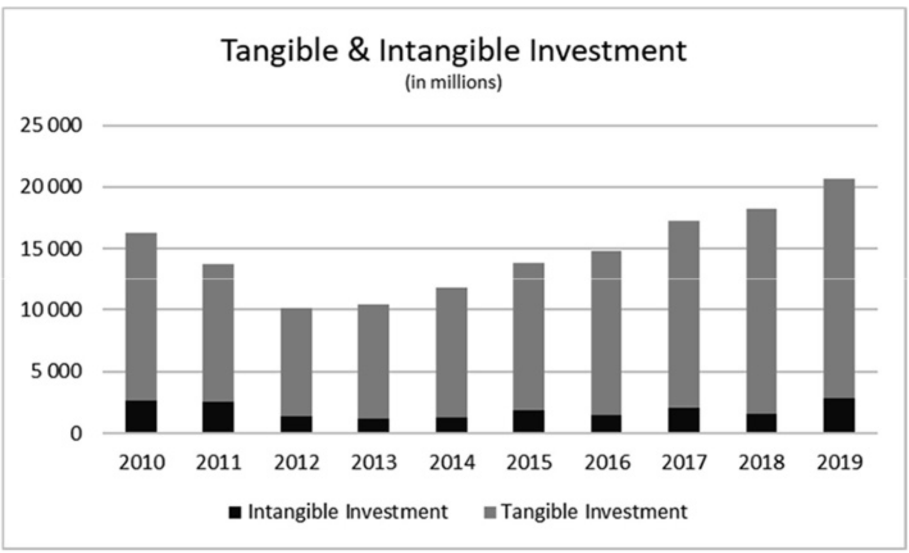
Graph 1: Tangible and intangible capital stock per year



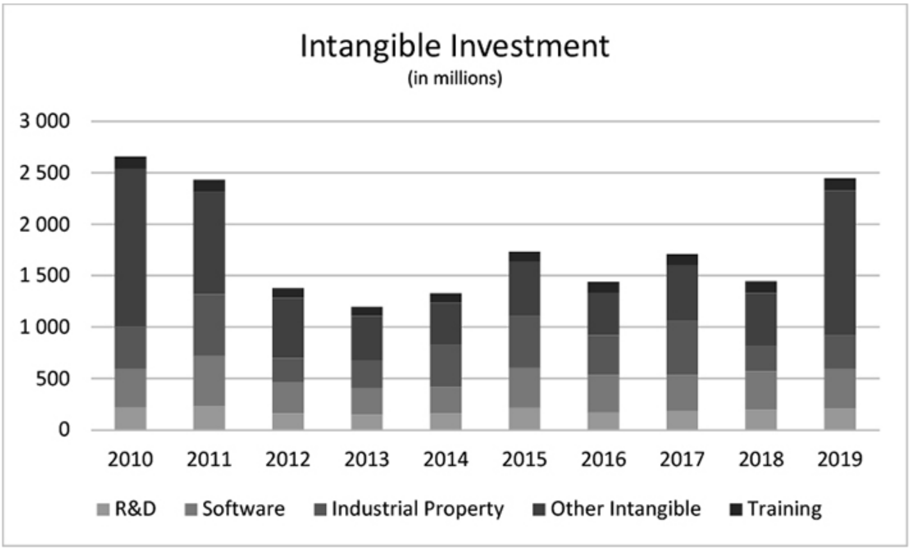
The investment of Portuguese companies had a "valley shape" behavior (in Graph 2), hitting bottom in 2012, with a recovery starting the next year. The 2017 investments exceeded the investments made in 2010. The largest investments in intangibles were made in 2010, 2011 and 2019.

In Graph 3 we see the evolution of investment in intangibles by type and note that investment in other intangibles is the largest item, while the spending on training was the smallest. In this graph we can better analyze the evolution of the total investment in intangibles and note some fluctuations between 2012 and 2018. The investment in intangibles in 2019 reached the amount of 2011 but not of 2010.

Graph 2: Tangible and intangible investment per year

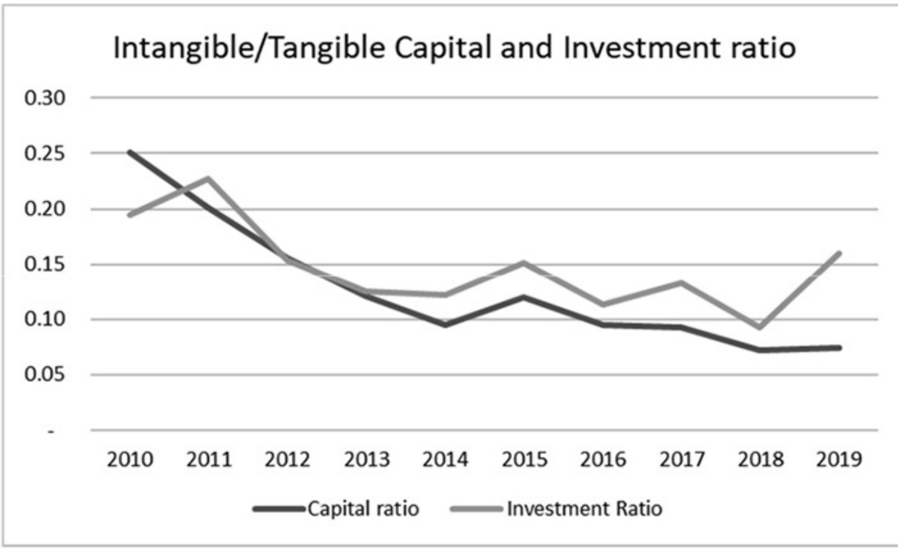


Graph 3: Evolution of intangible investment



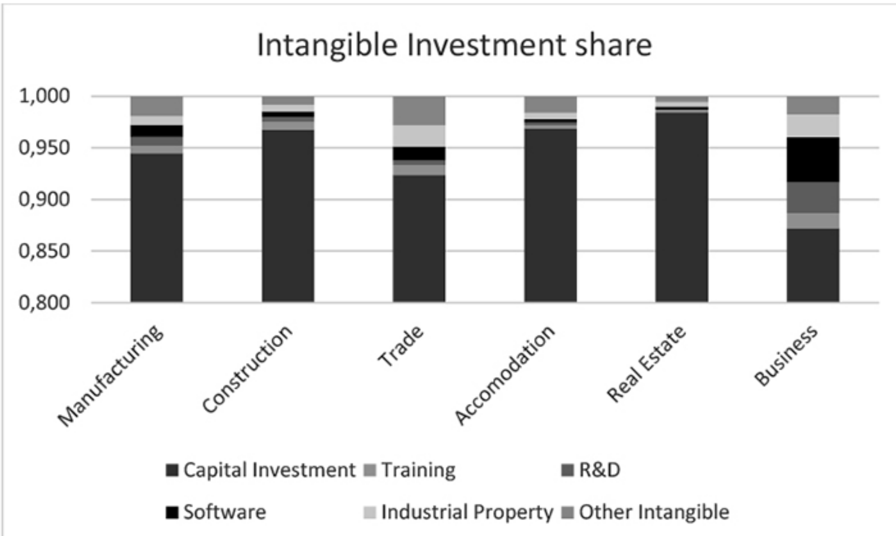
The evolution of investment and capital ratio between intangible and tangible can be seen in Graph 4. In both cases we have a descending trend, with a slight recovery in the investment ratio in 2019. The investment ratio starts slightly below 0.20 and ends above 0.15, with some fluctuation, and throughout the period in question we have a clear favoring of tangible investment over non-tangible investment. Regarding investments, the growth rates of tangibles were higher, compared to intangible growth rates; only in 4 years (2011, 2015, 2017 and 2019) the intangible growth rates were larger than tangible growth rates, but not the amount invested. The values of the capital ratio have a greater amplitude, starting at 0.25 and standing at 0.075, always falling, due to lower growth rates of investments in intangibles relative to tangibles, and the replaced intangible capital was not enough to recover depreciated capital (intangibles have higher depreciation rates than tangibles). The values of the capital ratios are very low even compared to other advanced economies, like that of Germany (see Roth et al., 2022) and USA (see Nakamura, 2010) which records ratios around 1.

Graph 4: Evolution of ratio of intangible over tangible for capital and investment



Finally, we analyzed the investment made by the companies in each sector (in Graph 5). In all sectors, the fixed capital factor share is the largest. The share of intangible investment is the largest In the Business and Trade sectors, and only in Business Services is it greater than 10%. The share of intangible investments in Manufacturing was 5.5%, with the other sectors being Construction (3.2%), Trade (7.6%), Accommodation (3.1%), Real Estate (1.6%) and Business (12,8%). In the Manufacturing sector, the item other intangibles have the largest share within intangibles, and the same holds true for Construction, Trade, Accommodation and Real Estate, while in the Business Sector, the largest share of intangible investment belonged to software.

Graph 5: Share of intangible investment across industries



4.2. Results and discussion

In presenting the results, we discuss them briefly and compare them with other studies (Marrocu et al., 2012; Crass and Peters, 2014; Roth et al., 2022).

Table 3 shows the estimation results for the entire sample between 2010 and 2019 (regression 1), which is then divided into two sub-periods, the first 2010-2014 (regression 2) – the period corresponding to the Great Recession – and the subsequent period of Economic Recovery in 2015–2019 (regression 3). The results with the amount of investment in disaggregated intangibles are also presented corresponding to regressions 4, 5, and 6. The Wald test demonstrates the existence of returns to scale in all models. The elasticities across the various models are similar, with the elasticity of the labor factor ranging between 0.433 and 0.460, and of the intermediate consumption between 0.628 and 0.658. The elasticities of the tangible capital factor are always higher than that of intangible capital, with those of tangible capital varying between 0.029 and 0.042, and those of intangible capital ranging from 0.003 to 0.017. All elasticities have the magnitudes found in the literature mentioned above, except the one pertaining to intermediate consumptions, which has a larger coefficient. The elasticity of intangible capital is positive in all models; however, it is only shown to be significant in regressions 1 and 2, but not in regression 3, thus reflecting the lower and decreasing amount of intangible capital in the stock of Portuguese firms. Analyzing the disaggregated investment in intangibles, the training sector is the only one that shows positive and significant elasticity in all regressions; the other coefficients have negative elasticities and, in some cases, no statistical significance, and a coefficient close to zero.

Comparing our results with those of Marrocu et al., (2012) and Roth et al., (2022) in regard to these estimations, we find that our intangible capital coefficients are lower. In the study of Marrocu et al., (2012) the coefficient of intangible capital is 0.03 for France, 0.051 for Germany, 0.023 for Spain and 0.081 for the UK; the aggregate value is 0.038 in the four countries, and Roth et al., (2022) report a 0.034 intangible capital coefficient for Germany. Comparing our regressions with those of another country with results closer to the Portuguese reality, Spain in this case (Marrocu et al., 2012), the elasticity of tangible capital is 0.067, the elasticity of intangible capital is 0.023, and the elasticity of the labor factor is 0.381.

We interpret the low coefficients of capital, and negative in some cases for investment in intangibles, as a consequence of the low levels of investment made by Portuguese firms, which are not sufficient to increase their productivity and divert resources from other productivity-enhancing factors.

Table 3: Aggregate firm-level production function estimations for Portugal

Regression		(1)	(2)	(3)	(4)	(5)	(6)
Time period		2010-2019	2010-2014	2015-2019	2010-2019	2010-2014	2015-2019
Labour		0.448*** (0.002)	0.436*** (0.001)	0.460*** (0.001)	0.445*** (0.001)	0.433*** (0.001)	0.457*** (0.002)
Intermediate Consumptions		0.658*** (0.003)	0.651*** (0.002)	0.646*** (0.003)	0.628*** (0.001)	0.644*** (0.002)	0.631*** (0.006)
Tangible capital		0.042*** (0.004)	0.034*** (0.004)	0.029*** (0.003)	0.039*** (0.001)	0.042*** (0.006)	0.033*** (0.001)
Intangible Capital		0.017*** (0.005)	0.010*** (0.001)	0.003 (0.009)			
Investment	Training				0.009*** (0.000)	0.010*** (0.000)	0.008*** (0.000)
	R&D				-0.007*** (0.000)	-0.007*** (0.001)	-0.007*** (0.001)
	Software				-0.001** (0.000)	0.000 (0.000)	-0.002*** (0.000)
	Industrial Property				-0.006*** (0.001)	-0.007*** (0.001)	-0.006*** (0.000)
	Other Intangible Capital				-0.004*** (0.001)	-0.003** (0.001)	-0.006*** (0.000)
No. firms		340 760	256 893	281 737	340 760	256 893	281 737
No. observations		2 094 536	983 806	1 110 730	2 094 536	983 806	1 110 730
Wald test (χ^2 on CRS		38 931.47***	21 423.81***	19 287.35***	46 364.78***	26 444.74***	1 194.04***

Note: Standard errors in parentheses *p<0.10, **p<0.05, ***p<0.01. All estimates are calculated using the Levinshon and Petrin (2003) methodology. The estimation results include control for industry- and time- specific effects, as well as for firm age, dimension, exporter, and business cycle. CRS denotes constant returns to scale.

To better understand the impacts of production factors at the sector level, we also estimated production functions at the firm level for different sectors. The results of our estimation are presented in Table 4. The effects of intangible capital are smaller when compared to the results presented at the aggregate level (regressions 1 to 6). In some cases they have no statistical significance and/or are negative (regressions C1, T1, A1, RE1 and BS1); manufacturing is the only one in which the values are positive and significant (regression M1). The coefficients of investments in intangibles generally follow the signs of the aggregated models (regressions 4 to 6), being positive for training and negative for the other investment variables. We also highlight the positive effect of software investment in Trade (regression T2), Real Estate (regression RE2) and Business Services (regression BS2).

Moreover, the Accommodation (regression A2) and Real Estate (regression RE2) sectors have elasticities of labor that are relatively lower than the aggregate model, and the elasticities of intermediate consumption stand out from the other regressions by excess. Regarding tangible capital, the smallest elasticities are found in the Manufacturing industry (M2).

Finally, we disaggregate the manufacturing sector into its various component industries.

The lowest coefficient of labor elasticity is in the chemicals industry (regression MI11 and MI12), and largest in Textiles (regression MI3 and MI4); the elasticity of intermediate consumption varies between 0.474 in Textiles (regression MI4) and 0.736 in the Food industry (regression MI1); the tangible capital ranges between 0.006 in Chemicals (regression MI12) and 0.042 in Paper (regression MI10). In general, the coefficients of intangible capital and investment (Table 5.1 and 5.2) have lower absolute values relative to those in the aggregating models (Table 3 and 4). In most of the regressions, the coefficient of intangible capital is positive and not significant (regressions MI1, MI3, MI7, MI9, MI15, MI19, MI21 and MI23); in the regressions of Leather (regression MI5), Chemicals (regression MI11), Other Non-Metallic (regression MI13) and Electronic Equipment (regression MI5) industries the coefficient is negative. The coefficients for investments in intangibles are smaller than in the previous models and have the same sign as those of the aggregate models. The impact of training is positive for most regressions, except for Textiles (regression MI4), Paper (regression MI10) and Electronic Equipment (regression MI18). For R&D investment the impact is only positive for the Food industry regression (MI2); the software impact is non-negative for three industries, but the value is near zero (regressions MI18, MI22 and MI24), and the same is true for industrial property for the following regressions MI10, MI14 and MI22. Finally, the impact of other intangible capital is negative in most industries, being positive only in Leather (regression MI6), Wood (regression MI8), Paper (regression MI10) and Chemicals (regression MI12).

Table 4: Sector firm-level production function estimations for Portugal

	Regression	M1	M2	C1	C2	T1	T2	A1	A2	RE1	RE2	BS1	BS2
	(Macro) Industry	Manufacturing		Construction		Trade		Accommodation		Real Estate		Business Services	
	Labour	0.409*** (0.005)	0.409*** (0.003)	0.461*** (0.001)	0.461*** (0.005)	0.471*** (0.004)	0.468*** (0.002)	0.358*** (0.003)	0.357*** (0.002)	0.323*** (0.016)	0.320*** (0.014)	0.466*** (0.004)	0.457*** (0.002)
	Intermediate consumptions	0.613*** (0.005)	0.597*** (0.007)	0.576*** (0.009)	0.576*** (0.008)	0.685*** (0.008)	0.674*** (0.004)	0.713*** (0.000)	0.686*** (0.007)	0.687*** (0.011)	0.656*** (0.013)	0.599*** (0.002)	0.596*** (0.004)
	Tangible capital	0.027*** (0.002)	0.014*** (0.002)	0.022*** (0.003)	0.033*** (0.003)	0.042*** (0.005)	0.056*** (0.001)	0.040*** (0.001)	0.029*** (0.009)	0.039*** (0.005)	0.034*** (0.005)	0.029*** (0.007)	0.023*** (0.002)
	Intangible capital	0.005** (0.003)		0.008 (0.005)		-0.002 (0.005)		0.002 (0.002)		0.004 (0.005)		-0.000	
Investment	Training		0.003*** (0.000)		0.004*** (0.000)		0.010*** (0.001)		0.006*** (0.001)		0.011*** (0.002)		0.013*** (0.000)
Investment	R&D		-0.003*** (0.000)		-0.005** (0.002)		-0.006*** (0.002)		-0.009*** (0.001)		-0.005 (0.009)		-0.008*** (0.002)
Investment	Software		-0.005*** (0.000)		-0.001 (0.000)		0.003*** (0.001)		-0.002*** (0.001)		0.004 (0.003)		0.001*** (0.000)
Investment	Industrial Property		-0.004*** (0.001)		-0.003 (0.002)		-0.004*** (0.001)		-0.003 (0.003)		-0.022*** (0.007)		-0.015*** (0.001)
Investment	Other Intangible Capital		-0.003*** (0.002)		-0.003** (0.002)		-0.007*** (0.001)		-0.008*** (0.002)		-0.017** (0.007)		-0.012*** (0.002)
	No. firms	46 134	46 134	48 689	48 689	108 366	108 366	45 516	45 516	15 344	15 344	57 441	57 441
	No. observations	305 364	305 364	274 281	274 281	684 163	684 163	255 433	255 433	78 399	78 399	335 157	335 157
	Wald test (χ^2) on CRS	64.97***	3 808.19***	20 939.04***	587.62***	182.77***	24.02***	4 803.01***	52.49***	392.86***	114.29***	5 959.78***	13 872.45***

Note: Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. All estimates are calculated using the Levinshon and Petrin (2003) methodology. The model includes industry- and time-specific effects, as well as for firm age, dimension, exporter, and business cycle. CRS = Constant returns to scale.

Table 5.1: Manufacturing industries firm-level production function estimates for Portugal

Regression		MI1	MI2	MI3	MI4	MI5	MI6	MI7	MI8	MI9	MI10	MI11	MI12
Manufacturing Industry		Food		Textiles		Leather		Wood		Paper		Chemicals	
Labour		0.310*** (0.006)	0.311*** (0.003)	0.483*** (0.005)	0.483*** (0.003)	0.408*** (0.010)	0.407*** (0.007)	0.348*** (0.013)	0.347*** (0.010)	0.319*** (0.022)	0.317*** (0.006)	0.203*** (0.009)	0.203*** (0.029)
Intermediate consumptions		0.736*** (0.002)	0.726*** (0.005)	0.494*** (0.003)	0.474*** (0.009)	0.496*** (0.007)	0.486*** (0.007)	0.699*** (0.005)	0.693*** (0.006)	0.693*** (0.008)	0.699*** (0.008)	0.706*** (0.065)	0.701*** (0.009)
Tangible capital		0.039*** (0.006)	0.037*** (0.007)	0.015** (0.007)	0.010 (0.002)	0.016*** (0.005)	0.031*** (0.007)	0.025*** (0.002)	0.039*** (0.004)	0.040*** (0.007)	0.042*** (0.007)	0.012 (0.010)	0.006 (0.007)
Intangible capital		0.004 (0.004)		0.005 (0.006)		-0.007 (0.006)		0.001 (0.005)		0.002 (0.006)		-0.001 (0.005)	
Investment	Training		0.002** (0.001)		-0.000 (0.001)		0.003*** (0.001)		0.004*** (0.000)		-0.001 (0.001)		0.003*** (0.001)
	R&D		0.001 (0.001)		-0.005* (0.003)		-0.001 (0.003)		-0.001 (0.002)		-0.003* (0.001)		-0.003 (0.002)
	Software		-0.001 (0.001)		-0.004*** (0.000)		-0.004*** (0.001)		-0.001*** (0.000)		-0.003** (0.001)		-0.000 (0.002)
	Industrial Property		-0.005* (0.003)		-0.005*** (0.001)		-0.003 (0.002)		-0.016* (0.009)		0.000 (0.001)		-0.006*** (0.001)
	Other Intangible Capital		-0.002 (0.002)		-0.005*** (0.002)		0.000 (0.003)		0.001 (0.004)		0.001 (0.003)		0.004 (0.003)
No. firms		7 614	7 614	8 036	8 036	2 782	2 782	6 101	6 101	2 532	2 532	652	652
No. observations		51 622	51 622	49 717	49 717	17 276	17 276	39 934	39 934	17 865	17 865	4 345	4 345
Wald test (χ^2) on CRS		2 044.09***	363.60***	1 993.30***	45.97***	17.83***	313.03***	10.13***	36.41***	6.51**	67.11***	1 713.69***	10.89***

Note: Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. All estimates are calculated using the Levinshon and Petrin (2003) methodology. The model includes time-specific effects, as well as for firm age, dimension, exporter, and business cycle. CRS denotes constant returns to scale.

Table 5.2: Manufacturing industries firm-level production function estimates for Portugal

Regression		MI13	MI14	MI15	MI16	MI17	MI18	MI19	MI20	MI21	MI22	MI23	MI24
Manufacturing Industry		Other non-metallic		Metals		Electrical equipment		Machinery		Transport equipment		Other manufacturing	
Labour		0.267*** (0.005)	0.266*** (0.002)	0.388*** (0.011)	0.388*** (0.012)	0.302*** (0.018)	0.301*** (0.022)	0.382*** (0.023)	0.381*** (0.013)	0.300*** (0.019)	0.301*** (0.012)	0.416*** (0.021)	0.415*** (0.010)
Intermediate consumptions		0.680*** (0.009)	0.647*** (0.005)	0.618*** (0.002)	0.602*** (0.005)	0.623*** (0.008)	0.635*** (0.023)	0.631*** (0.007)	0.617*** (0.007)	0.668*** (0.006)	0.665*** (0.004)	0.582*** (0.008)	0.591*** (0.003)
Tangible capital		0.024*** (0.005)	0.011*** (0.002)	0.039*** (0.006)	0.025*** (0.004)	0.022*** (0.007)	0.027** (0.013)	0.023*** (0.007)	0.037*** (0.007)	0.009 (0.008)	0.009 (0.003)	0.020*** (0.005)	0.018*** (0.006)
Intangible capital		-0.003 (0.010)		0.001 (0.004)		-0.018*** (0.004)		0.005 (0.003)		0.003 (0.003)		0.001 (0.005)	
Investment	Training		0.003*** (0.001)		0.002*** (0.001)		-0.000 (0.003)		0.005*** (0.001)		0.002** (0.001)		0.007*** (0.001)
	R&D		-0.003*** (0.001)		-0.004*** (0.001)		-0.004 (0.004)		-0.003*** (0.001)		-0.005*** (0.001)		-0.005** (0.003)
	Software		-0.002 (0.001)		-0.001*** (0.000)		0.000 (0.001)		-0.002* (0.001)		0.000*** (0.000)		0.000 (0.001)
	Industrial Property		0.001 (0.002)		-0.005** (0.002)		-0.006*** (0.002)		-0.006*** (0.001)		0.001 (0.001)		-0.008*** (0.002)
	Other Intangible Capital		-0.001 (0.001)		-0.001 (0.001)		-0.013*** (0.004)		-0.003 (0.005)		-0.003 (0.003)		-0.012** (0.006)
No. firms		3 654	3 654	8 270	8 270	826	826	1 347	1 347	509	509	4 031	4 031
No. observations		25 676	25 676	56 203	56 203	5 535	5 535	9 236	9 236	3 412	3 412	24 543	24 543
Wald test (χ^2) on CRS		1 355.63***	366.04***	3.09*	0.02	160.71***	4.47**	4.76**	43.97***	8.93***	465.15***	53.13***	104.56***

Note: Standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. All estimates are calculated using the Levinshon and Petrin (2003) methodology. The model includes time-specific effects, as well as firm age, dimension, exporter, and business cycle. CRS denotes constant returns to scale.

The impact of intangibles on production is positive since companies that invest in them have higher productivity. However it is not enough just to invest in intangibles, it is necessary to invest in large quantities; in cases where companies invest little, the impact is unable to be positive, and in fact, in these cases investing in intangibles can be harmful to productivity because their impact is not immediate and can lead to a diversion of resources that could be invested in other more productive factors (remember that the distribution of investment in intangibles is highly right-skewed). This is evidenced by the case of training, which is the type of intangible receiving the most investment, and in most regressions, the impact is positive.

5. Conclusion

The objective of this work was to understand the evolution of intangible assets and their impact on the productivity of Portuguese firms in the last decade, in the context of an economic recession and a recovery period. To this end, a Cobb-Douglas production function was estimated.

The reality of Portuguese firms is characterized by the fact that most of them are SMEs, which certainly affects the investment policy in intangible assets. Few firms, about 18.0% of those in the sample, invest in intangible assets, with investment in training and software being the most common type of investment. The aggregate levels of investment in intangible assets are low, with most firms investing little or not at all, and a few of them investing large amounts. The investment trend in intangible capital over the decade has been gradually negative, with the stock capital in 2019 already less than half of what it was in 2010.

Of course, over time, this development has brought firms a decreasing positive impact of intangible capital and investment on productivity, the latter being negative in many cases or insignificant in some industries. In manufacturing, intangible capital has the strongest impact on productivity. In the construction sector, the effect is positive but insignificant. For the business services and trade sectors, the effect is null and negative, respectively. When the manufacturing sector is split, the effect is larger in the textile industry and negative in the electrical industry. As in other studies, fixed capital has a greater effect on intangibles, and the elasticities observed at the aggregate level are larger than at the sector/industry level.

References

Aghion, P.; Howitt, P. (1992) A Model of growth through creative destruction. *Econometrica,* 60(2), 323-351.

Carreira, C.; Lopes, L. (2018) Regional knowledge spillovers: a firm-based analysis of non-linear effects. *Regional Studies,* 52(7), 948-958.

Corrado, C.; Haskel, J.; Jona-Lasinio, C.; Iommi, M. (2013) Innovation and intangible investment in europe, japan, and the United States. *Oxford Review of Economic Policy,* 29(2), 261-286.

Corrado, C.; Haskel, J.; Jona-Lasinio, C.; Iommi, M. (2016) *Intangible investment in the EU and US before and since the Great Recession and its contribution to productivity growth* (No. 2016/08). EIB Working Papers.

Corrado, C.; Hulten, C.; Sichel, D. (2009) Intangible capital and US economic growth. *Review of income and wealth,* 55(3), 661-685.

Crass, D.; Peters, B. (2014) Intangible assets and firm-level productivity. *ZEW-Centre for European Economic Research Discussion Paper*, (14-120).

Criscuolo, C.; Gal, P.; Leidecker, T.; Nicoletti, G. (2021) The Human Side of Productivity: Uncovering the Role of Skills and Diversity for Firm Productivity. *OECD Productivity Working Papers*, 29.

de Carvalho, P.; Serrasqueiro, Z.; Nunes, P. M. (2013) Profitability determinants of fitness SMEs: Empirical evidence from Portugal using panel data. *Amfiteatru Economic,* 15(34), 417.

Eberhardt, M.; Helmers, C. (2010) Untested assumptions and data slicing: A critical review of firm-level production function estimators. Economics Series Working Papers 513, University of Oxford, Department of Economics.

Grossman, G.; Helpman, E. (1991) Quality ladders and product cycles. *Quarterly Journal of Economics,* 106(2), 557-586.

Kaus, W.; Slavtchev, V.; Zimmermann, M. (2020) *Intangible capital and productivity: Firm-level evidence from German manufacturing* (No. 1/2020). IWH Discussion Papers.

Levinsohn, J.; Petrin, A. (2003) Estimating production functions using inputs to control for unobservables. R*eview of Economic Studies,* 70(2), 317-341.

Marrocu, E.; Paci, R.; Pontis, M. (2012) Intangible capital and firms' productivity. *Industrial and Corporate Change*, 21(2), 377-402.

Nakamura, L. (2010) Intangible assets and national income accounting. *Review of Income and Wealth,* 56, S135-S155.

Niebel, T.; O'Mahony, M.; Saam, M. (2017) The contribution of intangible assets to sectoral productivity growth in the EU. *Review of Income and Wealth*, 63(1), 49–67.

Nunes, P.; Almeida, A. (2009) The quadratic relationship between intangible assets and growth in Portuguese SMEs. *Amfiteatru Economic Journal,* 11(25), 151-158.

Oliveira, L.; Rodrigues, L.; Craig, R. (2010) Intangible assets and value relevance: Evidence from the Portuguese stock exchange. *British Accounting Review,* 42(4), 241-252.

Olley, S.; Pakes, A. (1992) The dynamics of productivity in the telecommunications equipment industry. *Econometrica,* 64(6), 1263-1297.

Piekkola, H. (2018) Broad-based intangibles as generators of growth in Europe. *Economics of Innovation and New Technology,* 27(4), 377-400.

Romer, P. (1990) Endogenous technological change. *Journal of Political Economy,* 98(5), S71-102.

Roth, F.; Sen, A.; Rammer, C. (2022) The role of intangibles in firm-level productivity: Evidence from Germany. *Industry and Innovation*, 1-23.

Rovigatti, G.; Mollisi, V. (2018) Theory and practice of total-factor productivity estimation: The control function approach using Stata. *Stata Journal*, 18(3), 618-662.

Silva, F.; Carreira, C. (2012) Do financial constraints threat the innovation process? Evidence from Portuguese firms. *Economics of Innovation and New Technology,* 21(8), 701-736.

Solow, R. (1957) Technical Change and the Aggregate Production Function. *Review of Economics and Statistics,* 39(3), 312-320.

Turovets, Y. (2021) Intangible assets and the efficiency of manufacturing firms in the age of digitalisation: The Russian case. *Engineering Management in Production and Services*, 13(1), 7-26.

Yallwe, A.; Buscemi, A. (2014) An era of intangible assets. *Journal of Applied Finance and Banking,* 4(5), 17.

Annex

Table A.1: Industry classification

NACE	Sector/Industry	Short name
10-33	Manufacturing	Manufacturing
10-11	Food products and beverages	Food
13-14	Textiles and wearing apparel	Textiles
15	Leather and leather products	Leather
16, 31	Wood and wood products; furniture	Wood
17-18	Pulp, paper, paper products and publishing	Paper
19-21	Chemical and chemical products	Chemicals
22-23	Rubber and plastic products; other non-metallic	Other non-metallic
24-25	Basic metals and fabricated metal products	Metals
26-27	Electronic and electrical equipment	Electrical equipment
28	Machinery and equipment	Machinery
29-30	Motor vehicles, trailers and other transport equip.	Transport equipment
32-33	Other manufacturing n.e.c. and recycling	Other manufacturing
41-43	Construction	Construction
45-47	Trade	Trade
55-56	Accommodation	Accommodation
68	Real estate	Real estate
62-63, 69-82	Business services	Business services

1 2 9 0

IMPRENSA DA UNIVERSIDADE DE COIMBRA
COIMBRA UNIVERSITY PRESS